

Olga D. Anosova, Vitaliy A. Kurlin

# Geometric Data Science

Moduli spaces of real data under equivalences

February 3, 2026

Springer Nature



# Preface

*Where there is Matter, there is Geometry*

— *Johannes Kepler (1571- 1630)*

*a key figure in the 17th-century Scientific Revolution.*

This book introduces the new research area of *Geometric Data Science*, where data can represent any real objects through geometric measurements. Some of the simplest inputs of real data objects are finite and periodic sets of unordered points.

For example, a molecule can be fully described by the positions of its atoms in a 3-dimensional space. However, many descriptions are highly ambiguous, especially to a computer, which operates only with numbers. For example, a photograph is ambiguous, because any object can have an astronomically large number of photographs.

All attempts to standardise photographs, as in passports, have shifted towards more reliable biometric data. Indeed, the identification of living organisms was dramatically improved due to the discovery of a DNA structure. However, *geometric structures* remained ambiguous for many objects, including proteins and materials, which are still represented by photograph-style inputs depending on arbitrary coordinate systems.

The major obstacle to progress from trial-and-error in chemistry and biology to a justified design of materials and drugs was the absence of rigorous definitions and problem statements. Geometric Data Science fills this gap by developing foundations based on equivalences, invariants, distance metrics, and polynomial-time algorithms.

The main *geo-mapping problem* is to analytically describe moduli spaces of *geometric structures* that are classes of data objects modulo an equivalence relation. These moduli spaces are prototypes of ‘treasure maps’ containing all known objects of a certain type, as well as all not yet discovered ones. A discrete example is Mendeleev’s table of chemical elements, which was initially half-empty, but importantly guided an efficient search for new elements. A continuous example is a geographic map of the Earth, where any location is unambiguously identified by the latitude and longitude.

Geometric Data Science aims to develop universal geographic-style coordinates for all real data objects under practically important equivalences, such as rigid motion.

The first part of the book focuses on finite point sets. The most important result is a complete and continuous classification of all finite clouds of unordered points under rigid motion in any Euclidean space. The key challenge was to avoid the exponential complexity arising from permutations of the given unordered points. For a fixed dimension of the ambient Euclidean space, the times of all algorithms for the resulting invariants and distance metrics depend polynomially on the number of points.

The second part of the book advances a similar classification in the much more difficult case of periodic point sets, which model all periodic crystals at the atomic scale. The most significant result is the hierarchy of invariants from the ultra-fast to complete ones. The key challenge was to resolve the discontinuity of crystal representations that break down under almost any noise. Experimental validation on all major materials databases confirmed the *Crystal Isometry Principle*: any real periodic crystal has a unique location in a common moduli space of all periodic structures under rigid motion. The resulting moduli space contains all known and also all not yet discovered periodic crystals, and hence continuously extends Mendeleev's table to the full crystal universe.

The book was written for research students and professionals who work in mathematics and need rigorously justified and computationally efficient methods for real data, such as crystalline materials and molecules, including proteins. The pre-requisite knowledge is linear algebra, metric geometry, and calculus at the undergraduate level.

We finish by extending Johannes Kepler's quote from the 17th century to inspire a transformation from brute-force computations, which currently 'burn' our planet, to a 21st-century *Maths for Science* revolution: *Where there is Data, there is Geometry*.

**Acknowledgments.** We thank Tatiana Kurlina (University College London) for helpful comments on initial drafts, and all members of the Data Science Theory and Applications group at the Materials Innovation Factory (Liverpool) for their hard work.

Most significantly, Daniel Widdowson created the invariant-based *Crystal Geomap* to visualise in real time the crystal universe of all known materials. Yury Elkin substantially contributed by correcting past claims on computational complexities of nearest neighbour search and by parallelising computations for molecular structures. Matthew Bright produced the first geographic-style maps of 2.6+ million 2D lattices extracted from real periodic crystals. Philip Smith's software for density functions demonstrated the first impact by detecting a missed crystal, which was confused with a different one by the authors of the original work. The implementations of Alexey Gorelov, Ziqiu Jiang, and William Jeffcott exposed thousands of duplicate chains in the Protein Data Bank. Jonathan Balasingham adapted ultra-fast invariants for predicting material properties. Miloslav Torda, Jonathan McManus, Marjan Safi-Samghabadi, Surya Majumder, and Jack Gallimore validated Geometric Data Science methods on many real datasets.

We are grateful to Marjorie Senechal, Nikolai Dolbilin, Simon Billinge, Nick Kotov, Andy Cooper FRS, Sally Price FRS, Richard Catlow FRS, Graeme Day, Ram Seshadri, Shmuel Weinberger, and Yoshua Bengio for helpful support and useful discussions. The latest version is <http://kurlin.org/Geometric-Data-Science-book.pdf>.

United Kingdom,  
November 2025

*Olga D. Anosova  
Vitaliy A. Kurlin*

# Contents

<b>1</b>	<b>Introduction: from practical challenges to fundamental problems</b>	1
1.1	What questions should we ask about real data objects? . . . . .	1
1.2	Abstract and practical equivalence relations on data objects . . . . .	3
1.3	Distance metrics on invariant values and equivalence classes . . . . .	7
1.4	The geo-mapping problem for data objects under equivalences . . . . .	9
1.5	Solutions to the geo-mapping problem in the simplest cases . . . . .	13
1.6	Related areas and connections of Geometric Data Sciences . . . . .	16
1.7	The chapter plan: from easier to more challenging data . . . . .	17
	References . . . . .	19

## Part I Geometric Data Science of finite point sets

<b>2</b>	<b>Sequences of ordered points under rigid motion in Euclidean spaces</b>	23
2.1	Classical invariants and shape spaces of ordered points . . . . .	23
2.2	The geo-mapping problem for protein backbones in $\mathbb{R}^3$ . . . . .	27
2.3	Complete and bi-continuous Backbone Rigid Invariant . . . . .	29
2.4	Average invariant, diagrams, and barcodes of backbones . . . . .	32
2.5	A fast detection of duplicate chains in the Protein Data Bank . . . . .	33
	References . . . . .	35
<b>3</b>	<b>Complete and polynomial-time invariants of unordered points in <math>\mathbb{R}^n</math></b>	37
3.1	Towards complete and polynomial-time invariants for clouds . . . . .	37
3.2	The Principal Coordinates Invariant of unordered clouds in $\mathbb{R}^n$ . . . . .	39
3.3	A symmetrised metric on principally generic clouds in $\mathbb{R}^n$ . . . . .	41
3.4	A complete invariant for all clouds of unordered points in $\mathbb{R}^n$ . . . . .	43
3.5	Polynomial-time metrics for all clouds of unordered points in $\mathbb{R}^n$ . . . . .	45
	References . . . . .	47
<b>4</b>	<b>Persistence vs easier, faster, and stronger invariants of point clouds</b>	49
4.1	A dendrogram and its stable-under-noise mergegram . . . . .	49
4.2	The mergegram is stronger than the 0D persistence of a cloud . . . . .	52
4.3	Generic point clouds that are indistinguishable by persistence . . . . .	56

4.4	Geo-mapping problem for generic clouds in a metric space .....	59
4.5	Pointwise Distance Distributions of unordered points .....	62
4.6	Extending the side-side-side theorem from 3 to 4 points in $\mathbb{R}^n$ .....	64
	References .....	68
<b>5</b>	<b>Higher order distance distributions of unordered points in a metric space</b>	71
5.1	Simplexwise Distance Distributions of a cloud in a metric space .....	71
5.2	The expressiveness of Simplexwise Distance Distributions .....	75
5.3	Continuous metrics on Simplexwise Distance Distributions .....	84
5.4	Measured Simplexwise Distributions for metric-measure spaces .....	85
	References .....	88
<b>6</b>	<b>Complete and Lipschitz continuous invariants of unordered points in <math>\mathbb{R}^n</math></b>	89
6.1	Geo-mapping problem under rigid motion in $\mathbb{R}^n$ .....	89
6.2	Simplexwise Centered Distributions of a cloud in $\mathbb{R}^n$ .....	94
6.3	The Lipschitz continuous strength of a simplex in $\mathbb{R}^n$ .....	96
6.4	Algorithms for continuous metrics on complete invariants .....	97
	References .....	99

## Part II Geometric Data Science of periodic point sets

<b>7</b>	<b>One-periodic sequences in high-dimensional Euclidean spaces</b> .....	103
7.1	One-periodic sequences under various isometries in $\mathbb{R} \times \mathbb{R}^{n-1}$ .....	103
7.2	Invariants and continuous metrics for finite sequences in $\mathbb{R}^n$ .....	105
7.3	Discontinuity of a minimal period for 1-periodic sequences .....	109
7.4	Period-independent metrics for 1-periodic sequences .....	110
	References .....	114
<b>8</b>	<b>Moduli spaces of 2D lattices under isometry and rigid motion</b> .....	115
8.1	Representations of lattices by unit cells and reduced bases .....	115
8.2	Geo-mapping problem for lattices under equivalences in $\mathbb{R}^2$ .....	119
8.3	Invariants of an obtuse superbase of a 2-dimensional lattice .....	123
8.4	Complete invariants of 2D lattices under four equivalences .....	126
8.5	Inverse design and a spherical map of 2D lattices .....	129
8.6	Metrics on spaces of 2D lattices under all four equivalences .....	136
8.7	Real-valued chiral distances measure asymmetry of lattices .....	138
8.8	Bi-continuity of the map from obtuse superbases to invariants .....	144
	References .....	148
<b>9</b>	<b>Density functions of periodic sets of points in <math>\mathbb{R}^n</math> and intervals in <math>\mathbb{R}</math></b> .....	149
9.1	Periodic sets in $\mathbb{R}^n$ and ambiguity of cell-based representations .....	149
9.2	Continuity and generic completeness of density functions in $\mathbb{R}^3$ .....	152
9.3	Computing density functions via degree- $k$ Voronoi domains .....	157
9.4	A description of density functions of periodic sequences .....	159
9.5	Density functions of periodic sequences of intervals in $\mathbb{R}$ .....	165
9.6	Properties of density functions of periodic sequences of intervals .....	175

References .....	182
<b>10 Pointwise Distance Distributions of periodic point sets in <math>\mathbb{R}^n</math></b> .....	185
10.1 Pointwise Distance Distributions for lattices and $l$ -periodic sets .....	185
10.2 Asymptotic and time of Pointwise Distance Distributions .....	187
10.3 Lipschitz continuous metrics and local novelty distances .....	190
10.4 Generic completeness of PDDs for periodic point sets in $\mathbb{R}^n$ .....	193
10.5 Detecting near-duplicates in the world's largest databases .....	195
10.6 Structural novelty of crystals and navigating materials space .....	199
References .....	201
<b>11 Complete and continuous isosets of all periodic point sets in <math>\mathbb{R}^n</math></b> .....	205
11.1 Geo-mapping for periodic point sets under rigid motion in $\mathbb{R}^n$ .....	205
11.2 Isotree of local clusters of points in a periodic set in $\mathbb{R}^n$ .....	207
11.3 A complete isoset of periodic point sets under rigid motion in $\mathbb{R}^n$ .....	212
11.4 Continuous metrics on isometry classes of periodic sets in $\mathbb{R}^n$ .....	216
11.5 Algorithms to compute isosets and their approximate metrics .....	219
11.6 Comparisons of isosets with past invariants and experiments .....	222
References .....	229
<b>12 A history of developing the most significant results and open problems</b> ..	231
12.1 The most important results of Geometric Data Science by 2025 .....	231
12.2 New definitions of geometric structures and verified principles .....	233
12.3 Geometric approaches in crystallography until 2019 .....	234
12.4 Towards the Crystal Isometry Principle in 2019-2021 .....	238
12.5 Further history of Geometric Data Science in 2022 - 2026 .....	243
References .....	246
<b>Index</b> .....	251



## Chapter 1

# Introduction: from practical challenges to fundamental problems

**Abstract** This chapter discusses how practical challenges in object recognition and data comparison can be converted into formally stated mathematical problems. After introducing the necessary concepts of equivalences, invariants, and metrics, we state the general *geo-mapping problem* to continuously parametrise moduli spaces for any data under a given equivalence. With this foundation in place, further chapters examine specific types of data objects that allow recently developed solutions in this book.

### 1.1 What questions should we ask about real data objects?

The initial question that can be asked about any real data object is *what is it?* or (more formally) *how is it defined?* or (more deeply) *how can we make sense of this data?*

The first obstacle in achieving these goals is to embrace differences between real objects and their digital representations. For example, a car is a physical object that is very different from a pixel-based image of this car, which is only a matrix of integers.

The second obstacle is the ambiguity of digital representations in the sense that any real object can have many representations that look very different to a computer.

If measurements have continuous real values, the resulting space of representations is infinite. Even if we fix a finite resolution of physical measurements, all potential data values still live in a huge space. For example, all images of size  $2 \times 2$  pixels and greyscale intensities  $0, \dots, 255$  form a huge collection of  $256^4 > 4$  billion images. This *combinatorial explosion* (or the *curse of dimensionality*) has blocked many brute-force attempts to make sense of the data so a different scientific approach is called for [6].

All concepts and results in this book are introduced for very general data, such as discrete sets of points, and hence are relevant to many applied areas, e.g. point clouds in Computer Vision and Graphics. However, since our latest work is joint with chemists and biologists, our motivations and examples will include data at the atomic scale, including molecules, atomic clouds, and solid crystalline materials (periodic crystals).

Some physical objects can be exactly represented in a digital form, for example, by listing the coordinates of atoms. This atomistic representation much better describes a real molecule than pixel-based images of a car. Though we can exactly describe a molecule by the positions of its atoms, is this description unambiguous? To better understand the underlying obstacles, we split this question into more questions below.

**First question:** which objects are the *same or different*? Indeed, if we shift all atomic positions by a fixed vector, the digital representation changes, but does the underlying object remain *the same*? The question is deceptively simple, but because the definition of a crystal structure was incomplete in practice [11], this problem has attracted considerable attention, even appearing in the titles of papers [44]. The missing ingredient was the concept of an equivalence, which should accompany all newly defined objects.

**Second question:** *if different, by how much*? Indeed, all real data is uncertain at least due to measurement noise. Moreover, all atoms vibrate so their relative positions are always uncertain. This basic fact in Richard Feynman's first lecture on physics [25, Chapter 1 "Atoms in motion"] called for a continuous quantification of similarities. The resulting problem is algorithmically difficult even for macroscopic objects. Indeed, when walking or driving, our brains (but not computers) easily recognise obstacles whose visual representations change in our moving coordinate system. If a car moved or the wind slightly deformed a bush, humans can still identify them as perturbations of the original objects, while a computer program needs an exact formula for a distance.

**Third question:** *where can we find new objects*? Discovery sciences, such as molecular or materials synthesis, struggle to find or even recognise new objects in the vast chemical space. Indeed, all known molecules, say for a fixed number  $m$  of atoms, live in a common space of  $m$ -atom configurations. This space potentially contains unknown molecules, which have not yet been discovered, but where should we look for them? Humans faced similar challenges in their early exploration of our planet to discover new places to live and thrive. The slightly rephrased question *where do all real objects live?* requires us to build a geographic-style map on a space of all potential objects of a given type.

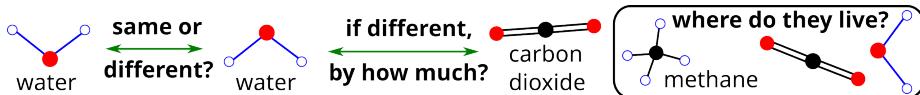
It took cartographers over two centuries (1400–1600), during the Age of Geographic Discoveries [4], to build a map of the Earth based on latitude and longitude coordinates. Scientists can build geographic-style maps of continuous spaces for other real objects.

In summary, *Geometric Data Science* (GDS) aims to mathematically formalise and answer the following questions for real data under practical equivalences, see Fig. 1.1.

**The first question:** Same or different?

**The second question:** If different, by how much?

**The third question:** Where do all (known and new) real objects live?



**Fig. 1.1** The main questions of Geometric Data Science are illustrated for molecules: H<sub>2</sub>O, CO<sub>2</sub>, CH<sub>4</sub>.

We finish this section by describing a few conventions used in the book. All spelling is British as we are based in the UK. All acronyms that are harder to guess than the UK are listed at the end of the book before the index. All environments are highlighted in the bold font and numbered according to sections, e.g. Definition 1.2.1 is followed by Example 1.2.2 in section 1.2. All figures and tables are numbered consecutively within each chapter, as Fig. 1.1. All environments have the following end symbols:

- for proofs;
- ▲ for definitions;
- ◆ for examples and remarks;
- ★ for problems and conjectures;
- for theorems, corollaries, propositions, and lemmas.

All new concepts in definitions are highlighted in the *italic* font, which is also used for emphasising keywords.  $\mathbb{R}^n$  denotes the Euclidean  $n$ -dimensional space with a fixed coordinate system of the standard orthonormal basis and origin 0. Any vector  $\mathbf{p}$  with real coordinates  $p_1, \dots, p_n$  can be positioned with the tail at  $0 \in \mathbb{R}^n$  and the head at the point  $p = (p_1, \dots, p_n) \in \mathbb{R}^n$ , i.e.  $p$  and  $\mathbf{p}$  are often used interchangeably.

## 1.2 Abstract and practical equivalence relations on data objects

This section recalls an equivalence relation and various types of invariant under a given equivalence. These concepts will help formalise the first question: *same or different?*

For any type of object, such as real numbers or all finite sets  $A$  of unordered points in  $\mathbb{R}^n$ , a binary relation describes ordered pairs  $(A, B)$  that satisfy this relation. If  $x, y$  are real numbers, one simple relation is the strict inequality  $x < y$ .

**Definition 1.2.1** (equivalence relation). A binary relation between objects of a given type is called an *equivalence* and denoted by  $\sim$  if the following axioms hold:

- (a) *reflexivity*: any object is equivalent to itself, i.e.  $A \sim A$ ;
- (b) *symmetry*: for any objects  $A, B$ , if  $A \sim B$ , then  $B \sim A$ ;
- (c) *transitivity*: for any objects  $A, B, C$ , if  $A \sim B$  and  $B \sim C$ , then  $A \sim C$ .

Any object  $A$  generates its *equivalence class*  $[A] = \{\text{all objects equivalent to } A\}$ . ▲

One widely used equivalence on real numbers is the usual equality ( $=$ ), which can be extended to vectors (points in  $\mathbb{R}^n$ ), matrices, and multisets of elements with multiplicities or weights. The axioms in Definition 1.2.1 justify a classification under a given equivalence as a splitting or *partition* into disjoint equivalence classes such that every objects belongs to exactly one class. Indeed, any such classes, say  $[A]$  and  $[C]$  share a common object  $B$ , i.e.  $A \sim B$  and  $B \sim C$ , then  $A \sim C$  and hence  $[A] = [C]$  due to the transitivity axiom. If a classification is based on a labelled dataset, often with labels produced by humans or a computer program, this finite classification is hard to extend to many other real objects, which are called “out-of-distribution”.

**Example 1.2.2** (non-equivalences). **(a)** The relation  $x < y$  (strict inequality) on real numbers is not an equivalence, because the reflexivity axiom fails:  $x < x$  is false.

**(b)** The relation  $x \leq y$  (non-strict inequality) on real numbers is not an equivalence. Though the reflexivity holds, the symmetry is also expected to hold for all real  $x, y$  but fails for any non-equal numbers:  $x \leq y$  does not imply that  $y \leq x$  for  $x \neq y$ .

**(c)** For any fixed real  $\varepsilon > 0$ , the relation  $|x - y| \leq \varepsilon$ , which can be called  $\varepsilon$ -closeness, satisfies the reflexivity and symmetry but fails the transitivity axiom. For instance, if  $x = -\varepsilon$ ,  $y = 0$ , and  $z = \varepsilon$ , then  $|x - y| = \varepsilon = |y - z|$ , but  $|x - z| = 2\varepsilon$ .  $\blacklozenge$

Example 1.2.2(c) illustrates the *sorites* paradox [29], which has been discussed since ancient times: “does a heap of sand remain a heap if grains of sand are removed one by one?” Removing one grain of sand plays the role of an  $\varepsilon$ -perturbation applied to a data object, such as a heap of millions of grains. Such a single grain can be considered similar to an outlier in data. If we are allowed to remove a point from a given set, such as an outlier, without noticing any difference, then all point sets can be made equivalent.

Similarly, if we assume that a given object remains the same (equivalent to the original one) under all perturbations up to any tiny threshold  $\varepsilon > 0$ , the transitivity axiom will imply that sufficiently many perturbations can make all objects equivalent. For instance, if we ignore slight deviations of vertices in a triangle, the resulting classification of triangles becomes trivial, consisting of a single class of all triangles. Hence, noise in real data cannot be ignored but should be properly measured.

A simple example of an equivalence (not restricted to a fixed dataset) is an equality for a specific property. For instance, two finite sets  $A, B$  can be called equivalent if they have the same size:  $|A| = |B|$ . However, many objects that share one property might differ in other properties. For finite objects, the equivalence relation defined by their size is *weak* in the sense that many substantially different objects have the same size, e.g. molecules of the same number of atoms, belong to the same equivalence class.

We will look for a *stronger* equivalence that better separates given objects. For objects that are more complicated than points in  $\mathbb{R}^n$ , such as sets of points, the identity relation is overkill in practice (too strong), because shifting all points of a cloud changes only its coordinate representation rather than physical properties.

Though many equivalence relations make sense for real objects, such as molecules or materials, one equivalence relation stands out in our world: a rigid motion usually preserves all meaningful properties and hence is the strongest relation for most applications. When comparing physical objects, the first thing people try to do is to superimpose them by rigid motion. Even if a given object, such as a human hand or a molecule, is intrinsically flexible (non-rigid), its different rigid conformations (classes under rigid motion) often have different properties and hence should be reliably distinguished.

Recall that a *basis* of  $\mathbb{R}^n$  consists of  $n$  vectors  $\mathbf{v}_1, \dots, \mathbf{v}_n$  such that any  $\mathbf{v} \in \mathbb{R}^n$  can be written as a linear combination  $\mathbf{v} = \sum_{i=1}^n t_i \mathbf{v}_i$  for some  $t_1, \dots, t_n \in \mathbb{R}$ . The basis vectors are *linearly independent* in the sense that if  $\sum_{i=1}^n t_i \mathbf{v}_i = 0$ , then  $t_1 = \dots = t_n = 0$ .

Later, we will discuss more technical concepts, such as the determinant, which has an algebraic definition. In the geometric spirit of this book, we only mention here that the *determinant*  $\det(\mathbf{v}_1, \dots, \mathbf{v}_n)$  of the  $n \times n$  matrix with columns  $\mathbf{v}_1, \dots, \mathbf{v}_n$  is the signed volume of the parallelepiped on the edge vectors  $\mathbf{v}_1, \dots, \mathbf{v}_n$ .

In particular, any vectors  $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$  form a linear basis of  $\mathbb{R}^n$  if and only if  $\det(\mathbf{v}_1, \dots, \mathbf{v}_n) \neq 0$ . Our default distance between any points  $\mathbf{a} = (a_1, \dots, a_n)$  and  $\mathbf{b} = (b_1, \dots, b_n)$  in  $\mathbb{R}^n$  is Euclidean, denoted as  $|\mathbf{a} - \mathbf{b}| = \sqrt{\sum_{i=1}^n (a_i - b_i)^2}$ .

**Example 1.2.3** (rigid motion, isometry, dilation, and homothety). **(a)** The *translation* along a fixed vector  $\mathbf{v} \in \mathbb{R}^n$  is the map  $T[\mathbf{v}] : \mathbb{R}^n \rightarrow \mathbb{R}^n$  such that  $T[\mathbf{v}](\mathbf{p}) = \mathbf{p} + \mathbf{v}$ .

**(b)** A rotation in  $\mathbb{R}^n$  is a linear map  $R[Q] : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $R[Q](\mathbf{p}) = Q\mathbf{p}$  represented by a *special orthogonal*  $n \times n$  matrix  $Q$  that has the determinant  $\det(Q) = 1$  and satisfies  $Q^T Q = I_n = Q^T Q$  is the identity matrix, where  $Q^T$  is the transpose of  $Q$ . All such matrices  $Q$  form the *special orthogonal* group  $\text{SO}(\mathbb{R}^n)$ , see [14, section 4.5].

**(c)** A *rigid motion*  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a composition of translations and rotations in  $\mathbb{R}^n$  and can be written as  $f(\mathbf{p}) = Q\mathbf{p} + \mathbf{v}$  for any  $\mathbf{p} \in \mathbb{R}^n$ , a fixed  $\mathbf{v} \in \mathbb{R}^n$  and  $Q \in \text{SO}(\mathbb{R}^n)$ . Any sets  $A, B \subset \mathbb{R}^n$  that are related by rigid motion are called *rigidly equivalent* (denoted by  $A \cong B$ ). All rigid motions in  $\mathbb{R}^n$  form the *Special Euclidean* group  $\text{SE}(\mathbb{R}^n)$ .

**(d)** The *mirror reflection* relative to an  $(n-1)$ -dimensional hyperspace  $H \subset \mathbb{R}^n$  with a normal vector  $\mathbf{v}$  is the map defined by  $f[H](p) = p$ , i.e. any point  $p \in H$  is fixed by  $f[H]$ , and  $f[H](\mathbf{u}) = -\mathbf{u}$  for any vector  $\mathbf{u}$  parallel to  $\mathbf{v}$ .

**(e)** A *Euclidean isometry* is any map  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  preserving Euclidean distance, i.e.  $|f(\mathbf{a}) - f(\mathbf{b})| = |\mathbf{a} - \mathbf{b}|$  for any vectors  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ . Alternatively, any *Euclidean isometry* is a composition of a rigid motion and a mirror reflection, and can be written as  $f(\mathbf{a}) = Q\mathbf{a} + \mathbf{v}$  for any  $\mathbf{a} \in \mathbb{R}^n$ , a fixed vector  $\mathbf{v} \in \mathbb{R}^n$ , and an *orthogonal* matrix  $Q$  satisfying  $Q^T Q = I_n = Q^T Q$ . Any subsets  $A, B \subset \mathbb{R}^n$  are related by isometry are called *isometric* (denoted by  $A \simeq B$ ). All isometries in  $\mathbb{R}^n$  form the *Euclidean* group  $\text{E}(\mathbb{R}^n)$ .

**(f)** For a fixed factor  $s > 0$ , the *uniform scaling* is the map  $u : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $u(\mathbf{a}) = s\mathbf{a}$  for any vector  $\mathbf{a} \in \mathbb{R}^n$ . A *dilation* is a composition of a rigid motion and a uniform scaling. A *homothety* is a composition of an isometry and a uniform scaling.  $\blacklozenge$

Any rigid motion  $f$  preserves *orientation* of  $\mathbb{R}^n$ , which can be defined as the sign of the determinant of the  $n \times n$  matrix consisting of the columns  $f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)$ , where  $\mathbf{v}_1, \dots, \mathbf{v}_n$  is a basis of  $\mathbb{R}^n$ . A mirror reflection, for example, changing the sign of the first coordinate (the reflection relative to the hyperspace  $a_1 = 0$ ) is not a rigid motion, because the orientation is changed. Hence, compositions or rigid motion with mirror reflections form a slightly wider collection of equivalences, which do not distinguish mirror images. The identities  $Q^T Q = I_n = Q^T Q$  imply that  $\det(Q) = \pm 1$ . All such orthogonal matrices form the *orthogonal* group  $O(\mathbb{R}^n)$ . We avoid the notation  $O(n)$ , which will be later used to denote a linear-time complexity of algorithms.

**Definition 1.2.4** (weaker vs stronger equivalences). For a fixed collection of objects, one equivalence relation  $\sim_1$  is (non-strictly) *weaker* than another  $\sim_2$  (then  $\sim_2$  is called

*stronger* than  $\sim_1$ ) if any objects equivalent under the stronger relation  $\sim_2$  are equivalent under the weaker relation, i.e.  $A \sim_2 B$  always implies that  $A \sim_1 B$ .  $\blacktriangle$

If one equivalence  $\sim_1$  is weaker than  $\sim_2$ , then the stronger equivalence  $\sim_2$  refines the partition into equivalence classes defined by the weaker equivalence  $\sim_1$ .

Rigid motions, isometries, dilations, and homotheties define equivalence relations in the sense of Definition 1.2.1. Among these four equivalences, rigid motion is the strongest ( $\cong$ ). Isometry ( $\simeq$ ) is slightly weaker because any pair of mirror images is in the same isometry class, not necessarily in the same class under rigid motion.

Dilation is weaker than rigid motion because all uniformly scaled objects belong to the same class. Homothety is the weakest of the four so that any objects related by isometry or dilation are homothetic. A substantially weaker equivalence is defined by bijection, which is a 1-1 map between all points of two objects.

After an equivalence is fixed, the next challenge is to classify all given objects under this equivalence. Such a classification should answer the first question (*same or different?*) by a practical algorithm that determines whether given objects are equivalent or not. A mathematically justified approach to any classification is to develop invariant descriptors that can reliably distinguish objects under a given equivalence, as defined below. Invariant values can be numbers, vectors, matrices, or more complicated objects in a metric space that should still be easier to compare than the original ones.

**Definition 1.2.5** (invariants and complete invariants). **(a)** Fix an equivalence on some objects. An *invariant*  $I$  is a function that takes the same value on all equivalent objects, i.e.  $A \sim B$  implies that  $I(A) = I(B)$ . Alternatively, if  $I(A) \neq I(B)$ , then  $A \not\sim B$ . In other words,  $I$  is a descriptor with **no false negatives** defined as pairs  $A, B$  that represent equivalent objects  $A \sim B$  but have different values of this descriptor.

**(b)** An invariant  $I$  is called *complete* if  $I$  distinguishes all non-equivalent objects, i.e. if  $A \not\sim B$ , then  $I(A) \neq I(B)$ . Alternatively, if  $I(A) = I(B)$ , then  $A \sim B$ , i.e.  $I$  takes the same value only on equivalent objects. In other words,  $I$  has **no false positives** defined as pairs of non-equivalent  $A \not\sim B$  that are indistinguishable by  $I$ , i.e.  $I(A) = I(B)$ .  $\blacktriangle$

A constant function  $I$  taking the same value on all objects satisfies Definition 1.2.5 but does not distinguish any objects. We will always assume that an invariant is not the same for all equivalence classes. Then the implication  $[I(A) \neq I(B)] \Rightarrow [A \not\sim B]$  allows us to distinguish complicated objects by using simpler invariants.

**Example 1.2.6** (invariants vs non-invariants). **(a)** A simple invariant of a finite set  $A$  under bijection or any stronger equivalence is the size of  $A$ , which we denote by  $|A|$ .

**(b)** Tom Leinster's *magnitude* of a finite metric space  $A$  extends the size  $|A|$  to the real-valued invariant of  $A$  under isometry [36, section 6.4], defined as the sum all elements in the inverse of the matrix consisting of  $e^{-d(p,q)}$  for all points  $p, q \in A$ .

**(c)** For any finite set  $A \subset \mathbb{R}^n$ , the *centre of mass*  $\bar{A} = \frac{1}{|A|} \sum_{p \in A} p$  is not invariant of  $A$  even under translations and rotations in  $\mathbb{R}^n$ , and hence under all weaker equivalences, including rigid motion, isometry, and bijection.

**(d)** For sets of two ordered points  $x, y \in \mathbb{R}$ , the difference  $x - y$  is invariant under rigid motion (only translations in  $\mathbb{R}$ ), but not under isometry that can swap the order of  $x, y$ . The Euclidean distance  $|x - y|$  is a complete invariant of two ordered point sets under isometry in  $\mathbb{R}$ , but not under rigid motion. Indeed, the ordered pairs  $(0, 1)$  and  $(1, 0)$  of numbers are not rigidly equivalent but have the same inter-point distance 1.  $\blacklozenge$

A complete invariant  $I$  fully answers the first main question (*same or different?*) by checking if  $I(A) = I(B)$ , which is equivalent to  $A \sim B$  by Definition 1.2.5.

Any function  $h$  generates its equivalence  $\sim_h$  on the domain where  $h$  is defined:  $A \sim_h B$  if and only if  $h(A) = h(B)$ . Then  $h$  is a complete invariant under its equivalence  $\sim_h$ .

For a fixed collection of objects, invariants can be compared by strength similar to equivalence relations in Definition 1.2.4. For ordered pairs  $(x, y)$ , the difference  $x - y$  is a stronger invariant than the distance  $|x - y|$ . For a fixed equivalence relation, a complete invariant is the strongest one among all invariants under this equivalence.

### 1.3 Distance metrics on invariant values and equivalence classes

To rephrase the second main question (*if different, by how much?*) in mathematical terms, this section introduces a distance metric between arbitrary objects, which can be equivalence classes or values of an invariant under a given equivalence.

**Definition 1.3.1** (metrics and pseudo-metrics). **(a)** A real-valued function  $d$  on pairs of objects under an equivalence relation  $\sim$  is a *metric* if these axioms hold:

- (1) *coincidence*:  $d(A, B) = 0$  if and only if  $A \sim B$ ;
- (2) *symmetry*:  $d(A, B) = d(B, A)$  for any objects  $A, B$ ;
- (3) *triangle inequality*:  $d(A, B) + d(B, C) \geq d(A, C)$  for any objects  $A, B, C$ .

**(b)** If axiom (1) is replaced with the weaker version (1')  $d(A, A) = 0$  for any  $A$ , then non-equivalent objects  $A \not\sim B$  can have  $d(A, B) = 0$ , and  $d$  is called a *pseudo-metric*.  $\blacktriangle$

The axioms in Definition 1.3.1(a) imply the non-negativity of a metric as follows:  $2d(A, B) = d(A, B) + d(B, A) \geq d(A, A) = 0$ . The word “metric” is often used in applications and evaluation functions that depend on a single object. We emphasise that all metrics measure a distance between two objects. The concept of a distance becomes more general if some of the three metric axioms are weakened [20].

**Definition 1.3.2** (metric spaces and clouds). **(a)** Any set  $M$  of objects with a metric  $d : M \times M \rightarrow \mathbb{R}$  is called a *metric space*.

**(b)** A *cloud* is any finite set of unordered points in a metric space. A *Euclidean cloud* is any finite set  $A \subset \mathbb{R}^n$  of unordered points with the Euclidean distance.  $\blacktriangle$

We will usually consider metrics on invariant values rather than on original objects. If  $I$  is a complete invariant under a given equivalence, then any metric on invariant values is the metric on equivalence classes of original objects so that the condition  $d(I(A), I(B)) = 0$  is equivalent to  $A \sim B$ . However, if an invariant  $I$  is incomplete,

then  $d(I(A), I(B)) = 0$  only guarantees that  $I(A) = I(B)$ , not necessarily  $A \sim B$ . Then any metric on (values of) an incomplete invariant defines only a pseudo-metric on equivalence classes of original objects in the sense of Definition 1.3.1(b). Example 1.3.3 introduces well-known metrics on vectors and arbitrary subsets in a metric space.

**Example 1.3.3** (Minkowski metrics, Hausdorff and bottleneck distances). **(a)** Fix a real parameter  $q \in [1, +\infty)$ . For any points  $\mathbf{a} = (a_1, \dots, a_n)$  and  $\mathbf{b} = (b_1, \dots, b_n)$  in  $\mathbb{R}^n$ , the *Minkowski metric* is  $L_q(a, b) = \left( \sum_{i=1}^n |a_i - b_i|^q \right)^{1/q}$ . In the limit case  $q = +\infty$ , the metric is defined as  $L_\infty(a, b) = \max_{i=1, \dots, n} |a_i - b_i|$ , also called the *Chebyshev metric*.

**(b)** Let  $A, B$  be subsets of a space  $X$  with a metric  $d$ . The distance from  $a \in A$  to  $B$  is  $d(a, B) = \inf_{b \in B} d_X(a, b)$ . The *directed* distance is  $d_H(A, B) = \sup_{a \in A} d(a, B)$ . The *Hausdorff distance* is  $HD(A, B) = \max\{d_H(A, B), d_H(B, A)\}$ . The *bottleneck distance*  $BD(A, B) = \inf_{g: A \rightarrow B} \sup_{p \in A} d(g(p), p)$  is minimised for all bijections  $g : A \rightarrow B$ .  $\blacklozenge$

If there are no bijections  $A \rightarrow B$ , one can set  $BD(A, B) = +\infty$ , so  $BD$  is a well-defined metric only on subsets that allow bijections. In Example 1.3.3(a), the parameters  $q = 1, 2, +\infty$  define the metrics that are also called Manhattan (sum metric), Euclidean, and Chebyshev (max metric), respectively. We will often consider Minkowski metrics for all parameters  $q \in [1, +\infty]$ , including the limit case  $q = +\infty$ .

For any  $\varepsilon \geq 0$  and a subset  $C$  of a metric space  $M$ , the  $\varepsilon$ -offset of  $C$  consists of all points  $q \in M$  at a maximum distance  $\varepsilon$  from  $C$ , i.e.  $d(q, C) \leq \varepsilon$ .

In other words, the  $\varepsilon$ -offset of  $C$  is the union of closed balls with the radius  $\varepsilon$  and centres at all points  $p \in C$ . Then the Hausdorff distance  $HD(A, B)$  can be defined as the minimal  $\varepsilon \geq 0$  such that the  $\varepsilon$ -offset of  $A$  covers  $B$  and the  $\varepsilon$ -offset of  $B$  covers  $A$ .

The Hausdorff distance  $HD$  is illustrated in terms of  $\varepsilon$ -offsets in Figure 1.2, where a single ball around a blue or green point can cover a cluster of several points from a different subset of another colour. The bottleneck distance  $BD$  is stricter by measuring a minimum required deviation for a bijective matching of points, as in Figure 1.2 (right).



**Fig. 1.2** **Left:** in the Euclidean line  $\mathbb{R}$ , the clouds  $A$  of 4 green points and  $B$  of 4 blue points have a small Hausdorff distance  $HD$ . **Right:** the same clouds  $A, B \subset \mathbb{R}$  have a large bottleneck distance  $BD$  based on a bijection  $g : A \rightarrow B$  (shown by red arrows), which minimises the maximum deviation of points, see Example 1.3.3(b).

Definition 1.3.1 allows a *discrete* metric that takes a constant value on all non-equivalent objects, e.g.  $d(A, B) = 1$  for all  $A \not\sim B$ , and  $d(A, B) = 0$  for all  $A \sim B$ .

However, this discrete metric is purely theoretic because all real objects slightly differ due to noise, so this metric would almost always have the same value 1. Definition 1.3.4 formalises the practically useful continuity under perturbations.

For simplicity, we will consider a collection  $X$  of objects that allow bijections between each other. For finite sets, this restriction means that all clouds from  $X$  have the same size. Bijections always exist between any infinite *discrete subsets* of  $\mathbb{R}^n$  (that have a positive minimum inter-point distances), because such subsets are countable.

**Definition 1.3.4** (Lipschitz continuity). Let  $X$  be a space of objects with a distance metric  $d_X$ . Let  $I : X \rightarrow Y$  be a function to a space  $Y$  with a metric  $d_Y$ , e.g. an invariant under a given equivalence. Then  $I$  is *Lipschitz continuous* with a Lipschitz constant  $\lambda$  if, for any  $\varepsilon \geq 0$ , the  $\varepsilon$ -closeness  $d_X(A, B) \leq \varepsilon$  implies that  $d_Y(I(A), I(B)) \leq \lambda\varepsilon$ . ▲

For any discrete subsets in a metric space, the condition  $\text{BD}(A, B) \leq \varepsilon$  can be replaced with the following:  $B$  is obtained from  $A$  by perturbing every point of  $A$  within its  $\varepsilon$ -neighbourhood. For molecules and materials consisting of atoms, such  $\varepsilon$ -perturbations and hence the bottleneck distance are motivated by thermal vibrations and experimental noise under which atoms can slightly change their positions but cannot disappear. For other applications, e.g. to point clouds in Computer Vision, Definition 1.3.4 can be made stricter by requiring the Lipschitz continuity under less restrictive perturbations of input data in the Hausdorff distance.

The classical continuity in terms of  $\varepsilon, \delta$  is much weaker than the more practical Definition 1.3.4. For instance, the function  $y = \frac{1}{x}$  is continuous for all  $x \neq 0$  but has no Lipschitz constant because  $\frac{1}{x} \rightarrow \infty$  as  $x \rightarrow 0$ . Hence, almost any function in practice can be called continuous in the weak sense of  $\varepsilon, \delta$  away from singular points.

The Lipschitz continuity brings a physical meaning due to an explicit constant  $\lambda$ . For example, if any atom is perturbed up to  $\varepsilon$ , any inter-atomic distance changes up to  $2\varepsilon$  due to the triangle inequality of Euclidean distance. Hence, inter-atomic distances in physically meaningful units have Lipschitz constant  $\lambda = 2$ .

## 1.4 The geo-mapping problem for data objects under equivalences

This section introduces auxiliary concepts of a metric moduli space and a computational complexity before stating the main geo-mapping problem. Several non-trivial cases of this problem will be solved in later chapters for point clouds and periodic point sets.

**Definition 1.4.1** (metric moduli spaces). Let  $X$  be a collection of objects with the bottleneck distance  $\text{BD}$  and an equivalence relation  $\sim$ . The *metric moduli space* is the set  $Y = X/\sim$  of all equivalence classes  $[A]$  for  $A \in X$ , equipped with a distance metric  $d$  satisfying all metric axioms in Definition 1.3.1(b) so that the *class map*  $X \rightarrow Y$  defined by  $A \mapsto [A]$  is Lipschitz continuous in the sense of Definition 1.3.4. If an equivalence is defined by an action of a group  $G$  on  $X$ , the moduli space is also denoted by  $X/G$ . ▲

Moduli spaces are also called quotient spaces or spaces of orbits. The adjective *metric* means that a moduli space is equipped with a metric satisfying Definition 1.3.1(a).

Algebraic geometry studied moduli spaces in more general settings [48], usually for varieties defined by polynomial equations and considered under actions of linear groups, not involving permutations and metrics. However, the recently emerged area of Metric Algebraic Geometry [9] started to explore metrics on moduli spaces. Geometric Data Science (GDS) goes further by requiring polynomial-time algorithms for complete invariants with continuous metrics, and adds the realisability and Euclidean embeddability to parametrise moduli spaces similar to geographic maps of Earth.

**Definition 1.4.2** (cloud spaces  $\text{CIS}(\mathbb{R}^n; m)$  and  $\text{CRS}(\mathbb{R}^n; m)$ ). **(a)** For the collection  $S(\mathbb{R}^n; m)$  of all  $m$ -point sequences  $p_1, \dots, p_m \in \mathbb{R}^n$ , the moduli space  $S(\mathbb{R}^n; m)/\text{SE}(\mathbb{R}^n)$  was previously called a *shape space*  $\Sigma_n^m$  [32]. Under the extra action of the permutation group, the moduli space  $S(\mathbb{R}^n; m)/(\text{SE}(\mathbb{R}^n) \times S_m) = \Sigma_n^m/S_m$  will be called the *Cloud Rigid Space* and denoted by  $\text{CRS}(\mathbb{R}^n; m)$ .

**(b)** Under isometry, not distinguishing mirror images, the space  $S(\mathbb{R}^n; m)/(\text{E}(\mathbb{R}^n) \times S_m)$  will be called the *Cloud Isometry Space* and denoted by  $\text{CIS}(\mathbb{R}^n; m)$ .  $\blacktriangle$

Definition 1.4.2 is motivated by the fact that points are unordered (unlabelled) in most practical scenarios. Then the equivalence relation is defined by the actions of  $\text{E}(\mathbb{R}^n)$  and the permutation group  $S_m$ . Though atoms in molecules are labelled by chemical elements and sometimes electric charges, many simple molecules such as benzene  $\text{C}_6\text{H}_6$  consist of many indistinguishable atoms, whose permutation group  $S_6 \times S_6$  consists of  $(6!)^2$  (more than half a million) permutations. These challenges motivated geo-mapping problems for finite and periodic sets of unordered points.

For all 2-point clouds (sets of two distinct unordered points) in  $\mathbb{R}^n$ , their inter-point distance  $d$  is a complete invariant under isometry. In this case, the Cloud Isometry Space  $\text{CIS}(\mathbb{R}^n; 2)$  is the interval  $(0, +\infty)$  parametrised by the distance  $d$ .

For any Euclidean cloud  $A$ , its *input size*  $|A|$  is the number of points, because the required computer memory is proportional to  $|A|$ , for a fixed dimension  $n$ . If  $A$  is a subset of a metric space, then  $A$  can be given by a distance matrix of size  $|A|^2$ .

If  $A \subset \mathbb{R}^n$  is a periodic set of points, its input size  $|A|$  can be defined as the number of points in a minimal cell whose periodic translations define the infinite set  $A$ .

All computational complexities will be considered in the Random Access Memory (RAM) model, where any numerical value can be accessed in a constant time.

**Definition 1.4.3** (the big  $O$  notation for computational complexities). Let an algorithm have an input size  $m$ . For a function  $f(m)$ , an algorithm has the *computational complexity*  $O(f(m))$  if the total number of required operations, including additions, multiplications, and evaluations of elementary functions has an upper bound  $cf(m)$  for a constant  $c$  and all sufficiently large  $m$ . If  $f$  is a linear or polynomial function of  $m$ , the resulting algorithms have a *linear* or *polynomial* time, respectively.  $\blacktriangle$

For ordered points  $p_1, \dots, p_n \in \mathbb{R}^n$ , their distance matrix can be computed in time  $O(n^2)$ , because we need only  $\frac{n(n-1)}{2}$  distances  $|p_i - p_j|$  for  $1 \leq i < j \leq n$ .

**Definition 1.4.4** (homeomorphism and embedding). A *homeomorphism*  $f$  is a bi-continuous bijection, i.e. both  $f$  and  $f^{-1}$  are continuous. Then an *embedding*  $f : X \rightarrow Y$  is a homeomorphism on image, i.e.  $X \rightarrow f(X)$  is a homeomorphism.  $\blacktriangle$

Since all our spaces have metrics, the goal is to guarantee the Lipschitz continuity in the sense of Definition 1.3.4 so that all embeddings have Lipschitz constants.

While we state Problem 1.4.5 in full generality below, the reader can keep in mind the partial case of 3-point clouds (triangles) under Euclidean isometry in the plane  $\mathbb{R}^2$ . Our typical distance  $d_X$  on a space  $X$  is the bottleneck distance  $\text{BD}$  from Example 1.3.3(b), which can be infinite for periodic sets, so  $d_X$  may not be a metric satisfying Definition 1.3.1(a). Even for finite sets, the bottleneck distance  $\text{BD}$  is impractical to compute via exponentially many permutations. Problem 1.4.5 aims to replace a complicated moduli space  $X/\sim$  with a simple metric space  $M$  of *geocodes*, as formalised below.

**Problem 1.4.5** (Geo-Mapping Problem). For any space  $X$  of objects with a distance  $d_X$  and an equivalence relation  $\sim$ , design a *geocode* defined as an invariant  $I : X \rightarrow M$  with values in a metric space  $M$  satisfying the following conditions.

- (a) *Completeness*: objects  $A, B \in X$  are equivalent ( $A \sim B$ ) if and only if  $I(A) = I(B)$ .
- (b) *Reconstruction*: any object  $A \in X$  can be reconstructed from its invariant value  $I(A) \in M$ , uniquely under the given equivalence.
- (c) *Metric*: there is a metric  $d_M$  in the *invariant space*  $I\{X\} = \{I(A) \mid A \in X\} \subset M$ , satisfying all metric axioms in Definition 1.3.1(a).
- (d) *Continuity*:  $I$  is Lipschitz continuous in the sense of Definition 1.3.4: there is a constant  $\lambda > 0$  such that, for any  $\varepsilon \geq 0$ , if  $d_X(A, B) \leq \varepsilon$ , then  $d_M(I(A), I(B)) \leq \lambda\varepsilon$ .
- (e) *Inverse continuity*: there is a constant  $\mu > 0$  such that, for any  $\delta \geq 0$ , if  $d_M(I(A), I(B)) \leq \delta$ , there is an equivalence  $f$  satisfying  $d_X(f(A), B) \leq \mu\delta$ .
- (f) *Realisability*: the invariant space  $I\{X\} = \{I(A) \mid A \in X\}$  can be parametrised so that we can generate any value  $I(A) \in I\{X\}$  realisable by some object  $A \in X$ .
- (g) *Euclidean embedding*: the invariant space  $I\{X\} = \{I(A) \mid A \in X\}$  with the metric  $d_M$  allows a bi-Lipschitz embedding into a Euclidean space  $\mathbb{R}^N$  for an integer  $N$ .
- (h) *Computability*: fix a metric space containing all objects of  $X$ , then the invariant  $I(A)$  in (a), a reconstruction of  $A$  from  $I(A)$  in (b), the metric  $d_M(I(A), I(B))$  in (c), an equivalence  $f$  in (e), the generation of a new value in  $I\{X\}$ , and an embedding  $I\{X\} \subset \mathbb{R}^N$  can be algorithmically computed in polynomial times of the input size.  $\star$

Completeness in 1.4.5(a) formalises the first main question (*same or different?*) in Geometric Data Science by requiring that a geocode  $I$  is a complete invariant code unambiguously representing any given object. Then  $I$  defines a bijection between the moduli space  $X/\sim$  and the invariant space  $I\{X\} \subset M$ . Completeness alone is impractical because one can define  $I$  as the entire collection of images  $I(A) = \{f(A) \text{ for all equivalences } f\}$ , which is infinite for most equivalence relations.

The reconstruction in 1.4.5(b) is stronger than the completeness, because a complete invariant can be abstract or too complicated without an algorithmic reconstruction. For

instance, a human fingerprint and a genetic code are practically used for identifying humans but are insufficient (yet) to grow a genetic replica of a living person.

The metric requirements in 1.4.5(c) are justified by recently designed distances  $d$  on point clouds [43], which guarantee pre-determined outputs of several clustering algorithms, such as  $k$ -means and DBSCAN, if a distance  $d$  between points is allowed to fail the triangle inequality, even with any small additive error.

The Lipschitz continuity in 1.4.5(d) in the bottleneck distance is motivated for atomic-scale objects by the fact that atoms vibrate around their average positions.

The inverse continuity in 1.4.5(e) allows us to deform a geocode  $I(A)$  and continuously trace the evolution of reconstructed objects  $A$ . Exact values of Lipschitz constants  $\lambda, \mu$  are less important than their existence, because one can always scale down a metric  $d$  to make  $\lambda$  in 1.4.5(d) smaller, then the constant  $\mu$  in 1.4.5(e) will be larger.

Conditions 1.4.5(c,d,e) formalise the second main question (*if different, by how much?*) by requiring that a geocode  $I$  is a bi-continuous invariant.

For most objects in this book, the initial metric  $d_X$  will be the bottleneck distance BD. However, the continuity in conditions 1.4.5(d,e) for BD might be unrealistic for some infinite objects, such as periodic lattices. In this case, an initial space  $X$  will consist of finite inputs with bottleneck-type metrics, e.g. lattice bases under isometry.

The realisability in 1.4.5(f) justifies the name *geocode* as an analogue of geographic coordinates and requires an explicit description of all realisable values in the invariant space  $I\{X\}$  similar to all hospitable places on Earth. This realisability in 1.4.5(f) is stated in purely mathematical terms and can be extended for practical applications by additionally requiring that  $I(A)$  is realised by a physical object  $A$ . For instance, a distance  $d$  between atoms cannot be any positive number, so 2-atom molecules have this distance  $d$  in a small range within the full moduli space  $\text{CIS}(\mathbb{R}^3; 2) = (0, +\infty)$ .

The Euclidean embeddability in 1.4.5(f) converts  $I(A)$  into a vector in some in  $\mathbb{R}^N$  with usual Euclidean distance, which can be used as an input of machine learning algorithms. However, the space exploration, such as a deformation or sampling of invariant values, should be performed in the invariant space  $I\{X\}$ , because the complement  $\mathbb{R}^N \setminus I\{X\}$  consists of artificial values that are not realisable by any objects.

The polynomial-time computability in 1.4.5(h) glues all previous conditions and makes Geo-Mapping Problem 1.4.5 notoriously hard even for finite sets of  $m \geq 4$  unordered points under isometry in  $\mathbb{R}^2$ , which were classified only into discrete types such as squares and parallelograms, through they live in a continuous 5D space.

A full solution to Problem 1.4.5 enables a continuous exploration of a complicated moduli space  $X/\sim$  by directly sampling geocodes in the invariant space  $I\{X\}$ .

Machine learning often relies on *latent* spaces of descriptor values, which can be ambiguous due to non-invariance or incompleteness, or discontinuity under noise.

For any real data under a given equivalence, Geometric Data Science aims to replace all latent spaces with *invariant spaces*, which should be parametrised by fast geocodes.

In the geographic analogy, a example geocode of any position on Earth (considered as a round sphere) consists of the latitude and longitude coordinates in the realisable ranges  $[-90^\circ, 90^\circ]$  and  $(-180^\circ, +180^\circ]$ . More exactly, the (interior of the) rectangle  $R = [-90^\circ, 90^\circ] \times [-180^\circ, +180^\circ]$  continuously maps to sphere  $S^2$ . This parametrisation assumes that the horizontal edges of  $R$  (all geocodes with a latitude  $+90^\circ$  or  $-90^\circ$ ) map to the north and south pole, respectively. We should also glue the vertical edges of  $R$  (all geocodes with longitudes  $\pm 180^\circ$  and a fixed latitude) to a single meridian of  $S^2$ .

Apart from these boundary identifications, geocodes have real values in known ranges and have enabled navigation on Earth. Indeed, the shortest way from the US to Japan is to cross the International Date Line over the Pacific Ocean, where the longitude changes from  $-180^\circ$  to  $+180^\circ$ . Hence, complete invariants of data objects become much more valuable with a continuous metric to find shortest paths in a moduli space.

The vision of Geometric Data Science is to develop such geographic-style maps (briefly, *geomaps*) for moduli spaces of all real objects under practical equivalences.

These geomaps have analytically defined invariant coordinates and substantially differ from outputs of dimensionality reduction algorithms for the following reasons.

Firstly, many such algorithms are stochastic in the sense that their outputs for the same input data can differ on runs with random seeds or on different machines.

Secondly, even if a dimensionality reduction is deterministic, such as Principal Component Analysis, the underlying algorithm is data-driven in the sense that adding new data changes the output projection of all data. Moreover, the coordinates of the resulting projections are so complicated that it is impractical to write them down.

Thirdly, any dimensionality reduction as a function  $h : \mathbb{R}^m \rightarrow \mathbb{R}^n$  for  $m > n \geq 1$  is either discontinuous, i.e. makes close points distant, or collapses an unbounded region of  $\mathbb{R}^m$  to a single point, i.e. loses an infinite amount of data [35]. Hence, dimensionality reductions can produce nice pictures, but a justified analysis of similarities and differences should use invariants and distances in the original high-dimensional space.

When we choose 2 or 3 invariants for a low-dimensional projection of a geomap, we know all other skipped invariants and hence can expand any cluster or hot spot from the first projection in other coordinates. Most importantly, adding new data to geomaps keeps the locations of all past data similar to mapping new places on a geographic map, because the invariant coordinates are defined in a data-independent way.

We considered the name *Metric Data Science* since it is similar to *Metric Algebraic Geometry* [9]. The progress beyond metrics towards geographic-style maps [10] of moduli spaces motivated the extra prefix in the name of *Geometric Data Science*.

## 1.5 Solutions to the geo-mapping problem in the simplest cases

This section discusses Examples 1.5.1-1.5.4, which solve Problem 1.4.5 for finite sets of unordered points under rigid motion and isometry in  $\mathbb{R}^n$  for the simplest known cases of dimension  $n = 1$ , up to  $m = 3$  points in  $\mathbb{R}^n$ , and for cyclic polygons in  $\mathbb{R}^2$ .

**Example 1.5.1** (geocodes of finite point sets in  $\mathbb{R}$ ). In dimension  $n = 1$ , any finite set  $A \subset \mathbb{R}$  consists of naturally ordered points  $p_1 < p_2 < \dots < p_m$ .

(a) Since any rigid motion in  $\mathbb{R}$  is a translation, the first point  $p_1$  can be fixed at the origin  $0 \in \mathbb{R}$ . Then the sequence  $p_1 < \dots < p_m$  is uniquely determined by the geocode  $I(A) = (d_1, \dots, d_{m-1})$  of the  $m-1$  distances  $d_i = p_{i+1} - p_i$ , where the only realisability condition is  $d_i > 0$ ,  $i = 1, \dots, m-1$ . The Cloud Rigid Space is  $\text{CRS}(\mathbb{R}; m) = \mathbb{R}_+^{m-1}$ .

(b) Any isometry in  $\mathbb{R}$  is a translation or its composition with the reflection  $x \mapsto -x$ , which reverses the order of all points of  $A$  and the order of the distances  $d_i$ , so that  $(d_1, \dots, d_{m-1}) \mapsto (d_{m-1}, \dots, d_1)$ . Under isometry in  $\mathbb{R}$ , the geocode  $I(A)$  is the unordered pair of these distance vectors in  $\mathbb{R}_+^{m-1}$ . The Cloud Isometry Space is  $\text{CIS}(\mathbb{R}; m) = \mathbb{R}_+^{m-1}/\sim$ , where the equivalence relation  $\sim$  reverses the order of all coordinates. If  $m = 2$ , then  $\text{CIS}(\mathbb{R}; 2) = (0, +\infty) = \text{CRS}(\mathbb{R}; 2)$ . If  $m = 3$ , then  $\text{CIS}(\mathbb{R}; 2) = \{(x, y) \in \mathbb{R}^2 \mid 0 < x < y \text{ for } x = \min\{d_1, d_2\}, y = \max\{d_1, d_2\}\}$ .  $\blacklozenge$

**Example 1.5.2** (geocodes for  $m = 2$  points in  $\mathbb{R}^n$ ). (a) For pairs of unordered points  $p, q$  in  $\mathbb{R}^n$ , one complete invariant under isometry is the inter-point distance  $d = |p - q|$ , because we can fix  $p$  at the origin  $0 \in \mathbb{R}^n$  by translation and then apply rotation from  $O(\mathbb{R}^n)$  to the point  $q$  at the distance  $d$  in the positive 1st coordinate axis of  $\mathbb{R}^n$ .

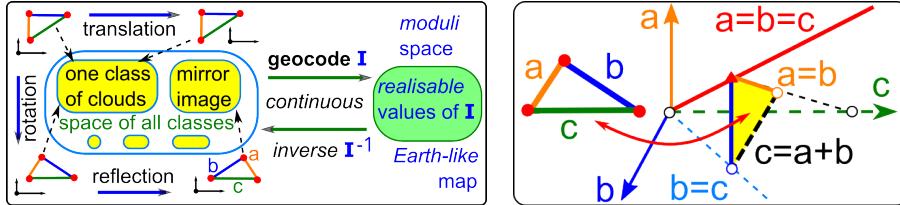
The distance  $d$  has Lipschitz constant  $\lambda = 2$ , because perturbing each of the points  $p, q$  up to  $\varepsilon$  changes their distance  $d = |p - q|$  up to  $2\varepsilon$  due to the triangle inequality.

To check the inverse continuity, let 2-point clouds  $A = \{p, q\}$  and  $B = \{u, v\}$  in  $\mathbb{R}^n$  have  $\delta$ -close distances  $d(A) = |\mathbf{p} - \mathbf{q}|$  and  $d(B) = |\mathbf{u} - \mathbf{v}|$  so that  $|d(A) - d(B)| = \delta$ . Let  $f$  be the isometry that translates the point  $p$  to  $u$  and then rotates the vector  $\mathbf{q} - \mathbf{p}$  around the point  $u$  with an orthogonal matrix from  $O(\mathbb{R}^n)$  to make  $f(\mathbf{q} - \mathbf{p})$  parallel to the fixed vector  $\mathbf{v} - \mathbf{u}$ . Since  $f(p) = u$ , the difference of parallel vectors can be estimated by the difference of their lengths:  $|f(\mathbf{q}) - \mathbf{v}| = ||f(\mathbf{q} - \mathbf{p})| - |\mathbf{v} - \mathbf{u}|| = |d(A) - d(B)| = \delta$ , so the image  $f(A)$  is  $\delta$ -close to  $B$ . We can even additionally shift the 2-point cloud  $f(A)$  along the straight line through the points of  $B$  to put each point of (the image of)  $A$  at a distance of  $\frac{\delta}{2}$  from its closest point of  $B$ . Hence, the Lipschitz constant in condition 1.4.5(e) is  $\mu = \frac{1}{2}$  so that  $\lambda\mu = 1$ . The realisability condition for an inter-point distance is  $d > 0$ . The moduli space  $\text{CIS}(\mathbb{R}^n; 2) = (0, +\infty)$  is embedded in  $\mathbb{R}$ .

(b) If the given points  $p, q$  are ordered, all conclusions in part (a) remain valid in dimensions  $n \geq 2$ , also under rigid motion instead of isometry, because any vectors in  $\mathbb{R}^n$  can be made parallel by rigid motion. In  $\mathbb{R}$ , the complete invariant of two ordered points  $p, q$  under rigid motion (translation) is the difference  $p - q$ . The Cloud Rigid Space  $\text{CRS}(\mathbb{R}; 2) = \mathbb{R} \setminus \{0\}$  excludes the degenerate case of identical points  $p = q$ .  $\blacklozenge$

Fig. 1.3 (left) illustrates Geo-Mapping Problem 1.4.5 by geocodes parametrising geographic-style maps for moduli spaces of 3-point clouds (triangles) under isometry.

**Example 1.5.3** (geocodes for  $m = 3$  points in  $\mathbb{R}^n$ ). (a) The side-side-side theorem in Euclidean geometry says that any triangles are congruent (isometric, in our language) if and only if they have the same triple of side lengths, under permutations.



**Fig. 1.3** **Left:** a geocode  $I$  from Problem 1.4.5 is illustrated for triangles (3-point clouds) whose isometry classes form a moduli space, which can be mapped like Earth. **Right:** the Cloud Isometry Space  $CIS(\mathbb{R}^n; 3)$  is continuously parametrised by triples of inter-point distances  $0 < a \leq b \leq c \leq a + b$ .

An isometry in the plane can reverse orientation, so the vertices of a triangle are unordered. Since a triangle is considered a cloud of 3 unordered points, its three inter-point distances can be written in increasing order, say  $0 < a \leq b \leq c$ .

The side-side-side theorem implies that the ordered triple  $(a, b, c)$  is a complete invariant of 3 unordered points under isometry in the plane and hence in any  $\mathbb{R}^n$ . Similar to Example 1.5.2, the distances  $a, b, c$  are continuous under perturbations with Lipschitz constant  $\lambda = 2$ . The inverse continuity in 1.4.5(e) is harder and will be tackled in forthcoming work. The only realisability condition is the single triangle inequality  $c \leq a + b$ , which is the upper bound for the largest distance. Then the Cloud Isometry Space  $CIS(\mathbb{R}^n; 3)$  is the triangular cone  $\{(a, b, c) \in \mathbb{R}^3 \mid 0 < a \leq b \leq c \leq a + b\}$ .

Under homothety (isometry composed with uniform scaling), this cone projects to the smaller moduli space represented by the yellow triangle in Fig. 1.3 (right).

The red diagonal  $\{a = b = c\}$  represents all equilateral triangles. The boundary planes  $\{a = b \leq c\}$  and  $\{a \leq b = c\}$  represent two types of isosceles triangles: “more horizontal” and “more vertical”, respectively. The third boundary plane  $\{c = a + b\}$  represents degenerate triangles of three points in a straight line.

**(b)** Under rigid motion in any  $\mathbb{R}^n$  for  $n \geq 3$ , all conclusions remain valid, because a mirror reflection in  $\mathbb{R}^2$  can be realised by a rotation in  $\mathbb{R}^n$  with a matrix from  $SO(\mathbb{R}^n)$ . Then, for  $n \geq 3$ , the Cloud Rigid Space  $CRS(\mathbb{R}^n; 3)$  is the same cone as in part (a).

In the plane  $\mathbb{R}^2$ , any rigid motion preserves the cyclic order of 3 points (vertices of a triangle). Then the complete invariant is a triple  $(a, b, c)$  of inter-point distances satisfying  $\max\{a, b\} \leq c \leq a + b$ , so the shorter distances  $a, b$  can be in any order. The resulting Cloud Rigid Space  $CRS(\mathbb{R}^2; 3)$  is obtained by gluing two copies of triangular cones  $CIS(\mathbb{R}^2; 3)$  along their boundaries, where 3-point clouds are mirror-symmetric.

While  $CIS(\mathbb{R}^n; 3)$  is embedded in  $\mathbb{R}^3$  for any  $n \geq 2$ , Euclidean embeddability of  $CRS(\mathbb{R}^2; 3)$  needs a higher-dimensional space  $\mathbb{R}^N$  in condition 1.4.5(g).  $\blacklozenge$

**Example 1.5.4** (geocodes of cyclic polygons in  $\mathbb{R}^2$ ). A polygon polygon in  $\mathbb{R}^2$  is *cyclic* if its vertex set  $A$  is a subset of a circle of a radius (say)  $r > 0$ . Then all points of  $A$  are cyclically ordered along this circle, say as  $p_1, \dots, p_m$ . Since the centre of the circle can be fixed at  $0 \in \mathbb{R}^2$ , the set  $A$  is determined, uniquely under rotation from  $SO(\mathbb{R}^2)$ , by the  $m$  cyclically ordered inter-point distances  $b_i = |p_{i+1} - p_i|$ ,  $i = 1, \dots, m$ , where  $p_{m+1} =$

$p_1$ . The geocode of  $A$  under rigid motion in  $\mathbb{R}^2$  is the sequence  $I(A) = (b_1, \dots, b_m)$  under cyclic permutations. [41, Theorem 1.8] provides the realisability condition:  $b_j < \sum_{i \neq j} b_i$  for  $j = 1, \dots, m$ . Under isometry in  $\mathbb{R}^2$ , the sequence  $(b_1, \dots, b_m)$  should also be considered under reversing the order:  $(b_1, \dots, b_m) \mapsto (b_m, \dots, b_1)$ .  $\blacklozenge$

Euclid might have drawn a geomap of triangles from Example 1.5.3(a) on sand more than 2000 years ago. Hence, it was surprising that even the case of  $m = 4$  unordered points in  $\mathbb{R}^2$  remain opened until complete, Lipschitz continuous, and polynomial-time isometry invariants were developed in 2023 for  $m$  unordered points in any  $\mathbb{R}^n$  [49].

## 1.6 Related areas and connections of Geometric Data Sciences

This section briefly relates Geometric Data Science to other areas in mathematics, data science, and computer science. Later chapters will review past work on specific data.

Our classifications of geometric objects by invariants were inspired by the famous question *Can we hear the shape of a drum?* [31], which has the negative answer in terms of 2D polygons indistinguishable by spectral invariants [26, 27]. Problem 1.4.5 went beyond complete classifications for even better invariants that satisfy extra conditions in 1.4.5(b-h) to allow us not only ‘hear’ but more fully ‘sense’ geometric shapes, e.g. equivalence classes under rigid motion in any  $\mathbb{R}^n$ . Problem 1.4.5 can be informally rephrased as a short question: *can we sense the shape of a real object?*

Though Geo-Mapping Problem 1.4.5 will be re-phrased for various discrete objects, the original statement covers all possible data under arbitrary equivalence relations. In this book, our data objects will be finite and periodic sets of unordered points, which represent atoms in molecules and materials. Other important objects include embedded graphs, polygonal surface meshes, and simplicial complexes. Our standard equivalences are rigid motion, isometry, and their compositions with uniform scaling. Weaker but still practical equivalences are defined by affine, projective, and conformal maps, or actions of specific linear groups on subsets of  $\mathbb{R}^n$  as in classical algebraic geometry.

The generality of Geo-Mapping Problem 1.4.5 and the recent progress in the practical cases of finite and periodic point sets justified the birth of Geometric Data Science [49, 34, 3] as a new area on the interface between metric geometry and data science.

Statistics, data analysis, and shape analysis considered a similar object-oriented approach [38], often for continuous shapes, such as curves and surfaces, under more complicated equivalences, including re-parametrisations or diffeomorphisms.

The carefully written book “Object-oriented data analysis” discussed the concepts of equivalence relations and classes (fibres or orbits) in [38, section 1.2.2] in the case of triangles under congruence as in Example 1.5.3(a), though without mapping the Cloud Isometry Space  $\text{CIS}(\mathbb{R}^n; 3)$  as in Fig. 1.3(b). Since the keyword *invariant* appeared once in [38, p.184], section 1.2 added more motivations and examples of invariants.

Several books on classical invariant theory [40, 33, 22] discuss invariants in the context of algebraic geometry for solutions of polynomial equations, usually under actions of linear groups with elements in  $\mathbb{Q}$  or  $\mathbb{C}$ , not involving translations or permutations, as discussed with B.Hasset [28] and F.Kirwan [39] in private communications.

In real algebraic geometry, the invariants are also studied for (semi-)algebraic sets [8, 45], including important computational aspects [18, 47]. The recent book ‘‘Metric Algebraic Geometry’’ [9] moved beyond invariants towards distance metrics after the earlier workshop on ‘‘Emerging applications of algebraic geometry’’ [42].

The key difference of this book from algebraic geometry is the focus on discrete sets of unordered points coming from real data, such as atomic configurations.

Discrete sets of points have been studied in discrete geometry [37] and rigidity theory [1], often for ordered points. This case also has practical applications in proteins, though unordered (unlabelled) sets are more common in real data.

Topological Data Analysis [23, 19] developed persistent homology summarising the evolution of complexes built on discrete data [15, 30]. For point clouds, the resulting persistence diagrams are invariants under isometry, which are usually computed in dimensions 0 and 1 due to the high complexity, and turned out to be weaker than previously anticipated [46] in comparison with isometry invariants in Chapter 4.

In Topological Robotics, M.Farber [24, Chapter 1] studied moduli spaces of linkages (or robot arms). A *linkage* is a polygonal line of straight segments with potential intersections in the plane or  $\mathbb{R}^n$ . If a linkage is a closed polygon, its vertices are cyclically ordered. If  $l_1, \dots, l_m$  are fixed lengths of straight segments between successive points, the topology of the resulting moduli space  $M(l_1, \dots, l_m)$  was described for generic lengths [24, Theorem 1.3] and for many singular cases [24, Theorem 1.6] in terms of homology [24, Theorems 1.7, 1.21] and cohomology [24, Theorems 1.24-1.29].

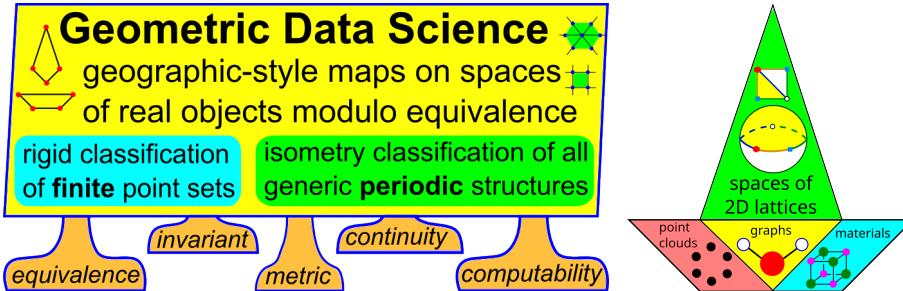
In Computer Science, the interest in the geometry of data has risen due to the influential area of ‘‘Geometric Deep Learning’’ [12, 13] advocating for the invariance of inputs or outputs under actions of  $SE(\mathbb{R}^3)$  or  $E(\mathbb{R}^3)$  in machine learning algorithms.

All these related developments essentially inspired the new area of Geometric Data Science, whose foundational concepts are highlighted in Fig. 1.4 (left).

The logo-style image in Fig. 1.4 (right) shows two quadrilaterals (a vertical kite in green and yellow, and a horizontal trapezium in red, yellow, and blue) whose vertex sets (clouds of 4 unordered points) are indistinguishable by 6 pairwise distances.

## 1.7 The chapter plan: from easier to more challenging data

This section outlines the plan of all further chapters, which are split into two big parts: finite point sets and periodic point sets, which will be mostly studied under isometry. In the first part, Chapters 2-6 solve partial cases of Problem 1.4.5 for finite point sets.



**Fig. 1.4** **Left:** the key concepts are introduced in Definitions 1.2.1, 1.2.5, 1.3.1, 1.3.4, and 1.4.3, all linked in Problem 1.4.5. **Right:** the main objects are finite and periodic sets of unordered points, including lattices in  $\mathbb{R}^2$  whose space under rigid motion was the first solution to Problem 1.4.5.

**Chapter 2** discusses complete, bi-continuous, and linear-time invariants [2] for finite sets of ordered points under rigid motion in  $\mathbb{R}^3$ , which exposed thousands of duplicate chains in the Protein Data Bank within a few hours on a modest desktop computer.

**Chapter 3** leverages Principal Component Analysis to get complete and polynomial-time invariants for finite clouds of unordered points under rigid motion in  $\mathbb{R}^n$ .

**Chapter 4** introduces Pointwise Distance Distribution. The PDD is a fast and generically complete isometry invariant of finite and periodic sets of unordered points. This chapter proves that the PDD is complete for all 4-point clouds under isometry in  $\mathbb{R}^n$ .

**Chapter 5** refines the isometry invariant PDD to a stronger Simplexwise Distance Distribution (SDD) for a finite cloud of unordered points in any metric space. Despite having a higher (polynomial-time) complexity, a simple definition of SDD allows us to distinguish all known non-isometric clouds in  $\mathbb{R}^3$  that have identical PDDs.

**Chapter 6** improves SDD to a Simplexwise Centred Distribution, which is a complete, Lipschitz continuous and polynomial time invariant of all point clouds under rigid motion in  $\mathbb{R}^n$ . The hardest obstacle in the proof of Lipschitz continuity was resolved by a *strength* of a simplex, which is a linear-growth analogue of the simplex volume.

In the second part, Chapters 7-11 solve partial cases of Problem 1.4.5 for periodic sets.

**Chapter 7** defines complete invariants for ordered sequences of points (under several versions of isometries in  $\mathbb{R} \times \mathbb{R}^{n-1}$ ) that are periodic along the first coordinate axis.

**Chapter 8** expands the classical approaches of Gauss, Lagrange [17], and Delone [7], who studied lattices via quadratic forms, and the more recent work of Conway and Sloane [16] to solve Problem 1.4.5 for all periodic lattices under rigid motion in  $\mathbb{R}^2$ .

**Chapter 9** discusses density functions, which extend the point density of periodic point sets to generically complete invariants under isometry in  $\mathbb{R}^3$ . These density functions will be analytically described for all periodic sequences of intervals within  $\mathbb{R}$ .

**Chapter 10** extends PDDs from finite to periodic point sets, proves their generic completeness under isometry in  $\mathbb{R}^n$ , and describes their asymptotic behaviour.

**Chapter 11** refines the seminal work of Dolbilin, Lagarias, and Senechal [21] to build a complete invariant *isoset* with Lipschitz continuous metrics, which can be approximated by polynomial-time algorithms for all periodic point sets in  $\mathbb{R}^n$ .

**Chapter 12** summarises the most significant results and experimentally verified principles, which inspired new concepts of geometric structures for crystals and molecules. A brief history of developments concludes with future directions and highlights several open problems, some of which can be accessible to school students and are partially inspired by V.I.Arnold’s “Problem for School Pupils” in [5, Chapter 6].

## References

1. Abdo Y Alfaik et al. *Euclidean distance matrices and their applications in rigidity theory*. Springer, 2018.
2. Olga Anosova et al. A complete and bi-continuous invariant of protein backbones under rigid motion. *MATCH Comm. Math. Comp. Chemistry*, 94:97–134, 2025.
3. Olga Anosova, Daniel Widdowson, and Vitaliy Kurlin. Recognition of near-duplicate periodic patterns by continuous metrics with approximation guarantees. *Pattern Recognition*, 171:112108, 2026.
4. David Arnold. *The age of discovery, 1400-1600*. Routledge, 2013.
5. V I Arnold. *Real algebraic geometry*, volume 66. Springer Science & Business Media, 2013.
6. Visar Berisha et al. Digital medicine and the curse of dimensionality. *NPJ digital medicine*, 4(1):153, 2021.
7. B.N.Delone, N. Padurov, and A. Aleksandrov. *Mathematical Foundations of Structural Analysis of Crystals*. State Technical-Theoretical Press, USSR, 1934.
8. Jacek Bochnak, Michel Coste, and Marie-Françoise Roy. *Real algebraic geometry*, volume 36. Springer Science & Business Media, 2013.
9. Paul Breiding, Kathlén Kohn, and Bernd Sturmfels. *Metric algebraic geometry*. Springer, 2024.
10. Matthew Bright et al. Geographic-style maps for 2D lattices. *Acta Cryst A*, 79:1–13, 2023.
11. C Brock. Change to the definition of “crystal” in the IUCr Online Dictionary of Crystallography. [www.iucr.org/news/newsletter/etc/articles?issue=151351&result\\_138339\\_result\\_page=17](http://www.iucr.org/news/newsletter/etc/articles?issue=151351&result_138339_result_page=17), 2021.
12. Michael Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: going beyond Euclidean data. *IEEE Signal Processing Magazine*, 34:18–42, 2017.
13. Michael M Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv:2104.13478*, 2021.
14. Jeff Calder and Peter J Olver. *Linear Algebra, Data Science, and Machine Learning*. Springer, 2025.
15. Gunnar Carlsson and Mikael Vejdemo-Johansson. *Topological data analysis with applications*. Cambridge University Press, 2021.
16. J Conway and N Sloane. Low-dimensional lattices. VI. Voronoi reduction of three-dimensional lattices. *Proceedings Royal Society A*, 436(1896):55–68, 1992.
17. Joseph Louis De Lagrange. Recherches d’arithmétique. *Nouveaux Mémoires de l’Académie de Berlin*, 1773.
18. Harm Derksen and Gregor Kemper. *Computational invariant theory*. Springer, 2001.
19. Tamal Krishna Dey and Yusu Wang. *Computational topology for data analysis*. Cambridge University Press, 2022.
20. Michel Marie Deza and Elena Deza. *Encyclopedia of distances*. Springer, 2009.
21. N. Dolbilin, J Lagarias, and M. Senechal. Multiregular point systems. *Discrete & Computational Geometry*, 20(4):477–498, 1998.
22. Igor Dolgachev. *Lectures on invariant theory*. Number 296. Cambridge University Press, 2003.
23. Herbert Edelsbrunner and John Harer. *Computational topology: an introduction*. AMS, 2010.

24. Michael Farber. *Invitation to topological robotics*. European Mathematical Society, 2008.
25. Richard Feynman. *The Feynman lectures on physics: the new millennium edition*, volume 1. 2011.
26. Carolyn Gordon, David Webb, and Scott Wolpert. Isospectral plane domains and surfaces via riemannian orbifolds. *Inventiones mathematicae*, 110(1):1–22, 1992.
27. Carolyn Gordon, David L Webb, and Scott Wolpert. One cannot hear the shape of a drum. *Bulletin of the American Mathematical Society*, 27(1):134–138, 1992.
28. Brendan Hassett, Alena Pirutka, and Yuri Tschinkel. Stable rationality of quadric surface bundles over surfaces. 220(2):341–365, 2018.
29. Dominic Hyde. The sorites paradox. In *Vagueness: A guide*, pages 1–17. Springer, 2011.
30. Parvaneh Joharinad and Jürgen Jost. *Mathematical principles of topological and geometric data analysis*. Springer, 2023.
31. Mark Kac. Can one hear the shape of a drum? *Amer. Math. Monthly*, 73(4P2):1–23, 1966.
32. David Kendall, Dennis Barden, Thomas Carne, and Huiling Le. *Shape and shape theory*. Wiley & Sons, 2009.
33. H Kraft and C Procesi. *Classical Invariant Theory. A Primer*. 2000.
34. Vitaliy Kurlin. Mathematics of 2D lattices. *Found. Comp. Mathematics*, 24:805–863, 2024.
35. Peter Landweber, Emanuel Lazar, and Neel Patel. On fiber diameters of continuous maps. *The American Mathematical Monthly*, 123(4):392–397, 2016.
36. Tom Leinster. *Entropy and diversity: the axiomatic approach*. Cambridge university press, 2021.
37. Leo Liberti and Carlile Lavor. *Euclidean distance geometry*. Springer, 2017.
38. James S Marron and Ian L Dryden. *Object oriented data analysis*. Chapman and Hall/CRC, 2021.
39. David Mumford, John Fogarty, and Frances Kirwan. *Geometric invariant theory*, volume 34. Springer Science & Business Media, 1994.
40. Peter J Olver. *Classical invariant theory*. Number 44. Cambridge University Press, 1999.
41. R Penner. *Decorated Teichmüller theory*, volume 1. European Mathematical Society, 2012.
42. Mihai Putinar and Seth Sullivant. Emerging applications of algebraic geometry. 2008.
43. Stefan Rass et al. Metricizing the euclidean space towards desired distance relations in point clouds. *IEEE Trans. Information Forensics and Security*, 19:7304–7319, 2024.
44. Pietro Saccchi et al. Same or different – that is the question: identification of crystal forms from crystal structure data. *Cryst Eng Comm*, 22(43):7170–7185, 2020.
45. Claus Scheiderer. *A course in real algebraic geometry*. Springer, 2024.
46. Philip Smith and Vitaliy Kurlin. Generic families of finite metric spaces with identical or trivial 1-dimensional persistence. *Journal of Applied and Computational Topology*, 8:839–855, 2024.
47. Thorsten Theobald. *Real algebraic geometry and optimization*, volume 241. AMS, 2024.
48. Yuri Tschinkel. Geometry over nonclosed fields. In *International Congress of Mathematicians*, volume 2, pages 637–651, 2006.
49. Daniel E Widdowson and Vitaliy A Kurlin. Recognizing rigid patterns of unlabeled point clouds by complete and continuous isometry invariants with no false negatives and no false positives. In *Computer Vision and Pattern Recognition*, pages 1275–1284, 2023.

**Part I**  
**Geometric Data Science of finite point sets**



## Chapter 2

# Sequences of ordered points under rigid motion in Euclidean spaces

**Abstract** This chapter discusses complete invariants of finite sets of ordered points under rigid motion in Euclidean space  $\mathbb{R}^n$ . After discussing distance matrices, we adapt Geo-Mapping Problem 1.4.5 for protein backbones, which are non-degenerate polygonal chains in  $\mathbb{R}^3$ . Further sections describe the linear-time Backbone Rigid Invariant (BRI) and all-vs-all comparisons of chains in the Protein Data Bank.

## 2.1 Classical invariants and shape spaces of ordered points

This section reviews past approaches to classify sequences of ordered points  $p_1, \dots, p_m$  under rigid motion or isometry in  $\mathbb{R}^n$ . In the case of isometry, a complete invariant of the sequence  $p_1, \dots, p_m$ , known at least since 1935 [23], is the matrix of pairwise distances. An alternative (complete) isometry invariant is the Gram matrix of scalar products  $\mathbf{p}_i \cdot \mathbf{p}_j$  [25, chapter 2.9], which can be expressed in terms of the distance matrix and vice versa. Since these matrices do not distinguish mirror images, we state the Euclidean version of Problem 1.4.5 for any finite sets of ordered points below.

**Problem 2.1.1** (partial case of Problem 1.4.5 for sequences under rigid motion in  $\mathbb{R}^n$ ). Design a map  $I$  on finite sets of ordered points in  $\mathbb{R}^n$  satisfying the conditions below.

- (a) *Completeness*: any sequences  $A, B \subset \mathbb{R}^n$  are related by rigid motion ( $A \cong B$ ) in  $\mathbb{R}^n$  if and only if  $I(A) = I(B)$ .
- (b) *Reconstruction*: any sequence  $A \subset \mathbb{R}^n$  of ordered points can be reconstructed from its invariant value  $I(A)$ , uniquely under rigid motion.
- (c) *Metric*: there is a distance  $d$  on the invariant space  $\{I(A) \mid \text{sequences } A \subset \mathbb{R}^n\}$  satisfying all metric axioms in Definition 1.3.1(a).
- (d) *Continuity*: there is a constant  $\lambda$  such that, for any  $\varepsilon > 0$ , if  $B$  is obtained from  $A$  by perturbing every point of  $A$  up to Euclidean distance  $\varepsilon$ , then  $d(I(A), I(B)) \leq \lambda\varepsilon$ .

**(e) Computability:** for a fixed dimension  $n$ , the invariant  $I(A)$ , and the metric  $d(I(A), I(B))$  can be computed in times that depend polynomially on the maximum size  $\max\{|A|, |B|\}$  of sets  $A, B \subset \mathbb{R}^n$ . ★

Problem 2.1.1 will be solved in section 7.2. Further sections of this chapter will solve a non-degenerate version of Problem 2.1.1 for protein backbones in  $\mathbb{R}^3$ , see Problem 2.2.2. The following concepts will be essentially used later.

**Definition 2.1.2** (affine dimension). The *affine dimension*  $\text{aff}(A)$  of a cloud  $A \subset \mathbb{R}^n$  of points  $p_1, \dots, p_m$  is the maximum dimension of the vector space generated by all inter-point vectors  $\mathbf{p}_i - \mathbf{p}_j$  for  $i, j \in \{1, \dots, m\}$ . ▲

The affine dimension  $\text{aff}(A)$  is an isometry invariant and is independent of the order of points of  $A$ . Any cloud  $A$  of 2 distinct points has  $\text{aff}(A) = 1$ . Any cloud  $A$  of 3 points that are not in the same straight line has  $\text{aff}(A) = 2$ . Lemma 2.1.3 provides a criterion for a matrix to be realisable by squared distances of a cloud in  $\mathbb{R}^n$ .

**Lemma 2.1.3** (realisation of distances). **(a)** A symmetric  $m \times m$  matrix of  $s_{ij} \geq 0$  with  $s_{ii} = 0$  is realisable as a matrix of squared distances between  $p_0 = 0, p_1, \dots, p_{m-1} \in \mathbb{R}^n$  for some  $n$  if and only if the  $(m-1) \times (m-1)$  matrix  $G$  of  $g_{ij} = \frac{s_{0i} + s_{0j} - s_{ij}}{2}$  has only non-negative eigenvalues.

**(b)** If  $G$  has only non-negative eigenvalues, then  $\text{aff}(0, p_1, \dots, p_{m-1})$  equals the number  $k \leq m-1 \leq n$  of positive eigenvalues of  $G$ . In this case,  $g_{ij} = p_i \cdot p_j$  define the *Gram matrix* of the vectors  $p_1, \dots, p_{m-1} \in \mathbb{R}^n$ , which are uniquely determined in time  $O(m^3)$  under a map from  $O(n)$ .

**Proof of Lemma 2.1.3.** **(a,b)** We extend [5, Theorem 1] to the case  $m < n+1$  and find  $p_1, \dots, p_{m-1} \in \mathbb{R}^n$  in time  $O(m^3)$ , uniquely under an orthogonal map from  $O(n)$ .

The part *only if*  $\Rightarrow$ . Let a symmetric matrix  $S$  consist of squared distances between points the  $p_0 = 0, p_1, \dots, p_{m-1} \in \mathbb{R}^n$ . For  $i, j = 1, \dots, m-1$ , the matrix of

$$g_{ij} = \frac{s_{0i} + s_{0j} - s_{ij}}{2} = \frac{|p_i|^2 + |p_j|^2 - |p_i - p_j|^2}{2} = p_i \cdot p_j$$

is the Gram matrix, which can be written as  $G = P^T P$ , where the columns of the  $n \times (m-1)$  matrix  $P$  are the vectors  $p_1, \dots, p_{m-1}$ . For any vector  $v \in \mathbb{R}^{m-1}$ , we have

$$0 \leq |Pv|^2 = (Pv)^T (Pv) = v^T (P^T P) v = v^T G v.$$

Since the quadratic form  $v^T G v \geq 0$  for any  $v \in \mathbb{R}^{m-1}$ , the matrix  $G$  is positive semi-definite, i.e.  $G$  has only non-negative eigenvalues, see [9, Theorem 7.2.7].

The part *if*  $\Leftarrow$ . For any positive semi-definite matrix  $G$ , there is an orthogonal matrix  $Q$  such that  $Q^T G Q = D$  is the diagonal matrix, whose  $m-1$  diagonal elements are non-negative eigenvalues of  $G$ . The diagonal matrix  $\sqrt{D}$  consists of the square roots of the eigenvalues of  $G$ . The number of positive eigenvalues of  $G$  equals the dimension  $k = \text{aff}(\{0, p_1, \dots, p_{m-1}\})$  of the subspace that is linearly spanned by  $p_1, \dots, p_{m-1}$ .

We may assume that all  $k \leq n$  positive eigenvalues of  $G$  correspond to the first  $k$  coordinates of  $\mathbb{R}^n$ . Since  $Q^T = Q^{-1}$ , the given matrix  $G = QDQ^T = (Q\sqrt{D})(Q\sqrt{D})^T$  becomes the Gram matrix of the columns of  $Q\sqrt{D}$ . These columns become the reconstructed vectors  $p_1, \dots, p_{m-1} \in \mathbb{R}^n$ .

If there is another diagonalisation  $\tilde{Q}^T G \tilde{Q} = \tilde{D}$  for  $\tilde{Q} \in O(n)$ , then  $\tilde{D}$  differs from  $D$  by a permutation of eigenvalues, which is realized by an orthogonal map, so we set  $\tilde{D} = D$ . Then  $G = \tilde{Q}D\tilde{Q}^T = (\tilde{Q}\sqrt{D})(\tilde{Q}\sqrt{D})^T$  is the Gram matrix of the columns of  $\tilde{Q}\sqrt{D}$ . The new columns differ from the previously reconstructed vectors  $p_1, \dots, p_{m-1} \in \mathbb{R}^n$  by the orthogonal map  $\tilde{Q}\tilde{Q}^T$ . Hence the reconstruction is unique under  $O(n)$ -transformations. Computing eigenvectors  $p_1, \dots, p_{m-1}$  needs a diagonalisation of  $G$  in time  $O(m^3)$ , see section 11.5 in [20].  $\square$

Chapter 3 in [16] discusses realisations of a complete graph given by a distance matrix in  $\mathbb{R}^n$ . Lemma 2.1.4(a) holds for all clouds, including degenerate ones, e.g. for 3 points in a straight line. Any points  $p_1, \dots, p_{n-1} \in A$  have  $\text{aff}(p_1, \dots, p_{n-1}) \leq n-2$ . For example, any two distinct points in  $A \subset \mathbb{R}^3$  generate a straight line.

**Lemma 2.1.4** (sequence reconstruction). **(a)** Any sequence of ordered points  $p_1, \dots, p_m$  in  $\mathbb{R}^n$  can be reconstructed (uniquely under isometry) from the matrix of the Euclidean distances  $|p_i - p_j|$  in time  $O(m^3)$ . If all distances are divided by  $R = \max_{i=1, \dots, m} |p_i|$ , the reconstruction of  $p_1, \dots, p_m$  is unique under isometry and uniform scaling in  $\mathbb{R}^n$ .

**(b)** If  $m \leq n$ , the uniqueness of reconstructions in part (a) remains true if we replace isometry with rigid motion in  $\mathbb{R}^n$ .  $\blacksquare$

**Proof of Lemma 2.1.4.** **(a)** By translation, we can fix  $p_1$  at the origin. Let  $G$  be the  $(m-1) \times (m-1)$  matrix  $g_{ij} = \frac{|p_i|^2 + |p_j|^2 - |p_i - p_j|^2}{2} = p_i \cdot p_j$ , where  $i, j = 2, \dots, m$ , which is obtained from the squared distances between the points  $p_1 = 0, p_2, \dots, p_m$ .

By Lemma 2.1.3 if  $G$  has  $k \leq n$  positive eigenvalues, then  $p_1 = 0, \dots, p_m$  can be uniquely determined under isometry in  $\mathbb{R}^k \subset \mathbb{R}^n$  in time  $O(m^3)$ . If all distances are divided by the same radius  $R = \max_{i=1, \dots, m} |p_i|$ , the above construction guarantees uniqueness under isometry and uniform scaling.

**(b)** If  $m \leq n$ , any mirror images of  $p_1, \dots, p_m \in \mathbb{R}^n$ , after a suitable rigid motion, can be assumed to belong to an  $(n-1)$ -dimensional hyperspace  $H \subset \mathbb{R}^n$ , where they are matched by a mirror reflection  $H \rightarrow H$  with respect to an  $(n-2)$ -dimensional subspace  $S \subset H$ , which is realized by the  $180^\circ$  orientation-preserving rotation around  $S$ .  $\square$

Lemma 2.1.4(b) for  $m = n = 3$  implies that any triangle is determined by its sides, uniquely under rigid motion in  $\mathbb{R}^3$ . For example, sides 3, 4, 5 define a right-angled triangle whose mirror images are not related by rigid motion within a plane  $H \subset \mathbb{R}^3$ , but are matched by a rigid motion in  $H$  and a  $180^\circ$  rotation of  $\mathbb{R}^3$  around a line in  $H$ .

The difference between the matrices of distances or scalar products can be converted into a continuous metric by taking a matrix norm. These matrices are preserved under any mirror reflections. Hence, these invariants are incomplete under rigid motion.

One can define the sign of orientation on some (or all subsets of)  $n + 1$  points from a given sequence. This extra sign is discrete and vanishes for degenerate configurations for  $n + 1$  points that affinely span a  $k$ -dimensional subspace in  $\mathbb{R}^n$  for  $k \leq n - 1$ .

Another attempt to satisfy the Lipschitz continuity in Problem 1.4.5(d) is to multiply the sign of orientation by the volume of the simplex spanned by  $n + 1$  points, say  $p_1, \dots, p_{n+1}$ . In other words, one can take the signed volume of the parallelepiped spanned by the vectors  $\mathbf{p}_i - \mathbf{p}_1$ ,  $i = 2, \dots, n + 1$ . If the first  $n + 1$  points are degenerate, then the zero volume of their spanned parallelepiped does not give any extra information to distinguish mirror images of the full sequence.

More importantly, the resulting signed volume is not Lipschitz continuous already in dimension  $n = 2$ . Indeed, let us consider the triangle  $A(t)$  on the vertices  $(\pm l, 0)$  and  $(0, t\varepsilon)$ , where  $l, \varepsilon > 0$  are fixed constants ( $l$  is large,  $\varepsilon > 0$  is small), and  $t \in [-1, 1]$  is a time parameter. The signed area of  $A(t)$  is  $t\varepsilon$ , changing from  $-l\varepsilon$  at  $t = -1$  to  $l\varepsilon$  at  $t = 1$ . Then the Lipschitz constant cannot be smaller than  $\lambda = \frac{2l\varepsilon}{2} = l\varepsilon$ . Then the signed area can have a fixed Lipschitz constant only for bounded triangles, not for all triangles, because we can choose  $\varepsilon = \frac{1}{\sqrt{l}}$  to make  $\lambda = l\varepsilon = \sqrt{l}$  unbounded.

We will resolve this obstacle to Lipschitz continuity by a different function (the strength of a simplex) in a later chapter. This section finishes by noting that a naive extension of the complete isometry invariant (matrix of distances or scalar products) from the ordered to  $m$  unordered points requires  $m!$  permutations. This exponential complexity is ruled out by the polynomial-time requirement in 1.4.5(h). Indeed,  $m!$  becomes too large already for  $m = 4$  points:  $4! = 24$  matrices of size  $4 \times 4$ .

Each distance matrix is symmetric and has zeros on the diagonal, and hence can be represented by only 6 distances. However, the total number of  $24 \times 6 = 144$  distances seems overwhelmingly unnecessary to unambiguously and continuously encode 4 unordered points under isometry in  $\mathbb{R}^2$ . Chapter 4 will prove that a smaller  $4 \times 3$  matrix invariant is complete for any 4 unordered points under isometry in  $\mathbb{R}^n$ .

In 1977, Kendall [13] started to study configuration spaces of ordered points modulo rigid motion in  $\mathbb{R}^n$  under the name of *size-and-shape spaces* [14]. If we consider sequences equivalent also under uniform scaling, the smaller *shape space*  $\Sigma_2^m$  of  $m$  ordered points in  $\mathbb{R}^2$  can be described as a complex projective space  $\mathbb{C}P^{m-1}$  due to the group  $\text{SO}(\mathbb{R}^2)$  being identified with the unit circle in the complex space  $\mathbb{C}^1 = \mathbb{R}^2$ . However, there is no easy description of the moduli space  $\Sigma_3^m$  of  $m$ -point sequences in  $\mathbb{R}^3$ , which has no multiplicative group structure similar to  $\mathbb{R}^2 = \mathbb{C}^1$ .

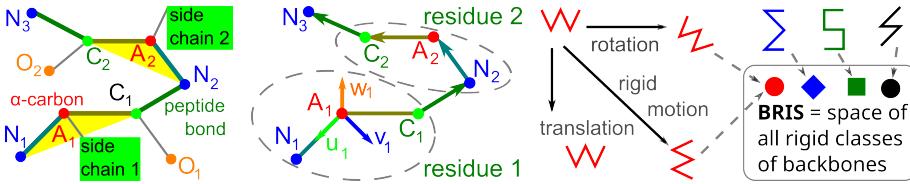
In a general metric space, let a sequence  $A$  of  $m$  ordered points be given by their  $m \times m$  distance matrix  $D$ . Multidimensional scaling [15] finds an embedding  $A \subset \mathbb{R}^k$  (if it exists) preserving all distances of  $M$  for a minimum dimension  $k \leq m$ . The underlying computation of  $m$  eigenvalues of the Gram matrix expressed via  $D$  needs  $O(m^3)$  time.

The resulting representation of an embedded sequence  $A \subset \mathbb{R}^k$  uses orthonormal eigenvectors whose ambiguity up to signs for potential comparisons leads to the time factor  $2^k$ , which can be close to  $2^m$  and hence exponential in the number  $m$  of points.

Further sections in this chapter follow papers [2, 26]. We are grateful to John Helliwell, Thérèse Malliavin, Elspeth Garman, Lauren Porter, Erica Flapan, and all our co-authors, especially Mariusz Jaskolski, Wladek Minor, and Alex Wlodawer for their expert advice and support of mathematical foundations in structural biology.

## 2.2 The geo-mapping problem for protein backbones in $\mathbb{R}^3$

A *protein* is a large biomolecule consisting of one or several chains of amino acid residues. The *primary structure (sequence)* of a protein chain is a string of residue labels (represented by one or three letters), each denoting one of (usually) 20 standard amino acids [24]. The *secondary structure* consists of frequent semi-rigid subchains such as  $\alpha$ -helices and  $\beta$ -strands [17]. A sequence of a protein is relatively easy to experimentally determine but important functional properties, such as interactions with drug molecules, depend on a 3-dimensional geometric *fold* (a *tertiary structure*) represented by an embedding of all its atoms in  $\mathbb{R}^3$ , see Fig. 2.1 (left).



**Fig. 2.1** **Left:** all main atoms  $N_i, A_i, C_i$  of a protein chain form a *backbone* embedded in  $\mathbb{R}^3$ . **Middle:** each triangle  $\triangle N_i A_i C_i$  defines an orthonormal basis  $u_i, v_i, w_i$ . The coordinates of the bond vectors  $\overrightarrow{C_i N_{i+1}}, \overrightarrow{N_{i+1} A_{i+1}}, \overrightarrow{A_{i+1} C_{i+1}}$  in this basis form the complete Backbone Rigid Invariant BRI. **Right:** All rigidly equivalent backbones form a single *rigid class*. All rigid classes form the *Backbone Rigid Space*. The image schematically illustrates four different classes of simple polygonal chains in  $\mathbb{R}^3$ .

In 1973, Nobel laureate Anfinsen conjectured that the sequence of any protein chain determines its 3D geometric fold [1]. Following this conjecture, neural networks such as AlphaFold2 and RosettaFold [11, 3] optimise millions of parameters to predict a protein fold from its sequence, but need re-training [10] on the growing Protein Data Bank (PDB), which is considered a ‘gold standard’ for experimental structures [4]. The reported accuracies of prediction are often based on the LDDT (Local Distance Difference Test) [18, p. 2728] and TM-score [27], which fail the metric axioms. Then clustering can produce pre-determined clusters and may not be trustworthy [21].

Backbones of the same length (number of residues) can be optimally aligned to minimise the Root Mean Square Deviation (RMSD) between corresponding atoms [8]. This RMSD is slow to compute for all pairs of proteins and gives only distances without mapping the protein universe (moduli space of proteins under rigid motion).

We develop a different approach by mapping the space of protein backbones in analytically defined coordinates similar to geographic-style maps of a new planet.

Any embedded protein in  $\mathbb{R}^3$  can be rigidly moved, which changes all atomic coordinates. However, the underlying structure remains the same in the sense that different images of a protein under rigid motion have the same properties in a fixed environment. Though proteins are flexible, it is important to distinguish their rigid structures that can interact differently [6] with other molecules, including medical drugs.

**Definition 2.2.1** (Backbone Rigid Space  $\text{BRIS}_m$ ). A protein *backbone* is a sequence of  $m$  ordered triplets of main chain atoms (nitrogen  $N_i$ ,  $\alpha$ -carbon  $A_i$ , and carbonyl carbon  $C_i$ ) given by their positions in  $\mathbb{R}^3$ . The *structure* of a backbone, or a protein chain (with all side chains), or a biomolecule consisting of several chains is the equivalence class of this geometric object under rigid motion in  $\mathbb{R}^3$ . For any  $m \geq 1$ , the classes of all backbones of  $m$  triplets under rigid motion form the *Backbone Rigid Space*  $\text{BRIS}_m$ .  $\blacktriangle$

Backbones were studied by incomplete invariants such as torsion angles, which allow *false positive* pairs of non-equivalent backbones  $S \not\cong Q$  with  $I(S) = I(Q)$ . Because all atoms in a backbone  $S$  are ordered, their distance matrix determines  $S \subset \mathbb{R}^3$ , uniquely under isometry, but has a large quadratic size in the number  $m$  of residues and fails to distinguish mirror images. Adding a sign of orientation creates discontinuity for polygonal chains that are almost mirror-symmetric.

Problem 2.2.2 adapts Geo-Mapping Problem 1.4.5 to protein backbones. The completeness in 1.4.5(a) is restricted to polygonal chains, where each triplet of atoms  $N_i, A_i, C_i$  is not in a straight line, as we have checked for all experimental structures in the PDB. The polynomial-time condition in 1.4.5(h) is strengthened to linear time.

**Problem 2.2.2** (geo-mapping for protein backbones). For any  $m \geq 1$ , design a map  $I : \text{BRIS}_m \rightarrow \mathbb{R}^N$  for some  $N$  satisfying the following conditions.

- (a) *Completeness*: any backbones  $S, Q \subset \mathbb{R}^3$  are rigidly equivalent if and only if  $I(S) = I(Q)$ , i.e.  $I$  has *no false negatives* and *no false positives*.
- (b) *Reconstruction*: any protein backbone  $S \subset \mathbb{R}^3$  can be reconstructed from its invariant value  $I(S)$  uniquely under rigid motion.
- (c) *Metric*: there is a distance  $d$  on invariant values satisfying all metric axioms in Definition 1.3.1(a).
- (d) *Continuity*: there is a constant  $\lambda$  such that, for any  $\varepsilon > 0$ , if  $Q$  is obtained from  $S$  by perturbing every atom up to Euclidean distance  $\varepsilon$ , then  $d(I(S), I(Q)) \leq \lambda\varepsilon$ .
- (e) *Atom matching*: there is a constant  $\mu$  such that, for any backbones  $S, Q$  with  $\delta = d(I(S), I(Q))$ , all their atoms can be matched up to a distance  $\mu\delta$  by a rigid motion.
- (f) *Realisability*: the invariant space  $I\{X\} = \{I(A) \mid A \in X\}$  can be parametrised so that we can generate any value  $I(A) \in I\{X\}$  realisable by some object  $A \in X$ .
- (g) *Respecting subchains*: for any subchain of residues  $R_i \cup \dots \cup R_{i+j}$  in a backbone  $S$ , the invariant  $I(R_i \cup \dots \cup R_{i+j})$  can be obtained from  $I(S)$  in linear time  $O(j)$  with respect to the length of the subchain.
- (h) *Linear-time computability*: the invariant  $I$ , the metric  $d$ , a reconstruction in (b), and a rigid motion in (e) can be computed in time  $O(m)$  for any backbone of  $m$  residues.  $\star$

The completeness in 2.2.2(a) means that  $I$  is the strongest invariant and hence distinguishes all protein backbones that cannot be exactly matched by rigid motion. The reconstruction in 2.2.2(b) is more practical because  $I$  may not allow an efficiently computable inverse map  $I^{-1}$  from an invariant value  $I(S)$  to a backbone  $S \subset \mathbb{R}^3$ .

The continuity in 2.2.2(d) fails for invariants based on principal directions that can discontinuously change in degenerate cases when eigenvalues become equal. The atom matching in 2.2.2(e) says that, after finding a rigid motion  $f$  in  $\mathbb{R}^3$ , any atom  $p \in S$  has Euclidean distance at most  $\mu\delta$  to the corresponding atom  $q \in f(Q)$ .

Conditions 2.2.2(d,e) guarantee the Lipschitz continuity of  $I$  and its inverse on the image  $I(\text{BRIS}_m) \subset \mathbb{R}^N$ . Though Lemma 2.1.3 gives a two-sided criterion for the realisability of distances by ordered points  $p_1, \dots, p_m \in \mathbb{R}^n$ , the space of distance matrices is highly singular and cannot be easily sampled. Since a random matrix of potential distances for  $m > n + 1$  is unlikely to be realisable by  $m$  ordered points in  $\mathbb{R}^n$ , the realisability condition in 2.2.2(g) is non-trivial for the distance matrix.

Since Problem 2.2.2 asked for an invariant  $I : \text{BRIS}_m \rightarrow \mathbb{R}^N$ , the Euclidean embeddability in 1.4.5(g) hold automatically and has been replaced with condition 2.2.2(g), motivated by secondary structures, which are subchains in full backbones.

The linear time in 2.2.2(h) makes all previous conditions practically useful because even the distance matrix needs  $O(m^2)$  time and space, substantially slower than linear time  $O(m)$  for thousands of residues.

Past work on similarities of proteins is reviewed in [2, section 2]. Section 2.3 introduces the *Backbone Rigid Invariant* BRI :  $\text{BRIS}_m \rightarrow \mathbb{R}^{9m-6}$  to solve Problem 2.2.2 by Theorems 2.3.4, 2.3.7, 2.3.9. The numerical components of BRI play the role of geocodes, which are geographic-style coordinates on the space  $\text{BRIS}_m$ , where any protein backbone has a uniquely defined location. Section 2.5 describes how BRI detected thousands of geometric duplicates in the PDB, some of which need updates.

## 2.3 Complete and bi-continuous Backbone Rigid Invariant

We start with the simpler *triangular invariant* that describes the rigid class of each residue triangle  $\triangle N_i A_i C_i$  on three main atoms: nitrogen  $N_i$ ,  $\alpha$ -carbon  $A_i$ , and carbonyl carbon  $C_i$ , for  $i = 1, \dots, m$ , see Fig. 2.1 (middle). For any points  $A, B \in \mathbb{R}^3$ , let  $|\overrightarrow{AB}|$  be the Euclidean length of the vector  $\overrightarrow{AB}$  from  $A$  to  $B$ . The *scalar* and *vector* products of vectors  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$  are denoted by  $\mathbf{u} \cdot \mathbf{v}$  and  $\mathbf{u} \times \mathbf{v}$ , respectively.

**Definition 2.3.1** (triangular invariant TRIN). Let a backbone  $S \subset \mathbb{R}^3$  have  $3m$  ordered atoms  $N_i, A_i, C_i, i = 1, \dots, m$ . In the plane of  $\triangle N_i A_i C_i$ , for the 2D basis obtained by Gaussian orthogonalisation of  $\overrightarrow{A_i N_i}, \overrightarrow{A_i C_i}$ , the vector  $\overrightarrow{A_i N_i}$  has the coordinates  $x(A_i N_i) = |\overrightarrow{A_i N_i}|$  and  $y(A_i N_i) = 0$ . Let  $\mathbf{x} = \frac{\overrightarrow{A_i N_i}}{|\overrightarrow{A_i N_i}|}$  be the unit vector. Then  $\overrightarrow{A_i C_i}$  has the coordinates  $x(A_i C_i) = \overrightarrow{A_i C_i} \cdot \mathbf{x}$  and  $y(A_i C_i) = |\overrightarrow{A_i C_i} - x(A_i C_i)\mathbf{x}|$  in the direction

orthogonal to  $\mathbf{x}$ . The *triangular invariant*  $\text{TRIN}(S)$  is the  $m \times 3$  matrix whose  $i$ -th row consists of the coordinates  $x(A_i N_i)$ ,  $x(A_i C_i)$ , and  $y(A_i C_i)$  for  $i = 1, \dots, m$ .  $\blacktriangle$

The  $i$ -th row of  $\text{TRIN}(S)$  uniquely determines the rigid class of  $\triangle N_i A_i C_i$ .

On May 4, 2024, the PDB had 213,191 entries with 1,091,420 chains. Protocol 2.3.2 below produced 104,688  $\approx 49\%$  entries with 707410  $\approx 65\%$  chains in 4 hours 48 min 11 sec. All experiments were run on CPU Core i7-11700 @ 2.50GHz RAM 32Gb.

**Protocol 2.3.2** (selecting a subset of 707K+ chains in the PDB). The PDB was filtered by removing the following entries and individual chains.

- (1) 4513 non-proteins (the entity is labeled as ‘not a protein’).
- (2) 178153 disordered chains, where some atoms have occupancies  $< 1$ .
- (3) 201648 chains with residues having non-consecutive indices.
- (4) 9941 incomplete chains missing one of the main atoms  $N_i, A_i, C_i$ .
- (5) 4364 chains with non-standard amino acids.

To guarantee new condition 2.2.2(e) respecting subchains, Definition 2.3.3 will represent atoms  $N_{i+1}, A_{i+1}, C_{i+1}$  in a basis of the previous  $i$ -th residue. The first residue needs only three invariants from Definition 2.3.1 to determine the rigid class of  $\triangle N_1 A_1 C_1$  in  $\mathbb{R}^3$ . Due to cleaning in Protocol 2.3.2, all consecutive atoms along any backbone have distances  $d \geq 0.01\text{\AA}$  and all angles in any residue triangle  $\triangle N_i A_i C_i$  are at least  $3^\circ$ , which makes the bases of all residue triangles well-defined in Definition 2.3.3 below.

**Definition 2.3.3** (backbone rigid invariant  $\text{BRI}(S)$  of a protein backbone  $S$ ). In the notations of Definition 2.3.1, define the orthonormal basis vectors  $\mathbf{u}_i = \frac{\overrightarrow{A_i N_i}}{|\overrightarrow{A_i N_i}|}, \mathbf{v}_i = \frac{\mathbf{h}_i}{|\mathbf{h}_i|}$  for

$\mathbf{h}_i = \overrightarrow{A_i C_i} - b_i \overrightarrow{A_i N_i}$ ,  $b_i = \frac{\overrightarrow{A_i C_i} \cdot \overrightarrow{A_i N_i}}{|\overrightarrow{A_i N_i}|^2}$ , and  $\mathbf{w}_i = \mathbf{u}_i \times \mathbf{v}_i$ . The *backbone rigid invariant*  $\text{BRI}(S)$  is the  $m \times 9$  matrix whose  $i$ -th row for  $i = 2, \dots, m$  contains the coefficients  $x, y, z$  of the vectors  $\overrightarrow{C_{i-1} N_i}, \overrightarrow{N_i A_i}, \overrightarrow{A_i C_i}$  in the basis  $\mathbf{u}_{i-1}, \mathbf{v}_{i-1}, \mathbf{w}_{i-1}$ . So, for  $i = 2, \dots, m$ , the nine columns of  $\text{BRI}(S)$  contain the coordinates  $x(C_{i-1} N_i), y(C_{i-1} N_i), z(C_{i-1} N_i)$  of  $\overrightarrow{C_{i-1} N_i}$ , followed by the three coordinates  $x(N_i A_i), y(N_i A_i), z(N_i A_i)$  of  $\overrightarrow{N_i A_i}$  and three coordinates  $x(A_i C_i), y(A_i C_i), z(A_i C_i)$  of  $\overrightarrow{A_i C_i}$ . In the exceptional case  $i = 1$ , the first row of  $\text{BRI}(S)$  has only three non-zero coordinates  $x(N_1 A_1), x(A_1 C_1)$  and  $y(C_1) = y(A_1 C_1)$  from the first row of the invariant  $\text{TRIN}(S)$  in Definition 2.3.1.  $\blacktriangle$

For a backbone of  $m$  residues, the first row of the  $m \times 9$  matrix  $\text{BRI}(S)$  contains only three non-zero coordinates. Hence the matrix  $\text{BRI}(S)$  can be considered a vector of length  $9(m - 1) + 3 = 9m - 6$ . The simplest metric on BRIs as vectors in  $\mathbb{R}^{9m-6}$  is  $L_\infty$  equal to the maximum absolute difference between all corresponding coordinates.

A small value  $\delta$  of  $L_\infty(\text{BRI}(S), \text{BRI}(Q))$  guarantees by Theorem 2.3.9 that backbones  $S, Q$  are closely matched by rigid motion. Another metric, such as Euclidean distance or its normalisation by the chain length, has no such guarantees and can be small even for a few outliers that can affect the rigid structure and hence functional properties of a protein. Theorem 2.3.4 proves conditions 2.2.2(a,b,c,e,h) in Problem 2.2.2.

All stated results below have references to the original papers with detailed proofs.

**Theorem 2.3.4** (completeness, reconstruction, and subchains [2, Theorem 3.5]). **(a)** Under any rigid motion in  $\mathbb{R}^3$ , the matrix  $\text{TRIN}(S)$  in Definition 2.3.1 is invariant, while  $\text{BRI}(S)$  in Definition 2.3.3 is a complete invariant, so any backbones  $S, Q \subset \mathbb{R}^3$  are matched by rigid motion if and only if  $\text{BRI}(S) = \text{BRI}(Q)$ .

**(b)** For any backbone  $S$  of  $m$  residues, the invariant  $\text{BRI}(S)$ , metric  $L_\infty$  on  $\text{BRI}$ s, and a reconstruction of  $S \subset \mathbb{R}^3$  from  $\text{BRI}(S)$  can be computed in time  $O(m)$ .

**(c)** Let  $Q$  be a subchain of  $j$  consecutive residues in a backbone  $S \subset \mathbb{R}^3$ . If  $Q$  includes the first residue of  $S$ , then  $\text{BRI}(Q)$  consists of the first  $j$  rows of  $\text{BRI}(S)$ . If  $Q$  starts from the  $i$ -th residue of  $S$  for  $i > 1$ , the rows  $2, \dots, j$  of  $\text{BRI}(Q)$  coincide with the rows  $i+1, \dots, i+j-1$  of  $\text{BRI}(S)$ . The 1st row of  $\text{BRI}(Q)$  is computed from the  $i$ -th row of  $\text{BRI}(S)$  in a constant time, so  $\text{BRI}(Q)$  is computed from  $\text{BRI}(S)$  in time  $O(j)$ .  $\blacksquare$

**Corollary 2.3.5** (completeness under isometry [2, Corollary 3.6]). Any mirror image  $\bar{S}$  of a backbone  $S \subset \mathbb{R}^3$  has the invariant  $\bar{\text{BRI}}(S) := \text{BRI}(\bar{S})$  obtained by reversing the signs in all  $z$ -columns of  $\text{BRI}(S)$ . The unordered pair of  $\text{BRI}(S)$  and  $\bar{\text{BRI}}(S)$  is complete under isometry.  $\blacksquare$

Since the realisability in condition 2.2.2(f) did not appear in [2, Problem 1.2], new Lemma 2.3.6 describes the geometric realisability of non-degenerate polygonal lines in  $\mathbb{R}^3$ . The physical realisability of protein backbones will be tackled in future work.

**Lemma 2.3.6** (realisability of  $\text{BRI}$ ). A sequence of  $m$  ordered triplets of points  $N_i, A_i, C_i, i = 1, \dots, m$ , is called *non-degenerate* if the vectors  $\overrightarrow{N_i A_i}$  and  $\overrightarrow{A_i C_i}$  are not parallel for  $i = 1, \dots, m-1$ . The invariant space  $I(\text{BRI})$ , i.e. the collection of  $\text{BRI}(S) \in \mathbb{R}^{9m-6}$  for all non-degenerate sequences  $S$  of  $m$  ordered triplets  $N_i, A_i, C_i$ , consists of any sequence of numbers  $l > 0, x, y \neq 0$ , followed by  $m-1$  triples of vectors  $\mathbf{a}_i, \mathbf{b}_i, \mathbf{c}_i, i = 2, \dots, m$ , such that  $\mathbf{a}_i$  and  $\mathbf{b}_i$  are not parallel for  $i = 2, \dots, m-1$ .  $\blacksquare$

*Proof.* The first three numbers  $l = |N_i A_i|$ ,  $x = x(A_i C_i)$ ,  $y = y(A_i C_i)$  form the triangular invariant  $\text{TRIN}(S)$  from Definition 2.3.1. The realisability conditions  $l > 0$  and  $y \neq 0$  mean that the vectors  $\overrightarrow{A_i N_i}$  and  $\overrightarrow{A_i C_i}$  are not parallel and hence define the orthonormal basis  $\mathbf{u}_1, \mathbf{v}_1, \mathbf{w}_1$  associated with the first residue triangle  $\triangle N_1 A_1 C_1$ . Similarly, every next pair of vectors  $\mathbf{a}_i = \overrightarrow{N_i A_i}$  and  $\mathbf{b}_i = \overrightarrow{A_i C_i}$  should not be parallel so that we can define an orthonormal basis of the  $(i+1)$ -st residue for  $i = 2, \dots, m-1$ .  $\square$

Theorem 2.3.7 will prove the Lipschitz continuity of  $\text{BRI}$  in condition 2.2.2(c). For a given backbone  $S$  and its perturbation  $Q$ , let  $l_{N,A}$  and  $L_{N,A}$  denote the minimum and maximum bond length between any  $\alpha$ -carbon  $A_i$  and nitrogen  $N_i$  in  $S, Q$ , respectively. The maximum bond lengths  $L_{A,C}, L_{C,N}$  are similarly defined for other types of bonds.

**Theorem 2.3.7** (Lipschitz continuity of  $\text{BRI}$ , [2, Theorem 4.1]). For any  $\varepsilon > 0$ , let  $Q$  be obtained from a backbone  $S \subset \mathbb{R}^3$  by perturbing every atom of  $S$  up to Euclidean distance  $\varepsilon$ . Let  $h = \min_i |y(A_i C_i)|$  be the minimum height in triangles  $\triangle N_i A_i C_i$  at  $C_i$  for all residues in the backbones  $S, Q$ . Set  $L = \max\{L_{C,N}, L_{N,A}, L_{A,C}\}$ ,  $K = \frac{1}{l_{N,A}} + \frac{2}{h} \left(1 + 2 \frac{L_{A,C}}{l_{N,A}}\right)$ , and  $\lambda = 2(1 + 2LK)$ . Then  $L_\infty(\text{BRI}(S), \text{BRI}(Q)) \leq \lambda\varepsilon$ .  $\blacksquare$

**Example 2.3.8** (continuity of BRI). For all 707K+ cleaned chains, the median upper bound for  $\lambda$  is about 34.5, but the real values are smaller as in the example below. Consider the backbone  $S$  of the chain A (141 residues) from the standard hemoglobin 2hhb in the PDB. We perturb  $S$  to  $Q$  by adding to each coordinate  $x, y, z$  of all atoms in  $S$  some uniform noise up to various thresholds  $\varepsilon = 0.01, 0.02, \dots, 0.1\text{\AA}$ . Fig. 2.2 (top left) shows how the distance  $L_\infty(\text{BRI}(S), \text{BRI}(Q))$  averaged over 20 perturbations depends on  $\varepsilon$ . As expected by Theorem 2.3.7, the metric  $L_\infty$  is perturbed linearly up to  $\lambda\varepsilon$  for  $\lambda \approx 4$ .  $\blacklozenge$

Since the metric  $L_\infty$  between invariants  $\text{BRI}$  ( $m \times 9$  matrices) can be computed in linear time  $O(m)$ , Theorem 2.3.7 also completes condition (2.2.2f) in Problem 2.2.2. Theorem 2.3.9 will prove condition in 2.2.2(d).

**Theorem 2.3.9** (inverse continuity of BRI, [2, Theorem 4.8]). For any  $\delta > 0$  and backbones  $S, Q \subset \mathbb{R}^3$  with  $L_\infty(\text{BRI}(S), \text{BRI}(Q)) < \delta$ , there is a rigid motion  $f$  of  $\mathbb{R}^3$  such that any atom of  $S$  is  $\mu\delta$ -close to the corresponding atom of  $f(Q)$  for  $\mu = \sqrt{3} \frac{(8LK)^{m-1} - 1}{8LK - 1}$ . Let  $\widehat{\text{BRI}}(S)$  be  $\text{BRI}(S)$  after multiplying the  $i$ -th row by  $\frac{(8LK)^{i-1} - 1}{8LK - 1}$  for  $i = 2, \dots, m$ . Then  $L_\infty(\widehat{\text{BRI}}(S), \widehat{\text{BRI}}(Q)) < \delta$  guarantees a rigid motion  $f$  of  $\mathbb{R}^3$  such that any atom of the backbone  $S$  is  $\sqrt{3}\delta$ -close to the corresponding atom of  $f(Q)$ .  $\blacksquare$

## 2.4 Average invariant, diagrams, and barcodes of backbones

This section simplifies the complete invariant  $\text{BRI}$  to its average vector in  $\mathbb{R}^9$  and introduces the diagram and barcode that visually represent the high-dimensional  $\text{BRI}$ .

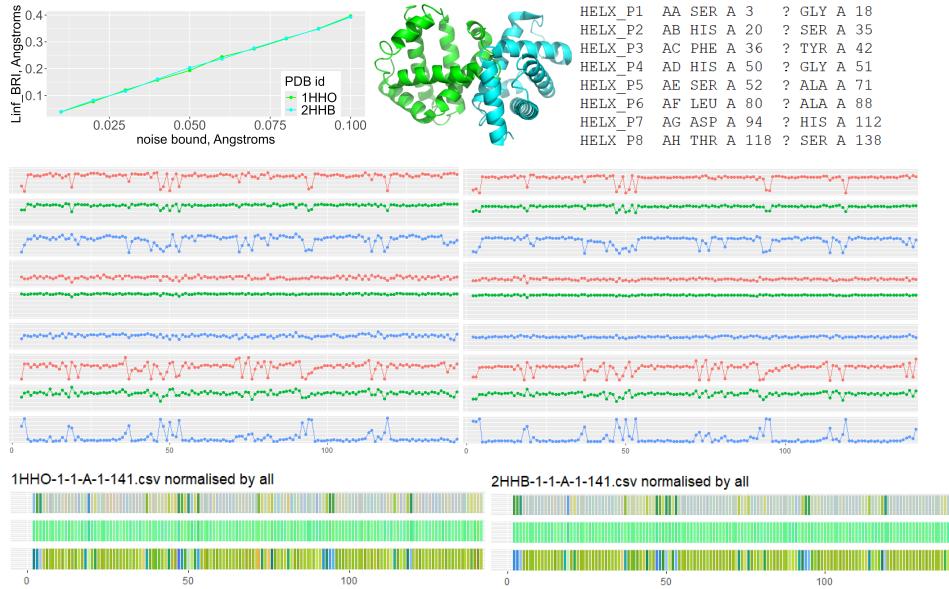
**Definition 2.4.1** (average invariant Brain, diagram BID, and barcode BIB). **(a)** For any protein backbone  $S$  of  $m$  residues, the *backbone rigid average invariant*  $\text{Brain}(S) \in \mathbb{R}^9$  is the vector of nine column averages in  $\text{BRI}(S)$  excluding the first row.

**(b)** The *backbone invariant diagram*  $\text{BID}(S)$  consists of nine polygonal curves going through the points  $(i, c(i))$ ,  $i = 2, \dots, m$ , where  $c$  is one of the coordinates (columns) of  $\text{BRI}(S)$ , see Fig. 2.2 (middle).

**(c)** For each atom type such as  $N$ , the coordinates  $(x(N_i), y(N_i), z(N_i))$  are linearly converted into the RGB color value for  $i = 1, \dots, m$ . The resulting color bars for the ordered atoms  $N, A, C$  form the *backbone invariant barcode*  $\text{BIB}(S)$ , see Fig. 2.2 (bottom).  $\blacktriangle$

**Example 2.4.2** (hemoglobins). The PDB contains thousands of hemoglobin structures. We consider here the structure 2hhb as a standard, and compare it with oxygenated 1hho, which contains an extra oxygen whose transport is facilitated by hemoglobin. In both cases, we considered the main chains (entity 1, model 1, chain A) of 141 residues.

The top left image in Fig. 2.2 (top) shows that the Lipschitz constant from Theorem 2.3.7 is  $\lambda \approx 4$  for both hemoglobins. Fig. 2.2 (middle) illustrates the complexity of



**Fig. 2.2** **Row 1:** the Lipschitz continuity of BRI from Theorem 2.3.7 is illustrated on the left by perturbing hemoglobins in Example 2.3.8, whose main chains A of 141 residues are shown in the middle (oxygenated 1hhb in green, standard 2hhb in cyan) and eight  $\alpha$ -helices found by [12] on the right. **Row 2:** the Backbone Invariant Diagram (BID) of the hemoglobins 1hhb vs 2hhb in the PDB, see Definition 2.4.1. **Row 3:** the Backbone Invariant Barcode (BIB), see Example 2.4.2.

identifying similar proteins with distant coordinates. The similarity under rigid motion becomes clear by comparing their diagrams and barcodes in Fig. 2.2 (rows 2, 3).

More importantly, a rigidly repeated pattern such as  $\alpha$ -helix or  $\beta$ -strand has constant invariants over several residue indices, which are easily detectable in BID and visible in BIB as intervals of uniform color. The PDB uses the baseline algorithm DSSP (Define Secondary Structure of Proteins) [12], which depends on several manual parameters and sometimes outputs  $\alpha$ -helices of only two residues.

For instance, the PDB entries 1hhb and 2hhb in Fig. 2.2 (right) include HELX\_P4 consisting of only residues 50 and 51, and HELX\_P5 of length 20 over residue indices  $i = 52, \dots, 71$ . Fig. 2.2 shows that a ‘constant’ interval of little noise appears only for  $i = 54, \dots, 70$ . Hence new invariants allow a more objective detection of secondary structures, which will be explored in future work.  $\blacklozenge$

## 2.5 A fast detection of duplicate chains in the Protein Data Bank

The linear time of the complete invariant  $BRI(S)$  has enabled all-vs-all comparisons for all tertiary structures in the PDB, which was additionally cleaned by Protocol 2.3.2. To speed up comparisons, Lemma 2.5.1 proves that the metric

$L_\infty(\text{BRI}(S), \text{BRI}(Q))$  between complete invariants is not smaller than the faster distance  $L_\infty(\text{Brain}(S), \text{Brain}(Q))$  between the averaged invariants (vectors of 9 coordinates) from Definition 2.4.1.

**Lemma 2.5.1** (metrics on BRI and Brain, [2, Lemma 6.1]). Any protein backbones  $S, Q$  of the same number of residues satisfy the inequality  $L_\infty(\text{Brain}(S), \text{Brain}(Q)) \leq L_\infty(\text{BRI}(S), \text{BRI}(Q))$ . ■

The complete invariants and their statistical summaries were computed in 3 hours 18 min 21 sec. After comparing all (888+ million) pairs of same-length backbones within 1 hour, we found 13907 pairs  $S, Q$  with the *exact zero-distance*  $L_\infty(\text{BRI}(S), \text{BRI}(Q)) = 0$  between complete invariants meaning that all these backbones  $S, Q$  are related by rigid motion, but they may not be geometrically identical.

However, 9366 of these pairs turned out to have  $x, y, z$  coordinates of all main atoms *identical to the last digit* despite many of them (763) coming from *different PDB entries*. Table 2.1 lists nine pairs whose geometrically identical chains unexpectedly differ in the sequences of amino acids. The duplicates from Table 2.1 were shown to the PDB validation team, who did not know about the found coincidences (in coordinates) and differences (in amino acids), because the PDB validation is currently done only for an individual protein (checking atom clashes, outliers etc).

**Table 2.1** Chains with identical backbones but different sequences of amino acid residues.

PDB id1 & chain	method and resolutions, Å	PDB id2 & chain	all atoms have identical $x, y, z$	different residues
1a0t-B	X-ray, 2.4, 2.4	1oh2-B	all $3 \times 413$	9
1ce7-A	X-ray, 2.7, 2.7	2mll-A	all $3 \times 241$	1, GLY≠HIS
1ruj-A	X-ray, 3, 3	4rhv-A	all $3 \times 237$	1, GLY≠SER
1gli-B/D	X-ray, 2.5, 1.7	3hhb-B/D	all $3 \times 146$	1, MET≠VAL
2hqe-A	X-ray, 2, 2	2o4x-A	all $3 \times 217$	1, GLN≠GLU
5adx-T	EM, 4, 8.2	5afu-Z	all $3 \times 165$	1, ILE≠VAL
5lj3-O	EM, 3.8, 10	5lj5-P	all $3 \times 252$	1, ALA≠VAL
8fdz-A	X-ray, 2.5, 2.2	8fe0-A	all $3 \times 200$	1, THR≠SER

In the row starting with 2hqe in Table 2.1, the chain IDs A, B refer to two pairs of duplicates: chain A of 2hqe is identical to chain A of 2o4x, similarly for B. The notation {B,D} in the row starting with 1gli means 4 duplicates: each of the chains B,D in 1gli is identical to each of the chains B,D in 3hhb.

The histograms in [2, Fig. 5] reveal about 220K pairs of near-duplicates among 707K+ cleaned chains up to  $L_\infty \leq 0.01\text{\AA}$ . The bound of  $0.01\text{\AA}$  is considered noise because the smallest inter-atomic distance is about 100 times larger at  $1\text{\AA} = 10^{-10} \text{ m}$ .

The physical meaning of distances follows from the bi-continuity conditions (c,d) in Problem 2.2.2. If every atom of a backbone  $S$  is shifted up to Euclidean distance  $\varepsilon$ , then  $\text{BRI}(S)$  changes up to  $\lambda\varepsilon$  in  $L_\infty$ . The Lipschitz constant  $\lambda$  was expressed in Theorem 2.3.7 and estimated as  $\lambda \approx 4$  for the hemoglobin chains in Example 2.4.2. So any small perturbation of atoms yields a small value of  $L_\infty$  in Angstroms.

The inverse Lipschitz continuity in 2.2.2(d) implies that a small Chebyshev distance  $L_\infty(\text{BRI}(S), \text{BRI}(Q)) = \delta$  guarantees that all atoms of the backbones  $S, Q$  can be matched (under a suitable rigid motion) up to Euclidean distance  $\mu\delta$  in Theorem 2.3.9.

One potential explanation of identical coordinates is the molecular replacement method [22], which uses an existing protein structure, often a previous PDB deposit or part thereof, to solve a new structure. If the newly calculated electron density map does not allow for further refinement then the coordinates may (reasonably) remain unchanged. The same coincidences can happen with lower-quality cryo-EM maps in which an existing PDB structure may be placed but where the resolution may not allow for further refinement of atomic coordinates [19, 7].

We have checked that the found duplicate backbones also have identical distance matrices on  $3m$  ordered atoms, which were slower to compute in time  $O(m^2)$  over two days on a similar machine. The widely used DALI server [8] also confirmed the found duplicates by the traditional Root Mean Square Deviation (RMSD) through optimal alignment. The DALI took about 30 min on average to find a short list of nearest neighbors of one chain in the whole PDB. Extrapolating this time to all pairwise comparisons for 707K+ cleaned chains yields 40+ years, slower by orders of magnitude than 6 hours needed for all comparisons of BRIs on the same desktop computer.

The ultra-fast speed of all-vs-all comparisons by BRI is explained by the hierarchical nature of this complete invariant. To find near-duplicates in the PDB, we first compared only average invariants  $\text{Brain}(S) \in \mathbb{R}^9$ . By Lemma 2.5.1 the full comparisons by BRI are needed only for a tiny proportion of backbones with the closest vectors  $\text{Brain}(S)$ . This hierarchical speed-up is unavailable for any distance without underlying invariants.

## References

1. Christian B Anfinsen. Principles that govern the folding of protein chains. *Science*, 181(4096):223–230, 1973.
2. Olga Anosova et al. A complete and bi-continuous invariant of protein backbones under rigid motion. *MATCH Comm. Math. Comp. Chemistry*, 94:97–134, 2025.
3. Minkyung Baek et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 373(6557):871–876, 2021.
4. Stephen K Burley, Helen M Berman, Gerard J Kleywegt, John L Markley, Haruki Nakamura, and Sameer Velankar. Protein Data Bank (PDB): the single global macromolecular structure archive. *Protein crystallography: methods and protocols*, pages 627–641, 2017.

5. Boris V Dekster and John B Wilker. Edge lengths guaranteed to form a simplex. *Archiv der Mathematik*, 49(4):351–366, 1987.
6. A Heifetz and Miriam Eisenstein. Effect of local shape modifications of molecular surfaces on rigid-body protein–protein docking. *Protein Engineering*, 16(3):179–185, 2003.
7. Maarten Hekkelman, Anastassis Perrakis, and Robbie Joosten. PDB-redo: updated and optimised crystallographic structures. <http://https://pdb-redo.eu>.
8. Liisa Holm. Dali: Protein structure comparison server. <http://ekhidna2.biocenter.helsinki.fi/dali>.
9. Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge University Press, 2012.
10. David T Jones and Janet M Thornton. The impact of AlphaFold2 one year on. *Nature methods*, 19(1):15–20, 2022.
11. John Jumper et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.
12. Wolfgang Kabsch and Christian Sander. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22(12):2577–2637, 1983.
13. David Kendall. The diffusion of shape. *Advances in Applied Probability*, 9(3):428–430, 1977.
14. David Kendall, Dennis Barden, Thomas Carne, and Huiling Le. *Shape and shape theory*. Wiley & Sons, 2009.
15. Joseph B Kruskal and Myron Wish. *Multidimensional scaling*. Number 11. Sage, 1978.
16. Leo Liberti and Carlile Lavor. *Euclidean distance geometry*. Springer, 2017.
17. Kaj Ulrik Linderstrøm-Lang. *Lane medical lectures: proteins and enzymes*, volume 6. Stanford University Press, 1952.
18. Valerio Mariani, Marco Biasini, Alessandro Barbato, and Torsten Schwede. Lddt: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics*, 29(21):2722–2728, 2013.
19. Garib Murshudov et al. Refmac5 for the refinement of macromolecular crystal structures. *Acta Cryst D*, 67(4):355–367, 2011.
20. William H Press, Saul A Teukolsky, William T Vetterling, and Brian P Flannery. *Numerical recipes: the art of scientific computing*. Cambridge University Press, 2007.
21. Stefan Rass et al. Metricizing the euclidean space towards desired distance relations in point clouds. *IEEE Trans. Information Forensics and Security*, 19:7304–7319, 2024.
22. Michael G Rossmann. The molecular replacement method. *Acta Cryst A*, 46(2):73–82, 1990.
23. Isaac Schoenberg. Remarks to Maurice Fréchet’s article “Sur la définition axiomatique d’une classe d’espaces distances vectoriellement applicable sur l’espace de Hilbert. *Annals of Mathematics*, pages 724–732, 1935.
24. L Ridgway Scott and Ariel Fernández. *A Mathematical Approach to Protein Biophysics*. Springer, 2017.
25. Hermann Weyl. *The classical groups: their invariants and representations*. Princeton Univ., 1946.
26. A Wlodawer et al. Duplicate entries in the protein data bank: how to detect and handle them. *Acta Cryst D*, 81:170–180, 2025.
27. Yang Zhang and Jeffrey Skolnick. Scoring function for automated assessment of protein structure template quality. *Proteins: Structure, Function, and Bioinform.*, 57(4):702–710, 2004.

## Chapter 3

# Complete and polynomial-time invariants of unordered points in $\mathbb{R}^n$

**Abstract** This chapter is the first in the book to focus on Euclidean clouds of unordered points under rigid motion in  $\mathbb{R}^n$ . We leverage Principal Component Analysis to construct a direction-based invariant of point clouds, whose continuity and completeness under isometry proved for principally generic clouds. This invariant is extended to a larger distribution that is complete for all clouds under rigid motion. The main novelty are polynomial-time algorithms for these invariants and distance metrics.

### 3.1 Towards complete and polynomial-time invariants for clouds

All sections in this chapter follow paper [9] with minor updates. Any finite chemical system, such as a molecule, can be represented as a cloud of atoms whose nuclei are real physical objects [12], while chemical bonds are not real sticks and only abstractly represent inter-atomic interactions. In the hardest scenario, all atoms are modelled as zero-sized points at all atomic centres without any labels such as chemical elements. For example, the  $C_{60}$  molecule [8] consists of 60 unordered carbons. Allowing different compositions enables a quantitative comparison of isomers, see Fig. 3.1.



**Fig. 3.1** Isomers of  $C_{20}$ , benzene  $C_6 H_6$ , phenyllithium  $C_6 H_5 Li$ , chlorobenzene  $C_6 H_5 Cl$  have many indistinguishable atoms.

This chapter studies finite clouds of unordered points in  $\mathbb{R}^n$  for a fixed dimension  $n$ . Problem 3.1.1 adjusts Geo-Mapping Problem 1.4.5 to unordered clouds under rigid motion. The stronger problem with Lipschitz continuity will be solved in Chapter 6.

**Problem 3.1.1** (complete and polynomial-time invariants of clouds in  $\mathbb{R}^n$ ). Design an invariant  $I$  of all clouds of unordered points in  $\mathbb{R}^n$  satisfying the conditions below.

(a) *Completeness*: any finite clouds  $A, B \subset \mathbb{R}^n$  of unordered points are related by rigid motion ( $A \cong B$ ) if and only if  $I(A) = I(B)$ .

(b) *Reconstruction*: any cloud  $A \subset \mathbb{R}^n$  of unordered points can be reconstructed from its invariant value  $I(A)$ , uniquely under rigid motion in  $\mathbb{R}^n$ .

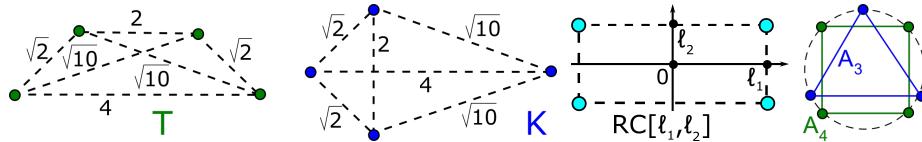
(c) *Metric*: there is a distance  $d$  on the space  $\{I(A) \mid \text{unordered clouds } A \subset \mathbb{R}^n\}$  satisfying all metric axioms in Definition 1.3.1(a).

(d) *Computability*: for a fixed dimension  $n$ , the invariant  $I(A)$ , a reconstruction of  $A \subset \mathbb{R}^n$  from  $I(A)$ , and the metric  $d(I(A), I(B))$  are computable in times that depend polynomially on the maximum size  $\max\{|A|, |B|\}$  of any clouds  $A, B \subset \mathbb{R}^n$ .  $\star$

Based on Principal Component Analysis, section 3.2 introduces the Principal Coordinates Invariant (PCI) to uniquely identify under isometry in  $\mathbb{R}^n$  all point clouds that allow a unique alignment by principal directions. Section 3.3 defines a symmetrised metric on PCIs, which is continuous under perturbations in general position and can be computed (for a fixed dimension  $n$ ) in a subquadratic time in the number of unordered points. Section 3.4 extends the PCI to the Weighted Matrices Invariant (WMI), which is complete for all point clouds under isometry in  $\mathbb{R}^n$ . Section 3.5 applies the Linear Assignment Cost and Earth Mover's Distance to define metrics on WMIs.

For a fixed dimension  $n$  of the ambient space  $\mathbb{R}^n$ , all these invariants and metrics have polynomial-time algorithms in the number  $m$  of the given points. For  $n = 2$ , the time  $O(m^{3.5} \log m)$  improves the time  $O(m^5 \log m)$  of the only previous exact algorithm [4] for the Hausdorff distance on isometry classes of clouds.

As a potential extension of the side-side-side theorem to  $m$  unordered points in  $\mathbb{R}^n$ , the seminal work [2] in 2004 proved that the total distribution of pairwise distances is a complete invariant under isometry in  $\mathbb{R}^n$  for generic clouds whose point coordinates are not solutions of a complicated polynomial equation. However, infinitely many counter-examples to the full completeness of this invariant were constructed even for  $m = 4$  points in  $\mathbb{R}^2$  [3]. The first two pictures of Fig. 3.2 show the simplest non-isometric clouds  $T \not\cong K$  of 4 points in  $\mathbb{R}^2$ . Other past work was reviewed in [9, section 2].



**Fig. 3.2 First and second**: non-isometric sets  $T \not\cong K$  of 4 points have the same 6 pairwise distances. **Third**: the vertex set  $RC[l_1, l_2]$  of a  $2l_1 \times 2l_2$  rectangle. **Fourth**: what is the distance between an equilateral triangle  $A_3$  and a square  $A_4$ ? See new invariants and metrics in Examples 3.2.4, 3.3.4, 3.5.7.

### 3.2 The Principal Coordinates Invariant of unordered clouds in $\mathbb{R}^n$

This section recalls Principal Component Analysis (PCA), whose principal directions [1] will be used to introduce the Principal Coordinates Invariant (PCI) in Definition 3.2.3. We assume that all coordinates in  $\mathbb{R}^n$  have the same units. In practice, we should first normalise all features given in different units.

Any cloud  $A \subset \mathbb{R}^n$  of  $m$  unordered points has the *centre of mass*  $O(A) = \frac{1}{m} \sum_{p \in A} p$ .

Shifting  $A$  by the vector  $-O(A)$  allows us to always assume that  $O(A)$  is the origin 0. Then Problem 3.1.1 reduces to invariants only under orthogonal maps from the orthogonal group  $O(\mathbb{R}^n)$  instead of the full Euclidean group  $E(\mathbb{R}^n)$ .

**Definition 3.2.1** (*covariance* matrix  $\text{Cov}(A)$  of a point cloud  $A$ ). If we arbitrarily order points  $p_1, \dots, p_m$  of a cloud  $A \subset \mathbb{R}^n$ , we get the sample  $n \times m$  matrix (or data table)  $P(A)$ , whose  $i$ -th column consists of  $n$  coordinates of the point  $p_i \in A$ ,  $i = 1, \dots, m$ .

The *covariance*  $n \times n$  matrix  $\text{Cov}(A) = \frac{P(A)P(A)^T}{n-1}$  is symmetric and positive semi-definite meaning that  $v^T \text{Cov}(A)v \geq 0$  for any vector  $v \in \mathbb{R}^n$ . Hence the matrix  $\text{Cov}(A)$  has real *eigenvalues*  $\lambda_1 \geq \dots \geq \lambda_n \geq 0$  satisfying  $\text{Cov}(A)\mathbf{v}_j = \lambda_j \mathbf{v}_j$  for an *eigenvector*  $\mathbf{v}_j \in \mathbb{R}^n$ , which can be scaled by any real  $s \neq 0$ .  $\blacktriangle$

If all eigenvalues of the covariance matrix  $\text{Cov}(A)$  are distinct and positive, there is an orthonormal basis of eigenvectors  $\mathbf{v}_1, \dots, \mathbf{v}_n$  ordered according to the decreasing eigenvalues  $\lambda_1 > \dots > \lambda_n > 0$ . This *eigenbasis* is unique under reflection  $\mathbf{v}_j \leftrightarrow -\mathbf{v}_j$  of each eigenvector,  $j = 1, \dots, n$ .

**Definition 3.2.2** (*principally generic* cloud). A point cloud  $A \subset \mathbb{R}^n$  is *principally generic* if, after shifting  $O(A)$  to the origin, the covariance matrix  $\text{Cov}(A)$  has distinct eigenvalues  $\lambda_1 > \dots > \lambda_n > 0$ . The  $j$ -th eigenvalue  $\lambda_j$  defines the  $j$ -th *principal direction* parallel to an eigenvector  $\mathbf{v}_j$ , which is uniquely determined under scaling.  $\blacktriangle$

The vertex set of any rectangle in  $\mathbb{R}^2$ , but not a square, is principally generic.

**Definition 3.2.3** (matrix PCM and invariant PCI). For  $n \geq 1$ , let  $A \subset \mathbb{R}^n$  be a principally generic cloud of points  $p_1, \dots, p_m$  with the centre of mass  $O(A)$  at  $0 \in \mathbb{R}^n$ . Then  $A$  has principal directions along unit eigenvectors  $\mathbf{v}_1, \dots, \mathbf{v}_n$ , defined up to a sign. In the orthonormal basis  $V = (\mathbf{v}_1, \dots, \mathbf{v}_n)^T$ , any point  $p_i \in A$  has the *principal coordinates*  $p_i \cdot \mathbf{v}_1, \dots, p_i \cdot \mathbf{v}_n$ , which can be written as a vertical column  $n \times 1$  denoted by  $Vp_i$ . The *Principal Coordinates Matrix* is the  $n \times m$  matrix  $\text{PCM}(A)$  whose  $m$  columns are the *coordinate sequences*  $Vp_1, \dots, Vp_m$ . Two such matrices are *equivalent* under changing signs of rows due to the ambiguity  $\mathbf{v}_j \leftrightarrow -\mathbf{v}_j$  of unit length eigenvectors in the basis  $V$ . The *Principal Coordinates Invariant*  $\text{PCI}(A)$  is an equivalence class of  $\text{PCM}(A)$ .  $\blacktriangle$

For simplicity, we skip the dependence on a basis  $V$  in the notation  $\text{PCM}(A)$ . The columns of  $\text{PCM}(A)$  are unordered, though we can write them according to any order of points in the cloud  $A$  considered as the vector  $(p_1, \dots, p_m)$ . Then  $\text{PCM}(A)$  can be viewed as the matrix product  $VA$  consisting of the  $m$  columns  $Vp_1, \dots, Vp_m$ . One can minimise the ambiguity under re-ordering of columns and switching signs  $\mathbf{v}_j \leftrightarrow -\mathbf{v}_j$

as follows. For each  $j = 1, \dots, n$ , choose a sign so that a coordinate with a largest value  $|p_i \cdot \mathbf{v}_j|$  is positive. Then write all columns in the lexicographically decreasing order:  $(c_{1j}, \dots, c_{nj}) > (c_{1k}, \dots, c_{nk})$  if a few first values (possibly none) coincide  $c_{ij} = c_{ik}$  and then  $c_{ij} > c_{ik}$  for the next index  $i$ .

**Example 3.2.4** (computing PCI). **(a)** For any  $l_1 > l_2 > 0$ , let the *rectangular cloud*  $\text{RC}[l_1, l_2]$  consist of the four vertices  $(\pm l_1, \pm l_2)$  of the rectangle  $[-l_1, l_1] \times [-l_2, l_2]$ . Then

$\text{RC}[l_1, l_2]$  has the centre at  $0 \in \mathbb{R}^2$  and the sample  $2 \times 4$  matrix  $P = \begin{pmatrix} l_1 & l_1 & -l_1 & -l_1 \\ l_2 & -l_2 & l_2 & -l_2 \end{pmatrix}$

whose columns are in a 1-1 correspondence with (arbitrarily) ordered points  $(l_1, l_2)$ ,

$(l_1, -l_2)$ ,  $(-l_1, l_2)$ ,  $(-l_1, -l_2)$ . The covariance matrix  $\text{Cov}(\text{RC}[l_1, l_2]) = \begin{pmatrix} 4l_1^2 & 0 \\ 0 & 4l_2^2 \end{pmatrix}$  has

eigenvalues  $\lambda_1 = 4l_1^2 > \lambda_2 = 4l_2^2$ . If we choose unit length eigenvectors  $\mathbf{v}_1 = (1, 0)$  and  $\mathbf{v}_2 = (0, 1)$ , then  $\text{PCM}(\text{RC}[l_1, l_2])$  coincides with the matrix  $P$  above. The invariant  $\text{PCI}(\text{RC}[l_1, l_2])$  is the equivalence class of all matrices obtained from  $P$  by changing signs of rows and re-ordering columns.

**(b)** The vertex set  $T$  of the trapezium in the first picture of Fig. 3.2 has four points

written in the columns of the sample matrix  $P(T) = \begin{pmatrix} 2 & 1 & -1 & -2 \\ -1/2 & 1/2 & 1/2 & -1/2 \end{pmatrix}$  so that

the centre of mass  $O(T)$  is the origin 0. Then  $\text{Cov}(T) = \begin{pmatrix} 10 & 0 \\ 0 & 1 \end{pmatrix}$  has eigenvalues 10, 1

with orthonormal eigenvectors  $(1, 0)$ ,  $(0, 1)$ , respectively. The invariant  $\text{PCI}(T)$  is the equivalence class of the matrix  $P(T)$  above. The vertex set  $K$  of the kite in the second picture of Fig. 3.2 consists of four points written in the columns of the sample matrix

$P(K) = \begin{pmatrix} 5/2 & -1/2 & -1/2 & -3/2 \\ 0 & 1 & -1 & 0 \end{pmatrix}$  so that the centre of mass  $O(K)$  is the origin 0. Then

$\text{Cov}(K) = \begin{pmatrix} 9 & 0 \\ 0 & 2 \end{pmatrix}$  has eigenvalues 9, 2 with orthonormal eigenvectors  $(1, 0)$ ,  $(0, 1)$ ,

respectively. The invariant  $\text{PCI}(K)$  is the equivalence class of the matrix  $P(K)$ .  $\blacklozenge$

All results in this chapter have details proofs in the original paper [9].

**Theorem 3.2.5** (generic completeness of PCI, [9, Theorem 3.5]). Any principally generic clouds  $A, B \subset \mathbb{R}^n$  of  $m$  unordered points are isometric if and only if their PCI invariants coincide as equivalence classes of matrices:  $\text{PCI}(A) = \text{PCI}(B)$ .  $\blacksquare$

**Lemma 3.2.6** (time complexity of PCI, [9, Lemma 3.6]). For a principally generic cloud  $A \subset \mathbb{R}^n$  of  $m$  points, a matrix  $\text{PCM}(A)$  from the invariant  $\text{PCI}(A)$  in Definition 3.2.3 can be computed in time  $O(n^2m + n^3)$ .  $\blacksquare$

Theorem 3.2.5 requires that clouds  $A, B$  are principally generic, which holds with 100% probability due to noise. If real clouds are close to symmetric configurations with equal eigenvalues, to avoid numerical instability, we should use the slower but always complete invariants from section 3.4.

### 3.3 A symmetrised metric on principally generic clouds in $\mathbb{R}^n$

This section defines a metric on PCI invariants, whose polynomial-time computation and continuity will be proved in Theorems 3.3.5 and 3.3.6. For any  $v = (x_1, \dots, x_n) \in \mathbb{R}^n$ , the *maximum norm* is  $\|v\|_\infty = \max_{i=1, \dots, n} |x_i|$ . Below we use the *Chebyshev distance*  $L_\infty(u, v) = \|\mathbf{u} - \mathbf{v}\|_\infty$  between points  $u, v \in \mathbb{R}^n$  and the bottleneck distance  $\text{BD}$  from Example 1.3.1(b) on matrices  $P$  interpreted as clouds  $[P]$  of column-vectors in  $\mathbb{R}^n$ .

**Definition 3.3.1** ( $m$ -point cloud  $[P] \subset \mathbb{R}^n$  of an  $n \times m$  matrix  $P$ ). For any  $n \times m$  matrix  $P$ , let  $[P]$  denote the unordered set of its  $m$  columns considered as vectors in  $\mathbb{R}^n$ . The set  $[P]$  of  $m$  columns can be interpreted as a cloud of  $m$  unordered points in  $\mathbb{R}^n$ .  $\blacktriangle$

For any  $n \times m$  matrices  $P, Q$ , let  $g : [P] \rightarrow [Q]$  be a bijection of columns indexed by  $1, 2, \dots, m$ . Then the Chebyshev distance  $L_\infty(v, g(v))$  between columns  $v \in [P]$  and  $g(v) \in [Q]$  is the maximum absolute difference of corresponding coordinates in  $\mathbb{R}^n$ . The minimisation over all column bijections  $g : [P] \rightarrow [Q]$  gives the bottleneck distance  $\text{BD}([P], [Q]) = \min_{g : [P] \rightarrow [Q]} \max_{v \in [P]} L_\infty(v, g(v))$  between the sets  $[P], [Q]$  considered as clouds of unordered points in  $\mathbb{R}^n$ .

An algorithm for detecting a potential isometry  $A \cong B$  will check if  $\text{SM}(A, B) = 0$  for the metric  $\text{SM}$  defined via changes of signs. A change of signs in  $n$  rows can be represented by a binary string  $\sigma$  in the product group  $\mathbb{Z}_2^n$ , where  $\mathbb{Z}_2 = \{\pm 1\}$ , 1 means no change,  $-1$  means a change.

For instance, the binary string  $\sigma = (1, -1) \in \mathbb{Z}_2^2$  acts on the matrix  $P = \text{PCM}(\text{RC}[l_1, l_2])$  from Example 3.2.4 as follows:

$$\sigma \begin{pmatrix} l_1 & l_1 & -l_1 & -l_1 \\ l_2 & -l_2 & l_2 & -l_2 \end{pmatrix} = \begin{pmatrix} l_1 & l_1 & -l_1 & -l_1 \\ -l_2 & l_2 & -l_2 & l_2 \end{pmatrix}.$$

**Definition 3.3.2** (symmetrised metric  $\text{SM}$  on matrices and clouds). For any  $n \times m$  matrices  $P, Q$ , the minimisation for  $2^n$  changes of signs represented by strings  $\sigma \in \mathbb{Z}_2^n$  acting on rows gives the *symmetrised metric*  $\text{SM}([P], [Q]) = \min_{\sigma \in \mathbb{Z}_2^n} \text{BD}([\sigma(P)], [Q])$ .

For any principally generic clouds  $A, B \subset \mathbb{R}^n$ , the *symmetrised metric* is  $\text{SM}(A, B) = \text{SM}([\text{PCM}(A)], [\text{PCM}(B)])$  for matrices  $\text{PCM}(A), \text{PCM}(B)$  in Definition 3.2.3.  $\blacktriangle$

If we denote the action of a column permutation  $g$  on a matrix  $P$  as  $g(P)$ , the matrix difference  $g(P) - Q$  has the maximum norm  $\max_{v \in [P]} L_\infty(v, g(v))$ . Then  $\text{BD}([P], [Q])$  will be computed by an efficient algorithm for bottleneck matching in Theorem 3.3.5.

**Lemma 3.3.3** (metric axioms for the symmetrised metric  $\text{SM}$ , [9, Lemma 4.4]). **(a)** The metric  $\text{SM}(P, Q)$  from Definition 3.3.2 is well-defined on equivalence classes of  $n \times m$  matrices  $P, Q$  considered under changes of signs of rows and permutations of columns, and satisfies all metric axioms.

**(b)** The metric  $\text{SM}(A, B)$  from Definition 3.3.2 is well-defined on isometry classes of principally generic clouds  $A, B$  and satisfies all axioms.  $\blacksquare$

**Example 3.3.4** (computing the symmetrised metric  $\text{SM}$ ). **(a)** By Example 3.2.4(a), the vertex set  $\text{RC}[l_1, l_2]$  of any rectangle with sides  $2l_1 > 2l_2$  in the plane has

$$\text{PCI represented by the matrix } \text{PCM}(\text{RC}[l_1, l_2]) = \begin{pmatrix} l_1 & l_1 & -l_1 & -l_1 \\ l_2 & -l_2 & l_2 & -l_2 \end{pmatrix}. \text{ The vertex set}$$

$\text{RC}[l'_1, l'_2]$  of any other rectangle has a similar matrix whose element-wise subtraction from  $\text{PCM}(\text{RC}[l_1, l_2])$  consists of  $\pm l_1 \pm l'_1$  and  $\pm l_2 \pm l'_2$ . Re-ordering columns and changing signs of rows minimises the maximum absolute value of these elements to  $\max\{|l_1 - l'_1|, |l_2 - l'_2|\}$ , which should equal  $\text{SM}(\text{RC}[l_1, l_2], \text{RC}[l'_1, l'_2])$ .

**(b)** The invariants PCI of the vertex sets  $T$  and  $K$  in Fig. 3.2 were computed in Example 3.2.4(b) and represented by these matrices in Definition 3.2.3:

$$\text{PCM}(T) = \begin{pmatrix} 2 & 1 & -1 & -2 \\ -1/2 & 1/2 & 1/2 & -1/2 \end{pmatrix}, \quad \text{PCM}(K) = \begin{pmatrix} 5/2 & -1/2 & -1/2 & -3/2 \\ 0 & 1 & -1 & 0 \end{pmatrix}.$$

The maximum absolute value of the element-wise difference of these matrices is  $|1 - (-\frac{1}{2})| = \frac{3}{2}$ , which cannot be smaller after permuting columns and changing signs of rows. The symmetrised metric equals  $\text{SM}(T, K) = \text{BD}(\text{PCM}(T), \text{PCM}(K)) = \frac{3}{2}$ .  $\blacklozenge$

**Theorem 3.3.5** (time of the metric  $\text{SM}$ , [9, Theorem 4.6]). **(a)** Given any  $n \times m$  matrices  $P, Q$ , the symmetrised metric  $\text{SM}(P, Q)$  in Definition 3.3.2 is computable in time  $O(m^{1.5}2^n \log^n m)$ . If  $n = 2$ , the time is  $O(m^{1.5} \log m)$ .

**(b)** The above conclusions hold for  $\text{SM}(A, B)$  of any principally generic  $m$ -point clouds  $A, B \subset \mathbb{R}^n$  represented by  $n \times m$  matrices  $\text{PCM}(A), \text{PCM}(B)$ .  $\blacksquare$

Theorem 3.3.6 proves the continuity in the weaker sense of  $\varepsilon, \delta$  because PCI is defined only for generic clouds anyway. Explicit estimates in the proof from [9, section 4] are based on recent bounds for perturbations of eigenvectors from [5, Theorem 3].

**Theorem 3.3.6** (continuity of SM, [9, Theorem 4.9]). For any principally generic cloud  $A \subset \mathbb{R}^n$  and any  $\varepsilon > 0$ , there is  $\delta > 0$  (depending on  $A$  and  $\varepsilon$ ) such that if any principally generic cloud  $B \subset \mathbb{R}^n$  has  $\text{BD}(A, B) < \delta$ , then  $\text{SM}(A, B) < \varepsilon$ .  $\blacksquare$

### 3.4 A complete invariant for all clouds of unordered points in $\mathbb{R}^n$

This section extends the PCI from Definition 3.2.3 to a complete WMI (Weighted Matrices Invariant) of all possible clouds, which may not be principally generic.

If a cloud  $A \subset \mathbb{R}^n$  is not principally generic, some of the eigenvalues  $\lambda_1 \geq \dots \geq \lambda_n \geq 0$  of the covariance matrix  $\text{Cov}(A)$  coincide or vanish. Let us start with the most singular case when all eigenvalues are equal to  $\lambda > 0$ . The case  $\lambda = 0$  means that  $A$  is a single point. Though  $A$  has no preferred (principal) directions,  $A$  still has the well-defined centre of mass  $O(A) = \frac{1}{m} \sum_{p \in A} p$ , which is at the origin  $0 \in \mathbb{R}^n$  as always. For  $n = 2$ , we consider  $m$  possible vectors from the origin  $0$  to every point of  $A \setminus \{0\}$ .

**Definition 3.4.1** (Weighted Matrices Invariant  $\text{WMI}(A)$  for clouds  $A \subset \mathbb{R}^2$ ). Let a cloud  $A$  of  $m$  points  $p_1, \dots, p_m$  in  $\mathbb{R}^2$  have the centre of mass at the origin  $0$ . For any point  $p_i \in A \setminus \{0\}$ , let  $\mathbf{v}_1$  be the unit length vector parallel to  $p_i \neq 0$ . Let  $\mathbf{v}_2$  be the unit length vector orthogonal to  $\mathbf{v}_1$  whose anti-clockwise angle from  $\mathbf{v}_1$  to  $\mathbf{v}_2$  is  $+\frac{\pi}{2}$ . The  $2 \times m$  matrix  $M(p_i)$  consists of the  $m$  pairs of coordinates of all points  $p \in A$

written in the orthonormal basis  $\mathbf{v}_1, \mathbf{v}_2$ , for example,  $\mathbf{p}_i = \begin{pmatrix} |\mathbf{p}_i| \\ 0 \end{pmatrix}$ . Each matrix  $M(p_i)$

is considered under re-ordering of columns. If one point  $p$  of  $A$  is the origin  $0$ , there is no basis defined by  $p = 0$ , let  $M(p)$  be the zero matrix in this centred case. If  $k > 1$  of the matrices  $M(p_i)$  are *equivalent* under re-ordering of columns, we collapse them into one matrix with the weight  $\frac{k}{m}$ . The unordered collection of the equivalence classes of  $M(p)$  with weights for all  $p \in A$  is called the *Weighted Matrices Invariant*  $\text{WMI}(A)$ .  $\blacktriangle$

In comparison with the generic case in Definition 3.2.3, for any fixed  $i = 1, \dots, m$ , if  $p_i \neq 0$ , then the orthonormal basis  $\mathbf{v}_1, \mathbf{v}_2$  is uniquely defined without the ambiguity of signs, which will re-emerge for higher dimensions  $n > 2$  in Definition 3.4.3 later. The vertex sets of regular polygons  $A_m$  have WMI consisting of a single matrix due to extra symmetries as shown below.

**Example 3.4.2** (regular clouds  $A_m \subset \mathbb{R}^2$ ). Let  $A_m$  be the vertex set of a regular  $m$ -sided polygon inscribed into a circle of a radius  $r$ , see the last picture in Fig. 3.2. Due to the  $m$ -fold rotational symmetry of  $A_m$ , the invariant  $\text{WMI}(A_m)$  consists of a single matrix

(with weight 1) whose columns are the vectors  $\begin{pmatrix} r \cos \frac{2\pi i}{m} \\ r \sin \frac{2\pi i}{m} \end{pmatrix}$ ,  $i = 1, \dots, m$ . For instance,

the vertex set  $A_3$  of the equilateral triangle has  $\text{WMI}(A_3) = \left\{ \begin{pmatrix} r & -r/2 & -r/2 \\ 0 & r\sqrt{3}/2 & -r\sqrt{3}/2 \end{pmatrix} \right\}$ .  
 The vertex set  $A_4$  of the square has  $\text{WMI}(A_4) = \left\{ \begin{pmatrix} r & 0 & 0 & -r \\ 0 & r & -r & 0 \end{pmatrix} \right\}$ .

Let  $B_m$  be obtained from  $A_m$  by adding the origin  $0 \in \mathbb{R}^2$ . Then  $\text{WMI}(B_m)$  has the matrix from  $\text{WMI}(A_m)$  with the weight  $\frac{m}{m+1}$  and the zero  $2 \times 4$  matrix with the weight  $\frac{1}{m+1}$  representing the added origin 0.  $\blacklozenge$

Definition 3.4.3 applies to all point clouds  $A \subset \mathbb{R}^n$  including the most singular case when all eigenvalues of the covariance matrix  $\text{Cov}(A)$  are equal, so we have no preferred directions at all.

**Definition 3.4.3** (Weighted Matrices Invariant WMI for any cloud  $A \subset \mathbb{R}^n$ ). Let a cloud  $A \subset \mathbb{R}^n$  of  $m$  points  $p_1, \dots, p_m$  have the centre of mass at the origin 0. For any ordered sequence of points  $p_1, \dots, p_{n-1} \in A$ , build an orthonormal basis  $\mathbf{v}_1, \dots, \mathbf{v}_n$  as follows. The first unit length vector  $\mathbf{v}_1$  is  $p_1$  normalised by its length. For  $j = 2, \dots, n-1$ , the unit length vector  $\mathbf{v}_j$  is  $p_j - \sum_{k=1}^{j-1} (p_j \cdot \mathbf{v}_k) \mathbf{v}_k$  normalised by its length.

Then every  $\mathbf{v}_j$  is orthogonal to all previous vectors  $\mathbf{v}_1, \dots, \mathbf{v}_{j-1}$  and belongs to the  $j$ -dimensional subspace spanned by  $p_1, \dots, p_j$ . Define the last unit length vector  $\mathbf{v}_n$  by its orthogonality to  $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}$  and the positive sign of the determinant  $\det(\mathbf{v}_1, \dots, \mathbf{v}_n)$  of the matrix with the columns  $\mathbf{v}_1, \dots, \mathbf{v}_n$ .

The  $n \times m$  matrix  $M(p_1, \dots, p_{n-1})$  consists of column vectors of all points  $p \in A$  in the basis  $\mathbf{v}_1, \dots, \mathbf{v}_n$ , for example,  $p_1 = (\|p_1\|_2, 0, \dots, 0)^T$ . If  $p_1, \dots, p_{n-1} \in A$  are affinely dependent, let  $M(p_1, \dots, p_{n-1})$  be the  $n \times m$  matrix of zeros in this centred case. If  $k > 1$  matrices are *equivalent* under re-ordering of columns, we collapse them into a single matrix with the weight  $\frac{k}{N}$ , where  $N = m(m-1) \dots (m-n+1)$ .

The *Weighted Matrices Invariant*  $\text{WMI}(A)$  is the unordered set of equivalence classes of matrices  $M(p_1, \dots, p_{n-1})$  with weights for all sequences of  $p_1, \dots, p_{n-1} \in A$ .  $\blacktriangle$

If  $\text{Cov}(A)$  has some equal eigenvalues,  $\text{WMI}(A)$  can be made smaller by choosing bases only for subspaces of eigenvectors with the same eigenvalue.

**Theorem 3.4.4** (completeness of WMI under rigid motion in  $\mathbb{R}^n$ , [9, Theorem 5.4]).

**(a)** Any clouds  $A, B \subset \mathbb{R}^n$  are related by rigid motion (orientation-preserving isometry) if and only if there is a bijection  $\text{WMI}(A) \rightarrow \text{WMI}(B)$  preserving all weights or, equivalently, some matrices  $P \in \text{WMI}(A)$ ,  $Q \in \text{WMI}(B)$  are related by re-ordering of columns. So  $\text{WMI}(A)$  is a complete invariant of  $A$  under rigid motion.

**(b)** Any mirror reflection  $f : A \rightarrow B$  induces a bijection  $\text{WMI}(A) \rightarrow \text{WMI}(B)$  respecting their weights and changing the sign of the last row of every matrix. This pair of WMIs is a complete invariant of  $A$  under isometry including reflections.  $\blacksquare$

It suffices to store in computer memory only one matrix  $M(p_1, \dots, p_{n-1})$  from the full  $\text{WMI}(A)$ . Any such matrix suffices to reconstruct a point cloud  $A$ , uniquely under rigid motion in  $\mathbb{R}^n$  by Theorem 3.2.5, as required in Problem 3.1.1(b). The full invariant  $\text{WMI}(A)$  can be computed from the reconstructed cloud  $A \subset \mathbb{R}^n$ .

**Lemma 3.4.5** (time of WMI, [9, Lemma 5.5]). For any cloud  $A \subset \mathbb{R}^n$  of  $m$  points and a fixed sequence of points  $p_1, \dots, p_{n-1} \in A$ , the matrix  $M(p_1, \dots, p_{n-1})$  from Definition 3.4.3 can be computed in time  $O(nm + n^3)$ . All  $N = m(m-1) \dots (m-n+1) = O(m^{n-1})$  matrices in the Weighted Matrices Invariant  $\text{WMI}(A)$  can be computed in time  $O((nm + n^3)N) = O(nm^n + n^3 m^{n-1})$ .  $\blacksquare$

### 3.5 Polynomial-time metrics for all clouds of unordered points in $\mathbb{R}^n$

This section introduces two metrics on Weighted Matrices Invariants (WMIs), which are computable in polynomial time by Theorems 3.5.3 and 3.5.6. Since any rigid motion  $f : A \rightarrow B$  induces a bijection  $\text{WMI}(A) \rightarrow \text{WMI}(B)$ , we will use a linear assignment cost [6] based on permutations of matrices.

**Definition 3.5.1** (Linear Assignment Cost LAC). Recall that Definition 3.3.2 introduced the bottleneck distance  $\text{BD}$  on matrices considered under re-ordering of columns. For any clouds  $A, B \subset \mathbb{R}^n$  of  $m$  points, consider the *Linear Assignment Cost*  $\text{LAC}(A, B) = \min_s \sum_{P \in \text{WMI}(A)} \text{BD}(P, g(P))$  minimised [6] over all bijections  $g : \text{WMI}(A) \rightarrow \text{WMI}(B)$  of full Weighted Matrices Invariants consisting of all  $N = m(m-1) \dots (m-n+1)$  equivalence classes of matrices.  $\blacktriangle$

**Lemma 3.5.2** (LAC on clouds, [9, Lemma 6.2]). **(a)** The Linear Assignment Cost from Definition 3.5.1 satisfies all metric axioms on clouds under rigid motion.

**(b)** Let  $O(A)$  be any mirror image of a cloud  $A \subset \mathbb{R}^n$ . Then  $\min\{\text{LAC}(A, B), \text{LAC}(O(A), B)\}$  is a metric on classes of clouds under isometry.  $\blacksquare$

**Theorem 3.5.3** (time complexity of LAC on WMIs, [9, Theorem 6.3]). For any clouds  $A, B \subset \mathbb{R}^n$  of  $m$  points, the invariants  $\text{WMI}(A), \text{WMI}(B)$  consists of at most  $N = m(m-1) \dots (m-n+1) = O(m^{n-1})$  matrices. Then the metric  $\text{LAC}(A, B)$  from Definition 3.5.1 can be computed in time  $O(m^{1.5}(\log^2 m)N^2 + N^3) = O(m^{2n-0.5} \log^2 m + m^{3n-3})$ . If  $n = 2$ , the time is  $O(m^{3.5} \log m)$ .  $\blacksquare$

The worst-case estimate  $N = O(m^{n-1})$  of the size (number of matrices in)  $\text{WMI}(A)$  is very rough. If the covariance matrix  $\text{Cov}(A)$  has equal eigenvalues,  $\text{WMI}(A)$  is often smaller due to extra symmetries of  $A$ .

However, for  $n = 2$ , even the rough estimate of the LAC time  $O(m^{3.5} \log m)$  improves the time  $O(m^5 \log m)$  for computing the exact Hausdorff distance between  $m$ -point clouds under Euclidean motion in  $\mathbb{R}^2$ .

Since real noise may include erroneous points, it is practically important to continuously quantify the similarity between close clouds consisting of different numbers of

points. The weights of matrices allow us to match them more flexibly via the Earth Mover's Distance [10] than via strict bijections  $\text{WMI}(A) \rightarrow \text{WMI}(B)$ . The Weighted Matrices Invariant  $\text{WMI}(A)$  can be considered as a finite distribution  $C = \{C_1, \dots, C_k\}$  of matrices (equivalent up to re-ordering columns) with weights.

**Definition 3.5.4** (Earth Mover's Distance on weighted distributions). Let  $C = \{C_1, \dots, C_k\}$  and  $D = \{D_1, \dots, D_l\}$  be finite unordered set of objects with weights  $w(C_i)$ ,  $i = 1, \dots, k$ , and  $w(D_j)$ ,  $j = 1, \dots, l$ , respectively such that  $\sum_{i=1}^k w(C_i) = 1 = \sum_{j=1}^l w(D_j)$ . Let  $d$  be a *ground metric* between any  $C_i$  and  $D_j$ . A *flow* from  $C$  to  $D$  is a  $k \times l$  matrix whose entry  $f_{ij}$  represents a *flow* from  $C_i$  to  $D_j$ . The *Earth Mover's Distance* is  $\text{EMD}(C, D) = \sum_{i=1}^k \sum_{j=1}^l f_{ij} d(C_i, D_j)$  minimised for  $f_{ij} \in [0, 1]$  subject to  $\sum_{j=1}^l f_{ij} \leq w(C_i)$ ,  $i = 1, \dots, k$ ,  $\sum_{i=1}^k f_{ij} \leq w(D_j)$ ,  $j = 1, \dots, l$ , and  $\sum_{i=1}^k \sum_{j=1}^l f_{ij} = 1$ .  $\blacktriangle$

The first condition  $\sum_{j=1}^l f_{ij} \leq w(C_i)$  means that not more than the weight  $w(C_i)$  'flows' into all objects  $D_j$  via  $f_{ij}$ ,  $j = 1, \dots, l$ . Similarly, the second condition  $\sum_{i=1}^k f_{ij} \leq w(D_j)$  means that all  $f_{ij}$  'flow' from  $C_i$ ,  $i = 1, \dots, k$  into  $D_j$  up to its weight  $w(D_j)$ .

The last condition  $\sum_{i=1}^k \sum_{j=1}^l f_{ij} = 1$  forces to 'flow' all  $C_i$  to all  $D_j$ . The EMD is a partial case of more general Wasserstein metrics [11] in transportation theory [7]. For finite distributions as in Definition 3.5.4, the metric axioms for EMD were proved in [10, appendix]. EMD can compare any weighted distributions of different sizes. Instead of the bottleneck distance BD on columns on PCM matrices, one can consider EMD on the distributions of columns (with equal weights) in these matrices.

**Lemma 3.5.5** (time complexity of EMD, [9, Lemma 6.5]). Any matrix  $P$  of a size  $n \times m(P)$  can be considered as a distribution of  $m(P)$  columns with equal weights  $\frac{1}{m(P)}$ . For two such matrices  $P, Q$  having the same number  $n$  of rows but potentially different numbers  $m(P), m(Q)$  of columns, measure the distance between any columns by the Chebyshev metric  $L_\infty$  in  $\mathbb{R}^n$ . For the matrices  $P, Q$  considered as weighted distributions of columns, the Earth Mover's Distance  $\text{EMD}(P, Q)$  can be computed in time  $O(m^3 \log m)$ , where  $m = \max\{m(P), m(Q)\}$ .  $\blacksquare$

**Theorem 3.5.6** (time of EMD on clouds, [9, Theorem 6.6]). Let clouds  $A, B \subset \mathbb{R}^n$  of up to  $m$  points have pre-computed invariants  $\text{WMI}(A), \text{WMI}(B)$  of sizes at most  $N \leq m(m-1) \dots (m-n+1) = O(m^{n-1})$ . Measure the distance between any matrices  $P \in \text{WMI}(A)$  and  $Q \in \text{WMI}(B)$  as  $\text{EMD}(P, Q)$  from Lemma 3.5.5. Then the Earth Mover's Distance  $\text{EMD}(\text{WMI}(A), \text{WMI}(B))$  from Definition 3.5.4 can be computed in time  $O(m^3(\log m)N^2 + N^3 \log N) = O((m^{2n+1} + nm^{3n-3}) \log m)$ .  $\blacksquare$

**Example 3.5.7** (EMD for a square and an equilateral triangle). Let  $A_4$  and  $A_3$  be the vertex sets of a square and equilateral triangle inscribed into the circle of a radius  $r$

in Example 3.4.2.  $\text{PCM}(A_3) = \begin{pmatrix} r & -r/2 & -r/2 \\ 0 & r\sqrt{3}/2 & -r\sqrt{3}/2 \end{pmatrix}$  and  $\text{PCM}(A_4) = \begin{pmatrix} r & 0 & 0 & -r \\ 0 & r & -r & 0 \end{pmatrix}$ .

Notice that switching the signs of the 2nd row keeps the PCI matrices the same up to permutation of columns. The weights of the three columns in  $\text{PCM}(A_3)$  are  $\frac{1}{3}$ . The weights of the four columns in  $\text{PCM}(A_4)$  are  $\frac{1}{4}$ . The EMD optimally matches the identical first columns of  $\text{PCM}(A_3)$  and  $\text{PCM}(A_4)$  with weight  $\frac{1}{4}$  contributing the cost

0. The remaining weight  $\frac{1}{3} - \frac{1}{4} = \frac{1}{12}$  of the first column  $\begin{pmatrix} r \\ 0 \end{pmatrix}$  in  $\text{PCM}(A_3)$  can be equally

distributed between the closest (in the  $L_\infty$  distance) columns  $\begin{pmatrix} 0 \\ \pm r \end{pmatrix}$  contributing the cost

$\frac{r}{12}$ . The column  $\begin{pmatrix} -r \\ 0 \end{pmatrix}$  in  $\text{PCM}(A_4)$  has equal distances  $L_\infty = \frac{r}{2}$  to the last columns

$\begin{pmatrix} -r/2 \\ \pm r\sqrt{3}/2 \end{pmatrix}$  in  $\text{PCM}(A_3)$  contributing the cost  $\frac{r}{8}$ . Finally, the distance  $L_\infty = \frac{r}{2}$  between

the columns  $\begin{pmatrix} 0 \\ \pm r \end{pmatrix}$  and  $\begin{pmatrix} -r/2 \\ \pm r\sqrt{3}/2 \end{pmatrix}$  with the common signs is counted with the weight  $\frac{5}{24}$

and contributes the cost  $\frac{5r}{48}$ . The final optimal flow  $(f_{jk})$  matrix

$$\begin{pmatrix} 1/4 & 1/24 & 1/24 & 0 \\ 0 & 5/24 & 0 & 1/8 \\ 0 & 0 & 5/24 & 1/8 \end{pmatrix}$$

gives  $\text{EMD}(\text{PCM}(A_3), \text{PCM}(A_4)) = \frac{r}{12} + \frac{r}{8} + \frac{5r}{48} = \frac{5r}{16}$ .  $\blacklozenge$

## References

1. Hervé Abdi and Lynne J Williams. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4):433–459, 2010.

2. Mireille Boutin and Gregor Kemper. On reconstructing n-point configurations from the distribution of distances or areas. *Advances in Applied Mathematics*, 32(4):709–735, 2004.
3. Terry Caelli. On generating spatial configurations with identical interpoint distance distributions. In *Proc. 7th Australian Conference on Combinatorial Mathematics*, pages 69–75, 1979.
4. Paul Chew, Michael Goodrich, Daniel Huttenlocher, Klara Kedem, Jon Kleinberg, and Dina Kravets. Geometric pattern matching under Euclidean motion. *Comp. Geom.*, 7(1-2):113–124, 1997.
5. Jianqing Fan, Weichen Wang, and Yiqiao Zhong. An eigenvector perturbation bound and its application to robust covariance estimation. *Journal of Machine Learning Research*, 18(207):1–42, 2018.
6. Roy Jonker and Anton Volgenant. A shortest augmenting path algorithm for dense and sparse linear assignment problems. *Computing*, 38(4):325–340, 1987.
7. Leonid Kantorovich. Mathematical methods of organizing and planning production. *Management science*, 6(4):366–422, 1960.
8. Harold W Kroto, James R Heath, Sean C O’Brien, Robert F Curl, and Richard E Smalley. C60: Buckminsterfullerene. *Nature*, 318(6042):162–163, 1985.
9. Vitaliy Kurlin. Polynomial-time algorithms for continuous metrics on atomic clouds of unordered points. *MATCH Comm. Math. Comp. Chemistry*, 91:79–108, 2024.
10. Yossi Rubner, Carlo Tomasi, and Leonidas Guibas. The Earth Mover’s Distance as a metric for image retrieval. *International Journal of Computer Vision*, 40(2):99–121, 2000.
11. Leonid Vaserstein. Markov processes over denumerable products of spaces, describing large systems of automata. *Probl. Per. Inf.*, 5(3):64–72, 1969.
12. Daniel Widdowson, Marco M Mosca, Angeles Pulido, Andrew I Cooper, and Vitaliy Kurlin. Average minimum distances of periodic point sets - foundational invariants for mapping all periodic crystals. *MATCH Commun. Math. Comput. Chem.*, 87:529–559, 2022.

## Chapter 4

# Persistence vs easier, faster, and stronger invariants of point clouds

**Abstract** This chapter adapts the general Geo-Mapping Problem to finite clouds under isometry in a metric space. We start by discussing Sorted Pairwise Distances, which distinguish all generic clouds under Euclidean isometry. Then we introduce the stronger invariant PDD (Pointwise Distance Distribution), which was recently proved to be complete for all 4-point clouds under isometry in any  $\mathbb{R}^n$ .

### 4.1 A dendrogram and its stable-under-noise mergegram

This section follows papers [9, 10] with slightly updated notations and connects persistent homology [8, 4, 7] in Topological Data Analysis (TDA) with a classical Minimum Spanning Tree (MST) and a new concept of a mergegram in Definition 4.1.5.

Recall that a (point) *cloud* is a finite set  $A$  of  $m$  unordered points in any metric space.

**Definition 4.1.1** (*graphs* and *metric graphs*). **(a)** An *graph* consists of a finite set of vertices  $A$  and a finite set of edges  $e$  that are unordered pairs of vertices  $p, q \in A$ , which are called *endpoints* of  $e$ . An *isomorphism* between graphs is a bijection  $f$  on vertices such that any vertices  $p, q$  are connected by the same number of edges as  $f(p), f(q)$ .

**(b)** If the vertex set  $A$  of a graph  $G$  is in a metric space  $M$  with a metric  $d_M$ , then an edge between vertices  $p, q \in A$  has the length  $d_M(p, q)$ . Then  $G$  is called a *metric graph*, considered under isometry of  $M$  that bijectively maps all vertices and edges. ▲

In general, a (metric) graph has no other structures, such as an embedding of its vertex set into  $\mathbb{R}^n$ . A (metric) graph can include *loops* (edges whose endpoints coincide) and *multiple edges* that share the same endpoints.

**Definition 4.1.2** (*paths*, *trees*, and a *Minimum Spanning Tree*  $\text{MST}(A)$  of a cloud). **(a)** In a graph  $G$  with a vertex set  $A$ , a *path* is a sequence of edges  $(p_1, p_2), (p_2, p_3), \dots, (p_{k-1}, p_m)$  for  $p_1, \dots, p_m \in A$ , where any consecutive edges share a common vertex.

**(b)** Any vertices of a graph  $G$  are called *path-connected* if there is path between them. A *connected component*  $E$  of  $G$  is any equivalence class under this relation on vertices, additionally with all edges of  $G$  between any vertices  $p, q$  of  $E$ .

**(c)** A path is called a *cycle* if the initial and final vertices coincide:  $p_1 = p_k$ , e.g. a loop is a cycle of one edge. A *tree* is a connected graph without cycles.

**(d)** For any point cloud  $A \subset M$  in a metric space, a *Minimum Spanning Tree*  $\text{MST}(A)$  is a metric tree that has the vertex set  $A$  and a minimum total length of edges. For any  $\alpha \geq 0$ , let  $\text{MST}(A; \alpha)$  denote  $\text{MST}(A)$  after removing all edges longer than  $\alpha$ .  $\blacktriangle$

If a point cloud  $A$  has pairs of points at equal distances, then a *Minimum Spanning Tree*  $\text{MST}(A)$  may not be unique under isomorphism. If  $A$  is the set of four vertices of a square in  $\mathbb{R}^2$ , then  $\text{MST}(A)$  can consist of any three of four edges of this square.

The subgraph  $\text{MST}(A; \alpha) \subset \text{MST}(A)$  may not be connected but still has no cycles.  $\text{MST}(A)$  is related to the single-linkage clustering below, which is an example of a more general dendrogram introduced below through partitions and merge sets.

**Definition 4.1.3** (the partition set  $\text{PS}(A)$ ). For any set  $A$ , a *partition* of  $A$  is a finite set of disjoint non-empty subsets  $A_1, \dots, A_k \subset A$  whose union is  $A$ . The *partition set*  $\text{PS}(A)$  consists of all partitions of  $A$ .  $\blacktriangle$

The partition set  $\text{PS}(A)$  of the abstract set  $A = \{0, 1, 2\}$  consists of the five partitions

$$(\{0\}, \{1\}, \{2\}), \quad (\{0, 1\}, \{2\}), \quad (\{0, 2\}, \{1\}), \quad (\{1, 2\}, \{0\}), \quad (\{0, 1, 2\}).$$

For example, the collections  $(\{0\}, \{1\})$  and  $(\{0, 1\}, \{0, 2\})$  are not partitions of  $A$ .

**Definition 4.1.4** (a *dendrogram*  $\Delta$  of merge sets and its *mergegram*  $\text{MG}(\Delta)$ ). A *dendrogram*  $\Delta$  over any set  $A$  is a function  $\Delta : [0, +\infty) \rightarrow \text{PS}(A)$  that depends on a *scale*  $\alpha \geq 0$  and satisfies the following conditions.

- (a)** There exists a scale  $\beta \geq 0$  such that  $\Delta(\beta)$  consists of the single set  $A$  for  $\alpha \geq \beta$ .
- (b)** If  $\alpha \leq \beta$ , then  $\Delta(\alpha)$  *refines*  $\Delta(\beta)$  so that any set from  $\Delta(\alpha)$  is a subset of some set from  $\Delta(\beta)$ . These inclusions of subsets induce the map  $\Delta[\alpha, \beta] : \Delta(\alpha) \rightarrow \Delta(\beta)$ .
- (c)** There are finitely many *merge scales*  $\alpha_i$  such that

$$\alpha_0 = 0 \text{ and } \alpha_{i+1} = \sup\{\alpha \mid \Delta[\alpha, \beta] \text{ is the identity for } \alpha' \in [\alpha_i, \alpha)\}, i = 0, \dots, m-1.$$

Since the map  $\Delta(\alpha_i) \rightarrow \Delta(\alpha_{i+1})$  is not the identity, there is a subset  $B \in \Delta(\alpha_{i+1})$  whose preimage has at least two subsets from  $\Delta(\alpha_i)$ . Then  $B \subset A$  is a *merge set* with the *birth scale*  $\alpha_i$ . All sets of  $\Delta(0)$  are merge sets at the birth scale 0. The *life*( $B$ ) is the interval  $[\alpha_i, \beta]$  from its birth scale  $\alpha_i$  to its *death scale*  $\beta = \sup\{\alpha \mid \Delta[\alpha_i, \beta](B) = B\}$ .  $\blacktriangle$

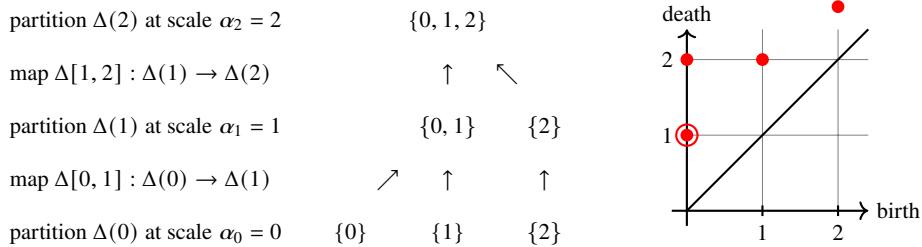
In Definition 4.1.4, condition (a) says that a partition of a set  $A$  is trivial for all large scales. Condition (b) means that if the scale  $\alpha$  is increasing, then any sets from a partition  $\Delta(\alpha)$  can only merge but cannot split into smaller subsets. Condition (c) implies that there are only finitely many mergers, when two or more subsets of  $A$  merge into a larger set. A dendrogram is often drawn as a tree whose nodes represent all sets

from the partitions  $\Delta(\alpha_i)$  at merge scales. Edges of this tree connect any set  $B \in \Delta(\alpha_i)$  with its preimages under  $\Delta(\alpha_i) \rightarrow \Delta(\alpha_{i+1})$ , see the middle picture in Fig. 4.1.

Similar to  $\text{MST}(A)$ , the isomorphism class of this tree can discontinuously change under perturbations of  $A$ . Definition 4.1.5 summarises a dendrogram into an isometry invariant of  $A$  whose Lipschitz continuity will be proved in Theorem 4.1.10.

**Definition 4.1.5** (the *mergegram*  $\text{MG}(\Delta)$  of a dendrogram  $\Delta$ ). The *mergegram*  $\text{MG}(\Delta) \subset \mathbb{R}^2$  of a dendrogram  $\Delta$  is the multiset that has one pair  $(\text{birth}, \text{death}) \in \mathbb{R}^2$  for each merge set  $B$  of  $\Delta$  with  $\text{life}(B) = [\text{birth}, \text{death}]$ . If any life interval appears  $k$  times, then the pair  $(\text{birth}, \text{death})$  has the multiplicity  $k$  in the mergegram  $\text{MG}(\Delta)$ .  $\blacktriangle$

**Fig. 4.1** A dendrogram  $\Delta$  on  $A = \{0, 1, 2\}$  and its mergegram  $\text{MG}(\Delta)$ , see Example 4.1.6.



**Example 4.1.6** (the mergegram of a dendrogram on  $A = \{0, 1, 2\}$ ). For the set  $A$  in Fig. 4.1, the partition  $\Delta(1)$  consists of  $\{0, 1\}$  and  $\{2\}$ . The maps  $\Delta[\alpha, \beta]$  induced by inclusions of sets respect the compositions in the sense that  $\Delta[\beta, \gamma] \circ \Delta[\alpha, \beta] = \Delta[\alpha, \gamma]$  for any  $\alpha \leq \beta \leq \gamma$ . For example,  $\Delta[0, 1](\{0\}) = \{0, 1\} = \Delta[0, 1](\{1\})$  and  $\Delta[0, 1](\{2\}) = \{2\}$ , so  $\Delta[0, 1]$  is a non-identity map from the partition  $\Delta(0)$  of 3 singleton sets to  $\Delta(1)$ . At the scale  $\alpha_0 = 0$  the merge sets  $\{0\}$  and  $\{1\}$  have life  $= [0, 1]$ , the merge set  $\{2\}$  has life  $= [0, 2]$ . At the scale  $\alpha_1 = 1$  the single merge set  $\{0, 1\}$  has life  $= [1, 2]$ . At the scale  $\alpha_2 = 2$  the single merge set  $\{0, 1, 2\}$  has life  $= [2, +\infty)$ . Hence, the mergegram  $\text{MG}(\Delta)$  in the last image of Fig. 4.1 consists of the pairs  $(0, 2)$ ,  $(1, 2)$ ,  $(2, +\infty)$ , which all have multiplicity 1 (shown by a red dot), and the pair  $(0, 1)$ , which has multiplicity 2 (shown by an extra red circle around a red dot).  $\blacklozenge$

Given a metric space  $(M, d)$  and a finite set  $A \subset M$ , the single-linkage dendrogram  $\Delta_{SL}\{A\}$  from Definition 4.1.7 below satisfies Definition 4.1.5(b) by [9, Lemma 3.3].

**Definition 4.1.7** (single-linkage dendrogram and mergegram  $\text{MG}(A)$  of a cloud  $A$ ). Let  $A$  be a finite set in a metric space  $M$  with a distance metric  $d : M \times M \rightarrow [0, +\infty)$ . For  $\alpha > 0$ , any points  $p, q \in A$  belong to one *single-linkage cluster* if and only if there is a finite sequence of points  $p = p_1, \dots, p_m = q \in A$  such that any two successive points have a distance at most  $\alpha$ , i.e.  $d(p_i, p_{i+1}) \leq \alpha$  for  $i = 1, \dots, m - 1$ .

Let  $\Delta_{SL}(\alpha)$  denote the set of all single-linkage clusters, which form a partition of  $A$  at the scale  $\alpha$  in the sense of Definition 4.1.5(a). The *single-linkage dendrogram* is

the map  $\Delta_{SL}\{A\} : [0, +\infty) \rightarrow \text{PS}(A)$ ,  $\alpha \mapsto \Delta_{SL}(\alpha)$ . The single-linkage mergegram  $\text{MG}(\Delta_{SL}\{A\})$  is briefly called the *mergegram*  $\text{MG}(A)$  of the cloud  $A$ .  $\blacktriangle$

[9, Algorithm 8.1] computes the mergegram  $\text{MG}(A)$  in the same asymptotic time as  $\text{MST}(A)$ . Recall that  $\alpha(n)$  is the very slow growing inverse Ackermann function [5].

**Theorem 4.1.8** (time of a mergegram, [9, Theorem 8.2]). For a fixed dimension  $n$  and any cloud  $A \subset \mathbb{R}^n$  of  $m$  unordered points, the mergegram  $\text{MG}(A)$  can be computed in the same time  $O(m \log m \alpha(m))$  as  $\text{MST}(A)$ .  $\blacksquare$

The big O notation in Theorem 4.1.8 hides some parameters that depend on the dimension  $n$ , but not on  $A$ , see details in [14].

Example 1.3.3(b) introduced the Hausdorff distance  $\text{HD}$  and the bottleneck distance  $\text{BD}$  between any sets  $A, B$  within a common metric space. Definition 4.1.9 adapts  $\text{BD}$  to multisets by considering bijections between all points repeated according to their multiplicities, which can be countable infinite, as for the set  $\mathbb{Z}$  of integers.

**Definition 4.1.9** (bottleneck distance  $\text{BD}$  between multisets in  $\mathbb{R}^2$ ). Let  $C, D$  be multisets of finitely many pairs  $(x, y) \in \mathbb{R}^2$ ,  $x < y$ , of finite multiplicities. Let  $\{x = y\}$  denote the set of diagonal pairs  $(x, x) \in \mathbb{R}^2$  of infinite multiplicity. For  $\delta \geq 0$ , a  $\delta$ -matching is a bijection  $h : C \cup \{x = y\} \rightarrow D \cup \{x = y\}$  such that  $L_\infty(h(a), a) \leq \delta$  in the Chebyshev metric for any point  $a \in C$ . The *bottleneck* distance is  $\text{BD}(C, D) = \inf\{\delta \mid \text{there is a } \delta\text{-matching } h : C \cup \{x = y\} \rightarrow D \cup \{x = y\}\}$ .  $\blacktriangle$

**Theorem 4.1.10** (stability of a mergegram, [10, Theorem 5.8]). The mergegrams of any point clouds  $A, B$  in a metric space satisfy  $\text{BD}(\text{MG}(A), \text{MG}(B)) \leq \text{HD}(A, B)$ .  $\blacksquare$

Theorem 4.1.10 implies that any small perturbation of  $A$  in the Hausdorff distance leads to a similarly small perturbation of  $\text{MG}(A)$  in the bottleneck distance  $\text{BD}$ .

## 4.2 The mergegram is stronger than the 0D persistence of a cloud

This section follows [10, section 4] and recalls the concept of 0D persistence. The main aim is to show that any 0D persistence and single-linkage dendrogram in general position can be reconstructed from the mergegram of a point cloud in any metric space.

**Definition 4.2.1** (persistence module  $\mathbb{V}$ ). A *persistence module*  $\mathbb{V}$  over the real numbers  $\mathbb{R}$  is a collection of vector spaces  $V_\alpha$ , which are indexed by  $\alpha \in \mathbb{R}$  and have linear maps  $v[\alpha, \beta] : V_\alpha \rightarrow V_\beta$ ,  $\alpha \leq \beta$  such that  $v[\alpha, \alpha]$  is the identity on  $V_\alpha$  and the composition is respected:  $v[\beta, \gamma] \circ v[\alpha, \beta] = v[\alpha, \gamma]$  for any  $\alpha \leq \beta \leq \gamma$ .  $\blacktriangle$

The set  $\mathbb{R}$  of reals can be considered as a category in the following sense. The objects of  $\mathbb{R}$  are all real numbers. Any real numbers  $a \leq b$  define a single morphism  $a \rightarrow b$ . The composition of morphisms  $a \rightarrow b$  and  $b \rightarrow c$  is the morphism  $a \leq c$ . Then a persistence module is a functor from  $\mathbb{R}$  to the category of vector spaces.

A simple example of a persistence module  $\mathbb{V}$  is an interval module below. An interval  $J$  between points  $p < q$  in the real line  $\mathbb{R}$  can be one of the following types: closed  $[p, q]$ , open  $(p, q)$ , half-open or half-closed  $[p, q)$  and  $(p, q]$ , all encoded as follows:

$$[p^-, q^+] := [p, q], \quad [p^+, q^-] := (p, q), \quad [p^+, q^+] := (p, q], \quad [p^-, q^-] := [p, q].$$

The endpoints  $p$  and  $q$  can have infinite values  $\pm\infty$ , but without superscripts.

**Example 4.2.2** (interval module  $\mathbb{I}(J)$ ). Let  $J \subset \mathbb{R}$  be an interval. The *interval module*  $\mathbb{I}(J)$  is defined by the following vector spaces  $I_\alpha$  and linear maps  $i[\alpha, \beta] : I_\alpha \rightarrow I_\beta$

$$I_\alpha = \begin{cases} \mathbb{Z}_2, & \text{for } \alpha \in J, \\ 0, & \text{otherwise;} \end{cases} \quad i[\alpha, \beta] = \begin{cases} \text{id}, & \text{for } \alpha \leq \beta \text{ within } J, \\ 0, & \text{otherwise.} \end{cases} \quad \blacklozenge$$

The *direct sum*  $\mathbb{W} = \mathbb{U} \oplus \mathbb{V}$  of persistence modules  $\mathbb{U}, \mathbb{V}$  is the persistence module with the vector spaces  $W_\alpha = U_\alpha \oplus V_\alpha$  and linear maps  $w[\alpha, \beta] = u[\alpha, \beta] \oplus v[\alpha, \beta]$ .

A persistence modules that is decomposable as a direct sum of interval modules can be described in a simple combinatorial way by a persistence diagram in  $\mathbb{R}^2$ .

**Definition 4.2.3** (persistence diagram  $\text{PD}(\mathbb{V})$ ). Let a persistence module  $\mathbb{V}$  be a finite direct sum of interval modules:  $\mathbb{V} \cong \bigoplus_{l \in L} i(x_l^*, y_l^*)$ , where  $*$  denotes a sign  $+$  or  $-$ . The *persistence diagram* of  $\mathbb{V}$  is the multiset  $\text{PD}(\mathbb{V}) = \{(x_l, y_l) \mid l \in L\} \setminus \{x = y\}$  of points above the diagonal  $\{x = y\}$  in  $\mathbb{R}^2$ , where a coordinate  $y_l$  can be  $+\infty$ .  $\blacktriangle$

Below we introduce the persistence module based on 0-dimensional homology, which suffices to define for graphs. As usual in Topological Data Analysis, we consider linear combinations with coefficients in  $\mathbb{Z}_2$ , which can be replaced with any field.

**Definition 4.2.4** (0D homology  $H_0(G)$  of a graph  $G$ ). For any graph  $G$ , the *0D homology*  $H_0(G)$  is the vector space (over  $\mathbb{Z}_2$ ) generated by connected components of  $G$ .  $\blacktriangle$

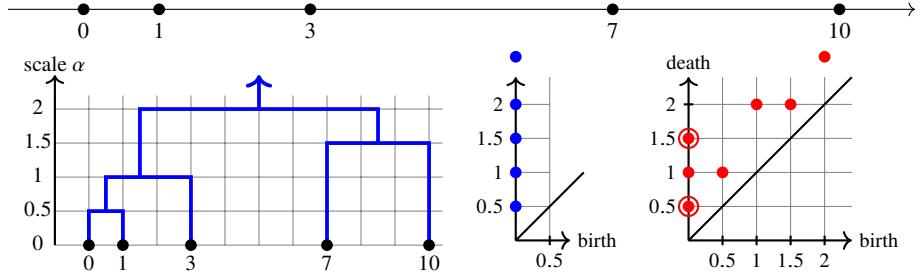
If  $G$  consists of  $m$  isolated vertices, then  $H_0(G)$  is the direct sum  $\mathbb{Z}_2^m$  of  $m$  copies of  $\mathbb{Z}_2$ . Any inclusion of graphs  $F \subset G$  induces a linear map  $H_0(F) \rightarrow H_0(G)$ .

**Definition 4.2.5** (0D persistence  $\text{PD}_0(A)$  of a point cloud  $A$ ). For any point cloud  $A$  in a metric space, the graphs  $\text{MST}(A; \alpha)$  form an ascending *filtration*  $\{\text{MST}(A; \alpha)\}_{\alpha \geq 0}$  such that  $\text{MST}(A; \alpha) \subset \text{MST}(A; \beta)$  for any  $\alpha \leq \beta$ . This filtration induces the persistence module  $H_0\{\text{MST}(A; \alpha)\}_{\alpha \geq 0}$  of vector spaces  $H_0(\text{MST}(A; \alpha))$ . The *0D persistence*  $\text{PD}_0(A)$  is the persistence diagram of the module  $H_0\{\text{MST}(A; \frac{\alpha}{2})\}_{\alpha \geq 0}$ .  $\blacktriangle$

The division by 2 in the scale  $\frac{\alpha}{2}$  is traditional for the following reasons. For any points  $p, q \in A$  at a distance  $d(p, q) = \alpha$ , the closed balls  $\bar{B}(p; r)$  and  $\bar{B}(q; r)$  with centres  $p, q$  merge when their radius  $r$  reaches  $\frac{\alpha}{2}$ . [20, Lemma 3.2] proved that  $\text{MST}(A; \alpha)$  has the minimum length among all graphs that span  $A$  within the  $\alpha$ -offset  $\bigcup_{p \in A} \bar{B}(p; \alpha)$ .

**Example 4.2.6** (0D persistence vs mergegram for a 5-point cloud). Fig. 4.2 illustrates how the 0D persistence and mergegram can be obtained from the single-linkage dendrogram for the point cloud  $A = \{0, 1, 3, 7, 10\}$  in the real line  $\mathbb{R}$ .

Imagine that we gradually blur original data points by growing disks of the same radius  $\alpha$  around the given points. The disks of the closest points 0, 1 start overlapping at the scale  $\alpha = 0.5$  when these points merge into one cluster  $\{0, 1\}$ . This merger is



**Fig. 4.2** Top: the 5-point cloud  $A = \{0, 1, 3, 7, 10\} \subset \mathbb{R}$ . Bottom from left to right: single-linkage dendrogram  $\Delta_{SL}\{A\}$  from Definition 4.1.7, the 0D persistence diagram  $PD_0\{A\}$  in Definition 4.2.3, the mergegram  $MG\{A\}$  from Definition 4.1.5, where double circles show pairs of multiplicity 2.

shown by blue arcs joined at the node at  $\alpha = 0.5$  in the single-linkage dendrogram, see the left picture at the bottom of Fig. 4.2. The persistence diagram  $PD_0\{A\}$  in the middle picture at the bottom of Fig. 4.2 represents this merger by the pair  $(0, 0.5)$  meaning that a singleton cluster of point  $1 \in A$  was born at the scale  $\alpha = 0$  and then died later at the scale  $\alpha = 0.5$  by merging into another cluster of point  $0 \in A$ .

When clusters  $\{0, 1, 3\}$  and  $\{7, 10\}$  merge at  $\alpha = 2$ , this merger was earlier encoded in the persistence diagram by the single pair  $(0, 2)$  meaning that one cluster inherited from (say) point 7 was born at  $\alpha = 0$  and died at  $\alpha = 2$ . The new mergegram in the bottom right picture of Fig. 4.2 represents the above merger by the following two pairs. The pair  $(1, 2)$  means that the cluster  $\{0, 1, 3\}$  is merging at the current scale  $\alpha = 2$  and was previously formed at the smaller scale  $\alpha = 1$ . The pair  $(1.5, 2)$  means that another cluster  $\{7, 10\}$  is merging at the scale  $\alpha = 2$  and was previously formed at  $\alpha = 1.5$ .

The 0D persistence diagram represents the cluster of the whole cloud  $A$  by the pair  $(0, +\infty)$ , because  $A$  was inherited from a singleton cluster starting from  $\alpha = 0$ . The mergegram represents the same cluster  $A$  by the pair  $(2, +\infty)$ , because  $A$  was formed during the last merger of  $\{0, 1, 3\}$  and  $\{7, 10\}$  at  $\alpha = 2$  and continues to live as  $\alpha \rightarrow +\infty$ . In  $\Delta_{SL}\{A\}$ , every vertical arc going up from a scale  $b$  to  $d$  contributes one pair  $(b, d)$  to the mergegram  $MG\{A\}$ . So both singleton clusters  $\{7\}$ ,  $\{10\}$  merging at  $\alpha = 1.5$  contribute one pair  $(0, 1.5)$  of multiplicity two shown by two red circles in Fig. 4.2. ◆

Let  $\beta$  be the length of a longest edge in a Minimum Spanning Tree  $MST(A)$ . For any  $\alpha \geq \beta$ , the graph  $MST(A; \alpha)$  coincides with  $MST(A)$  consisting of a single component. Hence, the 0D homology module  $H_0\{MST(A; \alpha)\}$  contains the interval module  $\mathbb{I}([\beta, +\infty))$  and then the 0D persistence  $PD_0\{A\}$  has the pair  $(\beta, +\infty)$ .

Lemma 4.2.7 is considered a folklore result, which is proved below, because we could not find a reference to a published proof.

**Lemma 4.2.7** ( $\Delta_{SL}\{A\}$  and  $MST(A)$  determine  $PD_0\{A\}$ ). For any point cloud  $A$  in a metric space, the 0D persistence  $PD_0\{A\}$  is uniquely determined by (a) the single-linkage dendrogram  $\Delta_{SL}\{A\}$ , and (b) a Minimum Spanning Tree  $MST(A)$ . ■

*Proof.* (a) Let  $0 < \alpha_1 < \dots < \alpha_m < \alpha_{m+1} = +\infty$  be all distinct merge scales in  $\Delta_{SL}\{A\}$  from Definition 4.1.7. If  $k \geq 2$  subsets of  $A$  merge into a larger subset of  $A$  at a scale

$\alpha_i$ , the multiplicity of  $\alpha_i$  equals  $\mu_i = k - 1$ . Then the persistence diagram  $\text{PD}_0(A)$  in Definition 4.2.5 consists of the pairs  $(0, \alpha_i)$  with multiplicities  $\mu_i$ ,  $i = 1, \dots, m + 1$ .

**(b)** Let  $\text{MST}(A)$  have edge-lengths  $\alpha_1 \leq \dots \leq \alpha_{n-1}$  with multiplicities  $\mu_i \geq 1$ . Then  $\Delta_{SL}\{A\}$  has merge scales  $\alpha_i$ , so the connected components of  $\text{MST}(A; \alpha)$  remain fixed between any successive merge scales. Then the persistence diagram  $\text{PD}_0(A)$  consists of the  $n - 1$  pairs  $(0, \frac{\alpha_i}{2})$  with multiplicities  $\mu_i$ , plus the infinite pair  $(0, +\infty)$ .  $\square$

Theorem 4.2.8 implies that the mergegram  $\text{MG}(A)$ , as an isometry invariant of a point cloud  $A$ , is at least as strong as the 0D persistence diagram  $\text{PD}_0(A)$ .

**Theorem 4.2.8** (a mergegram determines 0D persistence, [9, Theorem 5.3]). For any point cloud  $A$  in a metric space, let the mergegram  $\text{MG}(A)$  be a multiset  $\{(b_i, d_i)\}_{i=1}^k$  of pairs with multiplicities. Then the persistence diagram  $\text{PD}_0(A)$  is the difference of the multisets  $\{(0, d_i)\}_{i=1}^k - \{(0, b_i)\}_{i=1}^k$  containing each pair  $(0, \alpha)$  exactly  $\#b - \#d$  times, where  $\#b$  is the number of births  $b_i = \alpha$  and  $\#d$  is the number of deaths  $d_i = \alpha$ . All pairs  $(0, 0)$  are ignored, alternatively we take  $\{(0, d_i)\}_{i=1}^k$  only with  $d_i > 0$ .  $\blacksquare$

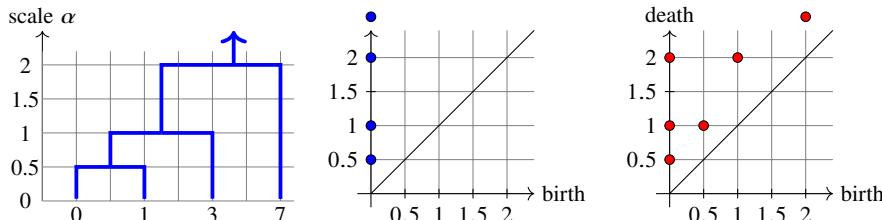
Theorem 4.2.8 is illustrated by Example 4.2.6, where  $A = \{0, 1, 3, 7, 10\}$  has the persistence diagram  $\text{PD}_0(A) = \{(0, 0.5), (0, 1), (0, 1.5), (0, 2), (0, +\infty)\}$ , which can be obtained from the mergegram

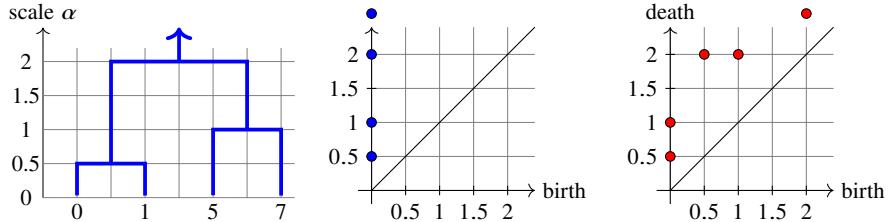
$$\text{MG}(A) = \{(0, 0.5), (0, 0.5), (0, 1), (0, 1.5), (0, 1.5), (0.5, 1), (1, 2), (1.5, 2), (2, +\infty)\}$$

as follows. The pair  $(0, 0.5) \in \text{PD}_0(A)$  of multiplicity 2 comes from two deaths and one birth scale  $\alpha = 0.5$  in  $\text{MG}(A)$ . Similarly each of the pairs  $(0, 1), (0, 1.5), (0, 2) \in \text{PD}_0(A)$  of multiplicity 2 comes from two deaths and one birth equal to one of the scales  $\alpha \in \{1, 1.5, 2\}$ .

**Example 4.2.9** (the mergegram is stronger than 0D persistence). Fig. 4.3 and 4.4 show the dendrograms, identical 0D persistence diagrams and different mergegrams for the non-isometric clouds  $A = \{0, 1, 3, 7\}$  and  $B = \{0, 1, 5, 7\}$  in  $\mathbb{R}$ . This example together with Theorem 4.2.8 justify that the mergegram  $\text{MG}(A)$  is strictly stronger than 0D persistence  $\text{PD}_0(A)$  as an isometry invariant of a point cloud  $A$ .

**Fig. 4.3** **Left:** single-linkage dendrogram  $\Delta_{SL}\{A\}$  for the cloud  $A = \{0, 1, 3, 7\} \subset \mathbb{R}$ . **Middle:** the 0D persistence diagram  $\text{PD}_0(A)$ . **Right:** the mergegram  $\text{MG}(A)$ , see Definition 4.1.5.





**Fig. 4.4** **Left:** the single-linkage dendrogram  $\Delta_{SL}\{B\}$  for the cloud  $B = \{0, 1, 5, 7\} \subset \mathbb{R}$ . **Middle:** the 0D persistence diagram  $\text{PD}_0(B)$ . **Right:** mergegram  $\text{MG}(B)$ .

Theorem 4.2.10 below can be contrasted with the weakness of 0D persistence  $\text{PD}_0(A)$  consisting of only pairs  $(0, \alpha)$  whose finite deaths are half-lengths of edges in a Minimum Spanning Tree  $\text{MST}(A)$ . In Example 4.2.6 these scales  $\alpha \in \{0.5, 1, 1.5, 2\}$  are insufficient to reconstruct the single-linkage dendrogram  $\Delta_{SL}\{A\}$  in Fig. 4.2. Such a unique reconstruction is possible by using the stronger mergegram below.

**Theorem 4.2.10** (the mergegram determines a dendrogram, [10, Theorem 4.6]). Let  $A$  be a point cloud in *general position* such that all merge scales of  $A$  in the single-linkage dendrogram  $\Delta_{SL}\{A\}$  from Definition 4.1.4 are different. Then  $\Delta_{SL}\{A\}$  is reconstructed from  $\text{MG}(A)$ , uniquely up to a permutation of nodes in  $\Delta_{SL}\{A\}$  at the scale  $\alpha = 0$ . ■

### 4.3 Generic point clouds that are indistinguishable by persistence

This section follows paper [21] and describes generic families of non-isometric clouds that have identical persistence in dimension 0 and 1, or empty 1D persistence.

Definition 4.3.1 extends the filtration of subgraphs  $\text{MST}(A; \alpha)$  of a Minimum Spanning Tree  $\text{MST}(A)$ , see Definition 4.2.5, to high-dimensional simplicial complexes.

**Definition 4.3.1** (a *filtration* of complexes  $\{C(A; \alpha)\}$ ). Let  $A$  be any finite set.

**(a)** A simplicial *complex*  $C$  on  $A$  is a finite set of subsets  $\sigma \subset A$  (called *simplices*) such that all subsets of  $\sigma \subset A$  and hence all intersections of simplices are simplices of  $C$ .

**(b)** The *dimension* of a simplex  $\sigma$  on  $k + 1$  points is  $k$ . We assume that all points of  $A$  are 0-dimensional simplices, also called *vertices* of  $C$ . A 1-dimensional simplex (or *edge*)  $e$  between points  $p, q \in A$  is the unordered pair denoted as  $[p, q]$ .

**(c)** An ascending *filtration*  $\{C(A; \alpha)\}$  is a family of complexes on the vertex set  $A$ , parametrised by a *scale*  $\alpha \in \mathbb{R}$  so that  $C(A; \alpha) \subseteq C(A; \beta)$  for any  $\alpha \leq \beta$ . ▲

Example 4.3.2 introduces the simplicial complexes  $\text{VR}(A; \alpha)$  and  $\check{\text{C}}\text{ech}(A; \alpha)$  on any finite set  $A$  inside an ambient metric space  $M$ , although  $A = M$  is possible. A Delaunay complex  $\text{Del}(A; \alpha) \subset \mathbb{R}^N$  will be defined for a finite set  $A$  only in  $\mathbb{R}^N$  because of extra complications arising if a point set  $A$  lives in a more general metric space [1].

**Example 4.3.2** (*geometric complexes* VR, Čech, and Del). Let  $A$  be any finite set in a metric space  $M$ . Fix a scale  $\alpha \geq 0$ . Each complex  $C(A; \alpha)$  below has the vertex set  $A$ .

(a) The *Vietoris-Rips* complex  $\text{VR}(A; \alpha)$  has all simplices on points  $p_1, \dots, p_k \in A$  whose pairwise distances are at most  $2\alpha$ , so  $d(p_i, p_j) \leq 2\alpha$  for  $i \neq j$  in  $\{1, \dots, k\}$ .

(b) The *Čech* complex  $\check{\text{Cech}}(A; \alpha)$  has all simplices on points  $p_1, \dots, p_k \in A$  such that the full intersection  $\cap_{i=1}^k \bar{B}(p_i; \alpha)$  is not empty.

(c) For any finite set of points  $A \subset \mathbb{R}^N$ , the convex hull of  $A$  is the intersection of all closed half-spaces of  $\mathbb{R}^N$  containing  $A$ . Each point  $p_i \in A$  has the *Voronoi domain*

$$V(p_i) = \{q \in \mathbb{R}^N \mid |q - p_i| \leq |q - p_j| \text{ for any point } p_j \in A, p_j \neq p_i\}.$$

The *Delaunay* complex  $\text{Del}(A; \alpha)$  has all simplices on points  $p_1, \dots, p_k \in A$  such that  $\cap_{i=1}^k (V(p_i) \cap \bar{B}(p_i; \alpha)) \neq \emptyset$  [6]. Alternatively, a simplex  $\sigma$  on points  $p_1, \dots, p_k \in A$  is called a *Delaunay* simplex if there is an  $(N-1)$ -dimensional sphere  $S^{N-1}$  that passes through the points  $p_1, \dots, p_k$  and does not enclose any points of  $A$  [18].

In a degenerate case, the smallest  $(k-2)$ -dimensional sphere  $S^{k-2}$  above can contain more than  $k$  points of  $A$ . If  $\sigma$  is enlarged to the convex hull  $H$  of all points in  $A \cap S^{k-2}$ , then  $\text{Del}(A; \alpha)$  becomes a polyhedral *Delaunay mosaic*.

For simplicity, we can choose any triangulation of  $H$  into Delaunay simplices. When the scale  $\alpha$  becomes too large,  $\text{Del}(A; \alpha) \subset \mathbb{R}^N$  stops growing and becomes a *Delaunay triangulation* of the convex hull of  $A$ , which is unique in general position.

The complexes of the types above will be called *geometric complexes* for brevity. ◆

Recall that all vector spaces are considered with coefficients in  $\mathbb{Z}_2$ .

**Definition 4.3.3** (*1D homology*  $H_1(C)$  of a complex). Let  $C$  be any simplicial complex. For any 1D simplex (edge)  $e$  with endpoints  $p, q$ , define the boundary operator  $\partial_1(e) = p + q$ . For any 2D simplex (face)  $\sigma$  with boundary edges  $e_1, \dots, e_m$ , define the boundary operator  $\partial_2(f) = \sum_{i=1}^m e_i$ . Both  $\partial_1, \partial_2$  extend to linear combinations by linearity.

The *1D homology*  $H_1(C)$  is the quotient space of all linear combinations  $E$  of edges with  $\partial_1(E) = 0$  over the subspace of  $\partial_2(F)$  for all linear combinations  $F$  of faces. ▲

Let a 2D simplex  $\sigma$  on 3 vertices  $v_1, v_2, v_3$  have 3 edges  $e_1, e_2, e_3$  with  $\partial_1(e_1) = v_2 + v_3$ ,  $\partial_1(e_2) = v_3 + v_1$ ,  $\partial_1(e_3) = v_1 + v_2$ . The space of 1D cycles  $E$ , which have  $\partial_1(E) = 0$  is generated by  $e_1 + e_2 + e_3$ . The space of faces generated by  $f$  maps to  $\partial_2(\sigma) = e_1 + e_2 + e_3$ . Hence, the 1D homology  $H_1(\sigma)$  is trivial as the quotient space in Definition 4.3.3. Geometrically, the only cycle  $e_1 + e_2 + e_3$  of  $\sigma$  is bounded by the triangle  $\sigma$ .

Any inclusion of complexes  $C(A; \alpha) \subset C(A; \beta)$  for  $\alpha \leq \beta$  in an ascending filtration induces a linear map in 1D homology:  $H_1(C(A; \alpha)) \rightarrow H_1(C(A; \beta))$ .

**Definition 4.3.4** (*1D persistence diagram*  $\text{PD}_1\{C(A; \alpha)\}$  of a filtration). For any filtration  $\{C(A; \alpha)\}$  of complexes on a cloud  $A$  in a metric space, a *homology* class  $\gamma \in H_1(C(A; \alpha_i))$  is *born* at  $\alpha_i = \text{birth}(\gamma)$  if  $\gamma$  is not in the full image under the induced homomorphism  $H_1(C(A; \alpha)) \rightarrow H_1(C(A; \alpha_i))$  for any  $\alpha < \alpha_i$ . The class  $\gamma$

*dies* at  $\alpha_j = \text{death}(\gamma) \geq \alpha_i$  when the image of  $\gamma$  under  $H_1(C(A; \alpha_i)) \rightarrow H_1(C(A; \alpha_j))$  merges into the image under  $H_1(C(A; \alpha)) \rightarrow H_1(C(A; \alpha_j))$  for some  $\alpha < \alpha_i$ .

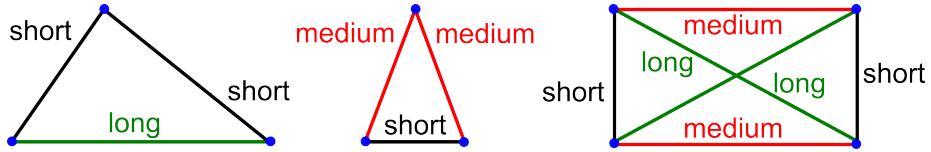
Let  $\alpha_1, \dots, \alpha_m$  be all scales when a homology class is born or dies in  $H_1(C(A; \alpha))$ . Let  $\mu_{ij}$  be the number of independent classes in  $H_1(C(A; \alpha))$  that are born at  $\alpha_i$  and die at  $\alpha_j$ . Then the 1D *persistence diagram*  $\text{PD}_1\{C(A; \alpha)\} \subset \mathbb{R}^2$  is defined as the multi-set consisting of the points  $(\alpha_i, \alpha_j)$  with integer multiplicities  $\mu_{ij} \geq 1$ .  $\blacktriangle$

**Definition 4.3.5** (short, medium, long edges in a filtration). Let  $\{C(A; \alpha)\}$  be any filtration of complexes on a finite vertex set  $A$ , see Definition 4.3.1. Let an edge  $e = [p, q]$  between points  $p, q \in A$  enter the complex  $C(A; \alpha)$  at the scale  $\alpha = d(p, q)/2$ .

(a) Consider the graph (1-dimensional skeleton)  $C'(A; \alpha)$  with vertex set  $A$  and all edges from  $C(A; \alpha)$  except the edge  $e$ . If the endpoints of  $e$  are in different connected components of  $C'(A; \alpha)$ , then the edge  $e$  is called *short* in the filtration  $\{C(A; \alpha)\}$ .

(b) The edge  $e$  is called *long* in  $\{C(A; \alpha)\}$  if  $A$  has a vertex  $v$  such that  $C(A; \alpha)$  has the 2-simplex  $\Delta pqv$  and both edges  $[p, v], [v, q]$  are in  $C(A; \alpha')$  for some  $\alpha' < \alpha$ .

(c) If  $e$  is neither short nor long, then the edge  $e$  is called *medium* in  $\{C(A; \alpha)\}$ .  $\blacktriangle$



**Fig. 4.5** Classes of edges by Definition 4.3.5 in Example 4.3.6.

**Example 4.3.6** (classes of edges on 3 and 4 points). (a) For any 3-point cloud  $A \subset \mathbb{R}^N$ , let the edges of  $A$  have lengths  $|e_1| \leq |e_2| < |e_3|$ . By Definition 4.3.5, in  $\{\text{VR}(A; \alpha)\}$  the edge  $e_3$  is long whilst the edges  $e_1, e_2$  are short, see Fig. 4.5 (left). If  $|e_1| < |e_2| = |e_3|$ , then the edge  $e_1$  is short but both edges  $e_2, e_3$  are medium, not long. If  $|e_1| = |e_2| = |e_3|$ , then all three edges are medium.

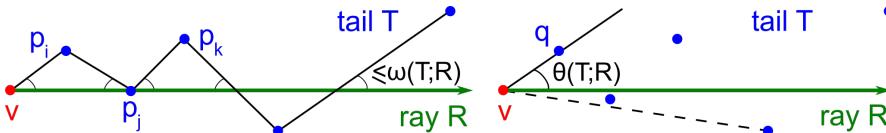
(b) Let  $C(A; \alpha)$  be any geometric complex from Example 4.3.2 on a point cloud  $A \subset \mathbb{R}^2$ . If  $A$  consists of four vertices of the unit square, all square sides are medium whilst both diagonals are long, see Fig. 4.5 (right). If  $A$  consists of four vertices of a rectangle that is not a square, the two shorter sides are short, the longer sides are medium and both diagonals are long.  $\blacklozenge$

**Definition 4.3.7** (a tail of points). For a fixed filtration  $\{C(A; \alpha)\}$  on a point cloud  $A$  from Definition 4.3.1, a *tail*  $T$  in a metric space is any ordered sequence  $T = \{p_1, \dots, p_n\}$ , where any edge  $[p_i, p_{i+1}]$  between successive points is short, and any edge  $[p_i, p_j]$  between non-successive points is long for all  $1 \leq i < j \leq n$ .  $\blacktriangle$

**Definition 4.3.8** (angular deviation  $\omega(T; R)$  and angular thickness  $\theta(T; R)$ ). In  $\mathbb{R}^N$ , a *ray* is any half-infinite line  $R$  going from a point  $v$  (the *vertex* of  $R$ ). Let  $T = \{p_1, \dots, p_n\}$  be any sequence of ordered points in  $\mathbb{R}^N$ .

(a) The *angular deviation*  $\omega(T; R)$  of the tail  $T$  with respect to the ray  $R$  is the maximum angle  $\angle(R, [p, q]) \in [0, \frac{\pi}{2}]$  over all distinct points  $p, q \in T$ .

(b) The *angular thickness*  $\theta(T; R)$  of the tail  $T$  with respect to the ray  $R$  is the maximum angle  $\angle(R, [p_1, p_i])$  for  $i = 2, \dots, n$ .  $\blacktriangle$



**Fig. 4.6** A tail  $T$  around a ray  $R$  with vertex  $v$  in  $\mathbb{R}^2$ , see Definition 4.3.8. **Left:** all marked angles are not greater than the angular deviation  $\omega(T; R)$ . **Right:** the angular thickness  $\theta(T; R)$  can be smaller than the angular deviation  $\omega(T; R)$ .

**Lemma 4.3.9** (tails in  $\mathbb{R}^N$ , [21, Lemma 3.5]). In  $\mathbb{R}^N$ , let  $R$  be a ray with a vertex  $v = p_1$  and  $T$  be any sequence of points  $p_1, \dots, p_n$  with an angular deviation  $\omega(T; R) < \frac{\pi}{4}$ .

(a) For any  $i < j < k$ , the angle  $\angle p_i p_j p_k$  is non-acute. The edge between the non-successive points  $p_i, p_k$  is long in any filtration  $\{C(T; \alpha)\}$  in Example 4.3.2.

(b) Any edge between successive points  $p_{j-1}, p_j, j = 2, \dots, n$ , is short in  $\{C(T; \alpha)\}$ .

Hence  $T$  has no medium edges in  $\{C(T; \alpha)\}$  and is a tail by Definition 4.3.7.  $\blacksquare$

**Theorem 4.3.10** (adding a tail to a long wedge keeps 1D persistence, [21, Theorem 4.4]). Let  $A \subset \mathbb{R}^N$  be a finite set,  $v \in A$  be on the boundary of the convex hull of  $A$ , and  $R$  be a ray with a vertex  $v$  so that  $\mu(R; A) = \min_{p \in A - \{v\}} \angle(R, [v, p]) \geq \frac{\pi}{2}$ . Let  $T$  be a tail with

the vertex  $v$  such that  $\mu(R; A) \geq \theta(T; R) + \frac{\pi}{2}$ . Assume that  $A \cup T$  is in general position such that every subset  $P \subset A \cup T$  of at most  $N + 1$  points is affinely independent, and no point of  $(A \cup T) \setminus P$  lies on the smallest  $(N - 1)$ -dimensional circumsphere of  $P$ . For any filtration from Example 4.3.2, we have that  $\text{PD}_1\{C(A \cup T; \alpha)\} = \text{PD}_1\{C(A; \alpha)\}$ .  $\blacksquare$

## 4.4 Geo-mapping problem for generic clouds in a metric space

This section continues to study finite clouds of unordered points under isometry in any metric space, and  $\mathbb{R}^n$  remains an important partial case. Problem 4.4.4 adjusts Geo-Mapping Problem 1.4.5 to include only realistically achievable conditions in an arbitrary metric space. One realistic condition is a general position, as formally defined below.

**Definition 4.4.1** (general position in a metric space). Let  $A$  be a set of  $m$  points in a metric space  $M$  with a metric  $d_M$ . A cloud  $A$  is called *generic* (or in a *general position* in  $M$ ) if all  $\frac{m(m-1)}{2}$  inter-point distances  $d_M(p, q)$  for  $p, q \in A$  are not solutions of a certain polynomial equation.  $\blacktriangle$

To give an example of a general position in Definition 4.4.1, recall the the *lexicographic order* on ordered pairs:  $(i, j) < (k, l)$  if  $i < j$  or  $i = j$  and  $k < l$ .

**Example 4.4.2** (a polynomial for a general position). Let  $p_1, \dots, p_m$  be all points of a set  $A$  in a space  $M$  with a metric  $d_M$ . Set  $P(A) = \prod_{(i,j) \neq (k,l)} (d_{i,j} - d_{k,l})$ , where  $d_{i,j} = d_M(p_i, p_j)$  for  $1 \leq i < j \leq m$ . If  $m = 3$ , then  $P(A) = (d_{1,2} - d_{1,3})(d_{1,2} - d_{2,3})(d_{1,3} - d_{2,3})$ . By Definition 4.4.1, the polynomial condition  $P(A) \neq 0$  describes the general position for all clouds  $A \subset M$  that have distinct distances.  $\blacklozenge$

The concept of a general position allows us to weaken the concept of a complete invariant from Definition 1.2.5(b) to a generically complete invariant below.

**Definition 4.4.3** (generically complete invariant). Let  $\sim$  be an equivalence relation on finite sets in a metric space  $M$ . An invariant  $I$  under this equivalence is *generically complete* if the implication  $I(A) = I(B) \Rightarrow A \sim B$  holds for all sets  $A, B \subset M$  in a general position for a certain polynomial on inter-point distances in Definition 4.4.1.  $\blacktriangle$

Instead of the full completeness in Problem 1.4.5, Problem 4.4.4 asks for a more realistic generic completeness under isometry in the sense of Definition 4.4.3.

**Problem 4.4.4** (geo-mapping for generic clouds under isometry in a metric space). For a space  $M$  with a metric  $d_M$ , find an isometry invariant  $I$  of generic clouds of unordered points in  $M$  with values in a metric space satisfying the following conditions.

**(a) Generic completeness:** any generic clouds  $A, B \subset M$  are isometric in  $M$ , i.e.  $A \simeq B$ , if and only if  $I(A) = I(B)$ .

**(b) Reconstruction:** any generic cloud  $A \subset M$  can be reconstructed from its invariant value  $I(A)$ , uniquely under isometry in  $M$ .

**(c) Metric:** there is a distance  $d$  on the *invariant space*  $\{I(A) \mid A \subset M\}$  satisfying all metric axioms in Definition 1.3.1(a).

**(d) Continuity:** there is a constant  $\lambda$  such that, for any  $\varepsilon > 0$ , if  $B$  is obtained from  $A$  by perturbing every point of  $A$  up to  $\varepsilon$  in the metric  $d_M$ , then  $d(I(A), I(B)) \leq \lambda\varepsilon$ .

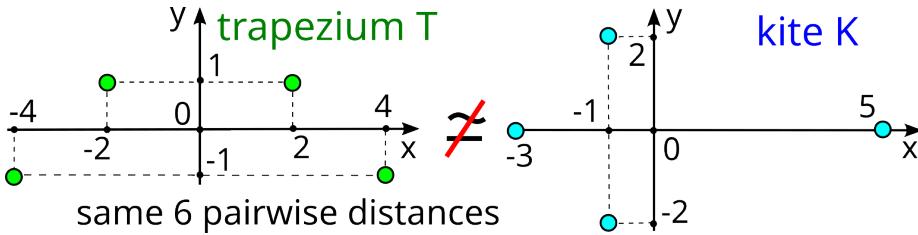
**(e) Computability:** for a fixed metric space  $M$ , the invariant  $I(A)$ , the metric  $d(I(A), I(B))$ , and a reconstruction of  $A \subset M$  from  $I(A)$  can be computed in a time that depends polynomially on the maximum size  $\max\{|A|, |B|\}$  of clouds  $A, B$ .  $\star$

Definition 4.4.5 introduces the invariant that nearly solved Problem 4.4.4 in 2004 by [2, Theorem 2.6] under Euclidean isometry in  $\mathbb{R}^n$ . Though this seminal work [2] talks about reconstructing point configurations, the main result actually proves the generic completeness of the following invariant under isometry in  $\mathbb{R}^n$  as stated in 4.4.4(a).

**Definition 4.4.5** (Sorted Pairwise Distances SPD). For any finite cloud  $A$  of unordered points in a metric space  $M$ , the vector  $\text{SPD}(A)$  of *Sorted Pairwise Distances* consists of all  $\frac{m(m-1)}{2}$  distances between all points of  $A$ , written in increasing order.  $\blacktriangle$

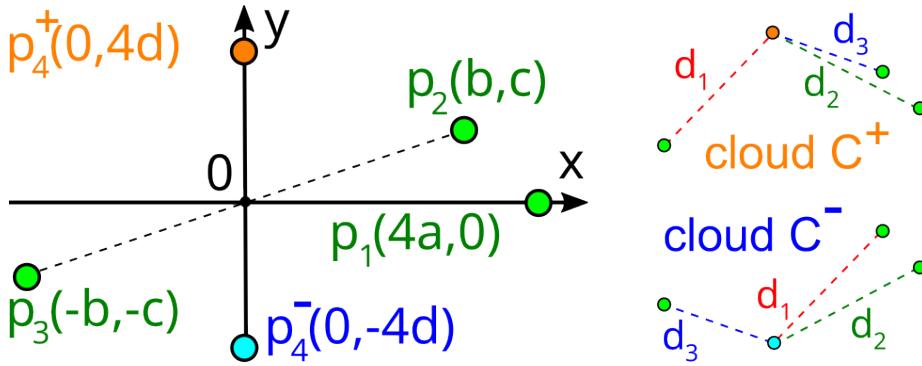
Any isometry in a metric space  $M$  preserves distances and hence  $\text{SPD}(A)$  for any cloud  $A \subset M$ . In the Euclidean case  $M = \mathbb{R}^n$ , if  $A$  consists of  $m = 3$  points,  $\text{SPD}(A)$

coincides with the geocode of three inter-point distances, which classified all triangles under isometry in Example 1.5.3(a). For a cloud  $A \subset \mathbb{R}^n$  of any  $m$  unordered points, [2, Theorem 2.6] proved that  $\text{SPD}(A)$  is generically complete under isometry in  $\mathbb{R}^n$ . We leave as an exercise that  $\text{SPD}(A)$  is Lipschitz continuous, for example, in any Minkowski metric  $L_q$ , because this Lipschitz continuity will be proved for stronger invariants in the next section. Since  $\text{SPD}(A)$  needs a quadratic time of the size  $|A| = m$ , this invariant solves Problem 4.4.4 for generic clouds in  $\mathbb{R}^n$ .



**Fig. 4.7** Non-isometric clouds of 4 points with the same 6 pairwise distances. **Left:** the trapezium  $T$  has the vertices  $(\pm 2, 1)$ ,  $(\pm 4, -1)$ . **Right:** the kite  $K$  has the vertices  $(5, 0)$ ,  $(-3, 0)$ ,  $(-1, \pm 2)$ .

However, infinitely many counterexamples to the completeness of  $\text{SPD}$  have been known at least since 1979 [3] even for  $m = 4$  points in  $\mathbb{R}^2$ . Fig. 4.7 shows the most famous pair of a trapezium and a kite, which inspired the flagship image of Geometric Data Science in Fig. 1.4 (right). Fig. 4.8 illustrates infinitely many non-isometric 4-point clouds in  $\mathbb{R}^2$ , which share three points  $p_1, p_2, p_3$  in green and differ only in points  $p_4^\pm$ , but share 3 distances  $d_1, d_2, d_3$  from  $p_4^\pm$  to three others.



**Fig. 4.8** Infinitely many non-isometric clouds  $C^+ \neq C^-$  depending on free parameters  $a, b, c, d > 0$  [3].

We can leverage the Euclidean structure of  $\mathbb{R}^n$  to introduce a simpler invariant below.

**Definition 4.4.6** (Sorted Radial Distances SRD). For any finite cloud  $A$  of unordered points in  $\mathbb{R}^n$ , a translation can fix the centre of mass  $\bar{A}$  of  $A$  at the origin  $0 \in \mathbb{R}^n$ . The vector  $\text{SRD}(A)$  of *Sorted Radial Distances* consists of all  $m$  Euclidean distances from  $\bar{A} = 0$  to all points of  $A$ , written in decreasing order.  $\blacktriangle$

The orders of distances in Definitions 4.4.5 and 4.4.6 are motivated by applications to molecules. The most characteristic inter-atomic distances are the smallest ones between bonded atoms at the beginning of the SPD, written in increasing order.

On another hand, the simpler invariant  $\text{SRD}(A)$  describes the global shape of a molecule by the largest distances to atoms from the centre of mass.

**Example 4.4.7** (SPD and SRD for 4-point clouds in Fig. 4.7). The vertex sets  $T, K$  of the trapezium and kite in Fig. 4.7 have  $\text{SPD} = (2\sqrt{2}, 2\sqrt{2}, 4, 2\sqrt{10}, 2\sqrt{10}, 8)$ , but are distinguished by  $\text{SRD}(T) = (\sqrt{17}, \sqrt{17}, \sqrt{5}, \sqrt{5})$  and  $\text{SRD}(K) = (5, 3, \sqrt{5}, \sqrt{5})$ .  $\blacklozenge$

The next two sections follow the finite (non-periodic) case of paper [25].

## 4.5 Pointwise Distance Distributions of unordered points

This section defines our main isometry invariant, which we first introduced for periodic point sets [24] in 2022 and only after that understood its importance in the finite case, where it was previously studied under the name of a *local distribution of distances* [15].

**Definition 4.5.1** (Pointwise Distance Distribution  $\text{PDD}(A; k)$  for a finite cloud  $A$ ). Let  $A = \{p_1, \dots, p_m\}$  be a finite cloud of unordered points in a metric space  $M$ .

Fix an integer  $k \geq 1$ . For every point  $p \in A$ , let  $d_1(p) \leq \dots \leq d_k(p)$  be the distances from  $p$  to its  $k$  nearest neighbours in  $A$ . The matrix  $D(A; k)$  has  $m$  rows consisting of the distances  $d_1(p_i), \dots, d_k(p_i)$  for  $i = 1, \dots, m$ . If any  $l \geq 2$  rows coincide, we collapse them into a single row with the weight  $\frac{l}{m}$ .

The resulting matrix of maximum  $m$  rows and  $k + 1$  columns, including the extra column of weights, is called the *Pointwise Distance Distribution*  $\text{PDD}(A; k)$ .  $\blacktriangle$

The rows of  $\text{PDD}(A; k)$  are unordered, though we might write them in a lexicographic order only for convenience. Hence  $\text{PDD}(A; k)$  can be considered a weighted distribution of (say)  $m$  rows of  $k$  distances. Each row can also be interpreted as a point in  $\mathbb{R}^k$ .

Then  $\text{PDD}(A; k)$  can be viewed as a cloud of  $m$  unordered points in  $\mathbb{R}^k$ . The crucial difference with the original cloud  $A$  under isometry in  $\mathbb{R}^n$  is the fixed coordinate system for  $\text{PDD}(A; k) \subset \mathbb{R}^k$ , not under any equivalence.

**Example 4.5.2** (PDD for 4-point clouds  $T, K$  in Fig. 4.7). Table 4.1 shows the  $4 \times 3$  matrices  $D(S; 3)$  from Definition 4.5.1. The matrix  $D(T; 3)$  in Table 4.1 has two pairs of identical rows, so the matrix  $\text{PDD}(T; 3)$  consists of two rows of weight  $\frac{1}{2}$  below.

The matrix  $D(K; 3)$  in Table 4.1 has only one pair of identical rows, so  $\text{PDD}(K; 3)$  has three rows of weights  $\frac{1}{2}, \frac{1}{4}, \frac{1}{4}$ . Then  $T, K$  are distinguished by PDDs even for  $k = 1$ .

**Table 4.1** Each point of  $T, K \subset \mathbb{R}^2$  in Figure 4.7 has three distances to other points in increasing order. After keeping only distances (not neighbours), the resulting PDDs distinguish  $T \neq K$ .

points of $T$	distance to neighbour 1	distance to neighbour 2	distance to neighbour 3
(-4, -1)	$2\sqrt{2}$ to (-2, +1)	$2\sqrt{10}$ to (+2, +1)	8 to (+4, -1)
(+4, -1)	$2\sqrt{2}$ to (+2, +1)	$2\sqrt{10}$ to (-2, +1)	8 to (-4, -1)
(-2, +1)	$2\sqrt{2}$ to (-4, -1)	8 to (+2, +1)	$2\sqrt{10}$ to (+4, -1)
(+2, +1)	$2\sqrt{2}$ to (+4, -1)	4 to (-2, +1)	$2\sqrt{10}$ to (-4, -1)
points of $K$	distance to neighbour 1	distance to neighbour 2	distance to neighbour 3
(-3, 0)	$2\sqrt{2}$ to (-1, -2)	$2\sqrt{2}$ to (-1, +2)	8 to (5, 0)
(+5, 0)	$2\sqrt{10}$ to (-1, -2)	$2\sqrt{10}$ to (-1, +2)	8 to (-3, 0)
(-1, -2)	$2\sqrt{2}$ to (-3, 0)	4 to (-1, +2)	$2\sqrt{10}$ to (5, 0)
(-1, +2)	$2\sqrt{2}$ to (-3, 0)	4 to (-1, -2)	$2\sqrt{10}$ to (5, 0)

$$\text{PDD}(T) = \begin{pmatrix} 1/2 & 2\sqrt{2} & 4 & 2\sqrt{10} \\ 1/2 & 2\sqrt{2} & 2\sqrt{10} & 8 \end{pmatrix} \neq \text{PDD}(K) = \begin{pmatrix} 1/4 & 2\sqrt{2} & 2\sqrt{2} & 8 \\ 1/2 & 2\sqrt{2} & 4 & 2\sqrt{10} \\ 1/4 & 2\sqrt{10} & 2\sqrt{10} & 8 \end{pmatrix}. \quad \blacklozenge$$

Since any isometry preserves distances,  $\text{PDD}(A; k)$  is an isometry invariant of  $A$ . The brute-force algorithm for  $\text{PDD}(A; k)$  needs only a quadratic time in the size  $|A| = m$ . In a general metric space with certain expansion constants, we found counter-examples [11] to past estimates for a parametrised complexity a nearest neighbour search and proved new linear-time complexities [12] with extra parameters depending, for example, on a dimension  $n$ . Hence, the invariant  $\text{PDD}(A; k)$  satisfies the computability in 4.4.4(e).

Interpreting  $\text{PDD}(A; k)$  as a discrete distribution of rows (or points in  $\mathbb{R}^k$ ) with weights as probabilities allows us to compare PDDs by many metrics on probability distributions. If we use the Earth Mover's Distance from Definition 3.5.4 with a ground metric  $L_q$  on rows of PDD, we denote the resulting metric by  $\text{EMD}_q$  for all parameters  $q \in [1, +\infty]$ . For  $\text{PDD}(A; k)$ , the notation  $\text{EMD}$  without any subscript means that the ground metric is the Root Mean Square RMS =  $\frac{L_2}{\sqrt{k}}$ .

The EMD satisfies all metric axioms [16, appendix], needs  $O(m^3 \log m)$  time for distributions of a maximum size  $m$ , and can be approximated in  $O(m)$  time [19].

**Theorem 4.5.3** (Lipschitz continuity of PDD for a finite cloud, [25, Theorem 4.2(a)]). Let  $A$  be a finite cloud in a space  $M$  with a metric  $d_M$ . For any  $\varepsilon > 0$ , let  $B$  be obtained from  $A$  by perturbing every point of  $A$  up to  $\varepsilon$  in the metric  $d_M$ . Fix any real

$q \in [1, +\infty]$  and an integer  $k \geq 1$ . Interpret  $\sqrt[k]{k}$  as 1 in the limit case  $q = +\infty$ . Then  $\text{EMD}_q(\text{PDD}(A; k), \text{PDD}(B; k)) \leq 2\epsilon \sqrt[k]{k}$ .  $\blacksquare$

For any cloud  $A \subset \mathbb{R}^n$  of  $m$  unordered points, the vector  $\text{SPD}(A)$  of Sorted Pairwise Distances obtained from  $\text{PDD}(A; m-1)$  by writing all distances in a single distribution and collapsing each pair of equal distances into one. Indeed, any distance  $|p_i - p_j|$  appears in both rows  $i, j$  of  $\text{PDD}(A; m-1)$ . Due to Example 4.5.2,  $\text{PDD}(A; m-1)$  is strictly stronger than  $\text{SPD}(A)$ . Due to this strength, the generic completeness of  $\text{PDD}(A; m-1)$  under isometry in  $\mathbb{R}^n$  is much easier to prove than for  $\text{SPD}(A)$ .

**Theorem 4.5.4** (generic completeness of PDD for a finite cloud, [25, Theorem 5.1]). Any cloud  $A \subset \mathbb{R}^n$  of  $m$  unordered points with distinct inter-point distances can be reconstructed from  $\text{PDD}(A; m-1)$ , uniquely under isometry.

*Proof.* Since all inter-point distances are distinct, every such distance  $|p - q|$  between points  $p, q \in A$  appears twice in  $\text{PDD}(A; m-1)$ : once in the row of  $p$  and once in the row of  $q$ . Hence, after choosing an arbitrary order of points, we can use  $\text{PDD}(A; m-1)$  to reconstruct the classical distance matrix on ordered points. This distance matrix determines  $A \subset \mathbb{R}^n$  uniquely under isometry [17].  $\square$

The following open conjecture should be understandable to schools students.

**Conjecture 4.5.5** (completeness of PDD in  $\mathbb{R}^2$ ). Any cloud  $C \subset \mathbb{R}^2$  of  $m$  unordered points can be reconstructed from  $\text{PDD}(C; m-1)$ , uniquely under isometry in  $\mathbb{R}^2$ .  $\star$

In  $\mathbb{R}^3$ , the known non-isometric clouds with the same PDD inspired the stronger invariant, which will distinguish all these examples in the next chapter. In a general metric space, Problem 4.4.4 is notoriously hard, but provides targets for further research.

## 4.6 Extending the side-side-side theorem from 3 to 4 points in $\mathbb{R}^n$

Many authors considered criteria of congruence for plane quadrilaterals [22], whose vertices are (cyclically) ordered. If a quadrilateral has fixed side lengths, the resulting moduli space under rigid motion in  $\mathbb{R}^2$  was explicitly described in [13, section 1.3].

In any  $\mathbb{R}^n$ , the  $m \times m$  matrix of pairwise distances [17] and the Gram matrix of scalar products [23] are complete and continuous invariants of  $m$  ordered points under isometry, all known at least since 1935. The extension of this approach to  $m$  unordered points leads to the exponential complexity because of  $m!$  permutations on  $m$  points.

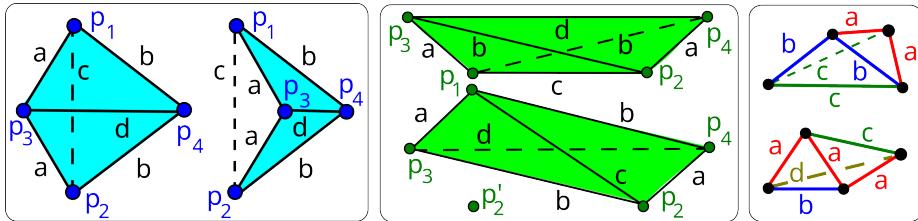
For  $m = 4$  unordered points, Theorem 4.6.1 will prove the completeness of  $\text{PDD}(C; m-1)$  under isometry in any  $\mathbb{R}^n$ . For any  $m$ , the invariant  $\text{PDD}(C; m-1)$  can be computed in quadratic time  $O(m^2)$ . For  $m = 4$ ,  $\text{PDD}(C; 3)$  contains only 12 numbers (6 pairs of distances between 4 points), while  $4! = 24$  distance matrices on 4 points contain at least 144 numbers if we consider only distances above the diagonal.

**Theorem 4.6.1** (completeness of PDD for  $m \leq 4$  points, [25, Theorem 5.3]). The Pointwise Distance Distribution  $\text{PDD}(C; m-1)$  from Definition 4.5.1 is a complete isometry invariant of all clouds  $C \subset \mathbb{R}^n$  of any  $m \leq 4$  unordered points.  $\blacksquare$

Since Theorem 4.6.1 finally extends the side-side-side criterion of congruence to  $m = 4$  unordered points, without relying on brute-force permutations, we include the detailed proof, which previously appeared only in supplementary materials of [25].

If a cloud  $A$  of  $m$  points has a line or plane of symmetry  $L$  in  $\mathbb{R}^2$  or  $\mathbb{R}^3$ , then all points  $A \setminus L$  split into pairs of points that are symmetric in  $L$  and hence have equal rows in  $\text{PDD}(A; m - 1)$ . Lemma 4.6.2 shows that the converse holds for  $m = 4$ .

**Lemma 4.6.2** (PDD detects symmetry of  $m = 4$  points, [25, Lemma SM3.5]). For any cloud  $A \subset \mathbb{R}^n$  of  $m = 4$  points for  $n = 2, 3$ , if  $\text{PDD}(A; 3)$  has two equal rows, then  $A$  is either (1) mirror symmetric in the plane passing through two points of  $A$  orthogonally to the line segment joining the other points of  $A$ , or (2) symmetric by the  $180^\circ$  degree rotation around the line through the mid-points of two pairs of points of  $A$ . If  $n = 2$ , then  $A$  defines a kite, or a parallelogram or an isosceles trapezoid; see Fig. 4.9. ■



**Fig. 4.9** **Left:** in  $\mathbb{R}^2$ , the convex and non-convex kites have two equal rows  $\{a, b, c\}$  in  $\text{PDD}(A; 3)$  and are distinguished by  $d = |p_3 - p_4|$ , see Lemma 4.6.2. **Middle:** an isosceles trapezoid and parallelogram have  $\text{PDD}(A; 3)$  with two pairs of equal rows  $\{a, b, c\}$  and  $\{a, b, d\}$ , e.g. a rectangle for  $c = d$ . **Top right:** a trisosceles cloud. **Bottom right:** a 3-chain-equal cloud, see Example 4.6.3.

*Proof of Lemma 4.6.2.* Let points  $p_1, p_2 \in A$  have the same row  $a \leq b \leq c$  in  $\text{PDD}(C; 3)$ . One of the distances  $a, b, c$  is between the points  $p_1, p_2$ . Without loss of generality, assume that  $|p_1 - p_2| = c$ . Then  $p_1, p_2$  have distances  $a, b$  to the points  $p_3, p_4 \in A \setminus \{p_1, p_2\}$ , but it is unknown which distance corresponds to which point.

*Isosceles case.* Let  $|p_1 - p_3| = a = |p_2 - p_3|$  and  $|p_1 - p_4| = b = |p_2 - p_4|$ , see Fig. 4.9 (left). Then  $A$  has two equal triangles  $\Delta p_1 p_3 p_4 = \Delta p_2 p_3 p_4$  and two isosceles triangles  $\Delta p_3 p_1 p_2$  and  $\Delta p_4 p_1 p_2$  with equal sides at  $p_3, p_4$ , respectively. Let  $L$  be the plane that passes through  $p_3, p_4$  and is orthogonal to the line segment  $[p_1, p_2]$ . Then the mirror reflection in  $L$  swaps  $p_1, p_2$ . If  $n = 2$ ,  $A$  defines a (non-)convex kite.

*Non-isosceles case.* Then  $|p_1 - p_3| = a = |p_2 - p_4|$  and  $|p_2 - p_3| = b = |p_1 - p_4|$ , see Fig. 4.9 (middle). Let  $L$  be the perpendicular bisector of the line segment  $[p_3, p_4]$ . The mirror reflection in  $L$  swaps  $p_3 \leftrightarrow p_4$  and either swaps  $p_1 \leftrightarrow p_2$  (then  $A$  defines an isosceles trapezoid in  $\mathbb{R}^2$ ) or maps  $p_2$  to  $p'_2$ , so that  $p_1, p'_2, p_3, p_4$  satisfy the previous case. In the latter case, the composition with the reflection in the plane through  $p_3, p_4$  orthogonal to  $[p_1, p'_2]$  is the  $180^\circ$  degree rotation that swaps the points as  $p_1 \leftrightarrow p_2$  and  $p_3 \leftrightarrow p_4$ . If  $n = 2$ , then  $A$  defines a parallelogram, see Fig. 4.9 (bottom middle). □

**Example 4.6.3** (trisosceles and 3-chain-equal clouds in  $\mathbb{R}^3$ ). Fig. 4.9 (right) shows *trisosceles* and *3-chain-equal* clouds that have 3 pairs of equal distances and a chain of

3 equal distances, their PDDs are  $\begin{pmatrix} a & a & c \\ a & b & b \\ a & b & c \\ b & c & c \end{pmatrix}$  and  $\begin{pmatrix} a & a & b \\ a & a & c \\ a & b & d \\ a & c & d \end{pmatrix}$ , respectively.  $\blacklozenge$

*Proof of Theorem 4.6.1. Case  $m = 2$ .* Any cloud  $A \subset \mathbb{R}^n$  of  $m = 2$  unordered points  $p_1, p_2$  (labelled only for convenience) has  $\text{PDD}(A; 1)$  consisting of the single distance  $|p_1 - p_2|$ , which uniquely determines  $A$  under isometry in any  $\mathbb{R}^n$ .

*Case  $m = 3$ .* Any cloud  $A \subset \mathbb{R}^n$  of  $m = 3$  unordered points with pairwise distances

$a \leq b \leq c$  has  $\text{PDD}(A; 2) = \begin{pmatrix} a & b \\ a & c \\ b & c \end{pmatrix}$ . The (lexicographically) first row of  $\text{PDD}(A; 2)$  gives us  $a \leq b$ . Each of the remaining two rows of  $\text{PDD}(A; 2)$  should contain at least one value of  $a$  or  $b$ , including in all degenerate cases such as  $a = b$ . Removing these repeated values from the other two rows gives us  $c$ , also in the case  $b = c$ . So  $\text{PDD}(A; 2)$  identifies  $a \leq b \leq c$  and hence  $A$ , uniquely under isometry in any  $\mathbb{R}^n$ .

*Case  $m = 4$ ,* then  $n \leq 3$ . For a cloud  $A \subset \mathbb{R}^3$  of  $m = 4$  unordered points,  $\text{PDD}(A; 3)$  is a  $4 \times 3$  matrix. Assume that  $\text{PDD}(A; 3)$  has two equal rows  $a \leq b \leq c$ .

*Isosceles case.* In the first case of Lemma 4.6.2 in Fig. 4.9 (left),  $\text{PDD}(A; 3)$  has two more rows  $\{a, a, d\}$  and  $\{b, b, d\}$  including two repeated distances (say,  $a, b$ ) among  $a, b, c$ . We can form two isosceles triangles with sides  $a, a, c$  and  $b, b, c$ , which can be rotated in  $\mathbb{R}^3$  around their common side of the length  $c$ , but their positions are fixed under isometry in  $\mathbb{R}^3$  by the distance  $d$  between their non-shared vertices.

*Non-isosceles case.* In the second case of Lemma 4.6.2 in Fig. 4.9 (middle),  $\text{PDD}(A; 3)$  has two pairs of equal rows of (unordered) distances  $\{a, b, c\}$  and  $\{a, b, d\}$ . Each of these triples uniquely determines a pair of equal triangles with a common side that are symmetric in the perpendicular bisector to this side. For example, if we start with a fixed position of  $[p_3, p_4]$  in  $\mathbb{R}^3$ , the union of equal triangles  $\Delta p_1 p_3 p_4 = \Delta p_2 p_3 p_4$  in Fig. 4.9 (middle) is uniquely determined under isometry by the length  $d$  of  $[p_1, p_2]$ . In  $\mathbb{R}^2$ , the parallelogram and isosceles trapezoid are distinguished by this distance  $d$ .

Now we can assume that all rows of  $\text{PDD}(A; 3)$  are different. Then all points can be uniquely labelled as  $p_1, p_2, p_3, p_4$  according to the lexicographic order of rows. Our aim is to get  $\text{PDD}(\{p_2, p_3, p_4\}; 2)$ , reconstruct  $\Delta p_2 p_3 p_4$ , and then uniquely add  $p_1$ .

*Case of a row with 3 equal distances.* Let  $\text{PDD}(A; 3)$  have a row of (say)  $p_1$  with 3 equal distances  $a$ . After removing the row of  $p_1$ , the distance  $a$  from the rows of  $p_2, p_3, p_4$ , we get  $\text{PDD}(\{p_2, p_3, p_4\}; 2)$ . This smaller  $3 \times 2$  matrix determines

$\triangle p_2p_3p_4$ , uniquely under isometry in  $\mathbb{R}^3$ . For a fixed  $\triangle p_2p_3p_4$ , the position of  $p_1$  in  $\mathbb{R}^3$  is determined by its distance  $a$  to  $p_2, p_3, p_4$ , uniquely under the mirror reflection relative to the plane of  $\triangle p_2p_3p_4$ . If  $n = 2$ , then  $p_1$  is the unique circumcenter of  $\triangle p_2p_3p_4$ .

*Case of a row with 3 unique distances.* Let  $\text{PDD}(C; 3)$  have a row of (say)  $p_1$ , where each of the distances  $a, b, c$  (say, to  $p_2, p_3, p_4$ ) appears in at most one other row (then  $a, b, c$  are distinct). After removing the row of  $p_1$ , the distance  $a$  from the row  $p_2$ , the distance  $b$  from the row of  $p_3$ , and the distance  $c$  from the row of  $p_4$ , we get  $\text{PDD}(\{p_2, p_3, p_4\}; 2)$ . This  $3 \times 2$  matrix determines  $\triangle p_2p_3p_4$ , uniquely under isometry in  $\mathbb{R}^3$ . Then the position of  $p_1$  in  $\mathbb{R}^3$  is determined by its distances  $a, b, c$  to  $p_2, p_3, p_4$ , respectively, under a mirror reflection relative to the plane of the triangle  $\triangle p_2p_3p_4$ .

*Case of one distance in 4 rows.* Then two pairs of points have disjoint edges of the

$$\text{same length, e.g. } |p_1 - p_2| = a = |p_3 - p_4|, \text{ so } \text{PDD}(A; 3) = \begin{pmatrix} a & b & c \\ a & d & e \\ a & b & d \\ a & c & e \end{pmatrix} \text{ for } b = |p_1 - p_3|,$$

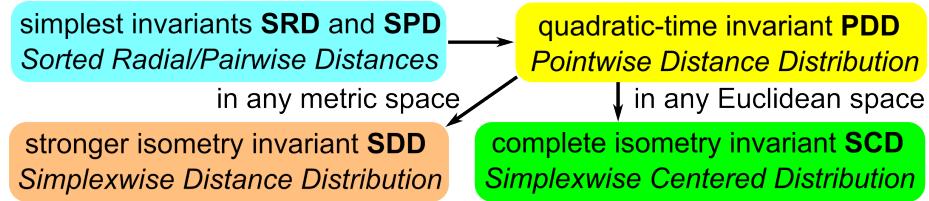
$c = |p_1 - p_4|, d = |p_2 - p_3|, e = |p_2 - p_4|$ . Then  $c \neq d$  and  $b \neq e$ , else  $\text{PDD}(A; 3)$  has two equal rows (considered above), similarly when  $b = c$  and  $d = e$ .

If  $a$  equals one of  $b, c, d, e$  (say,  $e$ ), then  $A$  is a 3-chain-equal cloud in Fig. 4.9 (bottom right) and the argument below still works. If  $b \neq c$ , we remove the row of  $p_1$ , the distance  $b$  from the only row of  $p_3$  containing  $b$ , the distance  $c$  from the only row of  $p_4$  containing  $c$ , and then remove  $a$  from the remaining row of  $p_2$ . This reduction to  $\text{PDD}(\{p_2, p_3, p_4\}; 2)$  allows us to reconstruct  $A$ , uniquely under isometry in  $\mathbb{R}^3$  as in the case of a row with 3 unique distances. If  $b = c$  but  $d \neq e$ , we remove the row of  $p_2$ , the distance  $d$  from the only row of  $p_3$  containing  $d$ , the distance  $e$  from the only row of  $p_4$  containing  $e$ , and then remove  $a$  from the remaining row of  $p_1$ , which allows us to uniquely reconstruct  $A$  as in the case of a row with 3 unique distances above.

*The final case:* no distance appears in all 4 distinct rows, but every row has a distance appearing in 3 rows, hence at least four times, including two times in the same row. Then  $A$  is a trisosceles cloud in Fig. 4.9 (top right). If any of the remaining distances  $a, b, c$  are equal,  $\text{PDD}(A; 3)$  has two equal rows (the case considered above). Then we remove any row (say  $a, b, b$ ) with two repeated distances, the distance  $b$  from the only two rows containing  $b$ , and the distance  $a$  from the remaining row.

This reduction to  $\text{PDD}(\{p_2, p_3, p_4\}; 2)$ , allows us to reconstruct  $\triangle p_2p_3p_4$ , uniquely under isometry in  $\mathbb{R}^3$ . Though  $p_1$  has equal distances to two of the vertices (say  $p_2, p_3$ ), the ambiguity of reconstructing  $p_1$  in  $\mathbb{R}^3$  by its distances to  $p_2, p_3, p_4$ , is only under the mirror reflections relative to the bisector plane of  $[p_2, p_3]$  and the plane of  $\triangle p_2p_3p_4$ .  $\square$

Chapters 5 and 6 will extend the PDD to the stronger invariants in a metric space (SDD) and complete invariant (SCD) under rigid motion in any  $\mathbb{R}^n$ , as shown in Fig. 4.10.



**Fig. 4.10** A hierarchy of invariants from the fastest (linear-time SRD and quadratic-time SPD) to the stronger PDD and SDD in Chapter 5 up to the Simplexwise Centered Distribution (SCD) in Chapter 6, which will satisfy the majority of conditions in Geo-Mapping Problem 1.4.5 for any finite  $n$ -dimensional clouds of unordered points under rigid motion in Theorems 6.2.3, 6.4.4, and 6.4.5.

## References

1. Jean-Daniel Boissonnat, Ramsay Dyer, Arijit Ghosh, and Nikolay Martynchuk. An obstruction to Delaunay triangulations in riemannian manifolds. *Discrete & Comp. Geometry*, 59(1):226–237, 2018.
2. Mireille Boutin and Gregor Kemper. On reconstructing  $n$ -point configurations from the distribution of distances or areas. *Advances in Applied Mathematics*, 32(4):709–735, 2004.
3. Terry Caelli. On generating spatial configurations with identical interpoint distance distributions. In *Proc. 7th Australian Conference on Combinatorial Mathematics*, pages 69–75, 1979.
4. Gunnar Carlsson and Mikael Vejdemo-Johansson. *Topological data analysis with applications*. Cambridge University Press, 2021.
5. Bernard Chazelle. A minimum spanning tree algorithm with inverse-ackermann type complexity. *Journal of the ACM*, 47(6):1028–1047, 2000.
6. Boris Delaunay. Sur la sphère vide. *Izv. Akad. Nauk USSR*, 7:793–800, 1934.
7. Tamal Krishna Dey and Yusu Wang. *Computational topology for data analysis*. Cambridge University Press, 2022.
8. Herbert Edelsbrunner and John Harer. *Computational topology: an introduction*. AMS, 2010.
9. Y. Elkin and V. Kurlin. The mergegram of a dendrogram and its stability. In *Proceedings of MFCS (Mathematical Foundations of Computer Science)*, pages 32:1–32:13, 2020.
10. Y. Elkin and V. Kurlin. Isometry invariant shape recognition of projectively perturbed point clouds by the mergegram extending 0d persistence. *Mathematics*, 9(17), 2021.
11. Yury Elkin and Vitaliy Kurlin. Counterexamples expose gaps in the proof of time complexity for cover trees introduced in 2006. In *Topological Data Analysis and Visualization*, pages 9–17, 2022.
12. Yury Elkin and Vitaliy Kurlin. A new near-linear time algorithm for k-nearest neighbor search using a compressed cover tree. In *Int. Conf. Machine Learning*, pages 9267–9311, 2023.
13. Michael Farber. *Invitation to topological robotics*. European Mathematical Society, 2008.
14. William B March, Parikshit Ram, and Alexander G Gray. Fast euclidean minimum spanning tree: algorithm, analysis, and applications. In *Proceedings of SIG KDD: Knowledge discovery and data mining*, pages 603–612, 2010.
15. Facundo Mémoli. Gromov–Wasserstein distances and the metric approach to object matching. *Foundations of Computational Mathematics*, 11(4):417–487, 2011.
16. Yossi Rubner, Carlo Tomasi, and Leonidas Guibas. The Earth Mover’s Distance as a metric for image retrieval. *International Journal of Computer Vision*, 40(2):99–121, 2000.
17. Isaac Schoenberg. Remarks to Maurice Fréchet’s article “Sur la définition axiomatique d’une classe d’espaces distances vectoriellement applicable sur l’espace de Hilbert. *Annals of Mathematics*, pages 724–732, 1935.
18. Jonathan Richard Shewchuk. Sweep algorithms for constructing higher-dimensional constrained Delaunay triangulations. In *Proceedings of SoCG*, pages 350–359, 2000.

19. S Shirdhonkar and D Jacobs. Approximate earth mover’s distance in linear time. In *Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.
20. Philip Smith and Vitaliy Kurlin. Skeletonisation algorithms with theoretical guarantees for unorganised point clouds with high levels of noise. *Pattern Recognition*, 115:107902, 2021.
21. Philip Smith and Vitaliy Kurlin. Generic families of finite metric spaces with identical or trivial 1-dimensional persistence. *Journal of Applied and Computational Topology*, 8:839–855, 2024.
22. Irvin E Vance. Minimum conditions for congruence of quadrilaterals. *School Science and Mathematics*, 82(5):403–15, 1982.
23. Hermann Weyl. *The classical groups: their invariants and representations*. Princeton Univ., 1946.
24. Daniel Widdowson and Vitaliy Kurlin. Resolving the data ambiguity for periodic crystals. *Advances in Neural Information Processing Systems*, 35:24625–24638, 2022.
25. Daniel Widdowson and Vitaliy Kurlin. Pointwise distance distributions for detecting near-duplicates in large materials databases. *SIAM Journal on Applied Mathematics* ([doi:10.1137/25M1736657](https://doi.org/10.1137/25M1736657), [arxiv:2108.04798](https://arxiv.org/abs/2108.04798)), 2026.



## Chapter 5

# Higher order distance distributions of unordered points in a metric space

**Abstract** This chapter presents further advances towards a solution of the geo-mapping problem under isometry in any metric space, as stated in the previous chapter. The Pointwise Distance Distribution (PDD) will be extended to stronger isometry invariant by collecting distances to  $h$ -point subsets. The resulting Simplexwise Distance Distribution (SDD) is Lipschitz continuous and computable in a polynomial time of the number of points, for a fixed order  $h$ . For  $h = 2$ , the SDD distinguishes all (infinitely many) known counter-examples to the completeness of the PDD under isometry in  $\mathbb{R}^3$ .

## 5.1 Simplexwise Distance Distributions of a cloud in a metric space

This chapter follows paper [8] and its extension [6] to metric spaces with measures.

We continue solving Problem 4.4.4 to find geocodes of finite clouds in any metric space. The first section extends the Pointwise Distance Distribution (PDD) from Definition 4.5.1 to a stronger invariant, which requires a few auxiliary definitions.

The key idea of a stronger invariant is to use a base sequence of  $h > 1$  ordered points instead of  $h = 1$  point in the PDD.

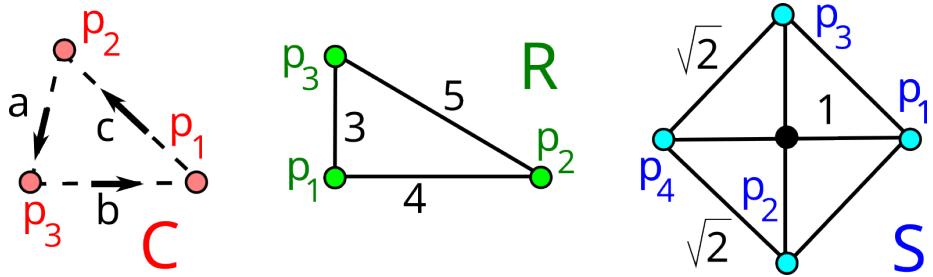
The *lexicographic* order  $\mathbf{u} < \mathbf{v}$  on vectors  $\mathbf{u} = (u_1, \dots, u_h)$  and  $\mathbf{v} = (v_1, \dots, v_h)$  means that if the first  $i$  coordinates (where  $i$  might be 0) of  $u, v$  coincide, then  $u_{i+1} < v_{i+1}$ . Let  $S_h$  denote the permutation group on indices  $1, \dots, h$ .

**Definition 5.1.1** (Relative Distance Distribution  $\text{RDD}(C; A)$ ). Let  $C$  be a cloud of  $m$  unlabelled points in a space with a metric  $d$ . A *base* sequence  $A = (p_1, \dots, p_h) \in C^h$  consists of  $1 \leq h < m$  distinct points. Let  $D(A)$  be the *triangular distance* matrix whose entry  $D(A)_{i,j-1}$  is  $d(p_i, p_j)$  for  $1 \leq i < j \leq h$ , all other entries are zeros.

Any permutation  $\xi \in S_h$  acts on  $D(A)$  by mapping  $D(A)_{ij}$  to  $D(A)_{kl}$ , where  $k \leq l$  is the pair of indices  $\xi(i), \xi(j) - 1$  written in increasing order. For any other point  $q \in C - A$ , write distances from  $q$  to  $p_1, \dots, p_h$  as a column. The  $h \times (m - h)$ -matrix  $R(C; A)$  is formed by these  $m - h$  lexicographically ordered columns. The action of  $\xi$  on  $R(C; A)$  maps any  $i$ -th row to the  $\xi(i)$ -th row, after which all columns can be

written in the lexicographic order. The *Relative Distance Distribution*  $\text{RDD}(C; A)$  is the equivalence class of the pair  $[D(A), R(C; A)]$  of matrices under permutations  $\xi \in S_h$ .  $\blacktriangle$

For  $h = 1$  and a base sequence  $A = (p_1)$ , the matrix  $D(A)$  is empty and  $R(C; A)$  is a single row of distances (in the increasing order) from  $p_1$  to all other points  $q \in C$ . For  $h = 2$  and a base sequence  $A = (p_1, p_2)$ , the matrix  $D(A)$  is the single number  $d(p_1, p_2)$  and  $R(C; A)$  consists of two rows of distances from  $p_1, p_2$  to all other  $q \in C$ .



**Fig. 5.1** **Left:** triangular cloud  $C$  of points  $p_1, p_2, p_3$  with inter-point distances  $a \leq b \leq c$ . **Middle:** right-angled cloud  $R$  of points  $(0, 0), (4, 0), (0, 3)$ . **Right:** square cloud  $S$  of points  $(1, 0), (-1, 0), (0, 1), (-1, 0)$ .

**Example 5.1.2** (RDD for a 3-point cloud  $C$ ). Let  $C \subset \mathbb{R}^2$  consist of  $p_1, p_2, p_3$  with inter-point distances  $a \leq b \leq c$  ordered counter-clockwise as in Fig. 5.1 (left). Then

$$\begin{aligned} \text{RDD}(C; p_1) &= [\emptyset; (b, c)], & \text{RDD}(C; \begin{pmatrix} p_2 \\ p_3 \end{pmatrix}) &= [a; \begin{pmatrix} c \\ b \end{pmatrix}], \\ \text{RDD}(C; p_2) &= [\emptyset; (a, c)], & \text{RDD}(C; \begin{pmatrix} p_3 \\ p_1 \end{pmatrix}) &= [b; \begin{pmatrix} a \\ c \end{pmatrix}], \\ \text{RDD}(C; p_3) &= [\emptyset; (a, b)], & \text{RDD}(C; \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}) &= [c; \begin{pmatrix} b \\ a \end{pmatrix}]. \end{aligned}$$

We have written  $\text{RDD}(C; A)$  for a base sequence  $A = (p_i, p_j)$  of ordered points represented by a column. Swapping the points  $p_1 \leftrightarrow p_2$  makes the last RDD above

$$\text{equivalent to } \text{RDD}(C; \begin{pmatrix} p_2 \\ p_1 \end{pmatrix}) = [c; \begin{pmatrix} a \\ b \end{pmatrix}]. \quad \blacklozenge$$

Though  $\text{RDD}(C; A)$  is defined up to a permutation  $\xi \in S_h$  of  $h$  points in  $A \in C^h$ , comparisons of RDDs will be practical for  $h = 2, 3$  with metrics independent of  $\xi$ .

**Definition 5.1.3** (Simplexwise Distance Distribution  $\text{SDD}(C; h)$ ). Let  $C$  be a cloud of  $m$  unlabelled points in a metric space. For an integer  $1 \leq h < m$ , the *Simplexwise Distance Distribution*  $\text{SDD}(C; h)$  of order  $h$  is the unordered set of  $\text{RDD}(C; A)$  for all unordered  $h$ -point subsets  $A \subset C$ .  $\blacktriangle$

For order  $h = 1$  and any  $m$ -point cloud  $C$ , the distribution  $\text{SDD}(C; 1)$  can be considered as a matrix of  $m$  rows of ordered distances from every point  $p \in C$  to all other  $m - 1$  points. If we lexicographically order these  $m$  rows and collapse any  $l > 1$  identical rows into a single one with the weight  $l/m$ , then we get the Pointwise Distance Distribution  $\text{PDD}(C; m - 1)$  introduced in Definition 4.5.1.

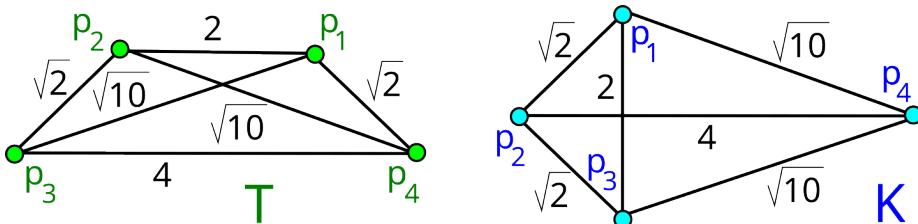
**Definition 5.1.4** (moments of a weighted distribution). Let  $A$  be any unordered set of real numbers  $a_1, \dots, a_m$  with weights  $w_1, \dots, w_m$ , respectively, such that  $\sum_{i=1}^m w_i = 1$ . The 1st moment (*average*) is the  $\mu_1(A) = \sum_{i=1}^m w_i a_i$ . The 2nd moment is  $\mu_2(A) = \sqrt{\frac{1}{m} \sum_{i=1}^m w_i a_i^2}$ . For  $t \geq 3$ , the  $t$ -th moment is  $\sqrt[m]{m^{1-t} \sum_{i=1}^m w_i a_i^t}$ , see [5, section 2.7].  $\blacktriangle$

The vector  $\text{SPD}(A)$  of Sorted Pairwise Distances was introduced in Definition 4.4.5.

**Definition 5.1.5** (Simplexwise Distance Moments SDM). For any  $m$ -point cloud  $C$  in a metric space, let  $A \subset C$  be a subset of  $h$  unordered points. The vector  $\mathbf{R}(C; A) \in \mathbb{R}^{m-h}$  is obtained from the  $h \times (m - h)$  matrix  $R(C; A)$  in Definition 5.1.1 by writing the vector of  $m - h$  column averages in increasing order.

The pair  $[\text{SPD}(A); \mathbf{R}(C; A)]$  is the *Average Distance Distribution*  $\text{ADD}(C; A)$  considered a vector of length  $\frac{h(h-3)}{2} + m$ . The unordered collection of  $\text{ADD}(C; A)$  for all  $\binom{m}{h}$  unordered subsets  $A \subset C$  is the *Average Simplexwise Distribution*  $\text{ASD}(C; h)$ .

The *Simplexwise Distance Moment*  $\text{SDM}(C; h, t)$  is the  $t$ -th moment of  $\text{ASD}(C; h)$  considered a probability distribution of  $\binom{m}{h}$  vectors, separately for each coordinate.  $\blacktriangle$



**Fig. 5.2** **Left:** trapezium cloud  $T$  of points  $(1, 1)$ ,  $(-1, 1)$ ,  $(-2, 0)$ ,  $(2, 0)$ . **Right:** kite cloud  $K$  of points  $(0, 1)$ ,  $(-1, 0)$ ,  $(0, -1)$ ,  $(3, 0)$ .

**Example 5.1.6** (SDD and SDM for the 4-point clouds  $T, K$ ). Fig. 5.2 shows the non-isometric 4-point clouds  $T, K$  with the same Ordered Pairwise Distances:  $\text{SPD} = \{\sqrt{2}, \sqrt{2}, 2, \sqrt{10}, \sqrt{10}, 4\}$ , see infinitely many examples in [1]. The arrows on the edges of  $T, K$  show orders of points in each pair of vertices for RDDs. Then  $T, K$  are distinguished under isometry by  $\text{SDD}(T; 2) \neq \text{SDD}(K; 2)$  in Table 5.1. The 1st coordinate of  $\text{SDM}(C; 2, 1) \in \mathbb{R}^3$  is the average of the six distances from  $\text{SPD}$  (the same for  $T, K$ ) but the other two coordinates (column averages from  $R(C; A)$  matrices) differ.  $\blacklozenge$

$\text{RDD}(T; A)$ in $\text{SDD}(T; 2)$	$\text{RDD}(K; A)$ in $\text{SDD}(K; 2)$
$[\sqrt{2}, \begin{pmatrix} 2 & \sqrt{10} \\ \sqrt{10} & 4 \end{pmatrix}] \times 2$	$[\sqrt{2}, \begin{pmatrix} 2 & \sqrt{10} \\ \sqrt{2} & 4 \end{pmatrix}] \times 2$
$[2, \begin{pmatrix} \sqrt{2} & \sqrt{10} \\ \sqrt{10} & \sqrt{2} \end{pmatrix}]$	$[2, \begin{pmatrix} \sqrt{2} & \sqrt{10} \\ \sqrt{2} & \sqrt{10} \end{pmatrix}]$
$[\sqrt{10}, \begin{pmatrix} \sqrt{2} & 4 \\ 2 & \sqrt{2} \end{pmatrix}] \times 2$	$[\sqrt{10}, \begin{pmatrix} \sqrt{2} & 2 \\ 4 & \sqrt{10} \end{pmatrix}] \times 2$
$[4, \begin{pmatrix} \sqrt{2} & \sqrt{10} \\ \sqrt{10} & \sqrt{2} \end{pmatrix}]$	$[4, \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ \sqrt{10} & \sqrt{10} \end{pmatrix}]$
$\text{ADD}(T; A)$ in $\text{ASD}(T; 2)$	$\text{ADD}(K; A)$ in $\text{ASD}(K; 2)$
$[\sqrt{2}, (\frac{2+\sqrt{10}}{2}, \frac{4+\sqrt{10}}{2})] \times 2$	$[\sqrt{2}, (\frac{2+\sqrt{2}}{2}, \frac{4+\sqrt{10}}{2})] \times 2$
$[2, (\frac{\sqrt{2}+\sqrt{10}}{2}, \frac{\sqrt{2}+\sqrt{10}}{2})]$	$[2, (\sqrt{2}, \sqrt{10})]$
$[\sqrt{10}, (\frac{2+\sqrt{2}}{2}, \frac{4+\sqrt{2}}{2})] \times 2$	$[\sqrt{10}, (\frac{2+\sqrt{10}}{2}, \frac{4+\sqrt{2}}{2})] \times 2$
$[4, (\frac{\sqrt{2}+\sqrt{10}}{2}, \frac{\sqrt{2}+\sqrt{10}}{2})]$	$[4, (\frac{\sqrt{2}+\sqrt{10}}{2}, \frac{\sqrt{2}+\sqrt{10}}{2})]$
$\text{SDM}_1 = \frac{3 + \sqrt{2} + \sqrt{10}}{3}$	$\text{SDM}_1 = \frac{3 + \sqrt{2} + \sqrt{10}}{3}$
$\text{SDM}_2 = \frac{6 + 2\sqrt{2} + 4\sqrt{10}}{12}$	$\text{SDM}_2 = \frac{8 + 5\sqrt{2} + 3\sqrt{10}}{12}$
$\text{SDM}_3 = \frac{16 + 4\sqrt{2} + 4\sqrt{10}}{12}$	$\text{SDM}_3 = \frac{16 + 3\sqrt{2} + 5\sqrt{10}}{12}$

**Table 5.1** Top: Relative Distance Distributions from Definition 5.1.1 for all 6 base sequences  $A$  in the 4-point clouds  $T, K$  in Fig. 5.2. The symbol  $\times 2$  indicates a doubled RDD. The three bottom rows show coordinates of  $\text{SDM}(C; 2, 1) \in \mathbb{R}^3$  from Definition 5.1.5 for  $h = 2, t = 1$  and  $C = T, K$ . Different elements are highlighted and imply that all invariants SDD, ADD, SDM distinguish  $T \neq K$ .

Some of the  $\binom{m}{h}$  RDDs in  $\text{SDD}(C; h)$  can coincide as in Example 5.1.6. If we collapse any  $l > 1$  identical RDDs into a single RDD with the weight  $l/\binom{m}{h}$ , SDD can be considered as a weighted probability distribution of RDDs.

In a general metric space, a point cloud  $C$  is usually given by a distance matrix on (arbitrarily ordered) points of  $C$ . Hence, we assume that the distance between any points of  $C$  is accessible in a constant time.

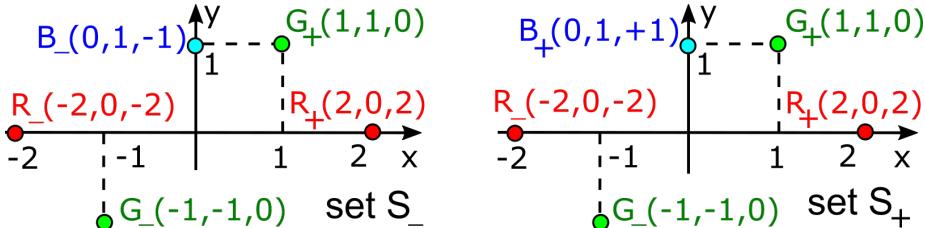
**Theorem 5.1.7** (invariance and time of SDD, [6, Theorem 3.6]). For any order  $h \geq 1$  and any cloud  $C$  of  $m$  unlabelled points in a metric space,  $\text{SDD}(C; h)$  is an isometry invariant, which can be computed in time  $O(m^{h+1}/(h-1)!)$ . For any  $t \geq 1$ , the invariant  $\text{SDM}(C; h, t) \in \mathbb{R}^{m+\frac{h(h-3)}{2}}$  has the same asymptotic time. ■

## 5.2 The expressiveness of Simplexwise Distance Distributions

This section shows that  $\text{SDD}(C; 2)$  distinguishes all infinitely many known pairs [3, Fig. S4] of non-isometric clouds  $S, Q \subset \mathbb{R}^3$  that have equal  $\text{PDD}(S) = \text{PDD}(Q)$

Examples 5.2.1 and 5.2.2 distinguish clouds of 5 points and 7 points, respectively, in  $\mathbb{R}^3$  by comparing their SDDs of order 2. In Example 5.2.3, the invariant  $\text{SDD}(C; 2)$  distinguishes 6-point clouds in a family of pairs depending on three parameters.

**Example 5.2.1** (5-point clouds). Fig. 5.3 shows the 5-point clouds  $S_{\pm} \subset \mathbb{R}^3$  taken from [3, Figure S4(A)]. The clouds  $S_{\pm}$  are not isometric, because  $S_+$  has the triple of points  $B_+, G_+, R_+$  with pairwise distances  $\sqrt{2}, \sqrt{6}, \sqrt{6}$ , but  $S_-$  has no such a triple.



**Fig. 5.3** See Example 5.2.1 **Left:**  $(x, y)$ -projection of the 5-point cloud  $S_- \subset \mathbb{R}^3$  consisting of the green points  $G_- = (-1, -1, 0)$  and  $G_+ = (1, 1, 0)$ , the red points  $R_- = (-2, 0, -2)$  and  $R_+ = (2, 0, 2)$ , and the blue point  $B_- = (0, 1, -1)$ . **Right:** to get  $S_+ \subset \mathbb{R}^3$  from the cloud  $S_-$ , replace the point  $B_-$  with another point  $B_+ = (0, 1, 1)$ .

Table 5.2 **highlights** differences between distance matrices. If we order distances to neighbours, the matrices in Table 5.3 differ only in one pair.

If we ignore the labels of all points in columns, Table 5.3 implies that  $S_{\pm}$  have identical Pointwise Distance Distribution (PDD). For easier visualisation, the matrix below is obtained by lexicographically ordering the rows in Table 5.3:

distances of $S_-$	$R_-$	$R_+$	$G_-$	$G_+$	$B_-$	distances of $S_+$	$R_-$	$R_+$	$G_-$	$G_+$	$B_+$
$R_-(-2, 0, -2)$	0	$\sqrt{32}$	$\sqrt{6}$	$\sqrt{14}$	$\sqrt{6}$	$R_-(-2, 0, -2)$	0	$\sqrt{32}$	$\sqrt{6}$	$\sqrt{14}$	$\sqrt{14}$
$R_+(+2, 0, +2)$	$\sqrt{32}$	0	$\sqrt{14}$	$\sqrt{6}$	$\sqrt{14}$	$R_+(+2, 0, +2)$	$\sqrt{32}$	0	$\sqrt{14}$	$\sqrt{6}$	$\sqrt{6}$
$G_-(-1, -1, 0)$	$\sqrt{6}$	$\sqrt{14}$	0	$\sqrt{8}$	$\sqrt{6}$	$G_-(-1, -1, 0)$	$\sqrt{6}$	$\sqrt{14}$	0	$\sqrt{8}$	$\sqrt{6}$
$G_+(+1, +1, 0)$	$\sqrt{14}$	$\sqrt{6}$	$\sqrt{8}$	0	$\sqrt{2}$	$G_+(+1, +1, 0)$	$\sqrt{14}$	$\sqrt{6}$	$\sqrt{8}$	0	$\sqrt{2}$
$B_-(0, +1, -1)$	$\sqrt{6}$	$\sqrt{14}$	$\sqrt{6}$	$\sqrt{2}$	0	$B_+(0, +1, +1)$	$\sqrt{14}$	$\sqrt{6}$	$\sqrt{6}$	$\sqrt{2}$	0

**Table 5.2** Distance matrices of the clouds  $S_{\mp} \subset \mathbb{R}^3$  in Fig. 5.3.

$S_-$ distances to	1st neighbour	2nd neighbour	3rd neighbour	4th neighbour
$R_- = (-2, 0, -2)$	$\sqrt{6}$	$\sqrt{6}$	$\sqrt{14}$	$\sqrt{32}$
$R_+ = (+2, 0, +2)$	$\sqrt{6}$	$\sqrt{14}$	$\sqrt{14}$	$\sqrt{32}$
$G_- = (-1, -1, 0)$	$\sqrt{6}$	$\sqrt{6}$	$\sqrt{8}$	$\sqrt{14}$
$G_+ = (+1, +1, 0)$	$\sqrt{2}$	$\sqrt{6}$	$\sqrt{8}$	$\sqrt{14}$
$B_- = (0, +1, -1)$	$\sqrt{2}$	$\sqrt{6}$	$\sqrt{6}$	$\sqrt{14}$

$S_+$ distances to	1st neighbour	2nd neighbour	3rd neighbour	4th neighbour
$R_- = (-2, 0, -2)$	$\sqrt{6}$	$\sqrt{14}$	$\sqrt{14}$	$\sqrt{32}$
$R_+ = (+2, 0, +2)$	$\sqrt{6}$	$\sqrt{6}$	$\sqrt{14}$	$\sqrt{32}$
$G_- = (-1, -1, 0)$	$\sqrt{6}$	$\sqrt{6}$	$\sqrt{8}$	$\sqrt{14}$
$G_+ = (+1, +1, 0)$	$\sqrt{2}$	$\sqrt{6}$	$\sqrt{8}$	$\sqrt{14}$
$B_+ = (0, +1, -1)$	$\sqrt{2}$	$\sqrt{6}$	$\sqrt{6}$	$\sqrt{14}$

**Table 5.3** For each point from the 5-point cloud  $S_+$  in Fig. 5.3, the distances to neighbours from Table 5.2 are ordered in each row.

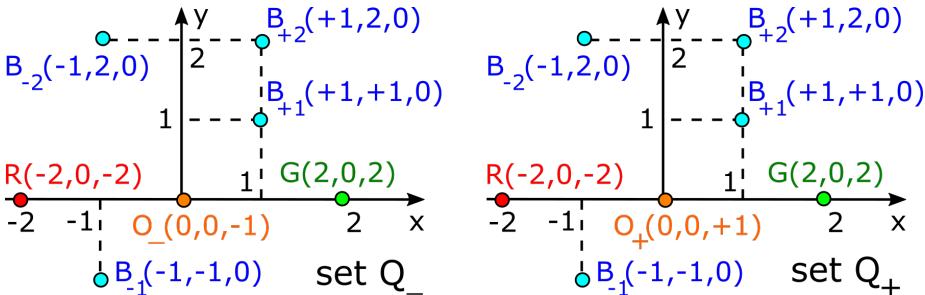
$$\text{PDD}(S_{\pm}) = \text{SDD}(S_{\pm}; 1) = \begin{pmatrix} \sqrt{2} & \sqrt{6} & \sqrt{6} & \sqrt{14} \\ \sqrt{2} & \sqrt{6} & \sqrt{8} & \sqrt{14} \\ \sqrt{6} & \sqrt{6} & \sqrt{8} & \sqrt{14} \\ \sqrt{6} & \sqrt{6} & \sqrt{14} & \sqrt{32} \\ \sqrt{6} & \sqrt{14} & \sqrt{14} & \sqrt{32} \end{pmatrix}.$$

Now we show that  $\text{SDD}(S_-; 2) \neq \text{SDD}(S_+; 2)$ . For  $h = 2$ , the Simplexwise Distance Distribution  $\text{SDD}(C; h)$  consists of  $\text{RDD}(C; A)$  for 2-point subsets  $A \subset C$ . Both sets  $S_{\pm}$  have a single pair of points  $(G_+, B_-)$  and  $(G_-, B_+)$  at distance  $\sqrt{2}$ . Hence it suffices to show that the Relative Distance Distributions differ for this pair:

$$\text{RDD}\left(S_-, \begin{pmatrix} G_+ \\ B_- \end{pmatrix}\right) = \left[ \sqrt{2}, \begin{pmatrix} \sqrt{8} & \sqrt{14} & \sqrt{6} \\ \sqrt{6} & \boxed{\sqrt{6}} & \boxed{\sqrt{14}} \\ G_- & R_- & R_+ \end{pmatrix} \right],$$

$$\text{RDD}\left(S_+, \begin{pmatrix} G_+ \\ B_+ \end{pmatrix}\right) = \left[ \sqrt{2}, \begin{pmatrix} \sqrt{8} & \sqrt{14} & \sqrt{6} \\ \sqrt{6} & \boxed{\sqrt{14}} & \boxed{\sqrt{6}} \\ G_- & R_- & R_+ \end{pmatrix} \right].$$

The last rows in the above  $3 \times 3$  matrices indicate a complementary point  $q \in C - A$  for indexing columns of the  $2 \times 3$  matrices  $R(C; A)$  in Definition 5.1.1. The resulting RDDs differ because any permutation of rows or columns of  $R(S_+; \{G_+, B_+\})$  keeps the pair  $\sqrt{6}, \sqrt{6}$  in the same column but  $R(S_-; \{G_+, B_+\})$  has no pair  $\sqrt{6}, \sqrt{6}$  in one column. Hence  $\text{SDD}(S_-; 2) \neq \text{SDD}(S_+; 2)$ .  $\blacklozenge$



**Fig. 5.4** See Example 5.2.2. **Left:**  $(x, y)$ -projection of the 7-point cloud  $Q_- \subset \mathbb{R}^3$ , which consists of the red point  $R = (-2, 0, -2)$ , green point  $G = (2, 0, 2)$ , four blue points  $B_{\pm 1} = (\pm 1, \pm 1, 0)$ ,  $B_{\pm 2} = (\pm 1, 2, 0)$ , orange point  $O_- = (0, 0, -1)$ . **Right:** to get the cloud  $Q_+$  from the cloud  $Q_- \subset \mathbb{R}^3$ , replace the point  $O_-$  with  $O_+ = (0, 0, +1)$ .

**Example 5.2.2** (7-point clouds). The clouds  $Q_{\pm}$  in Fig. 5.4 taken from [3, Figure S4(B)] have distances in Table 5.4. Both sets have only two pairs of points at distance  $\sqrt{6}$ . Hence it suffices to compare RDDs for these pairs below.

distances of $Q_-$	$R$	$G$	$B_{-1}$	$B_{+1}$	$B_{-2}$	$B_{+2}$	$O_-$
$R = (-2, 0, -2)$	0	$\sqrt{32}$	$\sqrt{6}$	$\sqrt{14}$	3	$\sqrt{17}$	$\sqrt{5}$
$G = (+2, 0, +2)$	$\sqrt{32}$	0	$\sqrt{14}$	$\sqrt{6}$	$\sqrt{17}$	3	$\sqrt{13}$
$B_{-1} = (-1, -1, 0)$	$\sqrt{6}$	$\sqrt{14}$	0	$\sqrt{8}$	3	$\sqrt{13}$	$\sqrt{3}$
$B_{+1} = (+1, +1, 0)$	$\sqrt{14}$	$\sqrt{6}$	$\sqrt{8}$	0	$\sqrt{5}$	1	$\sqrt{3}$
$B_{-2} = (-1, 2, 0)$	3	$\sqrt{17}$	3	$\sqrt{5}$	0	2	$\sqrt{6}$
$B_{+2} = (+1, 2, 0)$	$\sqrt{17}$	3	$\sqrt{13}$	1	2	0	$\sqrt{6}$
$O_- = (0, 0, -1)$	$\sqrt{5}$	$\sqrt{13}$	$\sqrt{3}$	$\sqrt{3}$	$\sqrt{6}$	$\sqrt{6}$	0

distances of $Q_+$	$R$	$G$	$B_{-1}$	$B_{+1}$	$B_{-2}$	$B_{+2}$	$O_+$
$R = (-2, 0, -2)$	0	$\sqrt{32}$	$\sqrt{6}$	$\sqrt{14}$	3	$\sqrt{17}$	$\sqrt{13}$
$G = (+2, 0, +2)$	$\sqrt{32}$	0	$\sqrt{14}$	$\sqrt{6}$	$\sqrt{17}$	3	$\sqrt{5}$
$B_{-1} = (-1, -1, 0)$	$\sqrt{6}$	$\sqrt{14}$	0	$\sqrt{8}$	3	$\sqrt{13}$	$\sqrt{3}$
$B_{+1} = (+1, +1, 0)$	$\sqrt{14}$	$\sqrt{6}$	$\sqrt{8}$	0	$\sqrt{5}$	1	$\sqrt{3}$
$B_{-2} = (-1, 2, 0)$	3	$\sqrt{17}$	3	$\sqrt{5}$	0	2	$\sqrt{6}$
$B_{+2} = (+1, 2, 0)$	$\sqrt{17}$	3	$\sqrt{13}$	1	2	0	$\sqrt{6}$
$O_+ = (0, 0, +1)$	$\sqrt{13}$	$\sqrt{5}$	$\sqrt{3}$	$\sqrt{3}$	$\sqrt{6}$	$\sqrt{6}$	0

**Table 5.4** The distance matrices of the 7-point clouds  $Q_{\mp}$  in Fig. 5.4 taken from [3, Figure S4(B)].

$$R \left( Q_- ; \begin{pmatrix} G \\ B_{+1} \end{pmatrix} \right) = \begin{pmatrix} \sqrt{32} & \sqrt{14} & \sqrt{17} & 3 & \sqrt{13} \\ \sqrt{6} & \sqrt{8} & \sqrt{5} & 1 & \sqrt{3} \\ R & B_{-1} & B_{-2} & B_{+2} & O_- \end{pmatrix},$$

$$R\left(Q_-; \begin{pmatrix} R \\ B_{-1} \end{pmatrix}\right) = \begin{pmatrix} \sqrt{32} & \sqrt{14} & 3 & \sqrt{17} & \sqrt{5} \\ \sqrt{14} & \sqrt{8} & 3 & \sqrt{13} & \sqrt{3} \\ G & B_{+1} & B_{-2} & B_{+2} & O_- \end{pmatrix}.$$

The pair above has submatrices  $\begin{pmatrix} 3 & \sqrt{13} \\ 1 & \sqrt{3} \end{pmatrix}$  and  $\begin{pmatrix} 3 & \sqrt{5} \\ 3 & \sqrt{3} \end{pmatrix}$  but the pair below has no such submatrices.

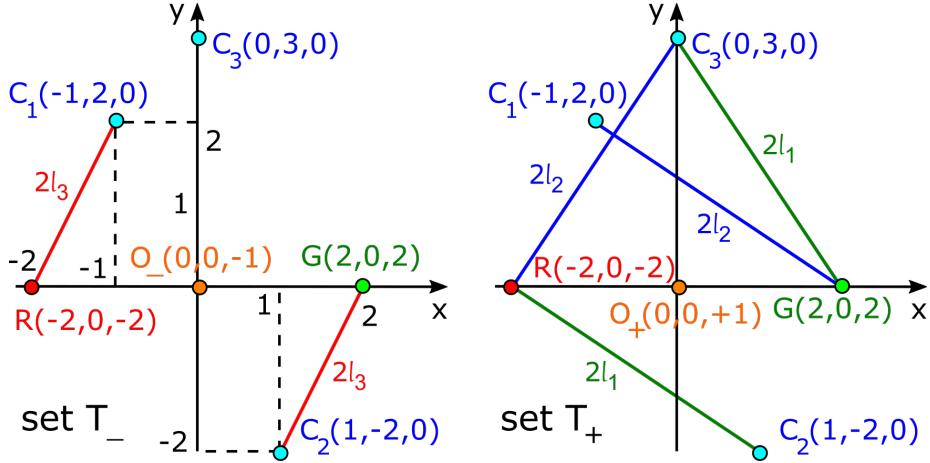
$$R\left(Q_+; \begin{pmatrix} G \\ B_{+1} \end{pmatrix}\right) = \begin{pmatrix} \sqrt{32} & \sqrt{14} & \sqrt{17} & 3 & \sqrt{5} \\ \sqrt{6} & \sqrt{8} & \sqrt{5} & 1 & \sqrt{3} \\ R & B_{-1} & B_{-2} & B_{+2} & O_+ \end{pmatrix},$$

$$R\left(Q_+; \begin{pmatrix} R \\ B_{-1} \end{pmatrix}\right) = \begin{pmatrix} \sqrt{32} & \sqrt{14} & 3 & \sqrt{17} & \sqrt{13} \\ \sqrt{14} & \sqrt{8} & 3 & \sqrt{13} & \sqrt{3} \\ G & B_{+1} & B_{-2} & B_{+2} & O_+ \end{pmatrix}$$

The pair of  $\text{RDD}\left(Q_-; \begin{pmatrix} G \\ B_{+1} \end{pmatrix}\right)$  and  $\text{RDD}\left(Q_-; \begin{pmatrix} R \\ B_{-1} \end{pmatrix}\right)$  differs from the pair  $\text{RDD}\left(Q_+; \begin{pmatrix} G \\ B_{+1} \end{pmatrix}\right)$  and  $\text{RDD}\left(Q_+; \begin{pmatrix} R \\ B_{-1} \end{pmatrix}\right)$ . Hence  $\text{SDD}(Q_-; 2) \neq \text{SDD}(Q_+; 2)$ .  $\blacklozenge$

**Example 5.2.3** (6-point clouds). The clouds  $T_{\pm}$  in Fig. 5.5, which was motivated by [3, Figure S4(C)], have the points  $R, G, O_{\pm}$  from the clouds  $Q_{\pm}$  in Example 5.2.2 and three new points  $C_1(x_1, y_1, 0), C_2(x_2, y_2, 0), C_3(x_3, y_3, 0)$  such that  $|RC_1| = |GC_2|$ ,  $|RC_2| = |GC_3|$ ,  $|RC_3| = |GC_1|$ . Denote by  $2l_1, 2l_2, 2l_3$  the lengths of these three pairs of line segments after their projection to the  $xy$ -plane so that

$$(5.2.3.1) \quad \begin{cases} (x_2 + 2)^2 + y_2^2 = |RC_2|^2 - 4 = (2l_1)^2, \\ (x_3 - 2)^2 + y_3^2 = |GC_3|^2 - 4 = (2l_1)^2; \end{cases}$$



**Fig. 5.5** See Example 5.2.3. **Left:**  $(x, y)$ -projection of the 6-point cloud  $T_- \subset \mathbb{R}^3$  consisting of the red point  $R = (-2, 0, -2)$ , green point  $G = (2, 0, 2)$ , three blue points  $C_1 = (x_1, y_1, 0)$ ,  $C_2 = (x_2, y_2, 0)$ ,  $C_3 = (x_3, y_3, 0)$ , and orange point  $O_- = (0, 0, -1)$  so that  $|RC_1| = 2l_3 = |GC_2|$ ,  $|RC_2| = 2l_1 = |GC_3|$ ,  $|RC_3| = 2l_2 = |GC_1|$ . **Right:** to get the cloud  $T_+ \subset \mathbb{R}^3$  from the cloud  $T_-$ , replace the point  $O_-$  with  $O_+ = (0, 0, +1)$ .

$$(5.2.3.2) \begin{cases} (x_3 + 2)^2 + y_3^2 = |RC_3|^2 - 4 = (2l_2)^2, \\ (x_1 - 2)^2 + y_1^2 = |GC_1|^2 - 4 = (2l_2)^2; \end{cases}$$

$$(5.2.3.3) \begin{cases} (x_1 + 2)^2 + y_1^2 = |RC_1|^2 - 4 = (2l_3)^2, \\ (x_2 - 2)^2 + y_2^2 = |GC_2|^2 - 4 = (2l_3)^2. \end{cases}$$

Comparing the 1st part of (5.2.3.1) with the 2nd part of (5.2.3.3), we get  $(2l_1)^2 - 4x_2 = (2l_3)^2 + 4x_2$ , so  $x_2 = \frac{l_1^2 - l_3^2}{2}$ . Similarly,  $x_3 = \frac{l_2^2 - l_1^2}{2}$ ,  $x_1 = \frac{l_3^2 - l_2^2}{2}$  so that  $x_1 + x_2 + x_3 = 0$ . From the second part of (5.2.3.2), we get  $x_1^2 - 4x_1 + 4 + y_1^2 = 4l_2^2$ , so

$$|O_{\pm}C_1|^2 = x_1^2 + y_1^2 + 1 = 4l_2^2 + 4x_1 - 3 = 2l_2^2 + 2l_3^2 - 3,$$

$$\text{similarly } |O_{\pm}C_2|^2 = 2l_3^2 + 2l_1^2 - 3, |O_{\pm}C_3|^2 = 2l_1^2 + 2l_2^2 - 3.$$

$$\text{Then } |C_1C_2|^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 = x_1^2 + y_1^2.$$

The last columns in Tables 5.6 and 5.7 show the pairs of distances that distinguish  $T_+ \neq T_-$ . The distributions  $\overline{\text{SDD}}(T_{\pm}; 2)$  can differ only by  $\overline{\text{RDD}}$ s of the pairs  $\{R, O_{\pm}\}, \{G, O_{\pm}\}, \{R, C_i\}, \{G, C_i\}$ , where  $i \in \{1, 2, 3\}$  is considered modulo 3 so

distances of $T_-$	$R$	$G$	$C_1$	$C_2$	$C_3$	$O_-$
$R = (-2, 0, -2)$	0	$\sqrt{32}$	$2\sqrt{l_3^2 + 1}$	$2\sqrt{l_1^2 + 1}$	$2\sqrt{l_2^2 + 1}$	$\sqrt{5}$
$G = (+2, 0, +2)$	$\sqrt{32}$	0	$2\sqrt{l_2^2 + 1}$	$2\sqrt{l_3^2 + 1}$	$2\sqrt{l_1^2 + 1}$	$\sqrt{13}$
$C_1 = (x_1, y_1, 0)$	$2\sqrt{l_3^2 + 1}$	$2\sqrt{l_2^2 + 1}$	0	$ C_1 C_2 $	$ C_3 C_1 $	$\sqrt{2l_2^2 + 2l_3^2 - 3}$
$C_2 = (x_2, y_2, 0)$	$2\sqrt{l_1^2 + 1}$	$2\sqrt{l_3^2 + 1}$	$ C_1 C_2 $	0	$ C_2 C_3 $	$\sqrt{2l_3^2 + 2l_1^2 - 3}$
$C_3 = (x_3, y_3, 0)$	$2\sqrt{l_2^2 + 1}$	$2\sqrt{l_1^2 + 1}$	$ C_3 C_1 $	$ C_2 C_3 $	0	$\sqrt{2l_1^2 + 2l_2^2 - 3}$
$O_- = (0, 0, -1)$	$\sqrt{5}$	$\sqrt{13}$	$\sqrt{2l_2^2 + 2l_3^2 - 3}$	$\sqrt{2l_3^2 + 2l_1^2 - 3}$	$\sqrt{2l_1^2 + 2l_2^2 - 3}$	0

distances of $T_+$	$R$	$G$	$C_1$	$C_2$	$C_3$	$O_+$
$R = (-2, 0, -2)$	0	$\sqrt{32}$	$2\sqrt{l_3^2 + 1}$	$2\sqrt{l_1^2 + 1}$	$2\sqrt{l_2^2 + 1}$	$\sqrt{13}$
$G = (+2, 0, +2)$	$\sqrt{32}$	0	$2\sqrt{l_2^2 + 1}$	$2\sqrt{l_3^2 + 1}$	$2\sqrt{l_1^2 + 1}$	$\sqrt{5}$
$C_1 = (x_1, y_1, 0)$	$2\sqrt{l_3^2 + 1}$	$2\sqrt{l_2^2 + 1}$	0	$ C_1 C_2 $	$ C_3 C_1 $	$\sqrt{2l_2^2 + 2l_3^2 - 3}$
$C_2 = (x_2, y_2, 0)$	$2\sqrt{l_1^2 + 1}$	$2\sqrt{l_3^2 + 1}$	$ C_1 C_2 $	0	$ C_2 C_3 $	$\sqrt{2l_3^2 + 2l_1^2 - 3}$
$C_3 = (x_3, y_3, 0)$	$2\sqrt{l_2^2 + 1}$	$2\sqrt{l_1^2 + 1}$	$ C_3 C_1 $	$ C_2 C_3 $	0	$\sqrt{2l_1^2 + 2l_2^2 - 3}$
$O_+ = (0, 0, +1)$	$\sqrt{13}$	$\sqrt{5}$	$\sqrt{2l_2^2 + 2l_3^2 - 3}$	$\sqrt{2l_3^2 + 2l_1^2 - 3}$	$\sqrt{2l_1^2 + 2l_2^2 - 3}$	0

**Table 5.5** The distance matrices of the 6-point clouds  $T_{\pm}$  in Fig. 5.5 motivated by [3, Figure S4(C)].

that  $1 - 1 \equiv 3 \pmod{3}$ . In rows of corresponding pairs of points, some pairs of distances are the same in both  $\widetilde{\text{SDD}}(T_{\pm}; 2)$ , but other pairs differ. If  $l_1, l_2, l_3$  are pairwise distinct, the rows  $\{R, O_-\}, \{G, O_+\}$  include three different pairs of distances, so  $\widetilde{\text{SDD}}(T_-; 2) \neq \widetilde{\text{SDD}}(T_+; 2)$ .

Table 5.5 contains all pairwise distances between the points of  $T_{\pm}$ . We show that  $T_{\pm}$  differ by the simplified invariants  $\widetilde{\text{SDD}}(T_{\pm}; 2)$  below. In each column of  $R(C; A)$ , we additionally allow any permutation of elements independent of other columns, so we could order each column (a pair of distances) lexicographically. Denote the resulting simplification of RDD by  $\widetilde{\text{RDD}}$ . Then  $\widetilde{\text{SDD}}(T_{\pm}; 2)$  have identical  $\widetilde{\text{RDD}}$ s for the 2-point subsets  $A$  from the list  $\{R, G\}, \{O_{\pm}, C_i\}, \{C_i, C_j\}$  for distinct  $i, j = 1, 2, 3$ .

For example, both  $\widetilde{\text{RDD}}(T_{\pm}; \{R, G\})$  start with the distance  $|R - G| = \sqrt{32}$  and then include the same four pairs  $(\sqrt{5}, \sqrt{13}), (2\sqrt{l_i^2 + 1}, 2\sqrt{l_{i-1}^2 + 1})$  for  $i \in \{1, 2, 3\}$  modulo 3, which should be ordered and written lexicographically.

$T_-$ pair	distance	common pairs in $\widetilde{\text{SDD}}(T_{\pm}; 2)$	pairs that differ in $\widetilde{\text{SDD}}(T_-; 2)$
$\{R, O_-\}$	$\sqrt{5}$	$(\sqrt{13}, \sqrt{32})$ to $G$	$(2\sqrt{l_3^2 + 1}, \sqrt{2l_2^2 + 2l_3^2 - 3})$ to $C_1$ , $(2\sqrt{l_1^2 + 1}, \sqrt{2l_3^2 + 2l_1^2 - 3})$ to $C_2$ , $(2\sqrt{l_2^2 + 1}, \sqrt{2l_1^2 + 2l_2^2 - 3})$ to $C_3$
$\{G, O_-\}$	$\sqrt{13}$	$(\sqrt{5}, \sqrt{32})$ to $R$	$(2\sqrt{l_2^2 + 1}, \sqrt{2l_2^2 + 2l_3^2 - 3})$ to $C_1$ , $(2\sqrt{l_3^2 + 1}, \sqrt{2l_3^2 + 2l_1^2 - 3})$ to $C_2$ , $(2\sqrt{l_1^2 + 1}, \sqrt{2l_1^2 + 2l_2^2 - 3})$ to $C_3$
$\{R, C_{i+1}\}$	$2\sqrt{l_i^2 + 1}$	$(2\sqrt{l_{i-1}^2 + 1}, \sqrt{32})$ to $G$ , $(2\sqrt{l_{i+1}^2 + 1},  C_{i+1}C_{i-1} )$ to $C_{i-1}$ , $(2\sqrt{l_{i-1}^2 + 1},  C_iC_{i+1} )$ to $C_i$	$(\sqrt{5}, \sqrt{2l_{i-1}^2 + 2l_i^2 - 3})$ to $O_-$
$\{G, C_{i-1}\}$	$2\sqrt{l_i^2 + 1}$	$(2\sqrt{l_{i+1}^2 + 1}, \sqrt{32})$ to $R$ , $(2\sqrt{l_{i-1}^2 + 1},  C_{i+1}C_{i-1} )$ to $C_{i+1}$ , $(2\sqrt{l_{i+1}^2 + 1},  C_{i-1}C_i )$ to $C_i$	$(\sqrt{13}, \sqrt{2l_i^2 + 2l_{i+1}^2 - 3})$ to $O_-$

**Table 5.6** Pairs of distances in the simplified invariant  $\widetilde{\text{SDD}}(T_-; 2)$ . For comparison with  $\widetilde{\text{SDD}}(T_+; 2)$ , see Table 5.7. The highlighted differences imply that  $\text{SDD}(T_-; 2) \neq \text{SDD}(T_+; 2)$ , so  $T_- \neq T_+$ .

Hence, it makes sense to compare  $\widetilde{\text{SDD}}(T_{\pm}; 2)$  only by the remaining  $\widetilde{\text{RDD}}(T_{\pm}; A)$  for  $A$  from the list  $\{R, O_{\pm}\}, \{G, O_{\pm}\}, \{R, C_i\}, \{G, C_j\}$  in Tables 5.6 and 5.7.

Without loss of generality, assume that  $l_1 \geq l_2 \geq l_3$ . If all the lengths are distinct, then  $l_1 > l_2 > l_3$ . Then the rows for  $\{R, O_-\}$  and  $\{G, O_+\}$  differ in Tables 5.6 and 5.7 even after ordering each pair so that a smaller distance precedes a larger one, and after writing all pairs lexicographically. So  $\widetilde{\text{SDD}}(T_-; 2) \neq \widetilde{\text{SDD}}(T_+; 2)$  unless two of  $l_i$  are equal. If (say)  $l_1 = l_2$ , the lexicographically ordered rows of  $\{R, O_-\}$  and  $\{G, O_+\}$  coincide in  $\widetilde{\text{SDD}}(T_{\pm}; 2)$ , similarly for the rows of  $\{G, O_-\}$  and  $\{R, O_+\}$ . Hence, it suffices to compare only the six rows for the remaining pairs  $\{R, C_i\}, \{G, C_j\}$  in  $\widetilde{\text{SDD}}(T_{\pm}; 2)$ .

For  $l_1 = l_2$ , we get  $x_3 = \frac{l_2^2 - l_1^2}{2} = 0$  and  $x_1 = -x_2 = \frac{l_3^2 - l_2^2}{2}$ . In equation (5.2.3.3) the equality  $(x_1 + 2)^2 + y_1^2 = (x_2 - 2)^2 + y_2^2$  with  $x_1 = -x_2$  implies that  $y_1^2 = y_2^2$ . The more degenerate case  $l_1 = l_2 = l_3$ , means that  $x_1 = x_2 = x_3 = 0$  and  $y_1^2 = y_2^2 = y_3^2$ , hence at least two of  $C_1, C_2, C_3$  should coincide. The above contradiction means that it

$T_+$ pair	distance	common pairs in $\widetilde{\text{SDD}}(T_{\pm}; 2)$	pairs that differ in $\widetilde{\text{SDD}}(T_+; 2)$
$\{G, O_+\}$	$\sqrt{5}$	$(\sqrt{13}, \sqrt{32})$ to $R$	$(2\sqrt{l_2^2 + 1}, \sqrt{2l_2^2 + 2l_3^2 - 3})$ to $C_1$ , $(2\sqrt{l_3^2 + 1}, \sqrt{2l_3^2 + 2l_1^2 - 3})$ to $C_2$ , $(2\sqrt{l_1^2 + 1}, \sqrt{2l_1^2 + 2l_2^2 - 3})$ to $C_3$
$\{R, O_+\}$	$\sqrt{13}$	$(\sqrt{5}, \sqrt{32})$ to $G$	$(2\sqrt{l_3^2 + 1}, \sqrt{2l_2^2 + 2l_3^2 - 3})$ to $C_1$ , $(2\sqrt{l_1^2 + 1}, \sqrt{2l_3^2 + 2l_1^2 - 3})$ to $C_2$ , $(2\sqrt{l_2^2 + 1}, \sqrt{2l_1^2 + 2l_2^2 - 3})$ to $C_3$
$\{R, C_{i+1}\}$	$2\sqrt{l_i^2 + 1}$	$(2\sqrt{l_{i-1}^2 + 1}, \sqrt{32})$ to $G$ , $(2\sqrt{l_{i+1}^2 + 1},  C_{i+1}C_{i-1} )$ to $C_{i-1}$ , $(2\sqrt{l_{i-1}^2 + 1},  C_iC_{i+1} )$ to $C_i$	$(\sqrt{13}, \sqrt{2l_{i-1}^2 + 2l_i^2 - 3})$ to $O_+$
$\{G, C_{i-1}\}$	$2\sqrt{l_i^2 + 1}$	$(2\sqrt{l_{i+1}^2 + 1}, \sqrt{32})$ to $R$ , $(2\sqrt{l_{i-1}^2 + 1},  C_{i+1}C_{i-1} )$ to $C_{i+1}$ , $(2\sqrt{l_{i+1}^2 + 1},  C_{i-1}C_i )$ to $C_i$	$(\sqrt{5}, \sqrt{2l_i^2 + 2l_{i+1}^2 - 3})$ to $O_+$

**Table 5.7** Pairs of distances in the simplified invariant  $\widetilde{\text{SDD}}(T_+; 2)$ . For comparison with  $\widetilde{\text{SDD}}(T_-; 2)$ , see Table 5.6. The highlighted differences imply that  $\widetilde{\text{SDD}}(T_-; 2) \neq \widetilde{\text{SDD}}(T_+; 2)$ , so  $T_- \neq T_+$ .

remains to consider the case  $l_1 = l_2 > l_3$  when  $x_1 = -x_2 \neq 0 = x_3$  and  $y_1 = \pm y_2$ , see Fig. 5.5.

If  $y_1 = y_2$ , the clouds  $T_{\pm}$  are isometric by  $(x, y, z) \mapsto (-x, y, -z)$ . If  $y_1 = -y_2$  and  $y_3 = 0$ , the clouds  $T_{\pm}$  are isometric by the isometry  $(x, y, z) \mapsto (-x, -y, -z)$ . If  $y_1 = -y_2$  and  $y_3 \neq 0$ , then  $C_1 = (x_1, y_1, 0)$ ,  $C_2 = (-x_1, -y_1, 0)$ ,  $C_3 \neq (0, 0, 0)$ . Then among the six remaining rows, only the rows of  $\{R, C_1\}$ ,  $\{G, C_2\}$  have points at the distance  $2\sqrt{l_3^2 + 1}$ , see Tables 5.6 and 5.7 for  $i = 3$  considered modulo 3. Then  $i+1 \equiv 1 \pmod{3}$ ,  $i-1 \equiv 2 \pmod{3}$ , so  $l_{i+1} = l_1 = l_2 = l_{i-1}$ .

Looking at the rows of  $\{R, C_1\}$ ,  $\{G, C_2\}$ , the three common pairs in each of  $\text{SDD}(T_{\pm}; 2)$  include the same distance  $2\sqrt{l_1^2 + 1} = 2\sqrt{l_2^2 + 1}$  but differ by  $|C_{i-1}C_i| = |C_2C_3| \neq |C_3C_1| = |C_iC_{i+1}|$  as  $C_1 = \pm C_2$ ,  $C_3 \neq (0, 0, 0)$ .

This couple of different rows implies that  $\text{SDD}(T_-; 2) \neq \text{SDD}(T_+; 2)$  due to the swapped distances  $\sqrt{5}, \sqrt{13}$  in the remaining pairs, see Tables 5.8 for the clouds  $T_\pm$  in Fig. 5.5 with  $l_1 = l_2 = \frac{\sqrt{13}}{2}, l_3 = \frac{\sqrt{5}}{2}$ .  $\blacklozenge$

$T_-$ pair	distance	dist. to neighb. 1	dist. to neighb. 2	dist. to neighb. 3	dist. to neighb. 4
$\{R, C_1\}$	3	$(\sqrt{2}, \sqrt{17})$ to $C_3$	$(\sqrt{5}, \sqrt{6})$ to $O_-$	$(\sqrt{17}, \sqrt{20})$ to $C_2$	$(\sqrt{17}, \sqrt{32})$ to $G$
$\{G, C_2\}$	3	$(\sqrt{6}, \sqrt{13})$ to $O_-$	$(\sqrt{17}, \sqrt{20})$ to $C_1$	$(\sqrt{17}, \sqrt{26})$ to $C_3$	$(\sqrt{17}, \sqrt{32})$ to $R$
$T_+$ pair	distance	dist. to neighb. 1	dist. to neighb. 2	dist. to neighb. 3	dist. to neighb. 4
$\{R, C_1\}$	3	$(\sqrt{2}, \sqrt{17})$ to $C_3$	$(\sqrt{6}, \sqrt{13})$ to $O_+$	$(\sqrt{17}, \sqrt{20})$ to $C_2$	$(\sqrt{17}, \sqrt{32})$ to $G$
$\{G, C_2\}$	3	$(\sqrt{5}, \sqrt{6})$ to $O_+$	$(\sqrt{17}, \sqrt{20})$ to $C_1$	$(\sqrt{17}, \sqrt{26})$ to $C_3$	$(\sqrt{17}, \sqrt{32})$ to $R$

**Table 5.8** The above rows show that  $\text{SDD}(T_-; 2) \neq \text{SDD}(T_+; 2)$  for the clouds  $T_\pm$  with  $C_1 = (-1, 2, 0), C_2 = (1, -2, 0), C_3 = (0, 3, 0)$  so that  $l_1 = l_2 = \frac{\sqrt{13}}{2}, l_3 = \frac{\sqrt{5}}{2}$  in Tables 5.6 and 5.7.

Examples 5.2.1, 5.2.2, and 5.2.3 motivate the following conjecture.

**Conjecture 5.2.4** (completeness of  $\text{SDD}(C; h)$  in  $\mathbb{R}^n$ ). For any  $n \geq 2$ , there is some  $2 \leq h \leq n$  such that the Simplexwise Distance Distribution  $\text{SDD}(C; h)$  from Definition 5.1.3 is a complete isometry invariant of all clouds  $C \subset \mathbb{R}^n$ .  $\star$

### 5.3 Continuous metrics on Simplexwise Distance Distributions

This section defines Lipschitz continuous metrics on SDDs, which can be computable in a polynomial time of the number  $m$  of points, for a fixed order  $h$ . The  $m - h$  permutable columns of the matrix  $R(C; A)$  in RDD from Definition 5.1.1 can be interpreted as  $m - h$  unlabelled points in  $\mathbb{R}^h$ . Since any isometry is bijective, the simplest metric respecting bijections is the bottleneck distance  $\text{BD}$  from Example 1.3.3(b).

**Definition 5.3.1** (the max metric  $M_\infty$  on RDDs). For any  $m$ -point clouds and ordered  $h$ -point base sequences  $A \subset C$  and  $A' \subset C'$ , set

$$d(\xi) = \max\{L_\infty(\xi(D(A)), D(A')), \text{BD}(\xi(R(C; A)), R(C'; A'))\}$$

for a permutation  $\xi \in S_h$  on  $h$  points. Then the *max metric* on Relative Distance Distributions is defined as  $M_\infty(\text{RDD}(C; A), \text{RDD}(C'; A')) = \min_{\xi \in S_h} d(\xi)$ .  $\blacktriangle$

We will use only  $h = n$  for Euclidean space  $\mathbb{R}^n$ , so the factor  $h!$  in Definition 5.3.1 is practically small for  $n = 2, 3$ . For  $h = 1$  and a 1-point sequence  $A \subset C$ , the matrix

$D(A)$  is empty, so  $d(\xi) = \text{BD}(\xi(R(C; A)), R(C'; A'))$ . The metric  $M_\infty$  on RDDs will be used for intermediate costs to get metrics on unordered collections of RDDs (SDDs) by using the standard tools in Definitions 3.5.4 and 5.3.2 below.

**Definition 5.3.2** (Linear Assignment Cost of a matrix [4]). For any  $k \times k$  matrix of costs  $c(i, j) \geq 0, i, j \in \{1, \dots, k\}$ , the *Linear Assignment Cost*  $\text{LAC} = \frac{1}{k} \min_g \sum_{i=1}^k c(i, g(i))$  is minimized for all bijections  $g$  on the indices  $1, \dots, k$ .  $\blacktriangle$

The normalisation factor  $\frac{1}{k}$  in LAC makes this metric better comparable with EMD in Definition 3.5.4 whose weights sum up to 1. For both LAC and EMD, the matrix of initial costs will consist of max metrics between all RDDs in two given SDDs.

Theorem 5.3.3(b) extends the  $O(m^{1.5} \log^n m)$  algorithm for fixed clouds of  $m$  unlabelled points in [2, Theorem 6.5] to the harder case of isometry classes but keeps the polynomial time in  $m$  for a fixed dimension  $n$ .

**Theorem 5.3.3** (time of metrics on SDDs, [6, Theorem 5.5]). **(a)** For any  $m$ -point clouds  $C, C'$  in their own metric spaces and  $h \geq 1$ , let the Simplexwise Distance Distributions  $\text{SDD}(C; h)$  and  $\text{SDD}(C'; h)$  consist of  $k = \binom{m}{h}$  RDDs with equal weights  $\frac{1}{k}$  without collapsing identical RDDs.

**(b)** Using the  $k \times k$  matrix of costs computed by the max metric  $M_\infty$  between RDDs from  $\text{SDD}(C; h)$  and  $\text{SDD}(C'; h)$ , the Linear Assignment Cost LAC from Definition 5.3.2 satisfies all metric axioms on SDDs and can be computed in time  $O(h!(h^2 + m^{1.5} \log^h m)k^2 + k^3 \log k)$ .

**(b)** Let  $\text{SDD}(C; h)$  and  $\text{SDD}(C'; h)$  have a maximum size  $l \leq k$  after collapsing identical RDDs. Using the same matrix of max metrics as in part (b), the EMD from Definition 3.5.4 satisfies all metric axioms on SDDs and can be computed in time  $O(h!(h^2 + m^{1.5} \log^h m)l^2 + l^3 \log l)$ .  $\blacksquare$

Theorem 5.3.4 substantially generalizes the fact that perturbing two points in their  $\varepsilon$ -neighbourhoods changes the distance between these points by at most  $2\varepsilon$ .

**Theorem 5.3.4** (Lipschitz continuity of SDDs, [6, Theorem 5.8]). In any metric space, let  $C'$  be obtained from a cloud  $C$  by perturbing every point of  $C$  within its  $\varepsilon$ -neighbourhood. For any order  $h \geq 1$ ,  $\text{SDD}(C; h)$  changes by at most  $2\varepsilon$  in the LAC and EMD metrics. The lower bound holds:  $\text{EMD}(\text{SDD}(C; h), \text{SDD}(C'; h)) \geq |\text{SDM}(C; h, 1) - \text{SDM}(C'; h, 1)|_\infty$ .  $\blacksquare$

## 5.4 Measured Simplexwise Distributions for metric-measure spaces

This section adapts Simplexwise Distance Distributions SDD to metric-measure spaces.

**Definition 5.4.1** (metric-measure space). A *metric-measure space*  $(X, d_X, \mu_X)$  is a compact space  $X$  with a metric  $d_X$  and a Borel measure  $\mu_X$  such that  $\mu_X(X) < +\infty$ . An *isomorphism* between metric-measure spaces is an isometry  $f : X \rightarrow Y$  that respects the measures in the sense that  $\mu_Y(U) = \mu_X(f^{-1}(U))$  for any subset  $U \subset Y$ .  $\blacktriangle$

Dividing  $\mu_X(U)$  by the measure  $\mu_X(X) < +\infty$  for any  $U \subset X$ , we can assume that  $\mu_X(X) = 1$ , so  $\mu_X$  is a probability measure. Any metric space  $X$  of  $m$  points can be considered a metric-measure space with the uniform measure  $\mu_X(p) = \frac{1}{m}$  for all  $p \in X$ . On two points  $0, 1$  in  $\mathbb{R}$ , the metric-measure spaces  $X = (\{0, 1\}, 1, \{\frac{1}{2}, \frac{1}{2}\})$  and  $Y = (\{0, 1\}, 1, \{\frac{1}{3}, \frac{2}{3}\})$  are isometric but not isomorphic because of different weights.

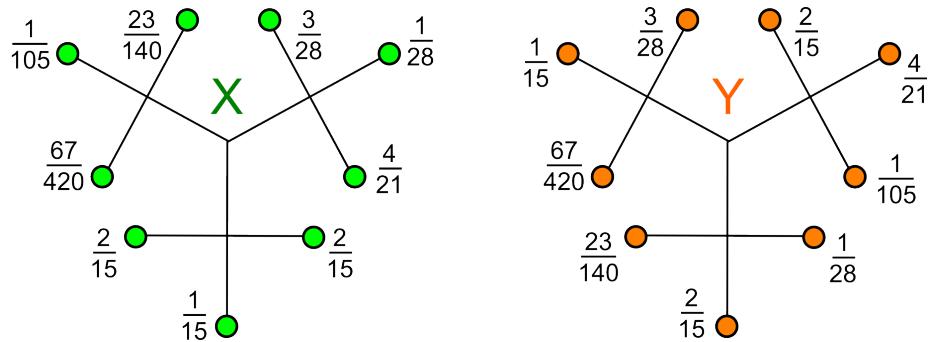
Problem 4.4.4 becomes much harder if we replace isometries between metric spaces with isomorphisms between metric-measure spaces because all known isometry invariants should be further refined to distinguish under isomorphism. Definition 5.4.2 extends the local distribution of distances from [7, Definition 5.5] to orders  $h > 1$ .

**Definition 5.4.2** (Measured Simplexwise Distribution MSD). Let  $(X, d_X, \mu_X)$  be any metric-measure space. For any base sequence  $A = (p_1, \dots, p_h)$  of  $h \geq 1$  ordered points of  $X$ , write the triangular distance matrix  $D(A)$  from Definition 5.1.1 row-by-row as the vector  $\mathbf{v}(A) \in \mathbb{R}_+^{h(h-1)/2}$  so that  $\mathbf{v}_k = d_X(p_i, p_j)$  for  $k = h(i-1) + j - 1$ ,  $1 \leq i < j \leq h$ . For a vector  $\mathbf{d} = (d_1, \dots, d_h) \in \mathbb{R}_+^h$  of distance thresholds, the vector  $\mathbf{m}(A; \mathbf{d}) \in \mathbb{R}_+^h$  consists of  $h$  values  $\mu_X(\{q \in X \mid d_X(q, p_i) \leq d_i\})$  for  $i = 1, \dots, h$ .

The *Measured Simplexwise Distribution* of order  $h \geq 1$  is the function  $\text{MSD}[X; h] : X^h \times \mathbb{R}_+^h \rightarrow \mathbb{R}_+^{h(h+1)/2}$  mapping any  $A \in X^h$  and  $\mathbf{d} \in \mathbb{R}_+^h$  to the pair  $[\mathbf{v}(A), \mathbf{m}(A; \mathbf{d})]$  considered as a concatenated vector in  $\mathbb{R}_+^{h(h+1)/2}$ .  $\blacktriangle$

For  $h = 1$ , the vector  $\mathbf{v}(A)$  is empty and the Measured Simplexwise Distribution of order  $h = 1$  coincides with the local distribution of distances [7, Definition 5.5]  $\text{MSD}[X; 1] : X \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$  mapping any point  $p \in X$  and a threshold  $d \in \mathbb{R}_+$  to the measure value  $\mu_X(\{q \in X \mid d_X(q, p) \leq d\})$ .

Any permutation  $\xi$  on indices  $1, \dots, h$  naturally permutes the components of  $\text{MSD}[X; h]$ . If  $X$  consists of  $m$  points,  $\text{MSD}[X; h]$  reduces to the finite collection of  $\binom{m}{h}$  vectors  $\text{VID}(A)$  paired with fields  $\text{VSM}(A; \mathbf{d}) : \mathbb{R}_+^h \rightarrow \mathbb{R}_+^h$  only for unordered  $h$ -point subsets  $A \subset X$ , which can be refined to a stronger analogue of SDD below.



**Fig. 5.6** Non-isomorphic metric-measure spaces  $X, Y$  from [7, Fig. 8] have equal local distributions of distances but are distinguished by the new Weighted Simplexwise Distribution of order 1 and the Measured Simplexwise Distributions of order 2, see details in Example 5.4.4. All edges have length  $\frac{1}{2}$ .

**Definition 5.4.3** (Weighted Simplexwise Distribution WSD). Let  $X$  be a finite metric-measure space whose any point  $p$  has a weight  $w(p)$ . For  $h \geq 1$  and a base sequence  $A = (p_1, \dots, p_h)$  of  $h$  ordered points of  $X$  in Definition 5.1.1, endow any distance  $d(p, q)$  in  $D(A)$  with the unordered pair  $w(p), w(q)$  of weights. For every point  $q \in X - A$ , put the weight  $w(q)$  in the extra  $(h + 1)$ -st row of the matrix  $M(X; A)$  whose columns are indexed by unordered  $q \in X - A$ . If  $h = 1$  and  $A = p_1$ , set  $D(A) = w(p_1)$ .

The *Weighted Distance Distribution*  $\text{WDD}(X; A)$  is the equivalence class of the pair  $[D(A); M(X; A)]$  under permutations  $\xi \in S_h$  acting on  $A$ . The *Weighted Simplexwise Distribution*  $\text{WSD}(X)$  is the unordered collection of  $\text{WDD}(X; A)$  for all subsets  $A \subset X$  of unordered  $h$  points.  $\blacktriangle$

For finite metric-measure spaces, a metric on WDDs can be defined similar to  $M_\infty$  from Definition 5.3.1 by combining the weights and distances. Then LAC and EMD from Definitions 5.3.2 and 3.5.4 can be computed as in Theorem 5.3.3.

**Example 5.4.4** (the strength of WSD for 9-point trees). Fig. 5.6 shows metric-measure spaces  $X, Y$  on 9 points visualised as trees [7, Fig. 8]. All edges have length  $\frac{1}{2}$  and induce the shortest-path metrics  $d_X, d_Y$ . The sum of weights in every small branch of 3 nodes is  $\frac{1}{3}$ . These metric-measure spaces  $X, Y$  have all inter-point distances only 1 and 2, and equal local distributions of distances  $\text{MSD}[X; 1] = \text{MSD}[Y; 1]$  by [7, Example 5.6].

Indeed, both MSDs can be considered the same set of 9 piecewise constant functions  $\mu(p)$  taking values  $w(p), \frac{1}{3}$ , and 1 on the intervals  $[0, 1), [1, 2), [2, +\infty)$ , respectively.

However, WSDs have more pointwise data:  $\text{WSD}[X; 1]$  has  $A(D) = w(p) = \frac{23}{140}$  and the following  $2 \times 8$  matrix

$$M(X; p) = \begin{pmatrix} 1 & 1 & 2 & 2 & 2 & 2 & 2 & 2 \\ \frac{1}{105} & \frac{67}{240} & \frac{2}{15} & \frac{1}{15} & \frac{2}{15} & \frac{4}{21} & \frac{1}{28} & \frac{3}{28} \end{pmatrix},$$

but  $\text{WSD}[Y; 1]$  has another matrix for  $w(p) = \frac{23}{140}$ .

$$M(Y; p) = \begin{pmatrix} 1 & 1 & 2 & 2 & 2 & 2 & 2 & 2 \\ \frac{2}{15} & \frac{1}{28} & \frac{1}{105} & \frac{4}{21} & \frac{2}{15} & \frac{3}{28} & \frac{1}{15} & \frac{67}{420} \end{pmatrix}.$$

The above matrices with freely permutable columns are different, so  $X, Y$  are distinguished by the Weighted Simplexwise Distribution WSD of order  $h = 1$ .

Also,  $\text{MSD}[X; 2] \neq \text{MSD}[Y; 2]$  because, for any base sequence  $A = (p, q) \in X^2$ , we have  $\text{VSM}[X; 2](A; d_1, d_2) = (w(p), w(q))$  for  $d_1, d_2 < 1$  since all other points have minimum distance 1 from  $p, q$ , similarly for  $Y$ . The unique points  $p, q$  of weights  $w(p) = \frac{23}{140}$  and  $w(q) = \frac{67}{420}$  have different distances  $d_X(p, q) = 1$  and  $d_Y(p, q) = 2$ .

Then  $\text{MSD}[X; 2] \neq \text{MSD}[Y; 2]$  differ by the uniquely identifiable fields mapping  $[0, 1)^2$  to the constant vector  $(w(p), w(q))$  with  $\text{VID}_X(A) = 1 \neq 2 = \text{VID}_Y(A)$ .  $\blacklozenge$

We conjecture that any metric-measure spaces  $X, Y$  are distinguished under isomorphism by Measured Simplexwise Distributions for a high enough  $h$  depending on  $X, Y$ .

## References

1. Mireille Boutin and Gregor Kemper. On reconstructing n-point configurations from the distribution of distances or areas. *Advances in Applied Mathematics*, 32(4):709–735, 2004.
2. Alon Efrat, Alon Itai, and Matthew J Katz. Geometry helps in bottleneck matching and related problems. *Algorithmica*, 31(1):1–28, 2001.
3. S. Pozdnyakov et al. Incompleteness of atomic structure representations. *Phys. Rev. Lett.*, 125:166001, 2020.
4. Michael L Fredman and Robert Endre Tarjan. Fibonacci heaps and their uses in improved network optimization algorithms. *Journal of the ACM*, 34(3):596–615, 1987.
5. Ernest Sydney Keeping. *Introduction to statistical inference*. Courier Corporation, 1995.
6. Vitaliy Kurlin. Simplexwise Distance Distributions for finite spaces with metrics and measures. *arXiv:2303.14161 (latest version at <http://kurlin.org/projects/cloud-isometry-spaces/SDD.pdf>)*, 2023.
7. Facundo Mémoli. Gromov–Wasserstein distances and the metric approach to object matching. *Foundations of Computational Mathematics*, 11(4):417–487, 2011.
8. Daniel E Widdowson and Vitaliy A Kurlin. Recognizing rigid patterns of unlabeled point clouds by complete and continuous isometry invariants with no false negatives and no false positives. In *Computer Vision and Pattern Recognition*, pages 1275–1284, 2023.

## Chapter 6

# Complete and Lipschitz continuous invariants of unordered points in $\mathbb{R}^n$

**Abstract** This chapter leverages the Euclidean structure of  $\mathbb{R}^n$  to improve the Simplexwise Distance Distribution to a smaller Simplexwise Centred Distribution (SCD) for any unordered points. The new invariant is complete under rigid motion and computable in polynomial time for a fixed dimension. The key ingredient of Lipschitz continuity is the new strength of a simplex, which is a linear-growth analogue of the simplex volume.

## 6.1 Geo-mapping problem under rigid motion in $\mathbb{R}^n$

This chapter follows paper [4] and its extension [2] to Euclidean spaces. Problem 6.1.1 adjusts Geo-Mapping Problem 1.4.5 to finite clouds of unordered points in  $\mathbb{R}^n$ .

The major difference with Problem 4.4.4, which was stated under isometry in a metric space, is the full completeness under the stronger equivalence of rigid motion in  $\mathbb{R}^n$ .

**Problem 6.1.1** (partial case of Problem 1.4.5 for clouds under rigid motion in  $\mathbb{R}^n$ ). Design a map  $I$  on finite clouds of unordered points in  $\mathbb{R}^n$  with values in a metric space satisfying the following conditions.

- (a) *Completeness*: any clouds  $A, B \subset \mathbb{R}^n$  are related by rigid motion ( $A \cong B$ ) in  $\mathbb{R}^n$  if and only if  $I(A) = I(B)$ .
- (b) *Metric*: there is a distance  $d$  on the *invariant space*  $\{I(A) \mid A \subset \mathbb{R}^n\}$  satisfying all metric axioms in Definition 1.3.1(a).
- (c) *Continuity*: there is a constant  $\lambda$  such that, for any  $\varepsilon > 0$ , if  $B$  is obtained from a cloud  $A \subset \mathbb{R}^n$  by perturbing every point up to Euclidean distance  $\varepsilon$ , then  $d(I(A), I(B)) \leq \lambda\varepsilon$ .
- (d) *Computability*: for a fixed dimension  $n$ , the invariant  $I(A)$  and the metric  $d(I(A), I(B))$  can be computed in times that depend polynomially on the maximum size  $\max\{|A|, |B|\}$  of clouds  $A, B \subset \mathbb{R}^n$ . ★

Problem 6.1.1 will be fully solved by the Oriented Simplexwise Distribution (OSD), which we introduce in Definition 6.1.3 after a few auxiliary concepts below.

**Definition 6.1.2** (matrices  $D(A)$  and  $M(C; A)$  for  $A \subset C$ ). Let  $C$  be a cloud of  $m$  unordered points in  $\mathbb{R}^n$  with a fixed orientation. Let  $A = (p_1, \dots, p_n)$  be a *base sequence* of  $n$  distinct ordered points of  $C$ . Let  $D(A)$  be the  $n \times n$  *distance matrix* whose entry  $D(A)_{i,j}$  is Euclidean distance  $|p_i - p_j|$  for  $1 \leq i < j \leq n$ , all other entries are zeros. For any other point  $q \in C - A$ , write distances from  $q$  to  $p_1, \dots, p_n$  as a column. Form the  $n \times (m - n)$ -matrix by these  $m - n$  lexicographically ordered columns.

At the bottom of the column of a  $q \in C - A$ , add the sign of the determinant consisting of the vectors  $q - p_1, \dots, q - p_n$ . The resulting  $(n + 1) \times (m - n)$ -matrix with signs in the bottom  $(n + 1)$ -st row is the *oriented relative distance matrix*  $M(C; A)$ .  $\blacktriangle$

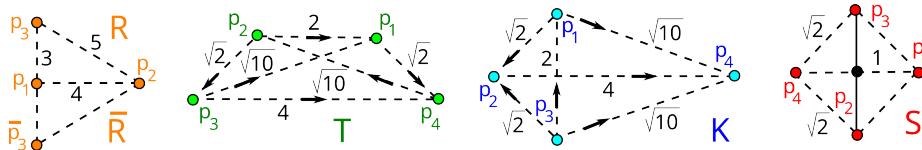
Let  $S_n$  denote the permutation group on indices  $1, \dots, n$ . Any permutation  $\xi \in S_n$  is a composition of some  $t$  transpositions  $i \leftrightarrow j$  and has sign  $\text{sign}(\xi) = (-1)^t$ .

**Definition 6.1.3** (oriented distributions  $\text{ORD}(C; A)$  and  $\text{OSD}(C)$  for a cloud  $C \subset \mathbb{R}^n$ ). Any permutation  $\xi \in S_n$  acts on  $D(A)$  by mapping  $D(A)_{ij}$  to  $D(A)_{kl}$ , where  $k \leq l$  is the pair  $\xi(i), \xi(j) - 1$  written in increasing order. Then the permutation  $\xi$  acts on  $M(C; A)$  by mapping any  $i$ -th row to the  $\xi(i)$ -th row and by multiplying the  $(n + 1)$ -st row by  $\text{sign}(\xi)$ , after which all columns are written in the lexicographic order.

The *Oriented Relative Distribution*  $\text{ORD}(C; A)$  is the equivalence class of the pair  $[D(A); M(C; A)]$  under all permutations  $\xi \in S_n$  acting on both  $D(A)$  and  $M(C; A)$ .

The *Oriented Simplexwise Distribution*  $\text{OSD}(C)$  is the unordered collection of  $\text{ORD}(C; A)$  for all  $\binom{m}{n}$  unordered subsets  $A \subset C$  of  $n$  points.  $\blacktriangle$

Any mirror reflection in  $\mathbb{R}^n$  reverses the sign of the  $n \times n$  determinant consisting of vectors  $v_1, \dots, v_n \in \mathbb{R}^n$ , hence reverses all signs in the  $(n + 1)$ -st rows of the matrices  $M(C; A)$  in Oriented Relative Distributions.  $\overline{\text{ORD}}(C; A)$  and  $\overline{\text{OSD}}(C)$  denote the ‘mirror images’ of  $\text{ORD}(C; A)$  and  $\text{OSD}(C)$ , respectively, with all signs reversed.



**Fig. 6.1** **1st:** the right-angled cloud  $R \subset \mathbb{R}^2$  consisting of points  $p_1 = (0, 0)$ ,  $p_2 = (4, 0)$ ,  $p_3 = (0, 3)$ , and its mirror image  $\bar{R}$  of  $p_1$ ,  $p_2$ , and  $\bar{p}_3 = (0, -3)$  with respect to the  $x$ -axis. **2nd:** the trapezium cloud  $T \subset \mathbb{R}^2$  consisting of points  $p_1 = (1, 1)$ ,  $p_2 = (-1, 1)$ ,  $p_3 = (-2, 0)$ ,  $p_4 = (2, 0)$ . **3rd:** the kite cloud  $K \subset \mathbb{R}^2$  consisting of points  $p_1 = (0, 1)$ ,  $p_2 = (-1, 0)$ ,  $p_3 = (0, -1)$ ,  $p_4 = (3, 0)$ . **4th:** the square cloud  $S \subset \mathbb{R}^2$  consisting of points  $p_1 = (1, 0)$ ,  $p_2 = (0, -1)$ ,  $p_3 = (-1, 0)$ ,  $p_4 = (0, 1)$ .

**Example 6.1.4** (OSD for mirror images on right-angled clouds). In  $\mathbb{R}^2$  with the counter-clockwise orientation, the right-angled cloud  $R$  on the vertices  $p_1 = (0, 0)$ ,  $p_2 = (4, 0)$ ,  $p_3 = (0, 3)$  of the triangle in Fig. 6.1 (1st) has the distribution  $\text{OSD}(R)$  consisting of

$$\text{ORD}(R; (p_1, p_2)) = [4, \begin{pmatrix} 3 \\ 5 \\ + \end{pmatrix}],$$

$$\text{ORD}(R; (p_2, p_3)) = [5, \begin{pmatrix} 4 \\ 3 \\ + \end{pmatrix}],$$

$$\text{ORD}(R; (p_3, p_1)) = [3, \begin{pmatrix} 5 \\ 4 \\ + \end{pmatrix}].$$

If we swap the points  $p_1 \leftrightarrow p_3$ , the last ORD above changes to the equivalent form

$$\text{ORD}(R; (p_1, p_3)) = [3, \begin{pmatrix} 4 \\ 5 \\ - \end{pmatrix}], \text{ without affecting others. If we reflect } R \text{ with respect to}$$

the  $x$ -axis, the mirror image  $\bar{R}$  of  $p_1, p_2, \bar{p}_3 = (0, -3)$  has  $\text{OSD}(\bar{R}) = \overline{\text{OSD}}(R)$  with

$$\text{ORD}(\bar{R}; (p_1, p_2)) = [4, \begin{pmatrix} 3 \\ 5 \\ - \end{pmatrix}],$$

$$\text{ORD}(\bar{R}; (p_2, \bar{p}_3)) = [5, \begin{pmatrix} 4 \\ 3 \\ - \end{pmatrix}],$$

$$\text{ORD}(\bar{R}; (p'_3, p_1)) = [3, \begin{pmatrix} 5 \\ 4 \\ - \end{pmatrix}],$$

which differs from  $\text{OSD}(R)$  even if we swap points in each pair.  $\blacklozenge$

**Example 6.1.5** (OSD for  $T, K$ ). Fix the *counter-clockwise* orientation on  $\mathbb{R}^2$  so that if a vector  $\mathbf{v}$  is obtained from  $\mathbf{u}$  by a counter-clockwise rotation, then  $\det(u, v) > 0$ . Table 6.1 shows the Oriented Simplexwise Distributions for the clouds  $T, K$  in Fig. 6.1. Each row contains the most similar ORDs whose differences are highlighted.  $\blacklozenge$

Though Problem 6.1.1 did not include the reconstruction condition as in 1.4.5(b), Lemma 6.1.6 below proves this reconstruction in more detail than [2, Lemma 3.6].

Recall that Definition 2.1.2 introduces the affine dimension of a base sequence  $A$  of  $n$  ordered points  $p_1, \dots, p_n$  as the maximum dimension of the vector space generated by all inter-point vectors  $\mathbf{p}_i - \mathbf{p}_j$  for  $i, j \in \{1, \dots, n\}$ .

**Lemma 6.1.6** (reconstruction from ORD). A cloud  $C \subset \mathbb{R}^n$  of  $m > n$  unordered points can be reconstructed, uniquely under rigid motion, from  $\text{ORD}(C; A)$  in Definition 6.1.3 for any base sequence  $A \subseteq C$  with  $\text{aff}(A) = n - 1$ .  $\blacksquare$

*Proof.* By Lemma 2.1.4(b), any base sequence  $A \subseteq C$  can be reconstructed, uniquely under rigid motion in  $\mathbb{R}^n$ , from the triangular distance matrix  $D(A)$  in Definition 6.1.3. We may assume that the first  $n$  points  $p_1, \dots, p_n$  of  $A \subseteq C$  span the subspace of the first  $n - 1$  coordinate axes of  $\mathbb{R}^n$ . We prove that any point  $q \in C - A$  has a unique location in  $\mathbb{R}^n$ , determined by the  $n$  distances  $|q - p_1|, \dots, |q - p_n|$  written in a column of the matrix  $\text{ORD}(C; A)$ . Since the points of  $A$  do not belong to any  $(n - 1)$ -dimensional affine subspace of  $\mathbb{R}^n$ , the  $n$  spheres  $S(p_i; |q - p_i|)$  of radii  $|q_i - p_i|$  and centres  $p_i$ ,  $i = 1, \dots, n$ , contain  $q$  and their full intersections consists of one or two points. We can uniquely choose  $q$  among these two options due to the sign of the determinant on the column vectors  $\mathbf{q} - \mathbf{p}_1, \dots, \mathbf{q} - \mathbf{p}_n$  in the bottom row of  $\text{ORD}(C; A)$ .  $\square$

Lemma 6.1.6 implies that  $\text{ORD}(C; A)$  can have identical columns only for degenerate subsets  $A \subset C$  with  $\text{aff}(A) < n - 1$ . For example, let  $n = 3$  and  $A$  consist of three points  $p_1, p_2, p_3$  in the same straight line  $L \subset \mathbb{R}^3$ . The three distances  $|q - p_i|$ ,  $i = 1, 2, 3$ , to any other point  $q \in C$  outside  $L$  define three spheres  $S(p_i; |q - p_i|)$  that share a common circle in  $\mathbb{R}^3$ , so the position of  $q$  is not uniquely determined in this case.

Though one  $\text{ORD}(C; A)$  with  $\text{aff}(A) = n - 1$  suffices to reconstruct  $C \subset \mathbb{R}^n$  up to rigid motion, the dependence on a subset  $A \subset C$  required us to consider the larger Oriented Simplexwise Distribution  $\text{OSD}(C)$  for all  $n$ -point subsets  $A \subset C$  to get a complete invariant in Theorem 6.1.7. Equality  $\text{OSD}(C) = \text{OSD}(C')$  is interpreted as a bijection  $\text{OSD}(C) \rightarrow \text{OSD}(C')$  matching all ORDs.

**Theorem 6.1.7** (completeness of OSD, [2, Theorem 3.7]). The Oriented Simplexwise Distribution  $\text{OSD}(C)$  in Definition 6.1.3 is a complete isometry invariant and can be computed in time  $O(m^{n+1}/(n - 3)!)$ . So any clouds  $C, C' \subset \mathbb{R}^n$  of  $m$  unlabelled points

ORDs in $\text{OSD}(T)$	ORDs in $\text{OSD}(K)$
$[\sqrt{2}, \begin{pmatrix} 2 & \sqrt{10} \\ \sqrt{10} & 4 \\ - & - \end{pmatrix}]$	$[\sqrt{2}, \begin{pmatrix} 2 & \sqrt{10} \\ \sqrt{2} & 4 \\ - & - \end{pmatrix}]$
$[\sqrt{2}, \begin{pmatrix} 2 & \sqrt{10} \\ \sqrt{10} & 4 \\ + & + \end{pmatrix}]$	$[\sqrt{2}, \begin{pmatrix} 2 & \sqrt{10} \\ \sqrt{2} & 4 \\ + & + \end{pmatrix}]$
$[2, \begin{pmatrix} \sqrt{2} & \sqrt{10} \\ \sqrt{10} & \sqrt{2} \\ - & - \end{pmatrix}]$	$[2, \begin{pmatrix} \sqrt{2} & \sqrt{10} \\ \sqrt{2} & \sqrt{10} \\ - & + \end{pmatrix}]$
$[\sqrt{10}, \begin{pmatrix} \sqrt{2} & 4 \\ 2 & \sqrt{2} \\ + & - \end{pmatrix}]$	$[\sqrt{10}, \begin{pmatrix} \sqrt{2} & 2 \\ 4 & \sqrt{10} \\ - & - \end{pmatrix}]$
$[\sqrt{10}, \begin{pmatrix} \sqrt{2} & 4 \\ 2 & \sqrt{2} \\ - & + \end{pmatrix}]$	$[\sqrt{10}, \begin{pmatrix} \sqrt{2} & 2 \\ 4 & \sqrt{10} \\ + & + \end{pmatrix}]$
$[4, \begin{pmatrix} \sqrt{2} & \sqrt{10} \\ \sqrt{10} & \sqrt{2} \\ + & + \end{pmatrix}]$	$[4, \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ \sqrt{10} & \sqrt{10} \\ + & - \end{pmatrix}]$

**Table 6.1** The Oriented Simplexwise Distributions OSDs from Definition 6.1.3 for the 4-point clouds  $T, K \subset$  in Fig. 6.1. Forgetting all signs in the bottom rows of ORDs gives SDDs in Table 5.1.

are related by rigid motion (isometry, respectively) if and only if  $\text{OSD}(C) = \text{OSD}(C')$  ( $\text{OSD}(C) = \text{OSD}(C')$  or its mirror image  $\overline{\text{OSD}}(C')$ , respectively).  $\blacksquare$

## 6.2 Simplexwise Centered Distributions of a cloud in $\mathbb{R}^n$

This section simplifies the OSD invariant to the Simplexwise Centred Distribution (SCD) in Definition 6.2.1. The Euclidean structure of  $\mathbb{R}^n$  allows us to translate the *centre of mass*  $\frac{1}{m} \sum_{p \in C} p$  of a given  $m$ -point cloud  $C \subset \mathbb{R}^n$  to the origin  $0 \in \mathbb{R}^n$ . Then Problem 6.1.1 reduces to only rotations around 0 from the orthogonal group  $O(\mathbb{R}^n)$ .

Definition 6.1.3 introduced the Oriented Simplexwise Distribution (OSD) as an ordered collection of  $ORD(C; A)$  for all  $\binom{m}{n}$  unordered subsets  $A \subset C$  of  $n$  points. Including the centre of mass allows us to consider the smaller number of  $\binom{m}{n-1}$  subsets  $A \subset C$  of  $n-1$  points instead of  $n$ .

Though the centre of mass is uniquely determined for any cloud  $C \subset \mathbb{R}^n$  of unordered points, real applications may offer one or several labelled points of  $C$  that substantially speed up metrics on invariants. For example, an atomic neighbourhood in a solid material is a cloud  $C \subset \mathbb{R}^3$  of atoms around a central atom, which may not be the centre of mass of  $C$ , but can be an extra base point in all constructions below.

For any base sequence  $A$  of  $n-1$  ordered points  $p_1, \dots, p_{n-1} \in C$ , add the origin  $0$  as the  $n$ -th point and consider the  $n \times n$  distance matrix  $D(A \cup \{0\})$  and the  $(n+1) \times (m-n)$  matrix  $M(C; A \cup \{0\})$  in Definition 6.1.2. Any  $n$  vectors  $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$  can be written as columns in the  $n \times n$  matrix whose determinant has a *sign*  $\pm 1$  or  $0$  (if the vectors  $\mathbf{v}_1, \dots, \mathbf{v}_n$  are linearly dependent). Any permutation  $\xi \in S_{n-1}$  of  $n-1$  points of  $A$  acts on  $D(A)$  by permuting the first  $n-1$  rows of  $M(C; A \cup \{0\})$  and by multiplying every sign in the  $(n+1)$ -st row by  $\text{sign}(\xi)$ .

**Definition 6.2.1** (Simplexwise Centred Distribution SCD). Let  $C \subset \mathbb{R}^n$  be any cloud of  $m$  unlabelled points. For any base sequence  $A$  of ordered  $p_1, \dots, p_{n-1}$  in a cloud  $C \subset \mathbb{R}^n$  with the center of mass at  $0 \in \mathbb{R}^n$ , the *Oriented Centred Distribution*  $OCD(C; A)$  is the equivalence class of pairs  $[D(A \cup \{0\}), M(C; A \cup \{0\})]$  considered up to permutations  $\xi \in S_{n-1}$  of points of  $A$ . The *Simplexwise Centred Distribution*  $SCD(C)$  is the unordered set of distributions  $OCD(C; A)$  for all  $\binom{m}{n-1}$  unordered  $(n-1)$ -point subsets  $A \subset C$ . The mirror image  $\overline{SCD}(C)$  is obtained from  $SCD(C)$  by reversing all signs.  $\blacktriangle$

Definition 6.2.1 needs no permutations for any  $C \subset \mathbb{R}^2$  as  $n-1=1$ . Columns of  $M(C; A \cup \{0\})$  can be lexicographically ordered without affecting future metrics.

Some of the  $\binom{m}{n-1}$  OCDs in  $SCD(C)$  can be identical as in Example 6.2.2(b). If we collapse any  $l > 1$  identical OCDs into a single OCD with the *weight*  $l/\binom{m}{h}$ , SCD can be considered as a weighted probability distribution of OCDs.

**Example 6.2.2** (Simplexwise Centered Distribution SCDs for clouds in Fig. 6.1). **(a)** Let  $R \subset \mathbb{R}^2$  be the right-angled cloud of the points  $p_1 = (0, 0)$ ,  $p_2 = (4, 0)$ ,  $p_3 = (0, 3)$  in Fig. 6.1 (1st). Though  $p_1 = (0, 0)$  is included in  $R$  and is not its centre of mass,  $SCD(R)$

still makes sense. In  $\text{OCD}(R; p_1) = \begin{bmatrix} 0, \begin{pmatrix} 4 & 3 \\ 4 & 3 \\ 0 & 0 \end{pmatrix} \end{bmatrix}$ , the matrix  $D(\{p_1, 0\})$  is  $|p_1 - 0| = 0$ ,

the top row has  $|p_2 - p_1| = 4$ ,  $|p_3 - p_1| = 3$ . In  $\text{OCD}(R; p_2) = \begin{bmatrix} 4, \begin{pmatrix} 4 & 5 \\ 0 & 3 \\ 0 & - \end{pmatrix} \end{bmatrix}$ , the first row has  $|p_1 - p_2| = 4$ ,  $|p_3 - p_2| = 5$ , the second row has  $|p_1 - 0| = 0$ ,  $|p_3 - 0| = 3$ ,

$\det \begin{pmatrix} -4 & 0 \\ 3 & 3 \end{pmatrix} < 0$ . In  $\text{OCD}(R; p_3) = \begin{bmatrix} 3, \begin{pmatrix} 3 & 5 \\ 0 & 4 \\ 0 & + \end{pmatrix} \end{bmatrix}$ , the first row has  $|p_1 - p_3| = 3$ ,

$|p_2 - p_3| = 5$ , the second row has  $|p_1 - 0| = 0$ ,  $|p_2 - 0| = 4$ ,  $\det \begin{pmatrix} 4 & 4 \\ -3 & 0 \end{pmatrix} > 0$ . So

$\text{SCD}(R)$  consists of the three OCDs above.

If we reflect  $R$  with respect to the  $x$ -axis, the new cloud  $\bar{R}$  of the points  $p_1, p_2, \bar{p}_3 = (0, -3)$  has  $\text{SCD}(\bar{R}) = \overline{\text{SCD}}(R)$  with

$$\text{OCD}(\bar{R}; p_1) = \text{OCD}(R), \text{OCD}(\bar{R}; p_2) = \begin{bmatrix} 4, \begin{pmatrix} 4 & 5 \\ 0 & 3 \\ 0 & + \end{pmatrix} \end{bmatrix}, \text{OCD}(\bar{R}; \bar{p}_3) = \begin{bmatrix} 3, \begin{pmatrix} 3 & 5 \\ 0 & 4 \\ 0 & - \end{pmatrix} \end{bmatrix}$$

whose signs changed under reflection, so  $\text{SCD}(R) \neq \text{SCD}(\bar{R})$ .

**(b)** Let  $S \subset \mathbb{R}^2$  consist of  $m = 4$  points  $(\pm 1, 0), (0, \pm 1)$  that are vertices of the square in Fig. 6.1 (4th). The centre of mass is  $0 \in \mathbb{R}^2$  and has a distance 1 to each point of  $S$ .

For each 1-point subset  $A = \{p\} \subset S$ , the distance matrix  $D(A \cup \{0\})$  on two points is the single number 1. The matrix  $M(S; A \cup \{0\})$  has  $m - n + 1 = 3$  columns.

For  $p_1 = (1, 0)$ , we have  $M\left(S; \begin{pmatrix} p_1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} \sqrt{2} & \sqrt{2} & 2 \\ 1 & 1 & 1 \\ - & + & 0 \end{pmatrix}$ , where the columns are ordered

according to  $p_2 = (0, -1)$ ,  $p_3 = (0, 1)$ ,  $p_4 = (-1, 0)$  in Fig. 6.1 (4th). The sign in the bottom right corner is 0 because the points  $p_1, 0, p_4$  are in a straight line. Due to the rotational symmetry,  $M(S; \{p_i, 0\})$  is independent of  $i = 1, 2, 3, 4$ . So  $\text{SCD}(S)$  can be

considered as one OCD =  $\left[ 1, M\left(S; \begin{pmatrix} p_1 \\ 0 \end{pmatrix}\right) \right]$  of weight 1.  $\blacksquare$

**Theorem 6.2.3** (completeness of SCD, [2, Theorem 3.10]). The Simplexwise Centred Distribution  $\text{SCD}(C)$  in Definition 6.2.1 is a complete isometry invariant of clouds  $C \subset \mathbb{R}^n$  of  $m$  unlabelled points with a centre of mass at the origin  $0 \in \mathbb{R}^n$ , and can be computed in time  $O(m^n/(n-4)!)$ . So any clouds  $C, C' \subset \mathbb{R}^n$  are related by rigid motion (isometry, respectively) if and only if  $\text{SCD}(C) = \text{SCD}(C')$  ( $\text{SCD}(C)$  equals  $\text{SCD}(C')$  or its mirror image  $\overline{\text{SCD}}(C')$ , respectively). For any  $m$ -point clouds  $C, C' \subset \mathbb{R}^n$ , let  $\text{SCD}(C)$  and  $\text{SCD}(C')$  consist of  $k = \binom{m}{n-1}$  OCDs.  $\blacksquare$

Corollary 6.2.4 follows from Lemma 6.1.6 by adding the centre of mass of  $C$  as an extra point to a base sequence  $A \subset C$ .

**Corollary 6.2.4** (reconstruction from OCD). A cloud  $C \subset \mathbb{R}^n$  of  $m > n$  unordered points with the centre of mass  $O(A)$  can be reconstructed, uniquely under rigid motion, from  $\text{OCD}(C; A)$  in Definition 6.1.3 for any base sequence  $A \subseteq C$  with  $\text{aff}(\{O(A)\} \cup A) = n - 1$ .  $\blacksquare$

Example 6.2.2(b) illustrates the key discontinuity challenge: if  $p_4 = (-1, 0)$  is perturbed, the corresponding sign can discontinuously change to +1 or -1.

To get a continuous metric on OCDs, we will multiply each sign by a continuous *strength* function, which vanishes for any zero sign, as defined in the next section.

### 6.3 The Lipschitz continuous strength of a simplex in $\mathbb{R}^n$

This section resolves the discontinuity of signs of determinants by introducing the multiplicative factor below.

**Definition 6.3.1** (*strength*  $\sigma(A)$  of a simplex). For a set  $A$  of  $n + 1$  points  $q = p_0, p_1, \dots, p_n$  in  $\mathbb{R}^n$ , let  $p(A) = \frac{1}{2} \sum_{i \neq j}^{n+1} |p_i - p_j|$  be half of the sum of all pairwise distances. Let  $V(A)$  denote the volume the  $n$ -dimensional simplex on the set  $A$ . Define the *strength*  $\sigma(A) = V^2(A)/p^{2n-1}(A)$ .  $\blacktriangle$

**Example 6.3.2** (strengths in dimensions 1, 2). **(a)** For  $n = 1$  and a set  $A = p_0, p_1 \subset \mathbb{R}$ , the volume is  $V(A) = |p_0 - p_1| = 2p(A)$ , so  $\sigma(A) = 2|p_0 - p_1|$  is the double length.

**(b)** For  $n = 2$  and a triangle  $A \subset \mathbb{R}^2$  with sides  $a, b, c$ , Heron's formula gives  $\sigma(A) = \frac{(p-a)(p-b)(p-c)}{p^2}$ ,  $p = \frac{a+b+c}{2} = p(A)$  is the half-perimeter of  $A$ .  $\blacklozenge$

The strength  $\sigma(A)$  depends only on the distance matrix  $D(A)$  from Definition 6.1.2, so the notation  $\sigma(A)$  is used only for brevity. In any  $\mathbb{R}^n$ , the squared volume  $V^2(A)$  is expressed by the Cayley-Menger determinant [3] in pairwise distances between points of  $A$ . The strength  $\sigma(A)$  vanishes when the simplex on a set  $A$  degenerates.

Theorem 6.4.5 will need the continuity of  $s\sigma(A)$ , when a sign  $s \in \{\pm 1\}$  from a bottom row of ORD changes while passing through a degenerate set  $A$ . In appendices, the proof of the continuity of  $\sigma(A)$  in Theorem 6.3.3 gives an explicit upper bound for a Lipschitz constant  $\lambda_n$  below.

**Theorem 6.3.3** (Lipschitz continuity of the strength  $\sigma$ , [2, Theorem 4.2]). Let a cloud  $A'$  be obtained from another  $(n+1)$ -point cloud  $A \subset \mathbb{R}^n$  by perturbing every point within its  $\varepsilon$ -neighbourhood. The strength  $\sigma(A)$  from Definition 6.3.1 is Lipschitz continuous so that  $|\sigma(A') - \sigma(A)| \leq 2\varepsilon\lambda_n$  for a Lipschitz constant  $\lambda_n$ .  $\blacksquare$

**Example 6.3.4** (approximates constants  $\lambda_n$  of strength). For  $n \geq 2$ , the proof of [2, Theorem 4.2] implies the following approximate values for upper bounds of the Lipschitz constant of strength:  $\lambda_2 = 2\sqrt{3}$ ,  $\lambda_3 \approx 0.43$ ,  $\lambda_4 \approx 0.01$ , which quickly tend to 0 due to the ‘curse of dimensionality’. The plots in [4, Fig. 4] illustrate that the strength  $\sigma(A)$  behaves smoothly under perturbations and the derivative  $|\frac{\partial\sigma}{\partial x}|$  is much smaller than the proved bounds of  $\lambda_n$  above.  $\blacklozenge$

## 6.4 Algorithms for continuous metrics on complete invariants

This section introduces Lipschitz continuous metrics on the invariants OSD and SCD by using the strength of a simplex. By Definition 6.1.3 an Oriented Relative Distribution ORD is a pair  $[D(A); M(C; A)]$  of matrices considered up to permutations  $\xi \in S_n$  of  $n$  ordered points of  $A$ . Any column of  $M(C; A)$  is a pair  $(v, s)$ , where  $s \in \{\pm 1, 0\}$  and  $v \in \mathbb{R}^n$  is a vector of distances from  $q \in C - A$  to  $p_1, \dots, p_n \in A$ .

For simplicity and similar to the case of a general metric space, we assume that a cloud  $C \subset \mathbb{R}^n$  is given by a matrix of pairwise Euclidean distances. If  $C$  is given by Euclidean coordinates of points, then any distance requires  $O(n)$  computations, and we should add the factor  $n$  in all complexities below, keeping all times polynomial in  $m$ . The  $m - n$  permutable columns of the matrix  $M(C; A)$  in ORD from Definition 6.1.3 can be interpreted as  $m - n$  unordered points in  $\mathbb{R}^n$ . Since any isometry is bijective, the simplest metric respecting bijections is the bottleneck distance BD from Example 1.3.1(b).

**Definition 6.4.1** (max metric on ORDs). Consider the bottleneck distance BD on the set of  $m - n$  permutable columns of  $M(C; A)$  as on a cloud of  $m - n$  unordered points  $(v, \frac{s}{\lambda_n}\sigma(A \cup \{q\})) \in \mathbb{R}^{n+1}$ . For another ORD' =  $[D(A'); M(C'; A')]$  and any permutation  $\xi \in S_n$  of indices  $1, \dots, n$  acting on  $D(A)$  and rows of  $M(C; A)$ , set

$$d_o(\xi) = \max\{L_\infty(\xi(D(A)), D(A')), \text{BD}(\xi(M(C; A)), M(C'; A'))\}.$$

Then the *max metric* is defined as  $M_\infty(\text{ORD}, \text{ORD}') = \min_{\xi \in S_n} d_o(\xi)$ .  $\blacktriangle$

The coefficient  $\frac{1}{\lambda_n}$  in front of the strength  $\sigma(A \cup \{q\})$  in Definitions 6.4.1 and 6.4.2 normalises the Lipschitz constant  $\lambda_n$  of  $\sigma$  to 1 in line with changes of distances by at most  $2\epsilon$  when points are perturbed within their  $\epsilon$ -neighbourhoods.

**Definition 6.4.2** (max metric on OCDs). Consider the bottleneck distance  $\text{BD}$  on the set of permutable  $m - n + 1$  columns of  $M(C; A \cup \{0\})$  as on a cloud of  $m - n + 1$  unordered points  $\left(v, \frac{s}{\lambda_n} \sigma(A \cup \{0, q\})\right) \in \mathbb{R}^{n+1}$ . For another OCD' =  $[D(A' \cup \{0\}); M(C'; A' \cup \{0\})]$  and any permutation  $\xi \in S_{n-1}$  of indices  $1, \dots, n - 1$  acting on  $D(A \cup \{0\})$  and the first  $n - 1$  rows of  $M(C; A \cup \{0\})$ , set  $d_o(\xi) = \max\{L, W\}$ ,

$$\text{where } L = L_\infty\left(\xi(D(A \cup \{0\})), D(A' \cup \{0\})\right),$$

$$W = \text{BD}\left(\xi(M(C; A \cup \{0\})), M(C'; A' \cup \{0\})\right).$$

Then the max metric is defined as  $M_\infty(\text{OCD}, \text{OCD}') = \min_{\xi \in S_{n-1}} d_o(\xi)$ . ▲

The max metric  $M_\infty$  is used for intermediate costs to get metrics on unordered collections OSDs and SCDs by using the metrics LAC and EMD from Definitions 5.3.2 and 3.5.4, respectively. Equality  $\text{OSD}(C) = \text{OSD}(C')$  between unordered collections of ORDs is best verified by checking if LAC or EMD between these OSDs is 0.

**Theorem 6.4.3** (times for metrics on OSDs, [2, Theorem 5.6]). **(a)** For the  $k \times k$  matrix of costs computed by the max metric  $M_\infty$  between ORDs from  $\text{OSD}(C)$  and  $\text{OSD}(C')$ , LAC from Definition 5.3.2 satisfies all metric axioms on OSDs and needs time  $O(n!(n^2 + m^{1.5} \log^{n+1} m)k^2 + k^3 \log k)$ .

**(b)** Let OSDs have a maximum size  $l \leq k$  after collapsing identical ORDs. The EMD from Definition 3.5.4 satisfies all metric axioms on OSDs and can be computed in time  $O(n!(n^2 + m^{1.5} \log^{n+1} m)l^2 + l^3 \log l)$ . ■

Equality  $\text{SCD}(C) = \text{SCD}(C')$  is interpreted as a bijection between unordered sets  $\text{SCD}(C) \rightarrow \text{SCD}(C')$  matching all OCDs, which is best verified by checking if the metrics between these SCDs vanish in Theorem 6.4.4.

**Theorem 6.4.4** (times for metrics on SCDs, [2, Theorem 5.7]). **(a)** For the  $k \times k$  matrix of costs computed by the max metric  $M_\infty$  between OCDs in  $\text{SCD}(C)$  and  $\text{SCD}(C')$ , the metric LAC from Definition 5.3.2 satisfies all metric axioms on SCDs and needs time  $O((n - 1)!(n^2 + m^{1.5} \log^n m)k^2 + k^3 \log k)$ .

**(b)** Let SCDs have a maximum size  $l \leq k$  after collapsing identical OCDs. Then EMD from Definition 3.5.4 satisfies all metric axioms on SCDs and can be computed in time  $O((n - 1)!(n^2 + m^{1.5} \log^n m)l^2 + l^3 \log l)$ . ■

If we estimate  $l \leq k = \binom{m}{n-1} = m(m-1)\dots(m-n+2)/n!$  as  $O(m^{n-1}/n!)$ , Theorem 6.4.4 gives time  $O(n(m^{n-1}/n!)^3 \log m)$  for metrics on SCDs, which is  $O(m^3 \log m)$  for  $n = 2$ , and  $O(m^6 \log m)$  for  $n = 3$ . Though the above estimates are very rough, the

time  $O(m^3 \log m)$  in  $\mathbb{R}^2$  is faster than the only past time  $O(m^5 \log m)$  for comparing  $m$ -point clouds by the Hausdorff distance minimized over isometries [1].

**Theorem 6.4.5** (continuity of OSD and SCD, [2, Corollary 6.1]). For any cloud  $C \subset \mathbb{R}^n$  of  $m$  unordered points, perturbing any point within its  $\varepsilon$ -neighbourhood changes  $\text{OSD}(C)$  and  $\text{SCD}(C)$  by at most  $2\varepsilon$  in the metrics LAC and EMD. ■

Theorems 6.2.3, 6.4.4, and 6.4.5 imply that the Simplexwise Centred Distribution fully solves Problem 6.1.1. In addition, Corollary 6.2.4 proves that  $\text{OCD}(C; A)$  from  $\text{SCD}(C)$  suffices to reconstruct  $C \subset \mathbb{R}^n$ , uniquely under rigid motion, if  $C$  has a base sequence  $A$  of  $n - 2$  points with  $\text{aff}(\{O(A)\} \cup A) = n - 1$ .

Forthcoming work will improve SCDs to better invariants that allow a reconstruction in all degenerate cases and satisfy the much harder realisability condition as in 1.4.5(f).

## References

1. Paul Chew, Michael Goodrich, Daniel Huttenlocher, Klara Kedem, Jon Kleinberg, and Dina Kravets. Geometric pattern matching under Euclidean motion. *Comp. Geom.*, 7(1-2):113–124, 1997.
2. Vitaliy Kurlin. The strength of a simplex is the key to a continuous isometry classification of Euclidean clouds of unlabelled points. *arXiv:2303.13486 (latest version at <http://kurlin.org/projects/cloud-isometry-spaces/SCD.pdf>)*, 2023.
3. Manfred Sippel and Harold Scheraga. Cayley-Menger coordinates. *PNAS*, 83:2283–2287, 1986.
4. Daniel E Widdowson and Vitaliy A Kurlin. Recognizing rigid patterns of unlabeled point clouds by complete and continuous isometry invariants with no false negatives and no false positives. In *Computer Vision and Pattern Recognition*, pages 1275–1284, 2023.



**Part II**  
**Geometric Data Science of periodic point sets**



## Chapter 7

# One-periodic sequences in high-dimensional Euclidean spaces

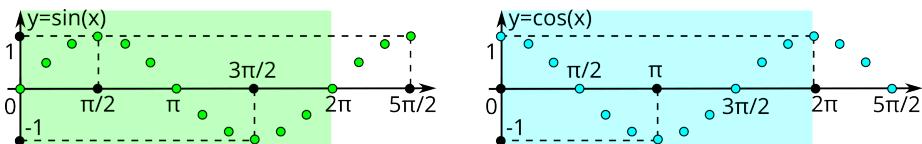
**Abstract** This chapter studies high-dimensional data that is periodic in one direction. These periodic sequences live in a high-dimensional space  $\mathbb{R} \times \mathbb{R}^{n-1}$  for any dimension  $n \geq 1$  and were indistinguishable by past invariants even in dimension  $n = 2$ . Experimental noise and atomic vibrations motivate a new continuous approach, because a minimal periodic pattern breaks down under almost any perturbation.

## 7.1 One-periodic sequences under various isometries in $\mathbb{R} \times \mathbb{R}^{n-1}$

All sections in this chapter follow paper [2] with minor updates.

**Definition 7.1.1** (1-periodic sequences in  $\mathbb{R} \times \mathbb{R}^{n-1}$ ). Let  $\mathbf{e}_1$  be the unit vector along the first axis in  $\mathbb{R} \times \mathbb{R}^{n-1}$  for  $n \geq 1$ . For a period  $l > 0$ , a *motif*  $M$  is a set of points  $p_1, \dots, p_m$  in the *slice*  $[0, l) \times \mathbb{R}^{n-1}$  of the width  $l > 0$ . We assume that the *time projections*  $t(p_1), \dots, t(p_m)$  under  $t : [0, l) \times \mathbb{R}^{n-1} \rightarrow [0, l)$  are distinct, while  $v(p_1), \dots, v(p_m)$  under the *value projection*  $v : [0, l) \times \mathbb{R}^{n-1} \rightarrow \mathbb{R}^{n-1}$  are arbitrary.

A *l-periodic sequence*  $S = M + l\mathbf{e}_1\mathbb{Z}$  is the infinite sequence of points  $p(i + mj) = p_i + j\mathbf{e}_1 \in \mathbb{R}^n$ , which are indexed by  $i + mj$ , where  $j \in \mathbb{Z}$  and  $i = 1, \dots, m$ .  $\blacktriangle$



**Fig. 7.1** The periodic sequences  $C, S \subset \mathbb{R} \times \mathbb{R}$  are sampled from the sine and cosine graphs. The motifs in the shaded slice  $[0, 2\pi) \times \mathbb{R}$  are non-isometric, but  $S, C$  are related by translation.

The slice  $[0, l) \times \mathbb{R}^{n-1}$  excludes all points with  $t = l$ , which are equivalent to points with  $t = 0$  by translation in the time factor  $\mathbb{R}$ . Then all points  $p_1, \dots, p_m \in [0, l) \times \mathbb{R}^{n-1}$  are counted once and ordered under the time projection  $t : [0, l) \times \mathbb{R}^{n-1} \rightarrow [0, l)$ .

**Example 7.1.2** (1-periodic sequences in  $\mathbb{R} \times \mathbb{R}$ ). Fig. 7.1 (left) shows the 1-periodic sequence  $S$  in  $\mathbb{R} \times \mathbb{R}$  (from the sine graph) with the period  $l = 2\pi$  and motif  $M_S$  of

$$(0, 0), \quad \left(\frac{\pi}{6}, \frac{1}{2}\right), \quad \left(\frac{\pi}{3}, \frac{\sqrt{3}}{2}\right), \quad \left(\frac{\pi}{2}, 1\right), \quad \left(\frac{2\pi}{3}, \frac{\sqrt{3}}{2}\right), \quad \left(\frac{5\pi}{6}, \frac{1}{2}\right), \\ (\pi, 0), \quad \left(\frac{7\pi}{6}, -\frac{1}{2}\right), \quad \left(\frac{4\pi}{3}, -\frac{\sqrt{3}}{2}\right), \quad \left(\frac{3\pi}{2}, -1\right), \quad \left(\frac{5\pi}{3}, -\frac{\sqrt{3}}{2}\right), \quad \left(\frac{11\pi}{6}, -\frac{1}{2}\right).$$

Fig. 7.1 (right) shows another sequence  $C$  with the same period  $l = 2\pi$  and a different motif  $M_C \neq M_S$ . However,  $S, C$  are identical under translation:  $\sin(x + \frac{\pi}{2}) = \cos(x)$ . ♦

Example 7.1.2 illustrates the ambiguity of digital representations when many real objects look different in various coordinate systems despite being equivalent as rigid objects. We adapt basic equivalences to sets in the product  $\mathbb{R} \times \mathbb{R}^{n-1}$ .

**Definition 7.1.3** (*cyclic* vs *dihedral* isometries and rigid motions). A *cyclic* isometry of  $\mathbb{R} \times \mathbb{R}^{n-1}$  is a composition of a translation in the time factor  $\mathbb{R}$  and an isometry in the value factor  $\mathbb{R}^{n-1}$ . If we allow compositions of a translation and symmetry  $x \mapsto -x$  in the time factor  $\mathbb{R}$ , the resulting isometry of  $\mathbb{R} \times \mathbb{R}^{n-1}$  is *dihedral*.

If we allow only isometries that preserve orientation in the value factor  $\mathbb{R}^{n-1}$ , the resulting equivalences are called *cyclic* and *dihedral rigid motions*, respectively. ▲

The adjectives *cyclic* and *dihedral* are motivated by the names of the cyclic group  $C_m$  and the dihedral group  $D_m$  consisting of orientation-preserving isometries and all isometries in  $\mathbb{R}^2$ , respectively, that map the regular polygon on  $m$  vertices to itself.

The equivalences in Definition 7.1.3 make sense for any finite sequence of points  $T \subset \mathbb{R} \times \mathbb{R}^{n-1}$ . However, the periodicity substantially worsens the ambiguity of representations based a period  $l$  and a motif  $M$  as follows. A translation in the time factor  $\mathbb{R}$  allows us to fix any point  $p$  of a motif  $M$  at  $t = 0$ , but this choice of  $p$  is arbitrary, so a motif  $M$  is defined only modulo cyclic permutations of its points.

The set of integers can be defined as  $\mathbb{Z}$  with period 1 or as  $\{0, 1\} + 2\mathbb{Z}$  with period 2, or with any integer period  $l > 0$ . For any sequence  $S = \{p_1, \dots, p_m\} + l\mathbf{e}_1\mathbb{Z}$ , we can choose a *minimal* period  $l$  such that  $S$  can not be represented with a smaller period.

This classical approach in crystallography leads to an *invariant*  $I$  based on a minimum period (primitive cell) and defined as a set of numerical properties preserved under any rigid motion. Fixing a minimum period creates the following discontinuity.

**Example 7.1.4** (discontinuity of a period). For any small  $\varepsilon > 0$  and integer  $m \geq 1$ , any point of  $\mathbb{Z}$  is  $\varepsilon$ -close to a unique point of the sequence  $\{0, 1 + \varepsilon, \dots, m + \varepsilon\} + (m + 1)\mathbb{Z}$ , though their minimum periods 1 and  $m + 1$  are arbitrarily different. Hence comparing periodic sequences by their minimal motifs can miss infinitely many near-duplicates. ♦

We assume that the input for a 1-periodic sequence  $S$  consists of a period  $l > 0$  and a motif of  $m = |S|$  points in the high-dimensional slice  $[0, l) \times \mathbb{R}^{n-1}$ .

**Problem 7.1.5** (invariants of 1-periodic sequences in  $\mathbb{R} \times \mathbb{R}^{n-1}$ ). Design an invariant  $I$  of all 1-periodic sequences of points in  $\mathbb{R} \times \mathbb{R}^{n-1}$  satisfying the following conditions.

(a) *Completeness*: any 1-periodic sequences  $S, Q \subset \mathbb{R} \times \mathbb{R}^{n-1}$  are related by cyclic isometry (denoted as  $S \cong Q$ ) in Definition 7.1.3 if and only if  $I(S) = I(Q)$ .

(b) *Reconstruction*: any 1-periodic sequence  $S \subset \mathbb{R} \times \mathbb{R}^{n-1}$  is reconstructable from its invariant value  $I(S)$ , uniquely under cyclic isometry.

(c) *Metric*: there is a distance  $d$  on the space  $\{I(S) \mid \text{all 1-periodic sequences } S \subset \mathbb{R} \times \mathbb{R}^{n-1}\}$  satisfying all metric axioms in Definition 1.3.1(a).

(d) *Continuity*: there is a constant  $\lambda > 0$ , such that, for all sufficiently small  $\varepsilon$ , if a 1-periodic sequence  $Q$  is obtained by perturbing every point of a 1-periodic sequence  $S \subset \mathbb{R} \times \mathbb{R}^n$  up to Euclidean distance  $\varepsilon$ , then  $d(I(S), I(Q)) \leq \lambda \varepsilon$ .

(e) *Computability*: the invariant  $I$ , a reconstruction of  $S \subset \mathbb{R} \times \mathbb{R}^{n-1}$  from  $I(S)$ , and the metric  $d(I(S), I(Q))$  can be computed in times that depend polynomially on the dimension  $n$  and the maximum motif size of 1-periodic sequences  $S, Q$ .  $\star$

Further sections will develop invariants that solve Problem 7.1.5 for all 1-periodic sequences under cyclic and dihedral isometries and rigid motions in  $\mathbb{R} \times \mathbb{R}^{n-1}$ .

Our constructions were inspired by the infinite family of 1-periodic sequences in [3, Fig. 4] that were not distinguished by past invariants, see a review in [2, section 2].

## 7.2 Invariants and continuous metrics for finite sequences in $\mathbb{R}^n$

This section studies complete invariants and metrics for isometry classes of finite sequences of ordered points in  $\mathbb{R}^n$ . These invariants are easily extendable to the 1-periodic sequences in  $\mathbb{R} \times \mathbb{R}^{n-1}$  and substantially differ from direction-based invariants in chapter 3 and the backbone invariants in chapter 2, which were defined only for non-degenerate sequences of triplets of points in  $\mathbb{R}^3$ .

**Definition 7.2.1** (distance matrices DM and CDM). Let  $T$  be an ordered sequence of  $m$  points  $p_1, \dots, p_m \in \mathbb{R}^n$ .

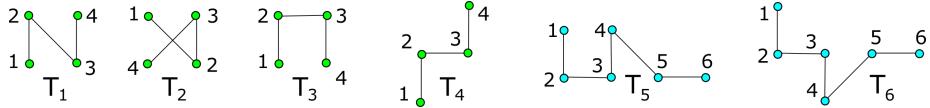
(a) In the *distance matrix*  $\text{DM}(T)$  of the size  $m \times m$ , each element  $\text{DM}_{ij}(T)$  is the Euclidean distance  $|p_j - p_j|$  for  $i, j \in \{1, \dots, m\}$ , so  $d_{ii} = 0$  for  $i = 1, \dots, m$ .

(b) To define the *cyclic distance matrix*  $\text{CDM}(T)$  of the size  $(m-1) \times m$ , set element  $\text{CDM}_{ij}(T)$  to 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,  $p_{m+1} = p_1$ .  $\blacktriangle$

Any  $m = 3$  points in  $\mathbb{R}^n$  with pairwise distances  $d_{ij}$  have the distance matrix

$$\text{DM} = \begin{pmatrix} 0 & d_{12} & d_{13} \\ d_{12} & 0 & d_{23} \\ d_{13} & d_{23} & 0 \end{pmatrix} \text{ and the cyclic distance matrix } \text{CDM} = \begin{pmatrix} d_{12} & d_{23} & d_{13} \\ d_{13} & d_{12} & d_{23} \end{pmatrix}. \text{ CDM}(T)$$

is obtained from  $\text{DM}(T)$  by removing the zero diagonal and cyclically shifting each column so that the first row of  $\text{CDM}(T)$  has distances from  $p_i$  to the next point  $p_{i+1}$ .



**Fig. 7.2** These sequences are distinguished by their cyclic distance matrices in Example 7.2.2.

**Example 7.2.2** (cyclic distance matrices). Fig. 7.2 shows the sequences  $T_1, \dots, T_6 \subset \mathbb{R}^2$  whose points are in the integer lattice  $\mathbb{Z}^2$  so that the minimum inter-point distance is 1. In each sequence, the points are connected by straight lines in the order  $1 \rightarrow 2 \rightarrow \dots \rightarrow m$ .

$$\text{CDM}(T_1) = \begin{pmatrix} 1 & \sqrt{2} & 1 & \sqrt{2} \\ 1 & 1 & 1 & 1 \\ \sqrt{2} & 1 & \sqrt{2} & 1 \end{pmatrix}, \text{CDM}(T_2) = \begin{pmatrix} \sqrt{2} & 1 & \sqrt{2} & 1 \\ 1 & 1 & 1 & 1 \\ 1 & \sqrt{2} & 1 & \sqrt{2} \end{pmatrix} \text{ are different but related}$$

by a cyclic shift of columns. This shift of indices in  $T_1$  gives a sequence isometric to

$$T_2. \text{ Then } \text{CDM}(T_3) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ \sqrt{2} & \sqrt{2} & \sqrt{2} & \sqrt{2} \\ 1 & 1 & 1 & 1 \end{pmatrix}, \text{CDM}(T_4) = \begin{pmatrix} 1 & 1 & 1 & \sqrt{5} \\ \sqrt{2} & \sqrt{2} & \sqrt{2} & \sqrt{2} \\ \sqrt{5} & 1 & 1 & 1 \end{pmatrix}. \text{ The CDMs of}$$

the sets  $T_5, T_6$  differ only by distances  $|p_1 - p_4| = 1$  in  $T_5$  and  $|p_1 - p_4| = \sqrt{5}$  in the highlighted cells below. If reduce the number  $m - 1$  of rows in CDM to the dimension  $n = 2$ , the smaller matrices fail to distinguish the non-isometric sequences  $T_5 \not\cong T_6$ .

$$T_5 : \begin{pmatrix} 1 & 1 & 1 & \sqrt{2} & 1 & \sqrt{10} \\ \sqrt{2} & \sqrt{2} & 1 & \sqrt{5} & \sqrt{5} & 3 \\ 1 & 2 & 2 & 1 & 2 & 2 \\ \sqrt{5} & 3 & \sqrt{2} & \sqrt{2} & 1 & \sqrt{5} \\ \sqrt{10} & 1 & 1 & 1 & \sqrt{2} & 1 \end{pmatrix} \text{ and } T_6 : \begin{pmatrix} 1 & 1 & 1 & \sqrt{2} & 1 & \sqrt{10} \\ \sqrt{2} & \sqrt{2} & 1 & \sqrt{5} & \sqrt{5} & 3 \\ \sqrt{5} & 2 & 2 & \sqrt{5} & 2 & 2 \\ \sqrt{5} & 3 & \sqrt{2} & \sqrt{2} & 1 & \sqrt{5} \\ \sqrt{10} & 1 & 1 & 1 & \sqrt{2} & 1 \end{pmatrix}. \quad \blacklozenge$$

Recall that Definition 6.3.1 introduced the strength  $\sigma(A) = \frac{V^2(A)}{p^{2n-1}(A)}$  of a simplex  $A$  on any set of  $n + 1$  points  $q_0, q_1, \dots, q_n \in \mathbb{R}^n$ , where  $V(A)$  is the volume of  $A$  and  $p(A) = \frac{1}{2} \sum_{0 \leq i < j \leq n} |q_i - q_j|$  is the half-perimeter.

**Definition 7.2.3** (*cyclic distances with signs* CDS). For any sequence  $T$  of  $p_1, \dots, p_m \in \mathbb{R}^n$  and  $i = 1, \dots, m$ , let  $\sigma_i(T)$  be the strength of the simplex on the points  $p_i, \dots, p_{i+n}$ , where all indices are modulo  $m$ . Let  $\text{sign}_i(T)$  be the sign ( $\pm 1$  or 0) of the  $n \times n$  determinant with the columns  $\mathbf{p}_{i+1} - \mathbf{p}_i, \mathbf{p}_{i+2} - \mathbf{p}_{i+1}, \dots, \mathbf{p}_{i+n} - \mathbf{p}_{i+n-1}$ . The matrix  $\text{CDS}(T)$  of *cyclic distances with signs* is obtained from  $\text{CDM}(T)$  in Definition 7.2.1 by attaching the extra  $m$ -th row  $\text{sign}(T) = (\text{sign}_1(T), \dots, \text{sign}_m(T))$ .  $\blacktriangle$

**Example 7.2.4** (strengths and signs). For the first sequence  $T_1$  in Fig. 7.2 with the points  $p_1 = (0, 0), p_2 = (0, 1), p_3 = (1, 0), p_4 = (1, 1)$ , the first  $2 \times 2$  determinant

with the columns  $p_2 - p_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  and  $p_3 - p_2 = (1, -1)$  is  $\det \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$  has sign

$-1$ . The further determinants for  $i = 2, 3, 4$  are  $\det \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} = +1, \det \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix} = -1$ ,

$\det \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} = +1$ , so the row of signs is  $\text{sign}(T_1) = (-1, +1, -1, +1)$ . All triangles on 4

triples  $p_i, p_{i+1}, p_{i+2}$  for  $i = 1, 2, 3, 4$  (with indices considered modulo 4) have the sides  $1, 1, \sqrt{2}$ , half-perimeter  $p = 1 + \frac{1}{\sqrt{2}}$ , area  $V = \frac{1}{2}$ , and strength  $\sigma = \frac{1}{\sqrt{2}(1 + \sqrt{2})^3}$ .  $\blacklozenge$

Since the sign of a determinant discontinuously changes when a point set passes through a degenerate configuration, this sign will be multiplied by the Lipschitz continuous strength to get a metric satisfying condition 7.1.5(d), see Theorem 7.2.8(d).

Section 7.4 will adapt the matrices from Definitions 7.2.1 and 7.2.3 to 1-periodic sequences whose motifs of points should be considered under cyclic permutations. The *cyclic* group  $C_m$  consists of  $m$  permutations on  $1, \dots, m$  generated by the *shift permutation*  $\gamma_m : (1, 2, \dots, m) \mapsto (2, \dots, m, 1)$ . The *dihedral* group  $D_m$  consists of  $2m$  permutations generated by  $\gamma_m$  and the *reverse permutation*  $\iota_m : (1, 2, \dots, m) \mapsto (m, \dots, 2, 1)$ .

**Lemma 7.2.5** (actions on vectors and matrices). The shift permutation  $\gamma_m \in C_m$  acts on the cyclic distance matrix  $\text{CDM}(T)$  by cyclically shifting its  $m$  columns and keeping all rows. The reverse permutation  $\iota_m \in D_m$  reverses the order of columns and rows in  $\text{CDM}(T)$ . These permutations act on the row of signs in Definition 7.2.3 as  $\gamma_m(s_1, s_2, \dots, s_m) = (s_2, \dots, s_m, s_1)$  and  $\iota_m(s_1, s_2, \dots, s_m) = (-1)^{[3n/2]}(s_m, \dots, s_2, s_1)$ . For any mirror image  $\bar{T}$  of  $T$ , the matrix  $\text{CDS}(\bar{T})$  is obtained from  $\text{CDS}(T)$  by reversing all signs in the last row. Any element of the groups  $C_m, D_m$  acts on any sequence of  $m$  numbers as a composition of  $\gamma_m, \iota_m$ .  $\blacksquare$

Any matrix  $k \times m$  can be rewritten row-by-row as a vector  $v \in \mathbb{R}^{km}$ . For any  $q \in [1, +\infty]$ , the Minkowski norm is  $\|v\|_q = \left( \sum_{i=1}^{km} |v_i|^q \right)^{1/q}$ , where the limit case is  $\|v\|_\infty = \max_{i=1, \dots, km} |v_i|$ . Any power  $a^{1/q}$  for  $a > 0$  is interpreted as 1 in the case  $q = +\infty$ .

**Definition 7.2.6** (metrics  $\text{MCD}_q$  and  $\text{MCS}_q$  for finite sequences). For any Minkowski norm with a parameter  $q \in [1, +\infty]$  and ordered sequences  $T, S \subset \mathbb{R}^{n-1}$  of  $m$  points, use the matrices from Definition 7.2.1 to define the *metric based on cyclic distances*  $\text{MCD}_q(S, T) = \frac{\|\text{CDM}(S) - \text{CDM}(T)\|_q}{(m(m-1))^{1/q}}$  and the *metric based on cyclic distances with signs*  $\text{MCS}_q(S, T) = \max \left\{ \text{MCD}_q(S, T), \frac{2}{\lambda_n} \max_{i=1, \dots, m} |\text{sign}_i(S)\sigma_i(S) - \text{sign}_i(T)\sigma_i(T)| \right\}$ .  $\blacktriangle$

**Example 7.2.7** (metric  $\text{MCD}_q$ ). For any  $q \in [1, +\infty)$ , we use cyclic distance matrices from Example 7.2.2 to compute

$$\text{MCD}_q(T_1, T_3) = \left(\frac{2}{3}\right)^{1/q}(\sqrt{2}-1),$$

$$\text{MCD}_q(T_3, T_4) = \left(\frac{1}{6}\right)^{1/q}(\sqrt{5}-1),$$

$$\text{MCD}_q(T_1, T_4) = \left(\frac{1}{2}(\sqrt{2}-1)^q + \frac{1}{6}(\sqrt{5}-\sqrt{2})^q\right)^{1/q}.$$

The triangle inequality holds for  $q \geq 1$  as follows:

$$\begin{aligned} & (\text{MCD}_q(T_1, T_3) + \text{MCD}_q(T_3, T_4))^q = \\ & \left( \left(\frac{2}{3}\right)^{1/q}(\sqrt{2}-1) + \left(\frac{1}{6}\right)^{1/q}(\sqrt{5}-1) \right)^q \geq \\ & \left( \left(\frac{1}{2}\right)^{1/q}(\sqrt{2}-1) + \left(\frac{1}{6}\right)^{1/q}(\sqrt{5}-\sqrt{2}) \right)^q \geq \\ & \frac{1}{2}(\sqrt{2}-1)^q + \frac{1}{6}(\sqrt{5}-\sqrt{2})^q = (\text{MCD}_q(T_1, T_4))^q \end{aligned}$$

due to  $(a+b)^q \geq a^q + b^q$  for  $a, b > 0$  and  $q \geq 1$ . For  $q = +\infty$ , the inequality becomes  $(\sqrt{2}-1) + (\sqrt{5}-1) \geq \sqrt{5}-\sqrt{2}$ . Then  $T_5 \not\cong T_6$  have  $\text{MCD}_q(T_5, T_6) = 2^{1/q}(\sqrt{5}-1)$ .  $\blacklozenge$

We use the extra factors  $(m(m-1))^{1/q}$  and  $\frac{2}{\lambda_n}$  in the definition above, where  $\lambda_n$  is a Lipschitz constant  $\sigma$  from Theorem 6.3.3, to guarantee the Lipschitz constant 2 for the new metrics. Indeed, perturbing any points up to  $\varepsilon$  changes the distance between them up to  $2\varepsilon$ . Instead of maxima in the formula for  $\text{MCS}_q(S, T)$ , one can consider other metric transforms from [1, section 4.1], for example, sums of metrics.

Theorem 7.2.8 solves Problem 2.1.1 and will help solve Problem 7.1.5.

**Theorem 7.2.8** (solution of Problem 2.1.1 for ordered points in  $\mathbb{R}^n$ , [2, Theorem 3.9]).  
**(a)** For any sequence  $T \subset \mathbb{R}^n$  of  $m$  ordered points, the matrices  $\text{CDM}(T)$  and  $\text{CDS}(T)$

are complete invariants of  $T \subset \mathbb{R}^n$  under isometry and rigid motion in  $\mathbb{R}^n$ , which are computable in times  $O(m^2n)$  and  $O(m^2n + mn^3)$ , respectively.

- (b) Any sequence  $T \subset \mathbb{R}^n$  of  $m$  points can be reconstructed from the invariants  $\text{CDM}(T)$  and  $\text{CDS}(T)$ , uniquely under isometry and rigid motion, respectively, in time  $O(m^3)$ .
- (c) For any sequences  $S, T \subset \mathbb{R}^n$  of  $m$  points,  $\text{MCD}_q(S, T)$ ,  $\text{MCS}_q(S, T)$  satisfy all metric axioms and are computable in time  $O(m^2)$  and  $O(m^2n + mn^3)$ , respectively.
- (d) If  $S$  is obtained from any sequence  $T \subset \mathbb{R}^n$  by perturbing every point up to Euclidean distance  $\varepsilon$ , then  $\text{MCD}_q(S, T) \leq 2\varepsilon$  and  $\text{MCS}_q(S, T) \leq 2\varepsilon$  for  $q \in [1, +\infty]$ .  $\blacksquare$

### 7.3 Discontinuity of a minimal period for 1-periodic sequences

The invariants and metrics from section 7.2 will be used for a motif of a 1-periodic sequence  $S$  projected to the value factor  $\mathbb{R}^{n-1}$ . To solve Problem 7.1.5, we first resolve the discontinuity of a period under perturbations by projecting  $S$  to the time factor  $\mathbb{R}$ .

**Definition 7.3.1** (time shift TS). Let  $S \subset \mathbb{R} \times \mathbb{R}^{n-1}$  be a 1-periodic sequence with a period  $l$  and a motif  $M$  of points  $p_1, \dots, p_m$ , which have ordered time projection  $t(p_1) < \dots < t(p_m)$  in  $[0, l)$  under  $t : \mathbb{R} \times \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ , see Definition 7.1.1. Set  $d_i = t(p_{i+1}) - t(p_i)$  for  $i = 1, \dots, m$ ,  $t(p_{m+1}) = t(p_1) + l$ . The *time shift* of the pair (motif, period) of the 1-periodic sequence  $S$  is  $\text{TS}(M; l) = (d_1, \dots, d_m)$ .  $\blacktriangle$

The sequences  $S_2 = \{0, 1\} + 3\mathbb{Z}$  and  $3 - S_2 = \{0, 2\} + 3\mathbb{Z}$  are related by translation but have different time shifts  $\text{TS}(\{0, 1\}; 3) = (1, 2)$  and  $\text{TS}(\{0, 2\}; 3) = (2, 1)$ . To get isometry invariants, these shifts are considered modulo cyclic or dihedral permutations.

**Definition 7.3.2** (cyclic and dihedral invariants under isometries). (a) For any 1-periodic sequence  $S = M + l\mathbf{e}_1\mathbb{Z} \subset \mathbb{R} \times \mathbb{R}^{n-1}$  with a minimum motif  $M$  of  $m$  points, let  $v(M) \subset \mathbb{R}^{n-1}$  be the image of  $M$  under the value projection  $v : \mathbb{R} \times \mathbb{R}^{n-1} \rightarrow \mathbb{R}^{n-1}$ .

(b) The *cyclic* and *dihedral isometry* invariants  $\text{CI}(S), \text{DI}(S)$  are the classes of the pair  $(\text{TS}(M; l), \text{CDM}(v(M)))$  under permutations  $\gamma$  from the groups  $C_m, D_m$ , respectively, acting simultaneously on the time shift  $\text{TS}(M; l)$  and the matrix  $\text{CDM}(v(M))$ .

(c) The *cyclic* and *dihedral rigid* invariants  $\text{CR}(S), \text{DR}(S)$  are the classes of the pair  $(\text{TS}(M; l), \text{CDS}(v(M)))$  under permutations  $\gamma$  from the groups  $C_m, D_m$ , respectively, acting simultaneously on the time shift  $\text{TS}(M; l)$  and the matrix  $\text{CDS}(v(M))$ .  $\blacktriangle$

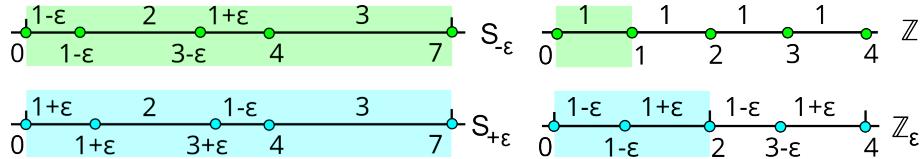
The matrices  $\text{CDM}, \text{CDS}$  are used for the projected motif  $v(M) \subset \mathbb{R}^{n-1}$  and do not depend on a period  $l$ , because a shift along the time direction  $\mathbf{e}_1$  keeps the value projection. For  $n = 1$ , when a periodic sequence  $S = \{p_1, \dots, p_m\} + l\mathbb{Z}$  is in the line  $\mathbb{R}$ , Definition 7.3.2 simplifies to a single time shift obtained by lexicographic ordering.

Recall the *lexicographic order* on vectors:  $(d_1, \dots, d_m) < (d'_1, \dots, d'_m)$  if  $d_1 = d'_1, \dots, d_i = d'_i$  for some  $0 \leq i < m$ , where  $i = 0$  means no identities, and  $d_{i+1} < d'_{i+1}$ .

**Definition 7.3.3** (time invariants CT, DT). Let  $S = \{p_1, \dots, p_m\} + l\mathbb{Z}$  be a 1-periodic sequence in  $\mathbb{R} \times \mathbb{R}^{n-1}$  with a minimum period  $l > 0$ . Set  $d_i = p_{i+1} - p_i$  for  $i = 1, \dots, m$ , where  $p_{m+1} = p_1 + l$ . Apply all permutations of cyclic group  $C_m$  to  $(d_1, \dots, d_m)$ , order all resulting lists lexicographically, and call the first (smallest) list the *cyclic time* invariant  $CT(S)$ . Similarly, define the *dihedral time* invariant  $DT(S)$  as the lexicographically smallest list obtained from  $(d_1, \dots, d_m)$  by the action of  $D_m$ .  $\blacktriangle$

**Example 7.3.4** (time invariants CT, DT). In  $\mathbb{R}$ , the periodic sequences  $S = \{0, 1, 3\} + 6\mathbb{Z}$  and  $Q = 6 - S = \{0, 3, 5\} + 6\mathbb{Z}$  are related by reflection  $x \mapsto 6 - x$  and not by translation. Their time shifts are  $TS(\{0, 1, 3\}; 6) = (1, 2, 3)$  and  $TS(\{0, 3, 5\}; 6) = (3, 2, 1)$ . So the dihedral time invariants of both  $S, Q$  are equal to  $DT = (1, 2, 3)$ , but their cyclic time invariants differ:  $CT(S) = (1, 2, 3) \neq (1, 3, 2) = CT(Q)$ .  $\blacklozenge$

Though the time invariants from Definition 7.3.3 can be proved to be complete for sequences in  $\mathbb{R}$ , Example 7.3.5 and Fig. 7.3 show their discontinuity under noise.



**Fig. 7.3** **Left:** the near-duplicate periodic sequences  $S_{\pm\varepsilon} = \{0, 1 \pm \varepsilon, 3 \pm \varepsilon, 4\} + 7\mathbb{Z}$  have distant time invariants from Definition 7.3.1, see Example 7.3.5. **Right:** the periodic sequence  $\mathbb{Z}$  and its  $\varepsilon$ -perturbation  $\mathbb{Z}_{\varepsilon}$  have incomparable time shifts  $TS(\{0\}; 1) = (1)$  and  $TS(\{0, 1 - \varepsilon\}; 2) = (1 - \varepsilon, 1 + \varepsilon)$  of different lengths sizes. This discontinuity motivates period-independent metrics in Definition 7.4.1.

**Example 7.3.5** (discontinuity of time shifts). In  $\mathbb{R}$ , the periodic sequence  $S_0 = \{0, 1, 3, 4\} + 7\mathbb{Z}$  has two perturbations  $S_{\pm\varepsilon} = \{0, 1 \pm \varepsilon, 3 \pm \varepsilon, 4\} + 7\mathbb{Z}$  for any small  $\varepsilon > 0$ . Rewriting the time shifts  $TS(\{0, 1 - \varepsilon, 3 - \varepsilon, 4\}; 7) = (1 - \varepsilon, 2, 1 + \varepsilon, 3)$  and  $TS(\{0, 1 + \varepsilon, 3 + \varepsilon, 4\}; 7) = (1 + \varepsilon, 2, 1 - \varepsilon, 3)$  in increasing order does not make them close, as the minimal distance  $1 - \varepsilon$  is followed by the different distances  $2 < 3$  in the nearly identical  $S_{\pm\varepsilon}$  for any  $\varepsilon > 0$ , see Fig. 7.3 (left). This discontinuity will be resolved by minimising over cyclic permutations, but there is one more obstacle below.  $\blacklozenge$

It seems natural to always use a minimum period  $l > 0$  of  $S = \{p_1, \dots, p_m\} + l\mathbf{e}_1\mathbb{Z} \subset \mathbb{R} \times \mathbb{R}^{n-1}$ . However, the time shift  $TS = (d_1, \dots, d_m)$  of a fixed size  $m$  cannot be directly used for comparing sequences that have different sizes of motifs, see Fig. 7.3 (right).

## 7.4 Period-independent metrics for 1-periodic sequences

Definition 7.4.1 introduces continuous metrics after extending motifs to a common size.

**Definition 7.4.1** (cyclic and dihedral metrics under isometry and rigid motion). **(a)** For any 1-periodic sequences  $S = M_S + l_S \mathbf{e}_1\mathbb{Z}$  and  $Q = M_Q + l_Q \mathbf{e}_1\mathbb{Z}$  in  $\mathbb{R} \times \mathbb{R}^{n-1}$ , let

$m = \text{lcm}(|M_S|, |M_Q|)$  be the lowest common multiple of their motif sizes (cardinalities).

For the integers  $k_S = \frac{m}{|M_S|}$  and  $k_Q = \frac{m}{|M_Q|}$ , the extended motifs defined as

$$k_S M_S = \bigcup_{i=1, \dots, k_S} (M_S + i l_S \mathbf{e}_1) \quad \text{and} \quad k_Q M_Q = \bigcup_{i=1, \dots, k_Q} (M_Q + i l_Q \mathbf{e}_1)$$

have the same number  $k_S |M_S| = m = k_Q |M_Q|$  of points.

Any permutation  $\gamma$  from  $C_m, D_m$  acts on the projected motif  $v(k_Q M_Q) \subset \mathbb{R}^{n-1}$  as in Lemma 7.2.5. For any parameter  $q \in [1, +\infty]$ , the *cyclic* and *dihedral isometry* metrics are  $\text{CIM}_q(S, Q) = \min_{\gamma \in C_m} \max \{d_t, d_v\}$  and  $\text{DIM}_q(S, Q) = \min_{\gamma \in D_m} \max \{d_t, d_v\}$ , where

$$d_t = m^{-1/q} \| \text{TS}(k_S M_S; k_S l_S) - \text{TS}(\gamma(k_Q M_Q); k_Q l_Q) \|_q,$$

$$d_v = \text{MCD}_q(v(k_S M_S), \gamma(v(k_Q M_Q))).$$

**(b)** The *cyclic* and *dihedral rigid* metrics  $\text{CRM}_q, \text{DRM}_q$  are defined by the same formulae as  $\text{CIM}_q, \text{DIM}_q$  after replacing  $\text{MCD}_q$  with  $\text{MCS}_q$  from Definition 7.2.6.  $\blacktriangle$

In the limit case  $q = +\infty$ , any factor  $a^{\pm 1/q}$  for  $a > 0$  is interpreted as  $\lim_{q \rightarrow +\infty} a^{\pm 1/q} = 1$ . In Definition 7.4.1, the extended periods  $k_S l_S$  and  $k_Q l_Q$  can be different. For simplicity, the metrics  $\text{MCD}_q, \text{MCS}_q$  were written via projected motifs as in Definition 7.2.6 but will be computable via the complete invariants from Definition 7.3.2.

For  $n = 1$ , the projected motifs are empty, so the cases of rigid motion and isometry in  $\mathbb{R}^0$  trivially coincide. In both cases, the metrics are obtained by minimizing only the differences  $d_t$  between time shifts under cyclic and dihedral permutations.

**Example 7.4.2** (invariant metrics). The periodic sequences  $S = \{0, 1\} + 3\mathbb{Z}$  and  $Q = \{0, 1, 3\} + 6\mathbb{Z}$  have motifs  $M_S = \{0, 1\}$  and  $M_Q = \{0, 1, 3\}$  of different sizes  $m_S = 2$  and  $m_Q = 3$  whose lowest common multiple is  $m = 6$ . In the notations of Definition 7.4.1, we get  $k_S = \frac{m}{|M_S|} = 3, k_Q = \frac{m}{|M_Q|} = 2$ . The extended motifs and periods are  $3M_S = \{0, 1, 3, 4, 6, 7\}, 3l_S = 9, 2M_Q = \{0, 1, 3, 6, 7, 9\}, 2l_Q = 12$ . Then  $\text{TS}(3M_S; 9) = (1, 2, 1, 2, 1, 2)$  and  $\text{TS}(2M_Q; 12) = (1, 2, 3, 1, 2, 3)$ . Any cyclic or dihedral permutation of the time shift  $\text{TS}(3M_S; 9)$  relative to  $\text{TS}(2M_Q; 12)$  gives the maximum component-wise distance  $|1 - 3| = 2$ , so  $\text{CIM}_{+\infty}(S, Q) = 2 = \text{DIM}_{+\infty}(S, Q)$ .

**Theorem 7.4.3** (solution to Problem 7.1.5 for 1-periodic sequences, [2, Theorem 4.8]).

**(a)** For any 1-periodic sequence  $S \subset \mathbb{R} \times \mathbb{R}^{n-1}$  with a motif of  $m$  points,  $\text{CI}(S), \text{DI}(S)$  from Definition 7.3.2 are complete invariants under cyclic and dihedral isometry in  $\mathbb{R} \times \mathbb{R}^{n-1}$ , respectively, and are computable in time  $O(m^3 n)$ . Then the invariants  $\text{CR}(S), \text{DR}(S)$  are complete under cyclic and dihedral rigid motion in  $\mathbb{R} \times \mathbb{R}^{n-1}$ , respectively, and are computable in time  $O(m^3 n + m^2 n^3)$ .

**(b)** Any 1-periodic sequence  $S \subset \mathbb{R} \times \mathbb{R}^{n-1}$  with a motif of  $m$  points can be reconstructed from its complete invariant under a relevant equivalence from part (a) in time  $O(m^3 n)$ .

**Table 7.1** Acronyms and references for the new invariants and metrics in sections 7.2, 7.3, 7.4.

$CDM(T)$	Cyclic Distance Matrix of a finite sequence $T \subset \mathbb{R}^n$	Definition 7.2.1
$CDS(T)$	matrix of Cyclic Distances and Signs of a sequence $T \subset \mathbb{R}^n$	Definition 7.2.3
$MCD_q$	Metric on Cyclic Distance matrices (CDM)	Definition 7.2.6
$MCS_q$	Metric on matrices of Cyclic distances and Signs (CDS)	Definition 7.2.6
$TS(M; l)$	Time Shift for a motif $M$ and period $l$ of a sequence	Definition 7.3.1
$CI(S)$	Cyclic Isometry invariant of a sequence $S \subset \mathbb{R} \times \mathbb{R}^{n-1}$	Definition 7.3.2
$DI(S)$	Dihedral Isometry invariant of a sequence $S \subset \mathbb{R} \times \mathbb{R}^{n-1}$	Definition 7.3.2
$CR(S)$	Cyclic Rigid invariant of a sequence $S \subset \mathbb{R} \times \mathbb{R}^{n-1}$	Definition 7.3.2
$DR(S)$	Dihedral Rigid invariant of a sequence $S \subset \mathbb{R} \times \mathbb{R}^{n-1}$	Definition 7.3.2
$CI(S)$	Cyclic Isometry invariant of a sequence $S \subset \mathbb{R} \times \mathbb{R}^{n-1}$	Definition 7.3.2
$DI(S)$	Dihedral Isometry invariant of a sequence $S \subset \mathbb{R} \times \mathbb{R}^{n-1}$	Definition 7.3.2
$CIM_q$	Cyclic Isometry Metric on 1-periodic sequences in $\mathbb{R} \times \mathbb{R}^{n-1}$	Definition 7.4.1
$DIM_q$	Dihedral Isometry Metric on 1-periodic sequences in $\mathbb{R} \times \mathbb{R}^{n-1}$	Definition 7.4.1
$CRM_q$	Cyclic Rigid Metric on 1-periodic sequences in $\mathbb{R} \times \mathbb{R}^{n-1}$	Definition 7.4.1
$DRM_q$	Dihedral Rigid Metric on 1-periodic sequences in $\mathbb{R} \times \mathbb{R}^{n-1}$	Definition 7.4.1

**(c)** The metrics in Definition 7.4.1 remain invariant if any 1-periodic sequence  $S = M + l\mathbf{e}_1\mathbb{Z}$  is alternatively represented by an extended motif  $kM$  and a period  $kl$  for any integer  $k > 0$ . For any 1-periodic sequences  $S, Q \subset \mathbb{R} \times \mathbb{R}^{n-1}$  with a lowest common multiple  $m$  of motif sizes, the distances  $CIM_q, DIM_q, CRM_q, DRM_q$  in Definition 7.4.1 satisfy all metric axioms and are computable in times  $O(m^3n)$  and  $O(m^3n + m^2n^3)$  for the cases of isometry and rigid motion, respectively.

**(d)** Let  $Q$  be a 1-periodic sequence  $S \subset \mathbb{R} \times \mathbb{R}^{n-1}$  after perturbing every point of  $S$  up to some Euclidean distance  $\varepsilon$  that is smaller than a half-distance between any points of  $t(S)$  and of  $t(Q)$ . Then  $CIM_q(S, T), DIM_q(S, Q), CRM_q(S, Q), DRM_q(S, Q) \leq 2\varepsilon$ . ■

**Example 7.4.4** (challenging 1-periodic sequences). The infinite family of counter-examples in [3, Fig. 4] to the completeness of past distance-based invariants includes the pairs of the 1-periodic sequences  $A^\pm \subset \mathbb{R} \times \mathbb{R}^2$  with a period  $l > 0$  and 6-point motifs  $M^+ = \{W', C_+, V, W, C'_+, V'\}$  and  $M^- = \{W', C_-, V, W, C'_-, V'\}$  with the points  $V = (v_x, v_y, 0)$ ,  $W = (\frac{l}{2}, w_y, w_z)$ ,  $C_\pm = (\frac{l}{4}, c_y, \pm c_z)$ , and free parameters  $l, w_y, w_z, c_y, c_z > 0$ , and  $v_x, v_y \in [0, \frac{l}{2}]$ . Any point denoted with a prime is obtained by  $g(x, y, z) = (x + \frac{l}{2}, y, -z)$ . The time projections are identical:  $t(M^\pm) = (0, \frac{l}{4}, v_x, \frac{l}{2}, \frac{3l}{4}, \frac{l}{2} + v_x)$ .

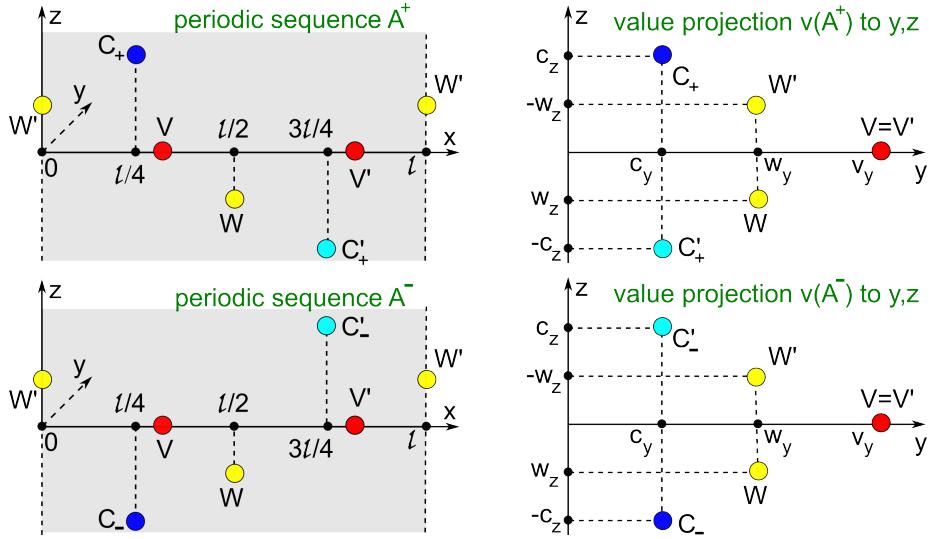


Fig. 7.4 These 1-periodic sequences  $A^\pm \subset \mathbb{R} \times \mathbb{R}^2$  from [3, Fig. 2] have identical past invariants.

Assuming that  $v_x \in (\frac{l}{4}, \frac{l}{2})$  as in Fig. 7.4, the time shifts are

$$\text{TS}(M^\pm; l) = \left( \frac{l}{4}, v_x - \frac{l}{4}, \frac{l}{2} - v_x, \frac{l}{4}, v_x - \frac{l}{4}, \frac{l}{2} - v_x \right).$$

Order value projections along the  $x$ -axis from  $\frac{l}{2}$  to the right:

$$v(M^\pm) = \{(w_y, -w_z), (c_y, \pm c_z), (v_y, 0), (w_y, w_z), (c_y, \mp c_z), (v_y, 0)\}.$$

The cyclic distance matrices of  $M^+$  and  $M^-$  are on the left and right, respectively:

$$\begin{pmatrix} d_{11} & d_{12} & d_{21} & d_{11} & d_{12} & d_{21} \\ d_{21} & d_{22} & d_{12} & d_{21} & d_{22} & d_{12} \\ 2|w_z| & 2|c_z| & 0 & 2|w_z| & 2|c_z| & 0 \end{pmatrix} \neq \begin{pmatrix} d_{22} & d_{12} & d_{21} & d_{22} & d_{12} & d_{21} \\ d_{21} & d_{11} & d_{12} & d_{21} & d_{11} & d_{12} \\ 2|w_z| & 2|c_z| & 0 & 2|w_z| & 2|c_z| & 0 \end{pmatrix}.$$

The differences in distances are highlighted in yellow:

$$d_{11} = \sqrt{(w_y - c_y)^2 + (w_z + c_z)^2}, \quad d_{12} = \sqrt{(c_y - v_y)^2 + c_z^2}, \\ d_{22} = \sqrt{(w_y - c_y)^2 + (w_z - c_z)^2}, \quad d_{21} = \sqrt{(w_y - v_y)^2 + w_z^2}.$$

The matrix difference has the norm  $\|\text{CDM}(M^+) - \text{CDM}(M^-)\|_\infty = |d_{11} - d_{22}| > 0$  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  $g(x, y, z) = (x + \frac{l}{2}, y, -z)$ . If both  $c_z, w_z \neq 0$ , then  $\text{CIM}_{+\infty}(A^+, A^-)$  is obtained by minimizing over 6 cyclic permutations  $\gamma \in C_6$ . The trivial permutation and the shift by 3 positions give  $|d_{11} - d_{12}|$ . Any other permutation gives  $d_t = \max\{v_x - \frac{l}{4}, \frac{l}{2} - v_x\}$  from comparing  $\text{TS}(M^+; l)$  with  $\gamma(\text{TS}(M^-; l))$  and  $d_v = \max\{|a - b|\}$  maximized for all pairs of parameters  $a, b \in \{d_{11}, d_{12}, d_{21}, d_{22}\}$ .

In all cases, the metric is positive:  $\text{CIM}_{+\infty}(A^+, A^-) \geq |d_{11} - d_{22}| > 0$ . Then the invariant CI from Definition 7.3.2 distinguishes these sequences  $A^+ \not\cong A^-$ .  $\blacklozenge$

Geo-Mapping Problem 1.4.5 becomes much harder for point sets that are periodic in two directions. The next chapter will solve the case of 2-dimensional lattices.

## References

1. Michel Marie Deza and Elena Deza. *Encyclopedia of distances*. Springer, 2009.
2. Vitaliy Kurlin. Complete and continuous invariants of 1-periodic sequences in polynomial time. *SIAM Journal on Mathematics of Data Science*, 7:1643–1663, 2025.
3. Sergey N Pozdnyakov and Michele Ceriotti. Incompleteness of graph neural networks for points clouds in three dimensions. *Machine Learning: Science and Technology*, 3(4):045020, 2022.

## Chapter 8

# Moduli spaces of 2D lattices under isometry and rigid motion

**Abstract** This chapter continuously parametrises moduli spaces of 2-dimensional lattices under Euclidean isometry, rigid motion, dilation, and homothety. The new root invariants have easily computable metrics and settle the past discontinuity of reduced bases. The moduli space of 2-dimensional lattices under rigid motion can be mapped to the sphere without one point. Hence, any geographic location on Earth can be associated with a canonical lattice. We also define chiral distances that continuously measure deviations from higher-symmetry lattices.

## 8.1 Representations of lattices by unit cells and reduced bases

All sections in this chapter follow papers [17, 6, 7] with minor updates of notations.

Recall that the most practical equivalence relations (rigid motion, isometry, dilation, homothety) on arbitrary subsets of  $\mathbb{R}^n$  were introduced in Example 1.2.3.

**Definition 8.1.1** (a basis and a primitive unit cell  $U(\mathbf{v}_1, \dots, \mathbf{v}_n)$  of a lattice  $\Lambda \subset \mathbb{R}^n$ ).

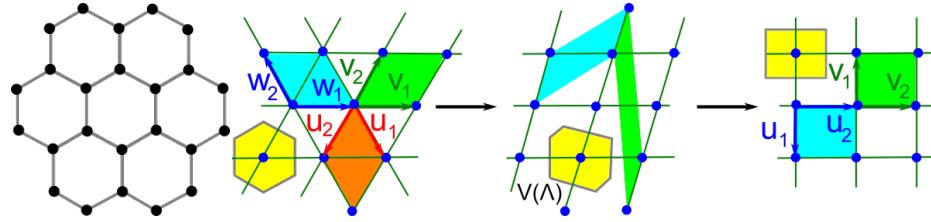
(a) Let vectors  $\mathbf{v}_1, \dots, \mathbf{v}_n$  form a linear *basis* in  $\mathbb{R}^n$ . A *lattice*  $\Lambda \subset \mathbb{R}^n$  consists of all linear combinations  $\sum_{i=1}^n c_i \mathbf{v}_i$  with integer coefficients  $c_i \in \mathbb{Z}$ . The parallelepiped

$$U(\mathbf{v}_1, \dots, \mathbf{v}_n) = \left\{ \sum_{i=1}^n t_i \mathbf{v}_i \mid t_i \in [0, 1) \right\} \subset \mathbb{R}^n \text{ is called a } \textit{primitive unit cell} \text{ of } \Lambda. \quad \blacktriangle$$

The inequalities  $0 \leq t_i < 1$  in Definition 8.1.1(a) guarantee that the copies of primitive unit cells  $U(\mathbf{v}_1, \dots, \mathbf{v}_n)$  translated by all  $\mathbf{v} \in \Lambda$  are disjoint and cover  $\mathbb{R}^n$ .

Recall that the *special linear* group  $SL(\mathbb{Z}_n)$  consists of all  $n \times n$  matrices with integer entries and determinant 1. Let vectors  $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$  form one basis of a lattice  $\Lambda \subset \mathbb{R}^n$ . Then, for any matrix  $A \in SL(\mathbb{Z}_n)$ , the vectors  $A\mathbf{v}_1, \dots, A\mathbf{v}_n$  form another basis of  $\Lambda$  with a very different primitive cell  $U(A\mathbf{v}_1, \dots, A\mathbf{v}_n)$ , see Fig. 8.1.

The past approach to tackling the ambiguity of lattice representations was to consider a reduced basis, briefly as rectangular as possible. In  $\mathbb{R}^3$ , [13] reviewed several reduced



**Fig. 8.1** **Left:** a 2-dimensional layer of graphene is formed by carbon atoms. **Right:** one can generate a hexagonal lattice (as any other) by infinitely many bases and continuously deform into a rectangular lattice (far right) whose bases  $\{v_1, v_2\}$  and  $\{u_1, u_2\}$  are related by an orientation-reversing map. The yellow Voronoi domain  $\bar{V}(\Lambda)$  of any point  $\mathbf{p}$  in a lattice  $\Lambda$  consists of all points  $q \in \mathbb{R}^2$  that are non-strictly closer to  $\mathbf{p}$  than to other points  $\Lambda \setminus \{\mathbf{p}\}$  in the lattice, see Definition 8.1.3(a).

bases. The most common is Niggli's cell [20], whose 2-dimensional version is introduced below. For any  $\mathbf{v}_1 = (a_1, a_2)$  and  $\mathbf{v}_2 = (b_1, b_2)$  in  $\mathbb{R}^2$ , let  $\det(\mathbf{v}_1, \mathbf{v}_2) = a_1 b_2 - a_2 b_1$

be the determinant of the matrix  $\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$  with the columns  $\mathbf{v}_1, \mathbf{v}_2$ .

**Definition 8.1.2** (reduced cell). **(a)** For a lattice  $\Lambda \subset \mathbb{R}^2$  under isometry, a basis and its unit cell  $U(\mathbf{v}_1, \mathbf{v}_2)$  are *reduced* (non-acute) if  $|\mathbf{v}_1| \leq |\mathbf{v}_2|$  and  $-\frac{1}{2}\mathbf{v}_1^2 \leq \mathbf{v}_1 \cdot \mathbf{v}_2 \leq 0$ .

**(b)** Under rigid motion, the conditions are weaker:  $|\mathbf{v}_1| \leq |\mathbf{v}_2|$  and  $|\mathbf{v}_1 \cdot \mathbf{v}_2| \leq \frac{1}{2}\mathbf{v}_1^2$ ,  $\det(\mathbf{v}_1, \mathbf{v}_2) > 0$ , and the new *special condition*: if  $|\mathbf{v}_1| = |\mathbf{v}_2|$  then  $\mathbf{v}_1 \cdot \mathbf{v}_2 \geq 0$ .  $\blacktriangle$

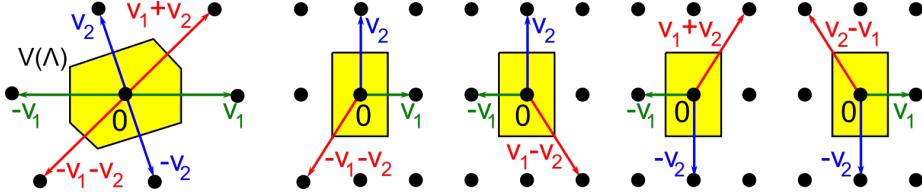
All bases in Fig. 8.1 are reduced under rigid motion. The condition  $|\mathbf{v}_1 \cdot \mathbf{v}_2| \leq \frac{1}{2}\mathbf{v}_1^2$  in Definition 8.1.2 geometrically means that  $\mathbf{v}_1, \mathbf{v}_2$  are close to being orthogonal: the projection of  $\mathbf{v}_2$  to  $\mathbf{v}_1$  is between  $\pm \frac{1}{2}|\mathbf{v}_1|$ . The conditions  $|\mathbf{v}_1| \leq |\mathbf{v}_2|$  and  $-\frac{1}{2}\mathbf{v}_1^2 \leq \mathbf{v}_1 \cdot \mathbf{v}_2 \leq 0$  in Definition 8.1.2 coincide with the conventional definition from [3, section 9.2.2] for type II (non-acute) cells in  $\mathbb{R}^3$  if we choose  $\mathbf{v}_3$  to be very long and orthogonal to  $\mathbf{v}_1, \mathbf{v}_2$ . Alternative type I cells with non-obtuse angles have  $0 \leq \mathbf{v}_1 \cdot \mathbf{v}_2 \leq \frac{1}{2}\mathbf{v}_1^2$ .

[17, Proposition 3.10(a)] proves the uniqueness of a reduced basis under isometry.

Another well-known cell of a lattice  $\Lambda \subset \mathbb{R}^n$  is the *Voronoi domain* [22], also called the *Wigner-Seitz cell*, *Brillouin zone* or *Dirichlet cell*. We use the word *domain* not to confuse it with a unit cell in Definition 8.1.1. Though the Voronoi domain can be defined for any point of a lattice, it suffices to consider only the origin 0.

**Definition 8.1.3** (Voronoi domain  $\bar{V}(\Lambda)$ ). **(a)** The *Voronoi domain* of a lattice  $\Lambda \subset \mathbb{R}^n$  is the neighbourhood  $\bar{V}(\Lambda) = \{\mathbf{p} \in \mathbb{R}^n : |\mathbf{p}| \leq |\mathbf{p} - \mathbf{v}| \text{ for any } \mathbf{v} \in \Lambda\}$  of 0 consisting of all  $\mathbf{p} \in \mathbb{R}^n$  that are non-strictly closer to 0 than to other points  $\mathbf{v} \in \Lambda$ .

**(b)** A vector  $\mathbf{v} \in \Lambda$  is called a *Voronoi vector* if the bisector hyperspace  $H(0, \mathbf{v}) = \{\mathbf{p} \in \mathbb{R}^n \mid \mathbf{p} \cdot \mathbf{v} = \frac{1}{2}\mathbf{v}^2\}$  between 0 and  $\mathbf{v}$  intersects  $\bar{V}(\Lambda)$ . If  $\bar{V}(\Lambda) \cap H(0, \mathbf{v})$  is an  $(n-1)$ -dimensional face of  $\bar{V}(\Lambda)$ , then  $\mathbf{v}$  is called a *strict Voronoi vector*.  $\blacktriangle$



**Fig. 8.2** **Left:** a generic lattice  $\Lambda \subset \mathbb{R}^2$  has a hexagonal Voronoi domain with an obtuse superbasis  $v_1, v_2, v_0 = -v_1 - v_2$ , which is unique under permutations and central symmetry. **Other pictures:** two pairs of obtuse superbases (related by reflection) for a rectangular lattice.

Fig. 8.2 shows how the Voronoi domain  $\bar{V}(\Lambda)$  can be obtained as the intersection of the closed half-spaces  $S(0, v) = \{\mathbf{p} \in \mathbb{R}^n \mid \mathbf{p} \cdot \mathbf{v} \leq \frac{1}{2}v^2\}$  whose boundaries  $H(0, v)$  are bisectors between 0 and all strict Voronoi vectors  $\mathbf{v} \in \Lambda$ . A generic lattice  $\Lambda \subset \mathbb{R}^2$  has a hexagonal Voronoi domain  $\bar{V}(\Lambda)$  with six Voronoi vectors.

Any lattice is determined by its Voronoi domain by [17, Lemma A.2]. However, the combinatorial structure of  $\bar{V}(\Lambda)$  is discontinuous under perturbations. Almost any perturbation of a rectangular basis in  $\mathbb{R}^2$  gives a non-rectangular basis generating a lattice whose Voronoi domain  $\bar{V}(\Lambda)$  is hexagonal, not rectangular. Hence, any integer-valued descriptors of  $\bar{V}(\Lambda)$ , such as the numbers of vertices or edges, are always discontinuous and unsuitable for continuous quantifications.

Lemma 8.1.4 shows how to find all Voronoi vectors of any lattice  $\Lambda \subset \mathbb{R}^n$ . The doubled lattice is  $2\Lambda = \{2\mathbf{v} \mid \mathbf{v} \in \Lambda\}$ . Vectors  $\mathbf{u}, \mathbf{v} \in \Lambda$  are called  $2\Lambda$ -equivalent if  $\mathbf{u} - \mathbf{v} \in 2\Lambda$ . Then any vector  $\mathbf{v} \in \Lambda$  generates its  $2\Lambda$ -class  $\mathbf{v} + 2\Lambda = \{\mathbf{v} + 2\mathbf{u} \mid \mathbf{u} \in \Lambda\}$ , which is  $2\Lambda$  translated by  $\mathbf{v}$  and containing  $-\mathbf{v}$ . All classes of  $2\Lambda$ -equivalent vectors form the quotient space  $\Lambda/2\Lambda$ . Any 1-dimensional lattice  $\Lambda$  generated by a vector  $\mathbf{v}$  has the quotient  $\Lambda/2\Lambda$  consisting of only two classes  $\Lambda$  and  $\mathbf{v} + \Lambda$ .

**Lemma 8.1.4** (criterion for Voronoi vectors [19], [8, Theorem 2]). For any lattice  $\Lambda \subset \mathbb{R}^n$ , a non-zero vector  $\mathbf{v} \in \Lambda$  is a Voronoi vector of  $\Lambda$  if and only if  $\mathbf{v}$  is a shortest vector in its  $2\Lambda$ -class  $\mathbf{v} + 2\Lambda$ . Also,  $\mathbf{v}$  is a strict Voronoi vector if and only if  $\pm\mathbf{v}$  are the only shortest vectors in the  $2\Lambda$ -class  $\mathbf{v} + 2\Lambda$ . ■

Any lattice  $\Lambda \subset \mathbb{R}^2$  generated by  $v_1, v_2$  has  $\Lambda/2\Lambda = \{v_1, v_2, v_1 + v_2\} + \Lambda$ . Notice that the vectors  $v_1 \pm v_2$  belong to the same  $2\Lambda$ -class. Assume that  $v_1, v_2$  are not longer than  $v_1 + v_2$ , which holds if the angle  $\angle(v_1, v_2) \in [60^\circ, 120^\circ]$ . If the sum  $v_1 + v_2$  is shorter than  $v_1 - v_2$  as in Fig. 8.2 (left), then  $\Lambda$  has three pairs of strict Voronoi vectors  $\pm v_1, \pm v_2, \pm(v_1 + v_2)$ . If  $v_1 \pm v_2$  have the same length, the unit cell spanned by  $v_1, v_2$  degenerates to a rectangle,  $\Lambda$  has four non-strict Voronoi vectors  $\pm v_1 \pm v_2$ .

The triple of vector pairs  $\pm v_1, \pm v_2, \mp(v_1 + v_2)$  in Fig. 8.2 motivates the concept of a superbasis with the extra vector  $v_0 = -v_1 - v_2$ , which extends to any dimension  $n$  by setting  $v_0 = -\sum_{i=1}^n v_i$ . For dimensions 2 and 3, [17, Theorem 2.9] proved that any lattice has an obtuse superbasis of vectors whose pairwise scalar products are non-positive and are called *Selling parameters* [21]. For any superbasis in  $\mathbb{R}^n$ , the negated parameters

$p_{ij} = -\mathbf{v}_i \cdot \mathbf{v}_j$  can be interpreted as conorms of lattice characters, which are functions  $\chi : \Lambda \rightarrow \{\pm 1\}$  satisfying  $\chi(\mathbf{u} + \mathbf{v}) = \chi(\mathbf{u})\chi(\mathbf{v})$ , see [8, Theorem 6]. So  $p_{ij}$  will be defined as *conorms* only for an obtuse superbase below.

**Definition 8.1.5** (obtuse superbase and conorms  $p_{ij}$ ). For any basis  $\mathbf{v}_1, \dots, \mathbf{v}_n$  of  $\mathbb{R}^n$ , the *superbase*  $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_n$  includes the vector  $\mathbf{v}_0 = -\sum_{i=1}^n \mathbf{v}_i$ . The *conorms*  $p_{ij} = -\mathbf{v}_i \cdot \mathbf{v}_j$  are the negative scalar products of the vectors above. The superbase is *obtuse* if all conorms  $p_{ij} \geq 0$ , so all angles between vectors  $\mathbf{v}_i, \mathbf{v}_j$  are non-acute for distinct indices  $i, j \in \{0, 1, \dots, n\}$ . The superbase is called *strict* if all  $p_{ij} > 0$ .  $\blacktriangle$

Formula (1) in [8] has a typo initially defining  $p_{ij}$  as exact Selling parameters, but later [8, Theorems 3, 7, 8] use the non-negative conorms  $p_{ij} = -\mathbf{v}_i \cdot \mathbf{v}_j \geq 0$ .

The indices of a conorm  $p_{ij}$  are distinct and unordered. We set  $p_{ij} = p_{ji}$  for all indices  $i, j$ . For  $n = 1$ , the 1-dimensional lattice generated by a vector  $\mathbf{v}_1$  has the obtuse superbase consisting of the two vectors  $\mathbf{v}_0 = -\mathbf{v}_1$  and  $\mathbf{v}_1$ , so the only conorm  $p_{01} = -\mathbf{v}_0 \cdot \mathbf{v}_1 = \mathbf{v}_1^2$  is the squared length of  $\mathbf{v}_1$ . Any superbase of  $\mathbb{R}^n$  has  $\frac{n(n+1)}{2}$  conorms  $p_{ij}$ , for example, three conorms  $p_{01}, p_{02}, p_{12}$  in dimension 2.

**Definition 8.1.6** (partial sums  $\mathbf{v}_S$  and vonorms  $\mathbf{v}_S^2$ ). Let a lattice  $\Lambda \subset \mathbb{R}^n$  have a superbase  $B = \{\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_n\}$ . For any proper subset  $S \subset \{0, 1, \dots, n\}$  of indices, consider its complement  $\bar{S} = \{0, 1, \dots, n\} \setminus S$  and the *partial sum*  $\mathbf{v}_S = \sum_{i \in S} \mathbf{v}_i$  whose squared lengths  $\mathbf{v}_S^2$  are called the *vonorms* of  $B$  and can be expressed as

$$\mathbf{v}_S^2 = \left( \sum_{i \in S} \mathbf{v}_i \right) \left( - \sum_{j \in \bar{S}} \mathbf{v}_j \right) = - \sum_{i \in S, j \in \bar{S}} \mathbf{v}_i \cdot \mathbf{v}_j = \sum_{i \in S, j \in \bar{S}} p_{ij}.$$

For  $n = 2$ , we get the following simple formulae

$$(8.1.6a) \quad \mathbf{v}_0^2 = p_{01} + p_{02}, \quad \mathbf{v}_1^2 = p_{01} + p_{12}, \quad \mathbf{v}_2^2 = p_{02} + p_{12}.$$

The above formulae allow us to express the conorms via vonorms as follows

$$(8.1.6b) \quad p_{12} = \frac{1}{2}(\mathbf{v}_1^2 + \mathbf{v}_2^2 - \mathbf{v}_0^2), \quad p_{01} = \frac{1}{2}(\mathbf{v}_0^2 + \mathbf{v}_1^2 - \mathbf{v}_2^2), \quad p_{02} = \frac{1}{2}(\mathbf{v}_0^2 + \mathbf{v}_2^2 - \mathbf{v}_1^2).$$

So  $p_{ij} = \frac{1}{2}(\mathbf{v}_i^2 + \mathbf{v}_j^2 - \mathbf{v}_k^2)$  for distinct  $i, j \in \{0, 1, 2\}$  and  $k = \{0, 1, 2\} - \{i, j\}$ .  $\blacktriangle$

Lemma 8.1.7 will later help to prove that a lattice is uniquely determined under isometry by an obtuse superbase, hence by its vonorms or, equivalently, conorms.

**Lemma 8.1.7** (Voronoi vectors  $\mathbf{v}_S$  [8, Theorem 3]). For any obtuse superbase  $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_n$  of a lattice, all partial sums  $\mathbf{v}_S$  from Definition 8.1.6 split into  $2^n - 1$  symmetric pairs  $\mathbf{v}_S = -\mathbf{v}_{\bar{S}}$ , which are Voronoi vectors representing distinct  $2\Lambda$ -classes in  $\Lambda/2\Lambda$ . All Voronoi vectors  $\mathbf{v}_S$  are strict if and only if all  $p_{ij} > 0$ .  $\blacksquare$

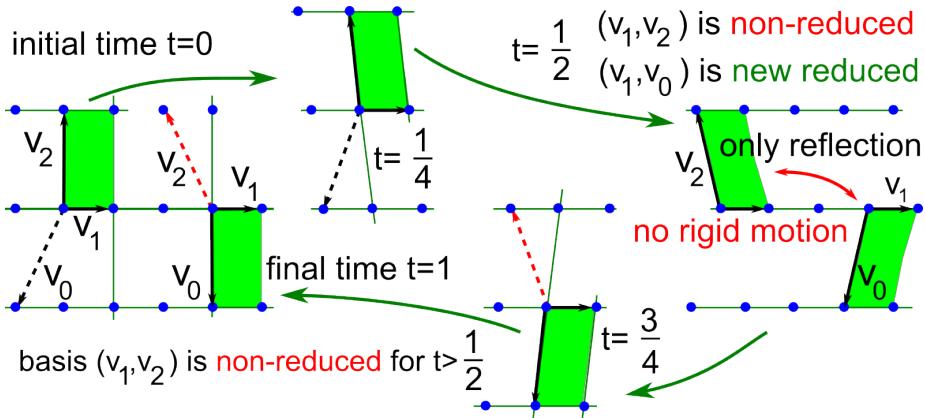
By Conway and Sloane [8, section 2], any lattice  $\Lambda \subset \mathbb{R}^n$  that has an obtuse superbase is called a *lattice of Voronoi's first kind*. Any lattice in dimensions 2 and 3 is of Voronoi's first kind due to [22, p. 277] for  $n = 2$  and [4, Section III.4.3] for  $n = 3$ .

**Theorem 8.1.8** (reduction to an obtuse superbase). Any lattice  $\Lambda$  in dimensions 2 and 3 has an obtuse superbase  $\{v_0, v_1, \dots, v_n\}$  so that  $v_0 = -\sum_{i=1}^n v_i$  and all conorms  $p_{ij} = -v_i \cdot v_j \geq 0$  for all distinct indices  $i, j \in \{0, 1, \dots, n\}$ .  $\blacksquare$

Conway and Sloane in [8, section 7] attempted to prove Theorem 8.1.8 for  $n = 3$  by example, which is corrected in [16]. Theorem 8.1.8 for  $n = 2$  is proved in [17, appendix A]. Proposition 8.3.8 will establish a 1-1 correspondence between obtuse superbases and reduced bases. The latter bases are implemented by many fast algorithms in crystallography [2]. So our lattice input will be any obtuse superbase.

## 8.2 Geo-mapping problem for lattices under equivalences in $\mathbb{R}^2$

This section shows how past approaches to lattice classifications remained discontinuous and states a suitable version of Geo-Mapping 1.4.5 for 2-dimensional lattices. Fig. 8.3 illustrates how both reduced basis and obtuse superbase discontinuously change under rigid motion. As for many other objects, we need invariants, because [23, Theorem 15] proved that any reduced basis is discontinuous under coordinate-wise comparisons.



**Fig. 8.3** Discontinuity of obtuse superbases under rigid motion. The obtuse superbase of the vectors  $v_1 = (1, 0)$ ,  $v_2(t) = (-t, 2)$ ,  $v_0(t) = (t - 1, -2)$  deforms for  $t \in [0, 1]$ . The initial and final superbases at  $t = 0$  and  $t = 1$  generate the same rectangular lattice but are not related by rigid motion.

Any lattice  $\Lambda \subset \mathbb{R}^2$  with a basis  $v_1, v_2$  defines the *positive quadratic form*

$$Q(x, y) = (xv_1 + yv_2)^2 = q_{11}x^2 + 2q_{12}xy + q_{22}y^2 \geq 0 \text{ for all } x, y \in \mathbb{R},$$

where  $q_{11} = \mathbf{v}_1^2$ ,  $q_{22} = \mathbf{v}_2^2$ ,  $q_{12} = \mathbf{v}_1 \cdot \mathbf{v}_2$ . Changing the basis  $\mathbf{v}_1, \mathbf{v}_2$  (possibly by reflection) is equivalent to replacing  $x, y$  by the linear combinations of the coordinates of  $x\mathbf{v}_1 + y\mathbf{v}_2$  in a new basis. Conversely, any positive quadratic form  $Q(x, y)$  can be written as a sum  $(a_1x + b_1y)^2 + (a_2x + b_2y)^2$ , see [11, Theorem 2 on p. 116], and defines the lattice with the basis  $\mathbf{v}_1 = (a_1, a_2)$ ,  $\mathbf{v}_2 = (b_1, b_2)$ .

In 1773, Lagrange [9] proved that any positive quadratic form can be rewritten so that  $0 < q_{11} \leq q_{22}$  and  $-q_{11} \leq 2q_{12} \leq 0$ . The resulting non-acute *reduced* basis  $\mathbf{v}_1, \mathbf{v}_2$  satisfies  $0 < \mathbf{v}_1^2 \leq \mathbf{v}_2^2$  and  $-\mathbf{v}_1^2 \leq 2\mathbf{v}_1 \cdot \mathbf{v}_2 \leq 0$  without the new special conditions in Definition 8.1.2. Alternatively,  $0 \leq 2q_{12} \leq q_{11}$  and  $0 \leq 2\mathbf{v}_1 \cdot \mathbf{v}_2 \leq \mathbf{v}_1^2$  define a non-obtuse reduced basis. The mirror images  $\Lambda^\pm$  of  $\Lambda(\frac{1}{4})$  in Fig. 8.3 (top) generated by the obtuse reduced bases  $\mathbf{v}_1 = (1, 0)$ ,  $\mathbf{v}_2^\pm = (-\frac{1}{4}, \pm 2)$  have the same reduced form  $Q(x, y) = x^2 - \frac{1}{2}xy + 4y^2$  not distinguishing  $\Lambda^\pm$  under rigid motion.

If  $\mathbf{v}_1, \mathbf{v}_2$  form a unique reduced basis, Lemma 8.5.2 shows that  $\pm\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_1 + \mathbf{v}_2\}$  are the three (pairs of) shortest Voronoi vectors. Then the *metric tensor*  $(\mathbf{v}_1^2, \mathbf{v}_1 \cdot \mathbf{v}_2, \mathbf{v}_2^2)$  is a complete isometry invariant but doesn't distinguish mirror images (enantiomorphs). Instead of one scalar product and two squared lengths, Delone used the homogeneous parameters [10, section 29] equal to the conorms  $p_{ij}$  from Definition 8.1.5:

$$p_{01} = q_{11} + q_{12} = \mathbf{v}_1^2 + \mathbf{v}_1 \cdot \mathbf{v}_2 = \mathbf{v}_1 \cdot (\mathbf{v}_1 + \mathbf{v}_2) = -\mathbf{v}_0 \cdot \mathbf{v}_1,$$

$$p_{02} = q_{22} + q_{12} = \mathbf{v}_2^2 + \mathbf{v}_1 \cdot \mathbf{v}_2 = \mathbf{v}_2 \cdot (\mathbf{v}_1 + \mathbf{v}_2) = -\mathbf{v}_0 \cdot \mathbf{v}_2,$$

$$p_{12} = -q_{12} = -\mathbf{v}_1 \cdot \mathbf{v}_2.$$

The quadratic form becomes a sum of squares:  $Q_\Lambda = p_{01}x^2 + p_{22}y^2 + p_{12}(x - y)^2$ . The inequalities for  $q_{ij}$  are equivalent to the simple ordering  $0 \leq p_{12} \leq p_{01} \leq p_{02}$ , which Definition 8.3.1 will use to introduce a more convenient root invariant.

Bi-continuity conditions in general Problem 1.4.5(d,e) become challenging for periodic sets already in dimension 1. Any lattice under rigid motion in  $\mathbb{R}$  is equivalent to a periodic sequence  $l\mathbb{Z}$  for a period  $l > 0$ . However, for any small  $\varepsilon > 0$ , the  $\varepsilon$  perturbation of  $l$  up to (say)  $l + \varepsilon$  makes the perturbed sequence  $(l + \varepsilon)\mathbb{Z}$  very different from  $l\mathbb{Z}$ . Indeed, extra  $\varepsilon$ -shifts gradually move points  $(l + \varepsilon)n$  further and further away from  $ln$  as  $n \in \mathbb{Z}$  increases. Nonetheless, we will prove bi-continuity of bijections between invariant spaces of 2D lattices and the following metric spaces of obtuse superbases.

**Definition 8.2.1** (spaces of obtuse superbases under equivalences). Let  $B = \{\mathbf{v}_i\}_{i=0}^n$  and  $B' = \{\mathbf{u}_i\}_{i=0}^n$  be any obtuse superbases in  $\mathbb{R}^n$ .

**(a)** Let  $\text{SBR}(\mathbb{R}^n)$  be the space of equivalence classes obtuse of *superbases under isometry* under the action of special orthogonal maps  $f \in \text{SO}(\mathbb{R}^n)$  with the *superbase rigid metric*  $\text{SRM}_\infty(B, B') = \min_{f \in \text{SO}(\mathbb{R}^n)} \max_{i=0, \dots, n} |f(\mathbf{u}_i) - \mathbf{v}_i|$ .

**(b)** Let  $\text{SBI}(\mathbb{R}^n)$  be the space of equivalence classes obtuse of *superbases under isometry* under the action of orthogonal maps  $f \in \text{O}(\mathbb{R}^n)$  with the *superbase isometry metric*  $\text{SIM}_\infty(B, B') = \min_{f \in \text{O}(\mathbb{R}^n)} \max_{i=0, \dots, n} |f(\mathbf{u}_i) - \mathbf{v}_i|$ .

(c) Let  $\text{SBD}(\mathbb{R}^n)$  be the space of equivalence classes obtuse of *superbases under dilation* under the action of dilation maps  $f \in \text{SO}(\mathbb{R}^n) \times \mathbb{R}_+$  with the *superbase dilation metric*  $\text{SDM}_\infty(B, B') = \min_{f \in \text{SO}(\mathbb{R}^n) \times \mathbb{R}_+} \max_{i=0, \dots, n} |f(\mathbf{u}_i) - \mathbf{v}_i|$ .

(d) Let  $\text{SBH}(\mathbb{R}^n)$  be the space of equivalence classes obtuse of *superbases under homothety* under the action of homothety maps  $f \in \text{O}(\mathbb{R}^n) \times \mathbb{R}_+$  with the *superbase homothety metric*  $\text{SHM}_\infty(B, B') = \min_{f \in \text{O}(\mathbb{R}^n) \times \mathbb{R}_+} \max_{i=0, \dots, n} |f(\mathbf{u}_i) - \mathbf{v}_i|$ .  $\blacktriangle$

Since any continuous function over a compact domain achieves its minimum value and  $\text{SO}(\mathbb{R}^n), \text{O}(\mathbb{R}^n)$  are compact, the minima in Definition 8.2.1(a,b) are achievable. For fixed superbases  $B, B'$ , one can restrict uniform scaling by reasonable bounds to guarantee the existence of minima in Definition 8.2.1(c,d).

Problem 8.2.2 is a case of Problem 1.4.5 for 2D lattices. We use isometry as the main equivalence, but moduli spaces of lattices will also be parametrised under rigid motion and their compositions with uniform scaling (dilation and homothety).

**Problem 8.2.2** (geo-mapping for 2D lattices). Design a *geocode* on the space of 2D lattices under isometry that is an invariant satisfying the following conditions.

(a) *Completeness*: any lattices  $\Lambda \simeq \Lambda'$  are isometric in  $\mathbb{R}^2$  if and only if  $I(\Lambda) = I(\Lambda')$ .

(b) *Reconstruction*: any lattice  $\Lambda \subset \mathbb{R}^2$  can be reconstructed from its invariant value  $I(\Lambda)$ , uniquely under isometry in  $\mathbb{R}^2$ .

(c) *Metric*: there is a metric  $d$  on the invariant space  $\{I(\Lambda) \mid \text{lattices } \Lambda \subset \mathbb{R}^2\}$ , which satisfies all axioms in Definition 1.3.1(a).

(d) *Continuity*: there is a constant  $\lambda > 0$  such that, for any  $\varepsilon > 0$ , if any lattices  $\Lambda, \Lambda' \subset \mathbb{R}^2$  have obtuse have superbases  $B, B'$ , respectively with  $\text{SIM}_\infty(B, B') \leq \varepsilon$ , then  $d(I(\Lambda), I(\Lambda')) \leq \lambda\sqrt{\varepsilon}$ .

(e) *Inverse continuity*: for any  $\varepsilon > 0$ , there is  $\delta > 0$  such that if lattices  $\Lambda, \Lambda' \subset \mathbb{R}^2$  satisfy  $d(I(\Lambda), I(\Lambda')) \leq \delta$ , then they have obtuse superbases  $B, B'$ , respectively, with  $\text{SIM}_\infty(B, B') \leq \varepsilon$ .

(f) *Realisability*: the invariant space  $\{I(\Lambda) \mid \text{lattices } \Lambda \subset \mathbb{R}^2\}$  can be parametrised so that we can generate any value  $I(\Lambda)$  realisable by some lattice  $\Lambda \subset \mathbb{R}^2$ .

(g) *Euclidean embedding*: the invariant space  $\{I(\Lambda) \mid \text{lattices } \Lambda \subset \mathbb{R}^2\}$  with the metric  $d$  allows a Lipschitz embedding into a suitable Euclidean space  $\mathbb{R}^N$  for some  $N$ .

(h) *Computability*: the invariant  $I$  and the metric  $d(I(\Lambda), I(\Lambda'))$  can be computed in a constant time  $O(1)$  from reduced bases of  $\Lambda, \Lambda'$ .  $\star$

A geocode that satisfies Problem 8.2.2 continuously parametrises the space of 2D lattices under isometries, similarly defined for other equivalences below.

**Definition 8.2.3** (moduli spaces of lattices under four equivalences in  $\mathbb{R}^n$ ). We consider all lattices  $\Lambda \subset \mathbb{R}^n$  below.

(a) The *Lattice Isometry Space*  $\text{LIS}(\mathbb{R}^n)$  is the space of lattices  $\Lambda$  under isometry.

(b) The *Lattice Rigid Space*  $\text{LRS}(\mathbb{R}^n)$  is the space of lattices under rigid motion.

- (c) The *Lattice Homothety Space*  $LHS(\mathbb{R}^n)$  is the space of lattices under homothety.
- (d) The *Lattice Dilation Space*  $LDS(\mathbb{R}^n)$  is the space of lattices under dilation.  $\blacktriangle$

The traditional approach to deciding if lattices are isometric is to compare their conventional or reduced cells. Though this comparison theoretically gives a complete invariant, in practice, all real lattices in periodic crystals are non-isometric due to noise in measurements. Since all atoms vibrate, any real lattice basis is always perturbed. The discontinuity of reduced bases under perturbations was experimentally known since 1965 [18, p. 80] and was proved for all potential reductions in [23, Theorem 15].

A more practical goal is to design a complete invariant that is continuous under any perturbations of (bases of) lattices. Such a *geocode*, which is more generally defined in Problem 1.4.5, will unambiguously parametrise the Lattice Rigid Space  $LRS(\mathbb{R}^n)$  consisting of infinitely many equivalence classes of lattices under rigid motion in  $\mathbb{R}^n$ . For example, our Earth is continuously parametrised by the latitude and longitude, very similar to the Lattice Dilation Space  $LDS(\mathbb{R}^2)$  as will become clear soon.

The Lattice Rigid Space  $LRS(\mathbb{R}^n)$  is continuous and connected because any two lattices can be joined by a continuous deformation of their bases as in Fig. 8.1. Such deformation can always be visualised as a continuous path in the space  $LRS(\mathbb{R}^n)$ .

The Euclidean embeddability in 8.2.2(g) raises Problem 8.2.2 above metric geometry to define a simpler Euclidean structure on  $LIS(\mathbb{R}^n)$ . It is easy to multiply any lattice by a fixed scalar, but a sum of any two lattices is harder to define in a meaningful way independent of lattice bases. We will overcome this obstacle by a natural embedding of invariant spaces of lattices into  $\mathbb{R}^3$  and  $\mathbb{R}^4$  to fully solve Problem 8.2.2.

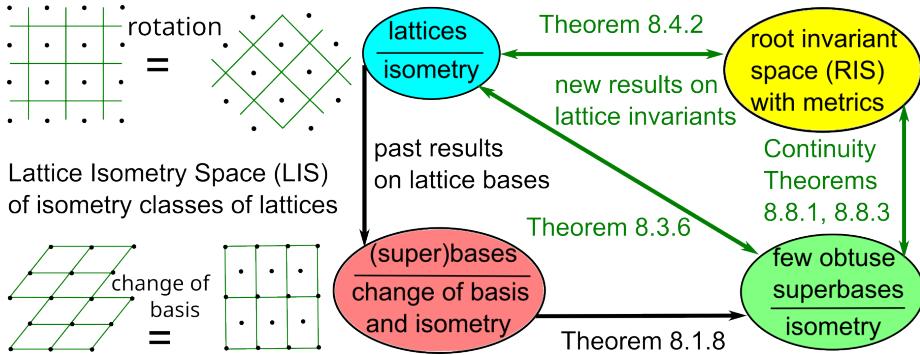
The isometry classification condition in 8.2.2(a) can be interpreted via group actions as follows, see [12] and [24]. Let  $\mathcal{B}_n$  be the space of all linear bases in  $\mathbb{R}^n$ .

Under a change of basis, all lattices in  $\mathbb{R}^n$  form the  $n^2$ -dimensional orbit space  $\mathcal{L}_n = \mathcal{B}_n/\text{GL}(\mathbb{Z}^n)$ , see [12, formula (1.38), p. 34]. Under homothety, the orbit space  $\mathcal{L}_n^d = \mathcal{L}_n/\mathbb{R}_+^\times$  becomes  $(n^2 - 1)$ -dimensional. Under orthogonal maps from the group  $O(\mathbb{R}^n)$ , the orbit space of lattices can be identified with the cone  $C_+(Q_n) = \mathcal{B}_n/O(\mathbb{R}^n)$  of positive quadratic forms, where  $Q_n$  denotes the space of real symmetric  $n \times n$  matrices, see [12, formula (1.67), p. 41]. The Lattice Isometry Space  $LIS(\mathbb{R}^n)$  was called the space of *intrinsic* lattices  $\mathcal{L}_n^o = C_+(Q_n)/\text{GL}(\mathbb{Z}^n)$  in [12, formula (1.70), p. 42].

Another approach to identify an intrinsic lattice (isometry class), say for  $n = 2$ , was to choose a fundamental domain of the action of  $\text{GL}(\mathbb{Z}^2)$  on the cone  $C_+(Q_2)$ . This choice is equivalent to a choice of a reduced basis, which can be discontinuous.

Mirror reflections of any lattice  $\Lambda$  correspond to quadratic forms  $q_{11}x^2 \pm 2q_{12}xy + q_{22}y^2$  that differ by a sign of  $q_{12}$ . To distinguish mirror images of lattices, Definition 8.3.4 will introduce  $\text{sign}(\Lambda)$ . Then continuous deformations of lattices become continuous paths in a space of invariants, see Remark 8.5.1.

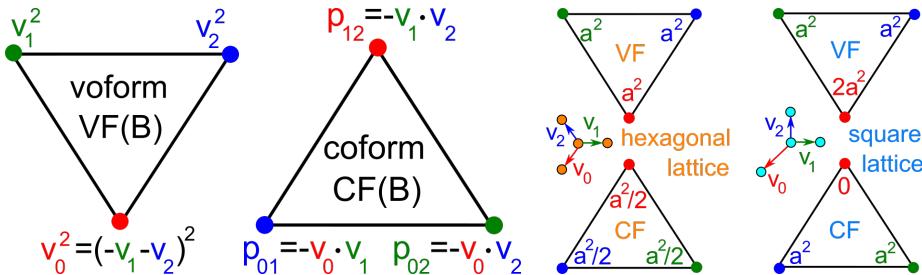
Fig. 8.4 summarises the past obstacles and a full solution to Problem 8.2.2. The Root Invariant Space  $RIS(\mathbb{R}^2)$  consists of ordered triples of square roots of conorms from Definition 8.1.5. Related invariants will continuously parametrise spaces of lattices under rigid motion, dilation, and homothety, as defined below.



**Fig. 8.4**  $\text{LIS}(\mathbb{R}^2)$  is bijectively and continuously mapped to the space of root invariants, which are ordered triples of square roots of scalar products of vectors of an obtuse superbase of a lattice  $\Lambda \subset \mathbb{R}^2$ .

### 8.3 Invariants of an obtuse superbase of a 2-dimensional lattice

Definition 8.3.1 introduces voforms VF and coforms CF, which are triangular cycles whose three nodes are marked by vonorms and conorms, respectively. We start from any obtuse superbase  $B$  of a lattice  $\Lambda \subset \mathbb{R}^2$  to define VF, CF, and a root invariant RI. Lemma 8.3.7(a) will justify that RI depends only on  $\Lambda$ , not on  $B$ .



**Fig. 8.5** **1st picture:** a voform  $\text{VF}(B)$  of a 2D lattice with an obtuse superbase  $B = \{v_0, v_1, v_2\}$ . **2nd picture:** nodes of a coform  $\text{CF}(B)$  are marked by conorms  $p_{ij}$ . **3rd and 4th pictures:** VF and CF of the hexagonal and square lattice with a minimum inter-point distance  $a$ .

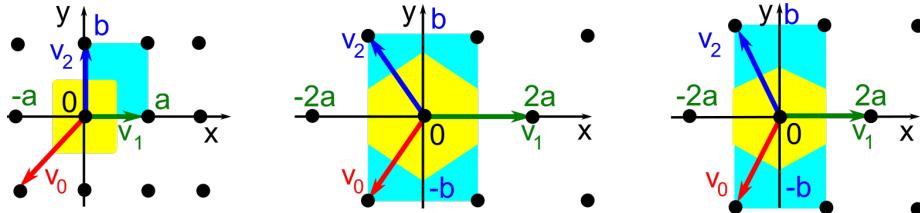
**Definition 8.3.1** (voform VF, coform CF, ordered *root invariant* RI). For any ordered obtuse superbase  $B$  in  $\mathbb{R}^2$ , the *voform*  $\text{VF}(B)$  is the cycle on three nodes marked by the vonorms  $v_0^2, v_1^2, v_2^2$ , see Fig. 8.5. The *coform*  $\text{CF}(B)$  is the cycle on three nodes marked by the conorms  $p_{12}, p_{02}, p_{01}$ . Since all conorms  $p_{ij} \geq 0$ , we can define the *root products*  $r_{ij} = \sqrt{p_{ij}}$ . The *root invariant*  $\text{RI}(B)$  is obtained by writing the three root products  $r_{12}, r_{01}, r_{02}$  in the increasing order.  $\blacktriangle$

By Theorem 8.1.8 any lattice  $\Lambda \subset \mathbb{R}^2$  has an obtuse superbase with all  $p_{ij} \geq 0$ . At least two root products  $r_{ij}$  should be positive, otherwise one vonorm vanishes, but there are no other restrictions on  $r_{ij} \geq 0$ . The vonorms  $\mathbf{v}_0^2, \mathbf{v}_1^2, \mathbf{v}_2^2 > 0$  should satisfy three triangle inequalities such as  $\mathbf{v}_0^2 \leq \mathbf{v}_1^2 + \mathbf{v}_2^2$ , only one of them can be an equality. The ordering  $r_{12} \leq r_{01} \leq r_{02}$  is equivalent to  $\mathbf{v}_1^2 \leq \mathbf{v}_2^2 \leq \mathbf{v}_0^2$  by formulae (8.1.6a). Root products have the same units as original coordinates of basis vectors, for example, Angstroms:  $1\text{\AA} = 10^{-10}\text{m}$ . The ordered root invariant  $\text{RI}(B)$  is more convenient than  $\text{VF}(B)$  and  $\text{CF}(B)$ , which depend on an order of vectors of  $B$ .

**Example 8.3.2.** (a) A lattice  $\Lambda$  with a rectangular cell of sides  $a \leq b$  has an obtuse superbase  $B$  with  $\mathbf{v}_1 = (a, 0)$ ,  $\mathbf{v}_2 = (0, b)$ ,  $\mathbf{v}_0 = (-a, -b)$ , and  $\text{RI}(B) = (0, a, b)$ .

(b) For any lattice  $\Lambda \subset \mathbb{R}^2$  whose Voronoi domain  $\bar{V}(\Lambda)$  is a mirror-symmetric hexagon, assume that the  $x$ -axis is its line of symmetry. Since  $\bar{V}(\Lambda)$  is centrally symmetric with respect to the origin 0, the  $y$ -axis is also its line of symmetry, see Fig. 8.6.

Then  $\Lambda$  has the centred rectangular (non-primitive) cell with sides  $2a \leq 2b$ . The obtuse superbase  $B$  with  $\mathbf{v}_1 = (2a, 0)$ ,  $\mathbf{v}_2 = (-a, b)$ ,  $\mathbf{v}_0 = (-a, -b)$  has  $\text{RI}(B) = (a\sqrt{2}, a\sqrt{2}, \sqrt{b^2 - a^2})$  for  $b \geq a\sqrt{3}$ . For  $a \leq b < a\sqrt{3}$ , we should swap  $r_{02} = \sqrt{b^2 - a^2}$  with  $r_{12} = a\sqrt{2}$  to get an ordered root invariant  $\text{RI}(B)$ .  $\blacklozenge$



**Fig. 8.6** Left:  $\Lambda$  has a rectangular cell and obtuse superbase  $B$  with  $\mathbf{v}_1 = (a, 0)$ ,  $\mathbf{v}_2 = (0, b)$ ,  $\mathbf{v}_0 = (-a, -b)$ , see Example 8.3.2 and Lemma 8.3.3. Other lattices  $\Lambda$  have a rectangular cell  $2a \times 2b$  and an obtuse superbase  $B$  with  $\mathbf{v}_1 = (2a, 0)$ ,  $\mathbf{v}_2 = (-a, b)$ ,  $\mathbf{v}_0 = (-a, -b)$ . Middle:  $\text{RI}(B) = (\sqrt{b^2 - a^2}, a\sqrt{2}, a\sqrt{2})$ ,  $a \leq b \leq a\sqrt{3}$ . Right:  $\text{RI}(B) = (a\sqrt{2}, a\sqrt{2}, \sqrt{b^2 - a^2})$ ,  $a\sqrt{3} \leq b$ .

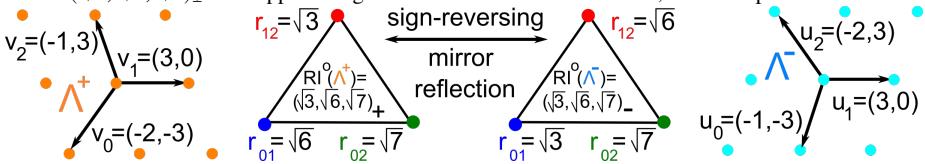
A lattice  $\Lambda \subset \mathbb{R}^n$  that can be mapped to itself by a mirror reflection with respect to a  $(n-1)$ -dimensional hyperspace can be called *mirror-symmetric* or *achiral*. Since a mirror reflection of any lattice  $\Lambda \subset \mathbb{R}^2$  with respect to a line  $L \subset \mathbb{R}^2$  can be realised by a rotation in  $\mathbb{R}^3$  around  $L$  through  $180^\circ$ , the term *achiral* sometimes applies to all 2D lattices and becomes non-trivial only for 3D lattices. This paper for 2D lattices uses the clearer adjective *mirror-symmetric*.

**Lemma 8.3.3** (root invariants of mirror-symmetric lattices  $\Lambda \subset \mathbb{R}^2$ , [17, Lemma 3.3]). An obtuse superbase  $B$  generates a *mirror-symmetric* lattice  $\Lambda(B)$  if and only if

- (a) the root invariant  $\text{RI}(B)$  contains a zero value and  $\Lambda(B)$  is rectangular, or
- (b)  $\text{RI}(B)$  has equal root products and the Voronoi domain of  $\Lambda(B)$  is a square or a hexagon whose symmetry group has two orthogonal axes of symmetry.  $\blacksquare$

**Definition 8.3.4** ( $\text{sign}(B)$ , the oriented root invariant  $\text{RI}^o(B)$ ). If an obtuse superbase  $B$  generates a mirror-symmetric lattice, set  $\text{sign}(B) = 0$ . Else all vectors of  $B$  have different lengths and angles not equal to  $90^\circ$  by Lemma 8.3.3. Let  $\mathbf{v}_1, \mathbf{v}_2$  be the shortest vectors of  $B$  so that  $|\mathbf{v}_1| < |\mathbf{v}_2|$ . Then  $\text{sign}(B) = \pm 1$  is the sign of the determinant  $\det(\mathbf{v}_1, \mathbf{v}_2)$  of the matrix with the columns  $\mathbf{v}_1, \mathbf{v}_2$ . The *oriented root invariant*  $\text{RI}^o(B)$  is obtained by adding  $\text{sign}(B)$  as a superscript to  $\text{RI}(B)$ , see Fig. 8.3.  $\blacktriangle$

**Fig. 8.7** The lattices  $\Lambda, \Lambda'$  are mirror reflections of each other and have oriented root invariants  $\text{RI}^o = (\sqrt{3}, \sqrt{6}, \sqrt{7})_{\pm}$  with opposite signs introduced in Definition 8.3.4, see Example 8.3.5.



If  $\text{sign}(B) = 0$ , this zero superscript in  $\text{RI}^o(B)$  can be skipped for simplicity, so  $\text{RI}^o(B) = \text{RI}(B)$  in this case. Theorem 8.3.6 will show that  $\text{sign}(B)$  can be considered as an invariant of a lattice  $\Lambda$  under dilation.

In Definition 8.3.4 the determinant  $\det(\mathbf{v}_1, \mathbf{v}_2)$  is the signed area of the unit cell  $U(\mathbf{v}_1, \mathbf{v}_2)$  equal to  $|\mathbf{v}_1| \cdot |\mathbf{v}_2| \sin \angle(\mathbf{v}_1, \mathbf{v}_2)$ , where the angle is measured from  $\mathbf{v}_1$  to  $\mathbf{v}_2$  in the anticlockwise direction around the origin  $0 \in \mathbb{R}^2$ . For a strict obtuse superbase  $B$ , all angles between its basis vectors are strictly obtuse. Then  $\text{sign}(B) = +1$  if  $\angle(\mathbf{v}_1, \mathbf{v}_2)$  is in the positive range  $(90^\circ, 180^\circ)$ , else  $\text{sign}(B) = -1$ .

**Example 8.3.5** (signs of lattices in Fig. 8.3). (a) The lattice  $\Lambda^+$  in the first picture of Fig. 8.3 has the obtuse superbase  $B$  with  $\mathbf{v}_1 = (3, 0)$ ,  $\mathbf{v}_2 = (-1, 3)$ ,  $\mathbf{v}_0 = (-2, -3)$  of lengths  $3, \sqrt{10}, \sqrt{13}$ , respectively, so  $\Lambda^+$  is not mirror-symmetric. Since  $\mathbf{v}_1, \mathbf{v}_2$  are the

two shortest vectors of  $B^+$  and  $\det(\mathbf{v}_1, \mathbf{v}_2) = \det \begin{pmatrix} 3 & -1 \\ 0 & 3 \end{pmatrix} > 0$ , we get  $\text{sign}(B^+) = +1$ .

The anticlockwise angle is  $\angle(\mathbf{v}_1, \mathbf{v}_2) = 180^\circ - \arcsin \frac{3}{\sqrt{10}} \approx 108^\circ$ .

(b) The lattice  $\Lambda^-$  in the last picture of Fig. 8.3 is obtained from  $\Lambda^+$  by a mirror reflection and has the obtuse superbase  $B^-$  with  $\mathbf{u}_1 = \mathbf{v}_1$ ,  $\mathbf{u}_2 = (-2, 3)$ ,  $\mathbf{u}_0 = (-1, -3)$  of lengths  $3, \sqrt{13}, \sqrt{10}$ , respectively, so  $\Lambda^-$  is not mirror-symmetric. Since  $\mathbf{u}_1, \mathbf{u}_0$  are the shortest

vectors,  $\det(\mathbf{u}_1, \mathbf{u}_0) = \det \begin{pmatrix} 3 & -1 \\ 0 & -3 \end{pmatrix} < 0$ , we get  $\text{sign}(B^-) = -1$ . The anticlockwise angle

is  $\angle(\mathbf{u}_1, \mathbf{u}_0) = \arcsin \frac{3}{\sqrt{10}} - 180^\circ \approx -108^\circ$ .  $\blacklozenge$

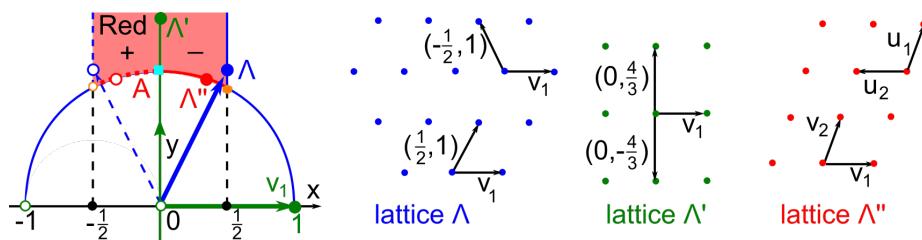
Theorem 8.3.6 below is crucial for a complete classification of 2D lattices in Theorem 8.4.2 and Corollary 8.4.5. Theorem 8.3.6 highlights that mirror-symmetric lattices

have more options for obtuse superbases under rigid motion. The same rectangular lattice can have two obtuse bases with  $\mathbf{v}_1 = (1, 0)$ ,  $\mathbf{v}_2 = (0, \pm 2)$ , which are related by reflection in the  $x$ -axis, not by rigid motion. This symmetry-related ambiguity is much harder to resolve for 3D lattices even under isometry, see [16].

**Theorem 8.3.6** (isometric obtuse superbases, [17, Theorem 3.7]). Any lattices  $\Lambda, \Lambda' \subset \mathbb{R}^2$  are isometric if and only if any obtuse superbases of  $\Lambda, \Lambda'$  are isometric. If  $\Lambda, \Lambda'$  are not rectangular, the same conclusion holds for rigid motion instead of isometry. Any rectangular (non-square) lattice has two obtuse superbases related by reflection. ■

**Lemma 8.3.7** (lattice invariants, [17, Lemma 3.8]). **(a)** For any obtuse superbase  $B$  of a lattice  $\Lambda \subset \mathbb{R}^2$ , the root invariant  $\text{RI}(B)$  is an isometry invariant of  $\Lambda$  and can be denoted by  $\text{RI}(\Lambda)$ . Similarly,  $\text{RI}^o(\Lambda)$  and  $\text{sign}(\Lambda)$  are invariants of a lattice  $\Lambda$  under rigid motion and dilation, respectively.

**(b)** A lattice  $\Lambda \subset \mathbb{R}^2$  is mirror-symmetric if and only if  $\text{sign}(\Lambda) = 0$ .



**Fig. 8.8 Left:** any reduced basis in Definition 8.1.2 can be mapped under dilation to the basis of  $\mathbf{v}_1 = (1, 0)$  and  $\mathbf{v}_2 = (x, y) \in \text{Red}$  from Proposition 8.3.8. **Right:** for each of the lattices  $\Lambda, \Lambda', \Lambda''$  represented by small blue, green, red circles/disks on the right, the conditions of Definition 8.1.2 choose one reduced basis among two bases that differ under rigid motion.

**Proposition 8.3.8** (reduced bases, [17, Proposition 3.10]). **(a)** Under isometry in  $\mathbb{R}^2$ , all reduced bases  $\mathbf{v}_1, \mathbf{v}_2$  from Definition 8.1.2 are in a 1-1 correspondence with all obtuse superbases  $B = \{\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2\}$  such that  $|\mathbf{v}_1| \leq |\mathbf{v}_2| \leq |\mathbf{v}_0|$ . Under isometry, any lattice  $\Lambda \subset \mathbb{R}^2$  has a unique reduced basis in the conditions of Definition 8.1.2, see Fig. 8.8.

(b) Under rigid motion, any lattice has a unique reduced basis in Definition 8.1.2. ■

## 8.4 Complete invariants of 2D lattices under four equivalences

Lemma 8.3.7 showed that  $\text{RI}(\Lambda)$ ,  $\text{RI}^o(\Lambda)$  are invariants of lattices under isometry and rigid motion, respectively. To prove completeness of the invariants in Theorem 8.4.2, Lemma 8.4.1 reconstructs an obtuse superbase of  $\Lambda$ . Corollary 8.4.5 will classify lattices under homothety by projected invariants introduced in Definition 8.4.4.

**Lemma 8.4.1** (superbase reconstruction, [17, Lemma 4.1]). An obtuse superbase  $B = \{\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2\}$  of a lattice  $\Lambda \subset \mathbb{R}^2$  can be uniquely reconstructed under isometry and under rigid motion from its root invariant  $\text{RI}(\Lambda)$  and its oriented root invariant  $\text{RI}^o(\Lambda)$ , respectively. If  $\text{RI}(\Lambda) = (r_{12}, r_{01}, r_{02})$ , the basis vectors  $\mathbf{v}_1, \mathbf{v}_2$  are determined by

$$|\mathbf{v}_1| = \sqrt{r_{12}^2 + r_{01}^2}, \quad |\mathbf{v}_2| = \sqrt{r_{12}^2 + r_{02}^2}, \quad \cos \angle(\mathbf{v}_1, \mathbf{v}_2) = \frac{-r_{12}^2}{\sqrt{r_{12}^2 + r_{01}^2} \sqrt{r_{12}^2 + r_{02}^2}},$$

and span a primitive unit cell of the area  $A(\Lambda) = \sqrt{r_{12}^2 r_{01}^2 + r_{12}^2 r_{02}^2 + r_{01}^2 r_{02}^2}$ .  $\blacksquare$

**Theorem 8.4.2** (completeness of root invariants, [17, Theorem 4.2]). **(a)** Any lattices  $\Lambda, \Lambda' \subset \mathbb{R}^2$  are isometric if and only if their root invariants coincide:  $\text{RI}(\Lambda) = \text{RI}(\Lambda')$ .

**(b)** Any lattices  $\Lambda, \Lambda'$  are related by rigid motion if and only if  $\text{RI}^o(\Lambda) = \text{RI}^o(\Lambda')$ .  $\blacksquare$

The above classification helps prove that some other isometry invariants of lattices are also complete and continuous. By (8.1.6ab) the voform  $\text{VF} = (\mathbf{v}_0^2, \mathbf{v}_1^2, \mathbf{v}_2^2)$  and coform  $\text{CF} = (p_{12}, p_{01}, p_{02})$  are both complete if considered under  $3!$  permutations. The root invariant  $\text{RI}$  is a uniquely ordered version of  $\text{CF}$  and deserves its own name. The square roots  $r_{ij} = \sqrt{p_{ij}}$  have original units of the vector coordinates.

Theorem 8.4.2 and Lemma 8.5.2 imply that, after taking square roots of vonorms, the ordered lengths, say  $|\mathbf{v}_1| \leq |\mathbf{v}_2| \leq |\mathbf{v}_0|$ , form a complete invariant that should satisfy the triangle inequality  $|\mathbf{v}_1| + |\mathbf{v}_2| \geq |\mathbf{v}_0|$ . This inequality is the only disadvantage of the complete invariant  $|\mathbf{v}_1| \leq |\mathbf{v}_2| \leq |\mathbf{v}_0|$  in comparison with ordered root products  $r_{12} \leq r_{01} \leq r_{02}$ , which are easier to visualise in Fig. 8.9, 8.10.

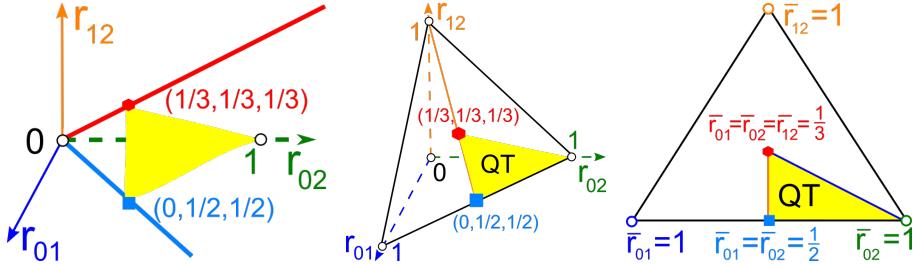
Classification Theorem 8.4.2 says that all isometry classes of lattices  $\Lambda \subset \mathbb{R}^2$  are in a 1-1 correspondence with all ordered triples  $0 \leq r_{12} \leq r_{01} \leq r_{02}$  of root products in  $\text{RI}(\Lambda)$ . Only the smallest root product  $r_{12}$  can be zero, two others  $r_{01} \leq r_{02}$  should be positive, otherwise  $\mathbf{v}_1^2 = r_{12}^2 + r_{01}^2 = 0$  by formulae (8.1.6a).

We explicitly describe the set of all possible root invariants, which will be later converted into metric spaces with continuous metrics in Definitions 8.6.1 and 8.6.3.

**Definition 8.4.3** (triangular cone  $\text{TC}$ ). All root invariants  $\text{RI}(\Lambda) = (r_{12}, r_{01}, r_{02})$  of lattices  $\Lambda \subset \mathbb{R}^2$  live in the *triangular cone*  $\text{TC} = \{0 \leq r_{12} \leq r_{01} \leq r_{02}\}$  within the octant  $\text{Oct} = [0, +\infty)^3$  excluding the axes in the coordinates  $r_{12}, r_{01}, r_{02}$ , see Fig. 8.9 (left).

The boundary  $\partial(\text{TC})$  of the cone  $\text{TC}$  consists of root invariants of all mirror-symmetric lattices from Lemma 8.3.3: the bisector planes  $\{r_{01} = r_{02}\}$  and  $\{r_{12} = r_{01}\}$  within  $\text{TC}$ . The orange line  $\{0 < r_{12} = r_{01} = r_{02}\} \subset \partial(\text{TC})$  in Fig. 8.9 (left) consists of root invariants of hexagonal lattices with a minimum inter-point distance  $r_{12}\sqrt{2}$ . The blue line  $\{r_{12} = 0 < r_{01} = r_{02}\} \subset \partial(\text{TC})$  consists of root invariants of square lattices with a minimum inter-point distance  $r_{01}$ .  $\blacktriangle$

To classify lattices under homothety, it is convenient to scale them by the *size*  $|\Lambda| = r_{12} + r_{01} + r_{02}$ . This sum is a simpler uniform measure of size than (say) the unit cell area  $A(\Lambda)$  from Lemma 8.4.1, which can be small even for long cells.



**Fig. 8.9** **Left:** the triangular cone  $TC = \{(r_{12}, r_{01}, r_{02}) \in \mathbb{R}^3 \mid 0 \leq r_{12} \leq r_{01} \leq r_{02} \neq 0\}$  represents the space RIS of all root invariants of 2D lattices, see Definition 8.4.3. **Middle:**  $TC$  projects to the quotient triangle  $QT = TC \cap \{r_{12} + r_{01} + r_{02} = 1\}$  representing the space LHS of 2D lattices under homothety, see Corollary 8.4.5. **Right:** the quotient triangle  $QT$  can be parametrised by  $x = \bar{r}_{02} - \bar{r}_{01} \in [0, 1]$  and  $y = 3\bar{r}_{12} \in [0, 1]$ , see QT also in Fig. 8.10.

**Definition 8.4.4** (projected invariants  $PI(\Lambda)$  and  $PI^o(\Lambda)$ ). The *triangular projection*  $TP : TC \rightarrow \{r_{12} + r_{01} + r_{02} = 1\}$  divides each coordinate by the size  $|\Lambda| = r_{12} + r_{01} + r_{02}$  and gives  $\overline{RI}(\Lambda) = (\bar{r}_{12}, \bar{r}_{01}, \bar{r}_{02}) = \frac{(r_{12}, r_{01}, r_{02})}{r_{12} + r_{01} + r_{02}}$  in  $TC \cap \{r_{12} + r_{01} + r_{02} = 1\}$ . Then we map  $(\bar{r}_{12}, \bar{r}_{01}, \bar{r}_{02})$  to the *projected invariant*  $PI(\Lambda) = (x, y)$  with  $x = \bar{r}_{02} - \bar{r}_{01} \in [0, 1]$  and  $y = 3\bar{r}_{12} \in [0, 1]$  in the *quotient triangle*

$$QT = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x < 1, 0 \leq y \leq 1, x + y \leq 1\},$$

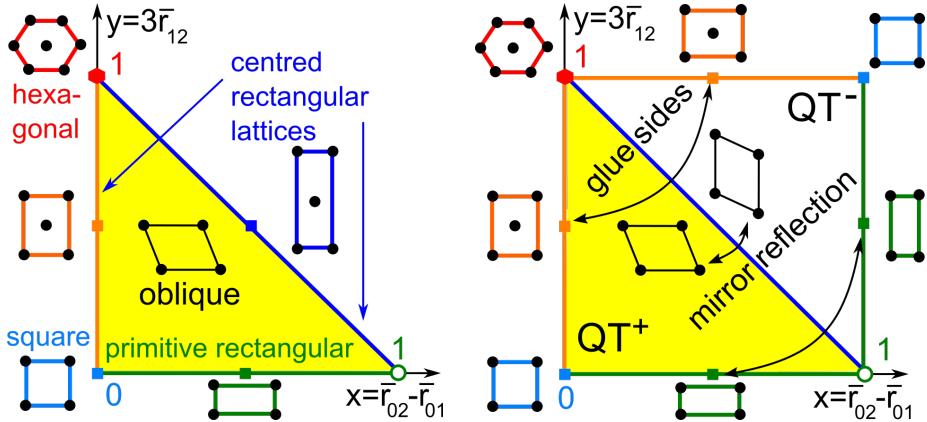
see Fig. 8.10. All oriented root invariants  $RI^o(\Lambda)$  live in the *doubled cone* DC that is the union of two triangular cones  $TC^\pm$ , where we identify any two boundary points representing the same root invariant  $RI(\Lambda)$  with  $\text{sign}(\Lambda) = 0$ . The *oriented projected invariant*  $PI^o(\Lambda) = (x, y)^\pm$  is  $PI(\Lambda)$  with the superscript from  $\text{sign}(\Lambda)$ .  $\blacktriangle$

The inequality  $1 \geq x + y = (\bar{r}_{02} - \bar{r}_{01}) + 3\bar{r}_{12}$  follows after multiplying both sides by the size  $|\Lambda|$ , because  $r_{12} + r_{01} + r_{02} \geq (\bar{r}_{02} - \bar{r}_{01}) + 3\bar{r}_{12}$  becomes  $r_{01} \geq r_{12}$ .

The set of oriented projected invariants  $PI^o$  is visualised in Fig. 8.10 (right) as the *quotient square* QS obtained by gluing the quotient triangle  $QT^+$  with its mirror image  $QT^-$ . The boundaries of both triangles excluding the vertex  $(x, y) = (1, 0)$  are glued by the diagonal reflection  $(x, y) \leftrightarrow (1 - y, 1 - x)$ . Any pair of points  $(x, y) \in QT^+$  and  $(1 - y, 1 - x) \in QT^-$  in Fig. 8.10 (right) represent mirror images of a lattice under homothety, see Corollary 8.4.5. So QS is a topological sphere without a single point and will be parameterised by geographic-style coordinates in [6].

Following Fig. 8.5, any square lattice has a root invariant  $RI = (0, a, a)$ , so its projected invariant  $PI = (0, 0)$  is at the bottom left vertex of  $QT$  in Fig. 8.10 (left), identified with top right vertex of QS in Fig. 8.10 (right). Any hexagonal lattice has a root invariant  $RI = (a, a, a)$ , so its projected invariant  $PI = (0, 1)$  is at the top left vertex of  $QT$  in Fig. 8.10 (left), identified with bottom right vertex of QS.

By Example 8.3.2(a) any rectangular lattice has  $RI = (0, a, b)$  for  $a < b$ , hence its projected invariant  $PI = (\frac{b-a}{a+b}, 0)$  belongs to the bottom edge of  $QT$  identified with the top edge of QS. By Example 8.3.2(b) any lattice with a mirror-symmetric Voronoi



**Fig. 8.10 Left:** all projected invariants  $\text{PI}(\Lambda)$  of lattices  $\Lambda \subset \mathbb{R}^2$  live in the quotient triangle  $QT$  from Fig. 8.9, which is parameterised by  $x = \bar{r}_{02} - \bar{r}_{01} \in [0, 1]$  and  $y = 3\bar{r}_{12} \in [0, 1]$ . **Right:** mirror reflections  $\Lambda^\pm$  of any non-mirror-symmetric lattice can be represented by a pair of points in the quotient square  $QS = QT^+ \cup QT^-$  symmetric in the diagonal  $x + y = 1$ .

domain has RI with 0 or two equal root products. Such lattices have a rhombic unit cell and form the centred rectangular Bravais class. Their projected invariants belong to the vertical edges and diagonal of  $QS$  in Fig. 8.10 (right). The companion paper [6] discusses Bravais classes of 2-dimensional lattices in detail.

**Corollary 8.4.5** (completeness of  $\text{PI}$ , [17, Corollary 4.6]). Any lattices  $\Lambda, \Lambda' \subset \mathbb{R}^2$  are related by homothety if and only if their projected invariants are equal:  $\text{PI}(\Lambda) = \text{PI}(\Lambda')$ . Any lattices  $\Lambda, \Lambda'$  are related by dilation if and only if  $\text{PI}^o(\Lambda) = \text{PI}^o(\Lambda')$ . ■

**Lemma 8.4.6** (criteria of mirror-symmetric lattices in  $\mathbb{R}^2$ , [17, Lemma 4.7]). A lattice  $\Lambda$  in  $\mathbb{R}^2$  is mirror-symmetric if and only if one of the following equivalent conditions holds:  $\text{sign}(\Lambda) = 0$  or  $\text{RI}(\Lambda) \in \partial\text{TC}$  or  $\text{PI}(\Lambda) \in \partial\text{QT}$ . So the boundaries of the triangular cone  $\text{TC}$  and the quotient triangle  $QT$  consist of root invariants and projected invariants, respectively, of all mirror-symmetric lattices  $\Lambda \subset \mathbb{R}^2$ . ■

## 8.5 Inverse design and a spherical map of 2D lattices

This section discusses lattices in terms of group actions, inversely designs lattices from invariants, and embeds the Lattice Rigid Space  $\text{LRS}(\mathbb{R}^2)$  in a 2-dimensional sphere.

In the theory of complex functions, any lattice  $\Lambda \subset \mathbb{R}^2$  can be considered as a subgroup of the complex plane  $\mathbb{C}$  whose quotient  $\mathbb{C}/\Lambda$  is a torus. By the Riemann mapping theorem any compact Riemann surface of genus 1 is conformally equivalent (holomorphically homeomorphic) to the quotient  $\mathbb{C}/\Lambda$  for some lattice  $\Lambda$ , see [15, Section 5.3]. These tori  $\mathbb{C}/\Lambda$  and  $\mathbb{C}/\Lambda'$  are *conformally equivalent* if and only if  $\Lambda, \Lambda'$  are related by homothety, see [14, Theorem 6.1.4]. The spaces  $\text{LHS}(\mathbb{R}^2)$  and  $\text{LDS}(\mathbb{R}^2)$

of all lattices  $\Lambda \subset \mathbb{C} = \mathbb{R}^2$  under homothety and dilation are the quotient triangle QT and the quotient square QS, respectively, see Fig. 8.10.

**Remark 8.5.1** (lattices via group actions). Another parametrisation of the Lattice Homothety Space  $\text{LHS}(\mathbb{R}^2)$  can be obtained from a fundamental domain of the action of  $\text{GL}(\mathbb{Z}^2) \times \mathbb{R}_+^\times$  on the cone  $C_+(Q_2)$  of positive quadratic forms. Recall that any lattice  $\Lambda \subset \mathbb{R}^2$  with a basis  $\mathbf{v}_1, \mathbf{v}_2$  defines the positive quadratic form

$$Q_\Lambda(x, y) = (x\mathbf{v}_1 + y\mathbf{v}_2)^2 = \mathbf{v}_1^2 x^2 + 2\mathbf{v}_1 \mathbf{v}_2 xy + \mathbf{v}_2^2 y^2 = q_{11}x^2 + 2q_{12}xy + q_{22}y^2 \geq 0$$

whose positivity for all  $(x, y) \in \mathbb{R}^2 - 0$  means that  $q_{12}^2 < q_{11}q_{22}$ . The cone  $C_+(Q_2)$  of all positive quadratic forms projects to the unit disk  $\xi^2 + \eta^2 < 1$  parameterised by  $\xi = \frac{q_{22} - q_{11}}{q_{11} + q_{22}}$  and  $\eta = \frac{-2q_{12}}{q_{11} + q_{22}}$ . Indeed, the positivity condition  $q_{12}^2 < q_{11}q_{22}$  for the form  $Q_\Lambda(x, y)$  is equivalent to  $\xi^2 + \eta^2 < 1$  in the coordinates above.

The quadratic form  $Q_\Lambda$  is in a *reduced* (non-acute) form if  $0 \leq -2q_{12} \leq q_{11} \leq q_{22}$  and  $q_{11} > 0$ , see [12, formula (1.130) on p. 75]. The above conditions define the fundamental domain  $T = \{0 \leq \xi < 1, 0 \leq \eta \leq \frac{1}{2}, \xi + 2\eta \leq 1\}$ , see [24, Fig. 8.1]. This non-isosceles triangle is one of the infinitely many triangular domains within the disk  $\xi^2 + \eta^2 < 1$  in [12, Fig. 1.2 on p. 82] or [24, Fig. 6.2]. Choosing one triangular domain is equivalent to choosing a reduced basis under isometry, not under rigid motion.

For instance, the mirror-symmetric bases  $\mathbf{v}_1 = (1, 0)$ ,  $\mathbf{v}_2^\pm = (-\frac{1}{2}, \pm 1)$  have the same reduced non-acute form  $x^2 - xy + \frac{5}{4}y^2$  represented only by  $(\xi, \eta) = (\frac{1}{9}, \frac{4}{9})$ . The above ambiguity under rigid motion is resolved by  $\text{sign}(\Lambda)$  in the twice larger space  $\text{LDS}(\mathbb{R}^2)$  visualised as the quotient square QS, see Example 8.8.2.

More importantly, the inverse map from  $\text{RI}^o(\Lambda)$  to a reduced basis is discontinuous at any rectangular lattice  $\Lambda$  with a unit cell  $a \times b$ . Indeed, slight perturbations of  $\Lambda$  have unique reduced bases that are not close to each other, being close to the distant bases  $(a, 0), (0, \pm b)$ , which are not equivalent under rigid motion for  $a < b$ . This discontinuity of lattice bases will emerge in  $\mathbb{R}^3$  even under isometry [16]. In  $\mathbb{R}^2$ , Corollary 8.8.4 will completely settle the basis discontinuity under rigid motion.

Another complete invariant is the ordered voform  $\mathbf{v}_1^2 \leq \mathbf{v}_2^2 \leq \mathbf{v}_0^2$  or the lengths  $|\mathbf{v}_1| \leq |\mathbf{v}_2| \leq |\mathbf{v}_0|$  of the three shortest Voronoi vectors from Lemma 8.5.2 below. However, this invariant doesn't extend even to dimension  $n = 3$  due to a 6-parameter family of pairs of non-isometric lattices  $\Lambda_1 \not\cong \Lambda_2$  that have the same lengths of seven shortest Voronoi vectors in  $\mathbb{R}^3$ , see [16]. The above reasons justify the choice of homogeneous coordinates  $r_{ij}$ , which easily extend to higher dimensions.  $\blacklozenge$

**Lemma 8.5.2** ([8, Theorem 7]). For any obtuse superbase  $(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2)$  of a lattice  $\Lambda \subset \mathbb{R}^2$ , the vonorms  $\mathbf{v}_0^2, \mathbf{v}_1^2, \mathbf{v}_2^2$  are squared lengths of shortest Voronoi vectors.  $\blacksquare$

The projected invariant  $\text{PI} = (x, y)$  obtained from  $\text{RI}$  is preferable to the coordinates  $(\xi, \eta)$ , which define a non-isosceles triangle, while the isosceles quotient triangle QT will lead to easier formulae for metrics in the next section. Since the metric tensor  $(\mathbf{v}_1^2, \mathbf{v}_1 \cdot \mathbf{v}_2, \mathbf{v}_2^2) = (q_{11}, q_{12}, q_{22})$  and its 3-dimensional analogue are more familiar to

crystallographers, we will rephrase key results from sections 8.6-8.7 by using these non-homogeneous coordinates in the companion paper [6].

**Proposition 8.5.3** (inverse design of 2D lattices, [17, Proposition 4.9]). For  $s > 0$  and any point  $(x, y)$  in the quotient triangle QT, there is a unique (under isometry) lattice  $\Lambda$  with the projected invariant  $\text{PI}(\Lambda) = (x, y)$  and size  $|\Lambda| = s = r_{12} + r_{01} + r_{02}$ . Then

$$(8.5.3a) \quad \text{RI}(\Lambda) = (r_{12}, r_{01}, r_{02}) = \left( \frac{s}{3}y, \frac{s}{6}(3 - 3x - y), \frac{s}{6}(3 + 3x - y) \right).$$

If  $(x, y)$  is in the interior of QT, the invariant RI defines a pair of lattices  $\Lambda^\pm$  that have opposite signs and unique (under isometry) reduced basis vectors  $\mathbf{v}_1, \mathbf{v}_2$  with the lengths  $|\mathbf{v}_1| = \sqrt{r_{12}^2 + r_{01}^2}$ ,  $|\mathbf{v}_2| = \sqrt{r_{12}^2 + r_{02}^2}$  and the anticlockwise angle

$$(8.5.3b) \quad \angle(\mathbf{v}_1, \mathbf{v}_2) = \arccos \frac{-4y^2}{\sqrt{(9x^2 + 5y^2 - 6y + 9)^2 - 36x^2(3 - y)^2}}. \quad \blacksquare$$

Example 8.5.4 shows the power of Proposition 8.5.3 based on Theorem 8.4.2 and Corollary 8.4.5 for inverse design by sampling the square QS at interesting places.

Fig. 8.11 (right) visualises the doubled cone DC of oriented root invariants  $\text{RI}^o$  from Definition 8.3.4 by uniting the triangular cone  $\text{TC} = \{0 \leq r_{12} \leq r_{01} \leq r_{02}\}$  with its mirror reflection in the vertical plane  $\{r_{01} = r_{02}\}$  including the  $r_{12}$ -axis.

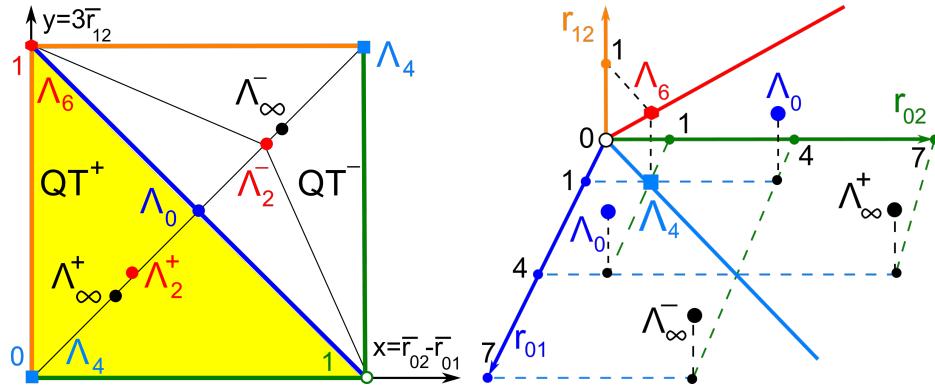
The lattice  $\Lambda_0$  with  $\text{RI} = (1, 1, 4)$  is represented by two boundary points of DC identified by  $(r_{01}, r_{02}) \leftrightarrow (r_{02}, r_{01})$ . The lattices  $\Lambda_\infty^\pm$  with the root invariant  $\text{RI} = (r_{12}, r_{01}, r_{02}) = (1, 4, 7)$  are represented by  $(1, 4, 7)$  and its mirror image  $(1, 7, 4)$  in DC related by the reflection in the vertical bisector plane  $r_{01} = r_{02}$  containing the root invariants of  $\Lambda_4, \Lambda_6$ . The superscript shows  $\text{sign}(\Lambda_\infty^\pm) = \pm 1$ .

**Example 8.5.4** (inverse design of 2D lattices). We will inversely design the lattices  $\Lambda_4, \Lambda_6, \Lambda_0, \Lambda_2^\pm, \Lambda_\infty^\pm$ , see their visualised invariants in Fig. 8.11 (right).

**( $\Lambda_4$ )** We design the square lattice  $\Lambda_4$  starting from its projected invariant at the origin  $\text{PI}(\Lambda_4) = (0, 0) \in \text{QT}$ , which is identified with the top right vertex  $(1, 1) \in \text{QS}$  in Fig. 8.11 (left). Formula (8.5.3a) for the size  $|\Lambda_4| = 2$  (only to get simplest integers) gives  $\text{RI}(\Lambda_4) = (0, 1, 1)$ . An obtuse superbase  $\{\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2\}$  can be reconstructed by Lemma 8.4.1. The vonorms are  $\mathbf{v}_1^2 = \mathbf{v}_2^2 = 0^2 + 1^2 = 1$ ,  $\mathbf{v}_0^2 = 1^2 + 1^2 = 2$ . We can choose the standard obtuse superbase  $\mathbf{v}_1 = (1, 0)$ ,  $\mathbf{v}_2 = (0, 1)$ ,  $\mathbf{v}_0 = (-1, -1)$ .

**( $\Lambda_6$ )** We design the hexagonal lattice  $\Lambda_6$  starting from the projected invariant at the top left vertex  $\text{PI}(\Lambda_6) = (0, 1) \in \text{QT}$ , which is identified with the bottom right vertex  $(1, 0) \in \text{QS}$  in Fig. 8.11 (left). Formula (8.5.3a) for the size  $|\Lambda_6| = 3$  (only to get simplest integers) gives  $\text{RI}(\Lambda_6) = (1, 1, 1)$ . To reconstruct an obtuse superbase  $\{\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2\}$  by Lemma 8.4.1, find the vonorms  $\mathbf{v}_1^2 = \mathbf{v}_2^2 = \mathbf{v}_0^2 = 1^2 + 1^2 = 2$ . Formula (8.5.3b) gives the angle  $\angle(\mathbf{v}_1, \mathbf{v}_2) = \arccos \frac{-4}{\sqrt{(5-6+9)^2}} = \arccos \left( -\frac{1}{2} \right) = 120^\circ$ . We can choose the superbase  $\mathbf{v}_1 = (\sqrt{2}, 0)$ ,  $\mathbf{v}_2 = \left( -\frac{1}{\sqrt{2}}, \frac{\sqrt{3}}{\sqrt{2}} \right)$ ,  $\mathbf{v}_0 = \left( -\frac{1}{\sqrt{2}}, -\frac{\sqrt{3}}{\sqrt{2}} \right)$ .

**( $\Lambda_0$ )** We inversely design the lattice  $\Lambda_0$  in Fig. 8.11 starting from  $\text{PI}(\Lambda_0) = (x, y)$  at the centre  $(\frac{1}{2}, \frac{1}{2}) \in \text{QS}$ . Formula (8.5.3a) for the size  $|\Lambda_0| = 6$  (only to get simplest



**Fig. 8.11** **Left:**  $QS = QT^+ \cup QT^-$  includes mirror-symmetric lattices  $\Lambda_4, \Lambda_6, \Lambda_0$  and non-mirror-symmetric lattices  $\Lambda_\infty^\pm$ , see Example 8.6.2 and Table 8.1 later. **Right:** the doubled cone DC is visualised as  $\{0 \leq r_{12} \leq \min\{r_{01}, r_{02}\} > 0\}$  bounded by the planes  $\{r_{12} = 0\}, \{r_{12} = r_{01}\}, \{r_{12} = r_{02}\}$  with the identifications  $(r_{12}, r_{01}, r_{02}) \leftrightarrow (r_{12}, r_{02}, r_{01})$  on the boundary  $\partial DC$ .

(integers) gives  $RI(\Lambda_0) = (1, 1, 4)$ . To reconstruct an obtuse superbase  $\{\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2\}$  by Lemma 8.4.1, find the vonorms  $\mathbf{v}_1^2 = 1^2 + 1^2 = 2, \mathbf{v}_0^2 = \mathbf{v}_2^2 = 1^2 + 4^2 = 17$ .

Formula (8.5.3b) gives the angle  $\angle(\mathbf{v}_1, \mathbf{v}_2) = \arccos \frac{-4}{\sqrt{(\frac{9}{4} + \frac{5}{4} - 3 + 9)^2 - 9(\frac{5}{4})^2}} = \arccos\left(-\frac{1}{\sqrt{34}}\right) \approx 99.9^\circ$ . We can choose the following superbase, see Fig. 8.12:  $\mathbf{v}_1 = (\sqrt{2}, 0), \mathbf{v}_2 = |\mathbf{v}_2|(\cos \angle(\mathbf{v}_1, \mathbf{v}_2), \sin \angle(\mathbf{v}_1, \mathbf{v}_2)) = \left(-\frac{1}{\sqrt{2}}, \frac{\sqrt{33}}{\sqrt{2}}\right), \mathbf{v}_0 = \left(-\frac{1}{\sqrt{2}}, -\frac{\sqrt{33}}{\sqrt{2}}\right)$ .

**(Λ<sub>2</sub>)** We inversely design the lattice  $\Lambda_2$  in Fig. 8.11 starting from their projected invariants  $PI(\Lambda_2) = \left(\frac{1}{2+\sqrt{2}}, \frac{1}{2+\sqrt{2}}\right)$ , which will maximise the chiral distance  $PC[D_2]$  in Theorem 8.7.5(a). Formula (8.5.3a) for the size  $|\Lambda_2| = 6$  (only to simplify the root invariant) gives  $RI(\Lambda_2) = (2-\sqrt{2}, 2\sqrt{2}-1, 5-\sqrt{2})$ . Since all root products are non-zero and distinct, by Lemma 8.3.3 there is a pair of lattices  $\Lambda_2^\pm$  with  $\text{sign}(\Lambda_2^\pm) = \pm 1$ . The lattices  $\Lambda_2^\pm$  are related by reflection, not by rigid motion.

To reconstruct an obtuse superbase  $\{\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2\}$  of  $\Lambda_2^\pm$  by Lemma 8.4.1, find  $\mathbf{v}_0^2 = (2\sqrt{2}-1)^2 + (5-\sqrt{2})^2 = (9-4\sqrt{2}) + (27-10\sqrt{2}) = 36-14\sqrt{2} \approx 16.2$ ,

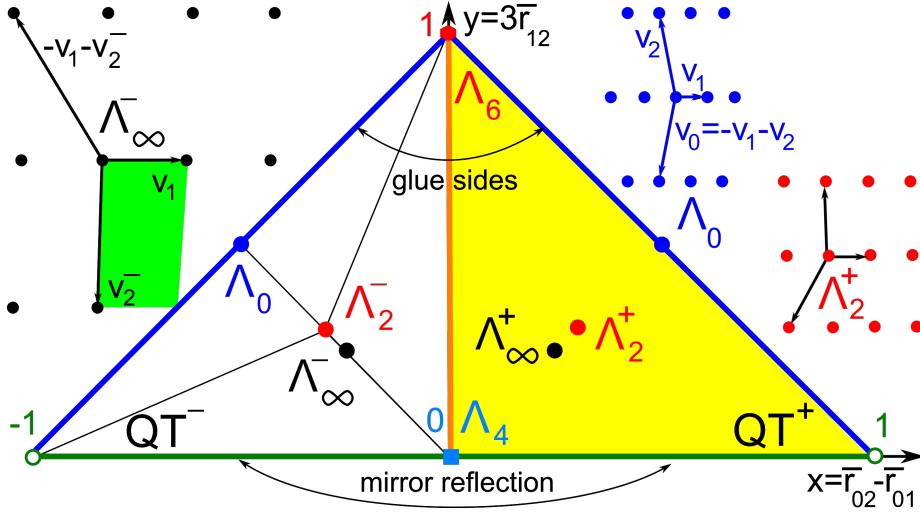
$\mathbf{v}_1^2 = (2-\sqrt{2})^2 + (2\sqrt{2}-1)^2 = (6-4\sqrt{2}) + (9-4\sqrt{2}) = 15-8\sqrt{2} \approx 3.7$ ,

$\mathbf{v}_2^2 = (2-\sqrt{2})^2 + (5-\sqrt{2})^2 = (6-4\sqrt{2}) + (27-10\sqrt{2}) = 33-14\sqrt{2} \approx 13.2$ ,

and the anticlockwise angle  $\angle(\mathbf{v}_1, \mathbf{v}_2) = \arccos \frac{-r_{12}^2}{|\mathbf{v}_1| \cdot |\mathbf{v}_2|} \approx 92.8^\circ$ . Then  $\Lambda_2^\pm$  have

the following obtuse superbases in Fig. 8.12:  $\mathbf{v}_1 = (\sqrt{15-8\sqrt{2}}, 0) \approx (1.9, 0), \mathbf{v}_2 = |\mathbf{v}_2|(\cos \angle(\mathbf{v}_1, \mathbf{v}_2), \sin \angle(\mathbf{v}_1, \mathbf{v}_2)) \approx (-0.18, 3.63), \mathbf{v}_0 \approx (-1.72, -3.63)$ .

**(Λ<sub>∞</sub>)** We inversely design the lattice  $\Lambda_\infty$  in Fig. 8.12 starting from  $PI(\Lambda_\infty) = (x, y)$  at the mid-point  $(\frac{1}{4}, \frac{1}{4})$  of the segment between  $PI(\Lambda_4), PI(\Lambda_0) \in QT$ . Formula (8.5.3a) for the size  $|\Lambda_\infty| = 12$  (only to simplify the root invariant) gives  $RI(\Lambda_\infty) = (1, 4, 7)$ .



**Fig. 8.12** The doubled cone DC in Fig. 8.11 (right) projects to the doubled triangle DT parameterised by  $x \in (-1, 1)$ ,  $y \in [0, 1]$  and obtained by gluing two copies  $QT^\pm$  of the quotient triangle along vertical sides instead of hypotenuses as in QS, see Example 8.5.4 and Table 8.1.

Since all root products are non-zero and distinct, by Lemma 8.3.3 there is a pair of lattices  $\Lambda_\infty^\pm$  of opposite signs  $\text{sign}(\Lambda_\infty^\pm) = \pm 1$ .

To reconstruct an obtuse superbasis  $\{\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2\}$  of  $\Lambda_\infty^\pm$  by Lemma 8.4.1, find the vonorms  $\mathbf{v}_0^2 = 4^2 + 7^2 = 65$ ,  $\mathbf{v}_1^2 = 1^2 + 4^2 = 17$ ,  $\mathbf{v}_2^2 = 1^2 + 7^2 = 50$ , and the anticlockwise angle  $\angle(\mathbf{v}_1, \mathbf{v}_2) = \arccos \frac{-r_{12}^2}{|\mathbf{v}_1| \cdot |\mathbf{v}_2|} = \arccos \left( -\frac{1}{\sqrt{850}} \right) \approx 92^\circ$ . Then  $\Lambda_\infty^\pm$  have the following obtuse superbases in Fig. 8.12:  $\mathbf{v}_1 = (\sqrt{17}, 0) \approx (4.12, 0)$ ,

$$\mathbf{v}_2^\pm = |\mathbf{v}_2|(\cos \angle(\mathbf{v}_1, \mathbf{v}_2), \sin \angle(\mathbf{v}_1, \mathbf{v}_2)) = \left( -\frac{1}{\sqrt{17}}, \pm \frac{\sqrt{849}}{\sqrt{17}} \right) \approx (-0.24, \pm 7.1),$$

$$\mathbf{v}_0^\pm = -\mathbf{v}_1 - \mathbf{v}_2^\pm = \left( -\frac{16}{\sqrt{17}}, \mp \frac{\sqrt{849}}{\sqrt{17}} \right) \approx (-3.88, \mp 7.1), \text{ see all forms in Table 8.1.} \quad \diamond$$

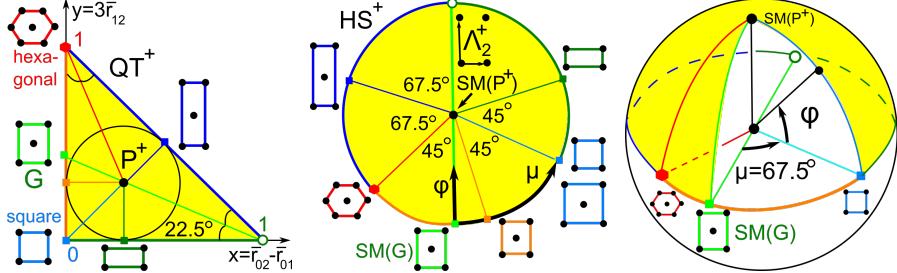
Since the quotient square  $QS = QT^+ \cup QT^-$  with identified sides is a punctured sphere, it is natural to visualise  $QS$  as the round surface of Earth with  $QT^\pm$  as the north/south hemispheres separated by the equator along their common boundary of  $QT$  represented by projected invariants  $\text{PI}(\Lambda)$  of all mirror-symmetric lattices  $\Lambda$ .

We can choose any internal point of the quotient triangle  $QT$  as the north pole. The most natural choice is the incentre  $P^+$  (pole), the centre of the circle inscribed into  $QT^+$  because the rays from  $P^+$  to the vertices of  $QT^+$  equally bisect the angles  $90^\circ, 45^\circ, 45^\circ$ . The incentre of  $QT^+$  has the coordinates  $(x, x)$ , where  $x = 1 - \frac{1}{\sqrt{2}} = \frac{1}{2+\sqrt{2}}$ . The lattice  $\Lambda_2^+$  with the projected invariant  $\text{PI}(\Lambda_2^+) = (x, x)$  has the basis  $v_1 \approx (1.9, 0)$ ,  $v_2 \approx (-0.18, 3.63)$  inversely designed in Example 8.5.4 ( $\Lambda_2$ ).

**Table 8.1** Various invariants of the lattices computed in Example 8.5.4, see Fig. 8.11 and 8.12.

$\Lambda$	$\Lambda_4$	$\Lambda_6$	$\Lambda_0$	$\Lambda_2^\pm$	$\Lambda_\infty^\pm$
$ \Lambda $	2	3	6	6	12
$\text{PI}(\Lambda)$	$(0, 0)$	$(0, 1)$	$\left(\frac{1}{2}, \frac{1}{2}\right)$	$\left(\frac{1}{2+\sqrt{2}}, \frac{1}{2+\sqrt{2}}\right)$	$\left(\frac{1}{4}, \frac{1}{4}\right)$
$\text{RI}^o(\Lambda)$	$(0, 1, 1)$	$(1, 1, 1)$	$(1, 1, 4)$	$(2 - \sqrt{2}, 2\sqrt{2} - 1, 5 - \sqrt{2})^\pm$	$(1, 4, 7)^\pm$
$\text{VF}(\Lambda)$	$(2, 1, 1)$	$(2, 2, 2)$	$(17, 2, 17)$	$(15 - 8\sqrt{2}, 33 - 14\sqrt{2}, 36 - 14\sqrt{2})$	$(65, 17, 50)$

**Fig. 8.13** **Left:** in  $\text{QT}^+$ , the Greenwich line goes from the ‘empty’ point  $(1, 0)$  through incentre  $P^+$  to the point  $G = (0, \sqrt{2} - 1)$ . **Middle:** the hemisphere  $\text{HS}^+$  has the north pole at  $P^+$ , the equator  $\partial\text{QT}^+$  of mirror-symmetric lattices. **Right:** the longitude  $\mu \in (-180^\circ, +180^\circ]$  anticlockwise measures angles from the Greenwich line, the latitude  $\varphi \in [-90^\circ, +90^\circ]$  measures angles from the equator to the north.



**Definition 8.5.5** (spherical lattice map  $\text{SLM} : \text{QS} \rightarrow S^2$ ). **(a)** The *spherical map*  $\text{SLM}$  sends the incentre  $P^+$  of  $\text{QT}$  to the north pole of the hemisphere  $\text{HS}^+$  and the boundary  $\partial\text{QT}$  to the equator of  $\text{HS}^+$ , see Fig. 8.5 (middle). Linearly map the line segment between  $P^+$  and any point  $(x, y)$  in the boundary  $\partial\text{QT}$  to the shortest arc connecting the north pole  $\text{SLM}(P^+)$  to  $\text{SLM}(x, y)$  in the equator of  $\text{HS}^+$ . Extend the *spherical map* to  $\text{SLM} : \text{QS} \rightarrow S^2$  by sending any pair of invariants  $\text{PI}^o(\Lambda^\pm)$  with  $\text{sign}(\Lambda^\pm) = \pm 1$  to the northern/southern hemispheres of the 2-dimensional sphere  $S^2$ , respectively.

**(b)** For any lattice  $\Lambda \subset \mathbb{R}^2$ , the *latitude*  $\varphi(\Lambda) \in [-90^\circ, +90^\circ]$  is the angle from the equatorial plane  $\text{EP}$  of  $S^2$  to the radius-vector to the point  $\text{SLM}(\text{PI}^o(\Lambda)) \in S^2$  in the upwards direction. Let  $v(\Lambda)$  be the orthogonal projection of this radius-vector to  $\text{EP}$ . Define the *Greenwich point* as  $G = (0, \sqrt{2} - 1) \in \partial\text{QT}$  in the line through  $P^+$  and  $(1, 0)$ . This  $G$  represents all centred rectangular lattices with a conventional unit cell  $2a \times 2b$  whose ratio  $r = \frac{b}{a}$  can be found from Example 8.3.2(b):  $\sqrt{2} - 1 = \frac{3\sqrt{b^2 - a^2}}{2a\sqrt{2} + \sqrt{b^2 - a^2}}$ . Setting  $s = \sqrt{r^2 - 1}$ , we get  $\sqrt{2} - 1 = \frac{3s}{2\sqrt{2} + s}$ ,  $s = \frac{4 - 2\sqrt{2}}{4 - \sqrt{2}}$ ,  $r = \sqrt{s^2 + 1} \approx 1.1$ . The *Greenwich meridian* is the great circle on  $S^2$  passing through the point  $\text{SLM}(G)$  in the equator  $E$ . The longitude  $\mu(\Lambda) \in (-180^\circ, 180^\circ]$  is the anticlockwise angle from the *Greenwich plane* through the Greenwich meridian to the vector  $v(\Lambda)$  above. ▲

For lattices with  $\text{PI}(\Lambda)$  in the straight-line segment between the excluded vertex  $(1, 0)$  and the incentre  $P^+$ , we choose the longitude  $\mu = +180^\circ$  rather than  $-180^\circ$ . Proposition 8.5.6 computes the longitude and latitude coordinates  $\mu(\Lambda), \varphi(\Lambda)$  via  $\text{PI}(\Lambda) = (x, y)$  in terms of the projected invariant  $\text{PI}(\Lambda) = (x, y)$ .

**Proposition 8.5.6** (formulae for spherical lattice map SLM, [6, Proposition 5.2]). For any lattice  $\Lambda \subset \mathbb{R}^2$  with  $\text{PI}(\Lambda) = (x, y) \in \text{QT}$ , if  $x \neq t = 1 - \frac{1}{\sqrt{2}}$ , then set  $\psi = \arctan \frac{y - t}{x - t}$ , otherwise  $\psi = \text{sign}(y - t)90^\circ$ .

(a) The longitude of the lattice  $\Lambda$  is  $\mu(\Lambda) = \begin{cases} \psi + 22.5^\circ & \text{if } x < t, \\ \psi - 157.5^\circ & \text{if } x \geq t, \psi \geq -22.5^\circ, \\ \psi + 202.5^\circ & \text{if } x \geq t, \psi \leq -22.5^\circ. \end{cases}$

(b) The latitude is  $\varphi(\Lambda) = \text{sign}(\Lambda) \cdot \begin{cases} \frac{x\sqrt{2}}{\sqrt{2}-1}90^\circ & \text{if } \mu(\Lambda) \in [-45^\circ, +67.5^\circ], \\ \frac{y\sqrt{2}}{\sqrt{2}-1}90^\circ & \text{if } \mu(\Lambda) \in [+67.5^\circ, +180^\circ], \\ \frac{1-x-y}{\sqrt{2}-1}90^\circ & \text{if } \mu(\Lambda) \in [-180^\circ, -45^\circ]. \end{cases}$

The incentres  $P^\pm \in \text{QT}^\pm$  have  $\psi = 0$  and  $\varphi = \pm 90^\circ$ , respectively,  $\mu$  is undefined. ■

**Example 8.5.7** (prominent lattices). Any mirror-symmetric lattice  $\Lambda \subset \mathbb{R}^2$  has  $\text{sign}(\Lambda) = 0$ , hence belongs to the equator  $E$  of  $S^2$  and has  $\varphi(\Lambda) = 0$  by (8.5.6b). Any square lattice  $\Lambda_4$  with  $\text{PI}(\Lambda_4) = (0, 0)$  has  $\mu(\Lambda_4) = \arctan 1 + 22.5^\circ = 67.5^\circ$  by (8.5.6a). Any hexagonal lattice  $\Lambda_6$  with  $\text{PI}(\Lambda_6) = (0, 1)$  has  $\mu(\Lambda_6) = \arctan \frac{1}{\sqrt{3}} + 22.5^\circ = -45^\circ$ . Any rectangular lattice  $\Lambda$  with  $\text{PI}(\Lambda) = (1 - \frac{1}{\sqrt{2}}, 0)$  has  $\mu(\Lambda) = -90^\circ + 202.5^\circ = 112.5^\circ$ . Any centered rectangular lattice  $\Lambda$  with  $\text{PI}(\Lambda) = (\frac{1}{2}, \frac{1}{2})$  at the mid-point of the diagonal of  $\text{QT}$  has  $\mu(\Lambda) = \arctan 1 - 157.5^\circ = -112.5^\circ$ . Any *Greenwich* lattice  $\Lambda_G$  with  $\text{PI}(\Lambda_G) = G = (0, \sqrt{2} - 1)$  has  $\mu(\Lambda_G) = \arctan(1 - \sqrt{2}) + 22.5^\circ = 0$ . ◆

In addition to the original paper [17], we add new Corollary 8.5.8 to fulfill the realisability and Euclidean embeddability conditions in Problem 8.2.2(f,g).

**Corollary 8.5.8** (Euclidean embeddings of lattice spaces). (a) For all lattices  $\Lambda \subset \mathbb{R}^2$  under isometry, the Root Invariant Space  $\text{RIS}(\mathbb{R}^2) = \{\text{RI}(\Lambda) \mid \text{lattices } \Lambda \subset \mathbb{R}^2\}$  consists of all ordered triples  $0 \leq r_{12} \leq r_{01} \leq r_{02}$ , where the smallest root product  $r_{12}$  can be zero. Then  $\text{RIS}(\mathbb{R}^2)$  is embedded into  $\mathbb{R}^3$  as the triangular cone  $\text{TC}$  in Definition 8.4.3.

(b) For all lattices  $\Lambda \subset \mathbb{R}^2$  under homothety, the invariant space  $\{\text{PI}(\Lambda) \mid \text{lattices } \Lambda \subset \mathbb{R}^2\}$  consists of all points  $(x, y)$  in the quotient triangle  $\text{QT} \subset \mathbb{R}^2$  from Definition 8.4.4.

(c) For all lattices  $\Lambda \subset \mathbb{R}^2$  under dilation, the invariant space  $\{\text{PI}^o(\Lambda) \mid \text{lattices } \Lambda \subset \mathbb{R}^2\}$  can be embedded onto  $S^2 \setminus \{\text{one point}\} \subset \mathbb{R}^3$  by the map SLM from Definition 8.5.5.

(d) For all lattices  $\Lambda \subset \mathbb{R}^2$  under rigid motion, the space  $\{\text{PI}^o(\Lambda) \mid \text{lattices } \Lambda \subset \mathbb{R}^2\}$  can be embedded onto  $(\mathbb{S}^2 \setminus \{\text{one point}\}) \times (0, +\infty) \subset \mathbb{R}^4$ . ■

*Proof.* Parts (a,b,c) follow Definitions 8.4.3, 8.4.4, and 8.5.5. Part (d) is obtained by extending the embedding  $S^2 \subset \mathbb{R}^3$  from part (c) by a scaling factor  $s \in (0, +\infty)$ . All embeddings are bi-Lipschitz because all invariant spaces can be compactified and all underlying maps are expressed via elementary functions as in 8.5.6.  $\square$

## 8.6 Metrics on spaces of 2D lattices under all four equivalences

All lattices  $\Lambda \subset \mathbb{R}^2$  are uniquely represented under isometry and homothety by their invariants  $\text{RI} \in \text{TC}$  and  $\text{PI} \in \text{QT}$ , respectively. Then any metric  $d$  on the triangular cone  $\text{TC} \subset \mathbb{R}^3$  or the quotient triangle  $\text{QT} \subset \mathbb{R}^2$  gives rise to a metric in Definition 8.6.1 on the spaces  $\text{LIS}$  and  $\text{LHS}$ , respectively. The oriented case in Definition 8.6.3 will be harder because of identifications on the boundary  $\partial\text{TC}$ .

**Definition 8.6.1** (root metrics  $\text{RM}$ , projected metrics  $\text{PM}$ ). Any metric  $d$  on  $\mathbb{R}^3$  defines the *root metric*  $\text{RM}(\Lambda_1, \Lambda_2) = d(\text{RI}(\Lambda_1), \text{RI}(\Lambda_2))$  on lattices  $\Lambda_1, \Lambda_2 \subset \mathbb{R}^2$  under isometry. The *Root Invariant Space*  $\text{RIS} = (\text{TC}, d)$  is the triangular cone with a fixed metric  $d$ . If we use the Minkowski norm  $M_q(v) = \|\mathbf{v}\|_q = (\sum_{i=1}^n |x_i|^q)^{1/q}$  of a vector  $\mathbf{v} = (x_1, \dots, x_n) \in \mathbb{R}^n$  for any real  $q \in [1, +\infty]$ , the root metric is denoted by  $\text{RM}_q(\Lambda_1, \Lambda_2) = \|\text{RI}(\Lambda_1) - \text{RI}(\Lambda_2)\|_q$ . The limit case  $q = +\infty$  uses  $\|\mathbf{v}\|_\infty = \max_{i=1, \dots, n} |x_i|$ . The *projected metric*  $\text{PM}(\Lambda_1, \Lambda_2) = d(\text{PI}(\Lambda_1), \text{PI}(\Lambda_2))$  is on lattices under homothety for any metric  $d$  on  $\mathbb{R}^2$ . The space of *projected invariants*  $\text{PIN} = (\text{QT}, d)$  is the quotient triangle with a metric  $d$ . The notation  $\text{PM}_q(\Lambda_1, \Lambda_2) = \|\text{PI}(\Lambda_1) - \text{PI}(\Lambda_2)\|_q$  includes a parameter  $q \in [1, +\infty]$  of  $M_q$ .  $\blacktriangle$

The Minkowski distance  $M_q$  for  $q = 2$  is Euclidean. The root metric  $\text{RM}_q$  can take any large values in original units of vector coordinates such as Angstroms. The projected metric  $\text{PM}_q$  is unitless and the space  $\text{PIN} = (\text{QT}, d)$  is bounded.

**Example 8.6.2** (metrics  $\text{RM}_q, \text{PM}_q$ ). Table 8.2 summarises metric computations for the lattices  $\Lambda_4, \Lambda_6, \Lambda_0, \Lambda_\infty^\pm$ , which were inversely designed in Example 8.5.4.  $\blacklozenge$

Since the mirror images  $\Lambda_\infty^\pm$  have the same root invariant  $\text{RI}(\Lambda_\infty^\pm) = (1, 4, 7)$ , for any lattice  $\Lambda$ , the distances  $\text{RM}(\Lambda, \Lambda_\infty^\pm)$  and  $\text{PM}(\Lambda, \Lambda_\infty^\pm)$  are independent of  $\text{sign}(\Lambda_\infty^\pm) = \pm 1$ . Any mirror images  $\Lambda^\pm$  have  $\text{RM}(\Lambda^+, \Lambda^-) = 0 = \text{PM}(\Lambda^+, \Lambda^-)$  because  $\Lambda^\pm$  are isometric to each other. The metric  $\text{RM}$  from Definition 8.6.1 is well-defined only for lattices under any isometry including reflections.

Definition 8.6.3 introduces the metric  $\text{RM}^o$  on lattices under rigid motion so that  $\text{RM}^o(\Lambda^+, \Lambda^-) > 0$  on mirror images of a non-mirror-symmetric lattice, see Fig. 8.14.

**Definition 8.6.3** (orientation-aware metrics  $\text{RM}^o, \text{PM}^o$ ). For lattices  $\Lambda_1, \Lambda_2 \subset \mathbb{R}^2$  with  $\text{sign}(\Lambda_1)\text{sign}(\Lambda_2) \geq 0$ , the *orientation-aware* root metric is  $\text{RM}^o(\Lambda_1, \Lambda_2) = \text{RM}(\Lambda_1, \Lambda_2)$  as in Definition 8.6.1. If any lattices  $\Lambda_1, \Lambda_2$  have opposite signs, set  $\text{RM}^o(\Lambda_1, \Lambda_2) = \inf_{\text{sign}(\Lambda_3)=0} (\text{RM}(\Lambda_1, \Lambda_3) + \text{RM}(\Lambda_2, \Lambda_3))$ . The *orientation-based* metric  $\text{PM}^o(\Lambda_1, \Lambda_2)$  is defined by the same formula, where we replace  $\text{RM}$  by  $\text{PM}$ .  $\blacktriangle$

**Table 8.2** Metrics  $\text{RM}_q$  and  $\text{PM}_q$  for the lattices from Example 8.6.2 and shown Fig. 8.11 and 8.12.

$\text{RM}_\infty$	$\Lambda_4 \ \Lambda_6 \ \Lambda_0 \ \Lambda_\infty^\pm$	$\text{PM}_\infty$	$\Lambda_4 \ \Lambda_6 \ \Lambda_0 \ \Lambda_\infty^\pm$
$\text{RI}(\Lambda_4) = (0, 1, 1)$	0 1 3 6	$\text{PI}(\Lambda_4) = (0, 0)$	0 1 $\frac{1}{2}$ $\frac{1}{4}$
$\text{RI}(\Lambda_6) = (1, 1, 1)$	1 0 3 6	$\text{PI}(\Lambda_6) = (0, 1)$	1 0 $\frac{1}{2}$ $\frac{3}{4}$
$\text{RI}(\Lambda_0) = (1, 1, 4)$	3 3 0 3	$\text{PI}(\Lambda_0) = (\frac{1}{2}, \frac{1}{2})$	$\frac{1}{2}$ $\frac{1}{2}$ 0 $\frac{1}{4}$
$\text{RI}(\Lambda_\infty^\pm) = (1, 4, 7)$	6 6 3 0	$\text{PI}(\Lambda_\infty^\pm) = (\frac{1}{4}, \frac{1}{4})$	$\frac{1}{4}$ $\frac{3}{4}$ $\frac{1}{4}$ 0

$\text{RM}_q$ for $q \in [1, +\infty)$	$\Lambda_4$	$\Lambda_6$	$\Lambda_0$	$\Lambda_\infty^\pm$
$\text{RI}(\Lambda_4) = (0, 1, 1)$	0	1	$(1 + 3^q)^{1/q}$	$(1 + 3^q + 6^q)^{1/q}$
$\text{RI}(\Lambda_6) = (1, 1, 1)$	1	0	3	$(3^q + 6^q)^{1/q}$
$\text{RI}(\Lambda_0) = (1, 1, 4)$	$(1 + 3^q)^{1/q}$	3	0	$3 \cdot 2^{1/q}$
$\text{RI}(\Lambda_\infty^\pm) = (1, 4, 7)$	$(1 + 3^q + 6^q)^{1/q}$	$(3^q + 6^q)^{1/q}$	$3 \cdot 2^{1/q}$	0

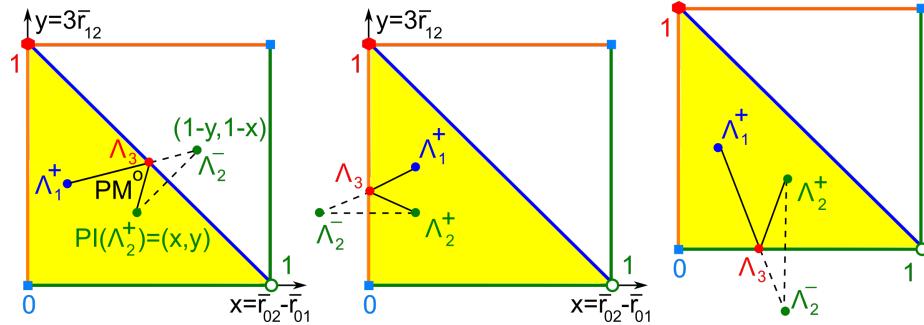
$\text{PM}_q$ for $q \in [1, +\infty)$	$\Lambda_4$	$\Lambda_6$	$\Lambda_0$	$\Lambda_\infty^\pm$
$\text{PI}(\Lambda_4) = (0, 0)$	0	1	$2^{(1/q)-1}$	$2^{(1/q)-2}$
$\text{PI}(\Lambda_6) = (0, 1)$	1	0	$2^{(1/q)-1}$	$\frac{1}{4}(1 + 3^q)^{1/q}$
$\text{PI}(\Lambda_0) = (\frac{1}{2}, \frac{1}{2})$	$2^{(1/q)-1}$	$2^{(1/q)-1}$	0	$2^{(1/q)-2}$
$\text{PI}(\Lambda_\infty^\pm) = (\frac{1}{4}, \frac{1}{4})$	$2^{(1/q)-2}$	$\frac{1}{4}(1 + 3^q)^{1/q}$	$2^{(1/q)-2}$	0

The infimum in  $\text{RM}^o(\Lambda_1, \Lambda_2)$  is the greatest lower bound defining a metric on a union of metric spaces glued by isometries. Theoretically, this bound may not be achieved over a non-compact domain. When using a Minkowski base metric  $M_q$ , Propositions 8.6.5-8.6.6 explicitly compute  $\text{RM}_q^o, \text{PM}_q^o$  for  $q = 2, +\infty$ , so the infimum in Definition 8.6.3 can be replaced by a minimum in practice.

The *oriented* root invariant space  $\text{RIS}^o$  and the space of *oriented* projected invariants  $\text{PIN}^o$  can be defined similarly to  $\text{RIS}$  and  $\text{PIN}$  in Definition 8.6.1 as the doubled cone  $\text{DC}$  and quotient square  $\text{QS}$  with any metrics from Definition 8.6.3. [17, Lemmas 5.3 and 5.5] prove the metric axioms for  $\text{RM}, \text{PM}$  and  $\text{RM}^o, \text{PM}^o$ , respectively. Lemma 8.6.4 speeds up computations in the oriented case, see Example 8.7.7.

**Lemma 8.6.4** (reversed signs, [17, Lemma 5.6]). If lattices  $\Lambda_1^\pm, \Lambda_2^\pm \subset \mathbb{R}^2$  have specified signs, then  $\text{RM}^o(\Lambda_1^+, \Lambda_2^-) = \text{RM}^o(\Lambda_1^-, \Lambda_2^+)$  and  $\text{PM}^o(\Lambda_1^+, \Lambda_2^-) = \text{PM}^o(\Lambda_1^-, \Lambda_2^+)$ . ■

If lattices  $\Lambda_1, \Lambda_2$  have non-opposite signs, so  $\text{sign}(\Lambda_1)\text{sign}(\Lambda_2) \geq 0$ , then the metrics  $\text{RM}_q^o$  and  $\text{PM}_q^o$  from Definition 8.6.3 coincide with the easily computable unoriented



**Fig. 8.14** By Definition 8.6.3, the projected metric  $PM_2(\Lambda_1^+, \Lambda_2^-)$  is the minimum sum  $PM_2(\Lambda_1^+, \Lambda_3^-) + PM_2(\Lambda_3 \Lambda_2^-)$  achieved in the left image, see computations in Proposition 8.6.5.

metrics  $RM_q, PM_q$  from Definition 8.6.1. Hence Propositions 8.6.5 and 8.6.6 compute  $RM_q^o(\Lambda_1, \Lambda_2)$  and  $PM_q^o(\Lambda_1, \Lambda_2)$  only for lattices of opposite signs.

**Proposition 8.6.5** (root metrics for  $q = 2, +\infty$ , [17, Proposition 5.8]). Let  $\Lambda_1, \Lambda_2 \subset \mathbb{R}^2$  be lattices of opposite signs with  $RI(\Lambda_1) = (r_{12}, r_{01}, r_{02}), RI(\Lambda_2) = (s_{12}, s_{01}, s_{02})$ .

(a)  $RM_2^o(\Lambda_1, \Lambda_2)$  is the minimum of the Euclidean distances from the point  $RI(\Lambda_1)$  to the three points  $(-s_{12}, s_{01}, s_{02}), (s_{01}, s_{12}, s_{02})$ , and  $(s_{12}, s_{02}, s_{01})$  in  $\mathbb{R}^3$ .

(b)  $RM_\infty^o(\Lambda_1, \Lambda_2) = \min\{d_0, d_1, d_2\}$ , where

$$d_0 = \max\{r_{12} + s_{12}, |r_{01} - s_{01}|, |r_{02} - s_{02}|\},$$

$$d_1 = \max\{MS(r_{12}, r_{01}, s_{12}, s_{01}), |r_{02} - s_{02}|\},$$

$$d_2 = \max\{|r_{12} - s_{12}|, MS(r_{01}, r_{02}, s_{01}, s_{02})\},$$

where  $MS(a, b, c, d) = \max\{|a - b|, |c - d|, \frac{1}{2}|a + b - c - d|\}$ .  $\blacksquare$

**Proposition 8.6.6** (projected metrics for  $q = 2, +\infty$ , [17, Proposition 5.9]). Let  $\Lambda_1, \Lambda_2$  be lattices with opposite signs and invariants  $PI(\Lambda_1) = (x_1, y_1), PI(\Lambda_2) = (x_2, y_2)$ .

(a)  $PM_2^o(\Lambda_1, \Lambda_2)$  is the minimum of the Euclidean distances from  $PI(\Lambda_1) = (x_1, y_1)$  to the three points  $(-x_2, y_2), (x_2, -y_2), (1 - y_2, 1 - x_2)$  in  $\mathbb{R}^2$ .

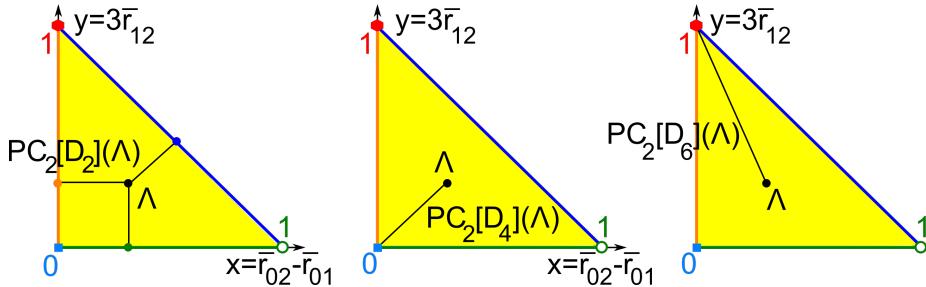
(b) For  $x_1 \leq x_2$ ,  $PM_\infty^o(\Lambda_1, \Lambda_2) = \min\{d_x, d_y, d_{xy}\}$  for  $d_x = \max\{x_2 - x_1, y_2 + y_1\}$ ,  $d_y = \max\{x_2 + x_1, |y_2 - y_1|\}$ ,  $d_{xy} = \max\{x_2 - x_1, 1 - x_2 - y_2 + |1 - y_1 - x_2|\}$ .  $\blacktriangle$

## 8.7 Real-valued chiral distances measure asymmetry of lattices

The classical concept of chirality is a binary property distinguishing mirror images of the same object such as a molecule or a periodic crystal. Continuous classifications in Theorem 8.4.2 and Corollary 8.4.5 imply that the binary chirality is discontinuous under almost any perturbations similar to other discrete invariants such as symmetry groups. To avoid arbitrary thresholds, it makes more sense to continuously quantify a deviation of a lattice from a higher-symmetry neighbour.

The term *chirality* often refers to 3-dimensional molecules or crystal lattices. One reason is the fact that in  $\mathbb{R}^2$  a reflection with respect to a line  $L$  is realised by the rotation in  $\mathbb{R}^3$  around  $L$  through  $180^\circ$ . However, if our ambient space is only  $\mathbb{R}^2$ , the concepts of isometry and rigid motion differ. For example, Lemma 8.3.3 described root invariants of all lattices that are related to their mirror images by rigid motion. Such lattices can be called *achiral*. We call them mirror-symmetric to avoid a potential confusion with the literature in crystallography. Definition 8.7.1 introduces the real-valued  $G$ -chiral distances of a lattice  $\Lambda \subset \mathbb{R}^2$ . Proposition 8.8.5 proves the continuity of these functions  $\text{RC}[G] : \text{LIS}(\mathbb{R}^2) \rightarrow \mathbb{R}$  and  $\text{PC}[G] : \text{LHS}(\mathbb{R}^2) \rightarrow \mathbb{R}$ .

Recall that the *crystallographic point group*  $G$  of a lattice  $\Lambda \subset \mathbb{R}^2$  containing the origin 0 consists of all symmetry operations that keep 0 and map  $\Lambda$  to itself. For example, any such group  $G$  includes the central symmetry with respect to  $0 \in \Lambda \subset \mathbb{R}^2$ . If  $G$  has no other non-trivial symmetries, we get  $G = C_2$  in Schonflies notations. All 2D lattices split into four crystal families by their point groups: oblique ( $C_2$ ), orthorhombic ( $D_2$ ), tetragonal or square ( $D_4$ ) and hexagonal ( $D_6$ ). Orthorhombic lattices split into rectangular and centred rectangular, see Fig. 8.10.



**Fig. 8.15** **Left:** by Definition 8.7.1, the projected  $D_2$  chiral distance  $\text{PC}_2[D_2](\Lambda)$  is the minimum Euclidean distance from  $\text{PI}(\Lambda) \in \text{QT}$  to the boundary  $\partial\text{QT}$ . **Middle:**  $\text{PC}_2[D_4](\Lambda)$  is the distance from  $\text{PI}(\Lambda)$  to  $(0, 0)$ . **Right:**  $\text{PC}_2[D_6](\Lambda)$  is the distance from  $\text{PI}(\Lambda)$  to  $(0, 1)$ .

**Definition 8.7.1** ( $G$ -chiral distances  $\text{RC}[G]$  and  $\text{PC}[G]$ ). For any crystallographic point group  $G$  in  $\mathbb{R}^2$ , let  $\text{LIS}[G] \subset \text{LIS}(\mathbb{R}^2)$  be the closure of the subspace of all (isometry classes of) lattices that have the crystallographic point group  $G$ . For  $G = D_2$  or  $G = D_4$  or  $G = D_6$ , the root and projected  $G$ -chiral distances are

$$\text{RC}[G](\Lambda) = \min_{\Lambda' \in \text{LIS}[G]} \text{RM}(\Lambda, \Lambda') \geq 0 \text{ and } \text{PC}[G](\Lambda) = \min_{\Lambda' \in \text{LIS}[G]} \text{PM}(\Lambda, \Lambda') \geq 0,$$

where  $\text{RM}$ ,  $\text{PM}$  are any metrics from Definition 8.6.1 with a base metric  $d$ . If  $d = M_q$  for  $q \in [1, +\infty]$ , denote the  $G$ -chiral distances by  $\text{RC}_q[G]$  and  $\text{PC}_q[G]$ .  $\blacktriangle$

Since any lattice  $\Lambda$  is symmetric with respect to the origin  $0 \in \Lambda$ , the closed subspace  $\text{LIS}[C_2]$  coincides with the 3-dimensional Lattice Isometry Space  $\text{LIS}(\mathbb{R}^2)$ . The 2-dimensional subspace  $\text{LIS}[D_2]$  consists of all mirror-symmetric lattices (rectangular

and centred-rectangular) represented by root invariants  $\text{RI}$  on the boundary  $\partial\text{TC}$  of the triangular cone in Definition 8.4.3, see Fig. 8.9. The 1-dimensional subspaces  $\text{LIS}[D_4], \text{LIS}[D_6] \subset \text{LIS}[D_2]$  can be viewed as the blue and orange rays  $\{r_{12} = 0 < r_{01} = r_{02}\}$  and  $\{0 < r_{12} = r_{01} = r_{02}\}$ , respectively.

The  $G$ -chiral distance  $\text{RC}[G]$  in Definition 8.7.1 measures a distance from  $\text{RI}(\Lambda)$  to the root invariant of a closest neighbour in the subspace  $\text{LIS}[G]$ . Any  $\text{RC}[G](\Lambda)$  is invariant under isometry and measures a distance from  $\Lambda$  to its nearest neighbour  $\Lambda' \in \text{LIS}[G]$ . The *signed chiral distances*  $\text{sign}(\Lambda)\text{RC}(\Lambda)$  and  $\text{sign}(\Lambda)\text{PC}(\Lambda)$  are invariant under rigid motion. Since  $\text{LIS}[G]$  is a closed subspace within  $\text{LIS}(\mathbb{R}^2)$ , the continuous distances  $\text{RM}, \text{PM}$  achieve their minima if their base distances  $d$  are continuous. If  $\text{LIS}[D_2]$  was defined as an open subspace of only lattices that have the point group  $D_2$  (not  $D_4$  or  $D_6$ ), then  $\text{RC}[G], \text{PC}[G]$  should be defined via infima instead of simpler minima. Indeed, any square or hexagonal lattice  $\Lambda$  can be approximated by infinitely many closer and closer orthorhombic lattices  $\Lambda'$ , but the expected distance  $\text{RM}(\Lambda, \Lambda') = 0$  will not be achieved on an open set.

For  $q = 2, +\infty$ , the distances  $\text{RC}_q, \text{PC}_q$  are computed in Propositions 8.7.4, 8.7.5.

**Lemma 8.7.2** (properties of chiral distances, [17, Lemma 6.2]). **(a)** A lattice  $\Lambda \subset \mathbb{R}^2$  is mirror-symmetric if and only if  $\text{RC}[D_2](\Lambda) = 0$  or, equivalently,  $\text{PC}[D_2](\Lambda) = 0$ .

**(b)** For any crystallographic point group  $G$  in  $\mathbb{R}^2$ , mirror reflections  $\Lambda^\pm \subset \mathbb{R}^2$  have equal  $G$ -chiral distances:  $\text{RC}[G](\Lambda^+) = \text{RC}[G](\Lambda^-), \text{PC}[G](\Lambda^+) = \text{PC}[G](\Lambda^-)$ . ■

**Lemma 8.7.3** (lower bounds, [17, Lemma 6.3]). **(a)** If lattices  $\Lambda_1, \Lambda_2$  have opposite signs, then  $\text{RM}^o(\Lambda_1, \Lambda_2) \geq \text{RC}[D_2](\Lambda_1) + \text{RC}[D_2](\Lambda_2)$  and  $\text{PM}^o(\Lambda_1, \Lambda_2) \geq \text{PC}[D_2](\Lambda_1) + \text{PC}[D_2](\Lambda_2)$ .

**(b)** For the mirror images  $\Lambda^\pm$  of any lattice  $\Lambda$ , the lower bounds in part (a) become equalities:  $\text{RM}^o(\Lambda^+, \Lambda^-) = 2\text{RC}[D_2](\Lambda)$  and  $\text{PM}^o(\Lambda^+, \Lambda^-) = 2\text{PC}[D_2](\Lambda)$ . ■

**Proposition 8.7.4** (chiral distances  $\text{RC}_q[G]$  for  $q = 2, +\infty$ , [17, Proposition 6.5]). Let a lattice  $\Lambda \subset \mathbb{R}^2$  have a root invariant  $\text{RI}(\Lambda) = (r_{12}, r_{01}, r_{02})$  with  $0 \leq r_{12} \leq r_{01} \leq r_{02}$ .

$$\begin{aligned} \text{(a)} \quad \text{RC}_2[D_2](\Lambda) &= \min \left\{ r_{12}, \frac{r_{01} - r_{12}}{\sqrt{2}}, \frac{r_{02} - r_{01}}{\sqrt{2}} \right\}; \\ \text{RC}_2[D_4](\Lambda) &= \sqrt{r_{12}^2 + \frac{1}{4}(r_{02} - r_{01})^2}; \\ \text{RC}_2[D_6](\Lambda) &= \sqrt{\frac{2}{3}(r_{12}^2 + r_{01}^2 + r_{02}^2 - r_{12}r_{01} - r_{12}r_{02} - r_{01}r_{02})}; \end{aligned}$$

$$\begin{aligned} \text{(b)} \quad \text{RC}_\infty[D_2](\Lambda) &= \min \left\{ r_{12}, \frac{r_{01} - r_{12}}{2}, \frac{r_{02} - r_{01}}{2} \right\}. \\ \text{RC}_\infty[D_4](\Lambda) &= \min \{r_{12}, \frac{r_{02} - r_{01}}{2}\}; \\ \text{RC}_\infty[D_6](\Lambda) &= \frac{r_{02} - r_{12}}{2}. \end{aligned}$$

▲

When considering lattices under homothety, the subspace  $\text{LHS}[D_4]$  consists of a single class of all square lattices, which are all equivalent under isometry and uniform

scaling. The subspace  $\text{LHS}[D_6]$  is also a single point representing all hexagonal lattices. Then  $\text{PC}[D_4]$  and  $\text{PC}[D_6]$  are distances to these single points.

**Proposition 8.7.5** (chiral distances  $\text{PC}_q$  for  $q = 2, +\infty$ , [17, Proposition 6.6]). Let a lattice  $\Lambda$  have  $\text{PI}(\Lambda) = (x, y) \in \text{QT}$  so that  $x \in [0, 1], y \in [0, 1], x + y \leq 1$ .

$$(a) \text{PC}_2[D_2](\Lambda) = \min \left\{ x, y, \frac{1-x-y}{\sqrt{2}} \right\},$$

$$\text{PC}_q[D_4](\Lambda) = (x^q + y^q)^{1/q} \text{ for any } q \in [1, +\infty),$$

$$\text{PC}_q[D_6](\Lambda) = (x^q + (1-y)^q)^{1/q} \text{ for any } q \in [1, +\infty);$$

$$(b) \text{PC}_\infty[D_2](\Lambda) = \min \left\{ x, y, \frac{1-x-y}{2} \right\},$$

$$\text{PC}_\infty[D_4](\Lambda) = x,$$

$$\text{PC}_\infty[D_6](\Lambda) = 1 - y.$$

(c) The upper bounds  $\text{PC}_2[D_2](\Lambda) \leq \frac{1}{2+\sqrt{2}}$ ,  $\text{PC}_\infty[D_2](\Lambda) \leq \frac{1}{4}$  hold for any  $\Lambda$ , achieved for lattices with  $\text{PI}(\Lambda_2) = (\frac{1}{2+\sqrt{2}}, \frac{1}{2+\sqrt{2}})$ ,  $\text{PI}(\Lambda_\infty) = (\frac{1}{4}, \frac{1}{4})$ , respectively. For  $q \in [1, +\infty]$ , the bound  $\text{PC}_q[D_4](\Lambda) \leq 1$  holds for any  $\Lambda$  and is achieved for any hexagonal lattice. For  $q \in [1, +\infty)$ , the upper bound  $\text{PC}_q[D_6](\Lambda) < 2^{1/q}$  holds for any  $\Lambda$  and is approached but not achieved as  $x \rightarrow 1$ . The bound  $\text{PC}_\infty[D_6](\Lambda) \leq 1$  holds for any  $\Lambda$  and is achieved for any square and rectangular lattice. ■

**Example 8.7.6** (distances  $\text{RC}_q, \text{PC}_q$ ). Table 8.3 shows the chiral distances computed by Propositions 8.7.4, 8.7.5 for the prominent lattices  $\Lambda_2^\pm, \Lambda_\infty^\pm$  in Example 8.5.4. ◆

**Example 8.7.7** (metrics  $\text{RM}_q^o, \text{PM}_q^o$ ). Table 8.4 has  $\text{RM}_q^o, \text{RM}_q^o$  for  $q = 2, +\infty$  and the prominent lattices  $\Lambda_2^\pm, \Lambda_\infty^\pm$ , which were inversely designed in Example 8.5.4.

If lattices have the same sign, then  $\text{RM}_q^o, \text{PM}_q^o$  coincide with their unoriented versions by Definition 8.6.3. For example,  $\text{PM}_q^o(\Lambda_2^+, \Lambda_\infty^+)$  is the distance  $M_q$  between the invariants  $\text{PI}(\Lambda_\infty) = (\frac{1}{4}, \frac{1}{4})$  and  $\text{PI}(\Lambda_2) = (\frac{1}{2+\sqrt{2}}, \frac{1}{2+\sqrt{2}}) = (1 - \frac{1}{\sqrt{2}}, 1 - \frac{1}{\sqrt{2}})$ , so  $\text{PM}_\infty^o(\Lambda_2^+, \Lambda_\infty^+) = \frac{3}{4} - \frac{1}{\sqrt{2}} \approx 0.04$  and  $\text{PM}_2^o(\Lambda_2^+, \Lambda_\infty^+) = \frac{3}{4}\sqrt{2} - 1 \approx 0.06$ .

Similarly,  $\text{RM}_q^o(\Lambda_2^+, \Lambda_\infty^+)$  is the  $M_q$  distance between the root invariants  $\text{PI}(\Lambda_\infty) = (1, 4, 7)$  and  $\text{RI}(\Lambda_2) = (2 - \sqrt{2}, 2\sqrt{2} - 1, 5 - \sqrt{2})$ , so  $\text{RM}_\infty^o(\Lambda_2^+, \Lambda_\infty^+) = \max\{\sqrt{2} - 1, 5 - 2\sqrt{2}, 2 + \sqrt{2}\} = 2 + \sqrt{2} \approx 3.41$  and  $\text{RM}_2^o(\Lambda_2^+, \Lambda_\infty^+) = \sqrt{6(7 - 3\sqrt{2})} \approx 4.1$ .

By Lemma 8.7.3(b) the distance between mirror images of the same lattice equals the doubled  $D_2$ -chiral distance. For example,  $\text{PM}_q^o(\Lambda_\infty^+, \Lambda_\infty^-) = 2\text{PC}_q[D_2](\Lambda_\infty) = \frac{1}{2}$  and  $\text{PM}_q^o(\Lambda_2^+, \Lambda_2^-) = 2\text{PC}_q[D_2](\Lambda_2) = \frac{2}{2+\sqrt{2}} = 2 - \sqrt{2} \approx 0.59$  for  $q = 2, +\infty$ .

**Table 8.3** Chiral distances  $PC_q, RC_q$  for the lattices  $\Lambda_2^\pm, \Lambda_\infty^\pm$  in Fig. 8.11 and 8.12, see Example 8.7.6.

$\Lambda$	$\Lambda_\infty$	$\Lambda_2$
$PI(\Lambda)$	$\left(\frac{1}{4}, \frac{1}{4}\right)$	$\left(\frac{1}{2+\sqrt{2}}, \frac{1}{2+\sqrt{2}}\right)$
$PC_2[D_2]$	$\frac{1}{4}$	$\frac{1}{2+\sqrt{2}}$
$PC_2[D_4]$	$\frac{\sqrt{2}}{4}$	$\sqrt{2}-1$
$PC_2[D_6]$	$\frac{\sqrt{10}}{4}$	$\sqrt{2}-\sqrt{2}$
$PC_\infty[D_2]$	$\frac{1}{4}$	$\frac{1}{2+\sqrt{2}}$
$PC_\infty[D_4]$	$\frac{1}{4}$	$\frac{1}{2+\sqrt{2}}$
$PC_\infty[D_6]$	$\frac{3}{4}$	$\frac{1}{\sqrt{2}}$

$\Lambda$	$\Lambda_\infty$	$\Lambda_2$
$RI(\Lambda)$	$(1, 4, 7)$	$(2-\sqrt{2}, 2\sqrt{2}-1, 5-\sqrt{2})$
$RC_2[D_2]$	1	$2-\sqrt{2}$
$RC_2[D_4]$	$\frac{\sqrt{13}}{2}$	$(2-\sqrt{2})\frac{\sqrt{13}}{2}$
$RC_2[D_6]$	$3\sqrt{2}$	$\sqrt{2(13-3\sqrt{2})}$
$RC_\infty[D_2]$	1	$2-\sqrt{2}$
$RC_\infty[D_4]$	1	$2-\sqrt{2}$
$RC_\infty[D_6]$	3	$\frac{3}{2}$

Lemma 8.7.3(b) and Table 8.3 also give  $RM_q^o(\Lambda_\infty^+, \Lambda_\infty^-) = 2RC_q[D_2](\Lambda_\infty) = 2$  and  $RM_q^o(\Lambda_2^+, \Lambda_2^-) = 2RC_q[D_2](\Lambda_2) = 2(2-\sqrt{2}) \approx 1.17$  for  $q = 2, +\infty$ .

Lemma 8.6.4 says that  $RM^o(\Lambda_2^+, \Lambda_\infty^-) = RM^o(\Lambda_2^-, \Lambda_\infty^+) = PM^o(\Lambda_2^+, \Lambda_\infty^-) = PM^o(\Lambda_2^-, \Lambda_\infty^+)$ . Using the above properties, it remains to find four distances.

Proposition 8.6.6(a) finds  $PM_2^o(\Lambda_2^+, \Lambda_\infty^-)$  as the minimum of the Euclidean distances from  $PI(\Lambda_2) = (\frac{1}{2+\sqrt{2}}, \frac{1}{2+\sqrt{2}}) = (1 - \frac{1}{\sqrt{2}}, 1 - \frac{1}{\sqrt{2}})$  to the three points  $(-\frac{1}{4}, \frac{1}{4}), (-\frac{1}{4}, \frac{1}{4}), (\frac{3}{4}, \frac{3}{4})$  obtained from  $PI(\Lambda_\infty) = (\frac{1}{4}, \frac{1}{4})$  by reflections in the edges of QT. The first two distances equal to  $\frac{\sqrt{25-16\sqrt{2}}}{2\sqrt{2}} \approx 0.54$  are larger than the third.

Given  $PI(\Lambda_2) = (x_1, y_1) = (1 - \frac{1}{\sqrt{2}}, 1 - \frac{1}{\sqrt{2}})$  and  $PI(\Lambda_\infty) = (x_2, y_2) = (\frac{1}{4}, \frac{1}{4})$ , Proposition 8.6.6(b) computes  $PM_\infty^o(\Lambda_2^+, \Lambda_\infty^-)$  for as the minimum of  $d_x = \max\{x_2 - x_1, y_2 + y_1\} = \frac{5}{4} - \frac{1}{\sqrt{2}}, d_y = \max\{x_2 + x_1, |y_2 - y_1|\} = \frac{5}{4} - \frac{1}{\sqrt{2}}, d_{xy} = \max\{x_2 - x_1, 1 - x_2 - y_2 + |1 - y_1 - x_2|\} = \frac{1}{4} + \frac{1}{\sqrt{2}}$ , so  $PM_\infty^o(\Lambda_2^+, \Lambda_\infty^-) = \frac{5}{4} - \frac{1}{\sqrt{2}} \approx 0.54$

Proposition 8.6.5(a) computes  $RM_2^o(\Lambda_2^+, \Lambda_\infty^-)$  as the minimum of the Euclidean distances from  $RI(\Lambda_2) = (2-\sqrt{2}, 2\sqrt{2}-1, 5-\sqrt{2})$  to the three points  $(-1, 4, 7), (4, 1, 7), (1, 7, 4)$  obtained from  $RI(\Lambda_\infty) = (1, 4, 7)$  by reflections in the boundaries of TC. The first distance is the smallest, so  $RM_2^o(\Lambda_2^+, \Lambda_\infty^-) = \sqrt{50-22\sqrt{2}} \approx 4.3$ .

Given  $RI(\Lambda_2) = (r_{12}, r_{01}, r_{02}) = (2-\sqrt{2}, 2\sqrt{2}-1, 5-\sqrt{2})$  and  $RI(\Lambda_\infty) = (s_{12}, s_{01}, s_{02}) = (1, 4, 7)$ , by Proposition 8.6.5(b)  $RM_\infty^o(\Lambda_2^+, \Lambda_\infty^-) = \min\{d_0, d_1, d_2\}$ . Using  $MS(a, b, c, d) = \max\{|a-b|, |c-d|, \frac{1}{2}|a+b-c-d|\}$ , we compute

$$\begin{aligned} d_0 &= \max\{r_{12} + s_{12}, |r_{01} - s_{01}|, |r_{02} - s_{02}|\} \\ &= \max\{3 - \sqrt{2}, 5 - 2\sqrt{2}, 2 + \sqrt{2}\} = 2 + \sqrt{2} \approx 3.4, \end{aligned}$$

**Table 8.4** Metrics  $\text{PM}_q^o$  and  $\text{RM}_q^o$  for the lattices given by their invariants in Table 8.3, see Fig. 8.11.

$\text{PM}_2^o$	$\Lambda_\infty^+$	$\Lambda_\infty^-$	$\Lambda_2^+$	$\Lambda_2^-$
$\Lambda_\infty^+$	0	$\frac{1}{2}$	$\frac{3}{4}\sqrt{2} - 1 \approx 0.06$	$\frac{\sqrt{25-16\sqrt{2}}}{2\sqrt{2}} \approx 0.54$
$\Lambda_\infty^-$	$\frac{1}{2}$	0	$\frac{\sqrt{25-16\sqrt{2}}}{2\sqrt{2}} \approx 0.54$	$\frac{3}{4}\sqrt{2} - 1 \approx 0.06$
$\Lambda_2^+$	$\frac{3}{4}\sqrt{2} - 1 \approx 0.06$	$\frac{\sqrt{25-16\sqrt{2}}}{2\sqrt{2}} \approx 0.54$	0	$2 - \sqrt{2} \approx 0.59$
$\Lambda_2^-$	$\frac{\sqrt{25-16\sqrt{2}}}{2\sqrt{2}} \approx 0.54$	$\frac{3}{4}\sqrt{2} - 1 \approx 0.06$	$2 - \sqrt{2} \approx 0.59$	0

$\text{PM}_\infty^o$	$\Lambda_\infty^+$	$\Lambda_\infty^-$	$\Lambda_2^+$	$\Lambda_2^-$
$\Lambda_\infty^+$	0	$\frac{1}{2}$	$\frac{3}{4} - \frac{1}{\sqrt{2}} \approx 0.04$	$\frac{5}{4} - \frac{1}{\sqrt{2}} \approx 0.54$
$\Lambda_\infty^-$	$\frac{1}{2}$	0	$\frac{5}{4} - \frac{1}{\sqrt{2}} \approx 0.54$	$\frac{3}{4} - \frac{1}{\sqrt{2}} \approx 0.04$
$\Lambda_2^+$	$\frac{3}{4} - \frac{1}{\sqrt{2}} \approx 0.04$	$\frac{5}{4} - \frac{1}{\sqrt{2}} \approx 0.54$	0	$2 - \sqrt{2} \approx 0.59$
$\Lambda_2^-$	$\frac{5}{4} - \frac{1}{\sqrt{2}} \approx 0.54$	$\frac{3}{4} - \frac{1}{\sqrt{2}} \approx 0.04$	$2 - \sqrt{2} \approx 0.59$	0

$\text{RM}_2^o$	$\Lambda_\infty^+$	$\Lambda_\infty^-$	$\Lambda_2^+$	$\Lambda_2^-$
$\Lambda_\infty^+$	0	2	$\sqrt{6(7-3\sqrt{2})} \approx 4.1$	$\sqrt{50-22\sqrt{2}}$
$\Lambda_\infty^-$	2	0	$\sqrt{50-22\sqrt{2}} \approx 4.3$	$\sqrt{6(7-3\sqrt{2})}$
$\Lambda_2^+$	$\sqrt{6(7-3\sqrt{2})}$	$\sqrt{50-22\sqrt{2}} \approx 4.3$	0	$2(2-\sqrt{2})$
$\Lambda_2^-$	$\sqrt{50-22\sqrt{2}}$	$\sqrt{6(7-3\sqrt{2})} \approx 4.1$	$2(2-\sqrt{2}) \approx 1.17$	0

$\text{RM}_\infty^o$	$\Lambda_\infty^+$	$\Lambda_\infty^-$	$\Lambda_2^+$	$\Lambda_2^-$
$\Lambda_\infty^+$	0	2	$2 + \sqrt{2} \approx 3.41$	3
$\Lambda_\infty^-$	2	0	3	$2 + \sqrt{2} \approx 3.41$
$\Lambda_2^+$	$2 + \sqrt{2} \approx 3.41$	3	0	$2(2-\sqrt{2}) \approx 1.17$
$\Lambda_2^-$	3	$2 + \sqrt{2} \approx 3.41$	$2(2-\sqrt{2}) \approx 1.17$	0

$$\begin{aligned}
d_1 &= \max\{\text{MS}(r_{12}, r_{01}, s_{12}, s_{01}), |r_{02} - s_{02}|\} = \\
&= \max\{\text{MS}(r_{12}, r_{01}, s_{12}, s_{01}), 2 + \sqrt{2}\} = \\
&= \max\{\text{MS}(2 - \sqrt{2}, 2\sqrt{2} - 1, 1, 4), 2 + \sqrt{2}\} = \\
&= \max\{\max\{3(\sqrt{2} - 1), 3, 2 - \frac{1}{\sqrt{2}}\}, 2 + \sqrt{2}\} = \max\{3, 2 + \sqrt{2}\} = 2 + \sqrt{2},
\end{aligned}$$

$$\begin{aligned}
d_2 &= \max\{|r_{12} - s_{12}|, \text{MS}(r_{01}, r_{02}, s_{01}, s_{02})\} = \\
&= \max\{\sqrt{2} - 1, \text{MS}(2\sqrt{2} - 1, 5 - \sqrt{2}, 4, 7)\} = \\
&= \max\{\sqrt{2} - 1, \max\{6 - 3\sqrt{2}, 3, \frac{7 - \sqrt{2}}{2}\}\} = 3, \text{ hence } \text{RM}_{\infty}^o(\Lambda_2^+, \Lambda_{\infty}^-) = 3. \quad \blacklozenge
\end{aligned}$$

## 8.8 Bi-continuity of the map from obtuse superbases to invariants

This section studies continuity of the bijection  $B \mapsto \Lambda(B)$ , where an obtuse superbase  $B$  and its lattice  $\Lambda(B)$  are considered under isometry, rigid motion, dilation, and homothety. Theorems 8.3.6 and 8.4.2 established the bijections  $\text{LIS} \rightarrow \text{SBI} \rightarrow \text{RIS}$ ,  $\Lambda \mapsto B \mapsto \text{RI}(B) = \text{RI}(\Lambda)$ , mapping any lattice  $\Lambda \subset \mathbb{R}^2$  to its obtuse superbase  $B$  (unique under isometry) and then to the complete invariant  $\text{RI}(\Lambda)$ . Hence, the Lattice Isometry Space  $\text{LIS}(\mathbb{R}^2)$  with a root metric  $\text{RM}$  can be identified with the Root Invariant Space  $\text{RIS} = (\text{TC}, d)$  with a metric  $d$  on the triangular cone  $\text{TC} \subset \mathbb{R}^3$ .

**Theorem 8.8.1** (continuity of  $\text{SBI} \rightarrow \text{LIS} = \text{RIS}$ , [17, Theorem 7.5]). **(a)** Let  $q \in [1, +\infty]$  and lattices  $\Lambda, \Lambda' \subset \mathbb{R}^2$  have obtuse superbases  $B$  and  $B'$  whose vectors have a maximum length  $l$ . If  $\text{SIM}_{\infty}(B, B') = \delta \geq 0$ , then  $\text{RM}_q(\Lambda, \Lambda') \leq 3^{1/q} \sqrt{2l\delta}$ . Hence the bijection  $\text{SBI}(\mathbb{R}^2) \rightarrow \text{LIS}(\mathbb{R}^2)$  is continuous in the metrics  $\text{SIM}_{\infty}$  and  $\text{RM}_q$ .

**(b)** In the conditions above, the projected metric satisfies  $\text{PM}_q(\Lambda, \Lambda') \leq 2^{1/q} 3 \sqrt{2\delta/l}$ , so the bijection  $\text{SBH}(\mathbb{R}^2) \rightarrow \text{LHS}(\mathbb{R}^2)$  is continuous in the metrics  $\text{SHM}_{\infty}, \text{PM}_q$ .

**(c)** In the oriented case, if  $\delta \rightarrow 0$ , then  $\text{RM}_q^o(\Lambda, \Lambda') \rightarrow 0$  and  $\text{PM}_q^o(\Lambda, \Lambda') \rightarrow 0$ , so both maps  $\text{SBR}(\mathbb{R}^2) \rightarrow \text{LRS}(\mathbb{R}^2)$  and  $\text{SBD}(\mathbb{R}^2) \rightarrow \text{LDS}(\mathbb{R}^2)$  are continuous. ■

Theorem 8.8.1 is proved for the metrics  $\text{RM}_q, \text{PM}_q$  only to give explicit upper bounds. A similar argument proves continuity for any metrics  $\text{RM}, \text{PM}$  in Definition 8.6.1 based on a metric  $d$  satisfying  $d(u, v) \rightarrow 0$  when  $u \rightarrow v$  coordinate-wise. All Minkowski norms in  $\mathbb{R}^n$  are topologically equivalent [1] due to the bounds  $\|v\|_q \leq \|v\|_r \leq n^{\frac{1}{q} - \frac{1}{r}} \|v\|_q$  for any  $1 \leq q \leq r$ , hence continuity for one value of  $q$  is enough. Theorem 8.8.1 implies continuity of  $\text{SBR} \rightarrow \text{RIS}^o$ , because closeness of superbases under rigid motion is a stronger condition than under isometry.

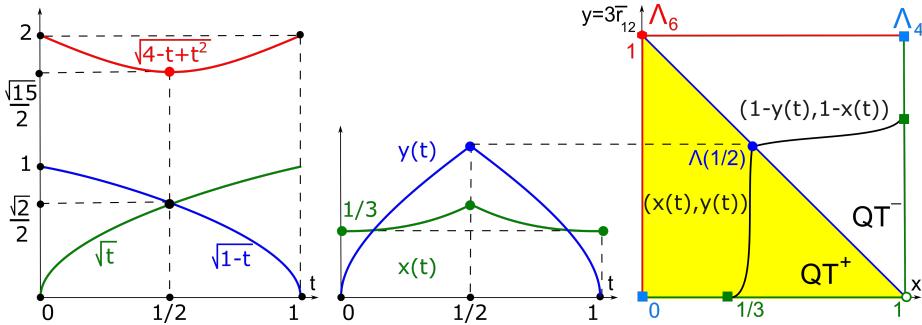
Example 8.8.2 illustrates Theorem 8.8.1 and shows that the root invariant changes continuously for a deformation when a reduced basis changes discontinuously.

**Example 8.8.2** (continuity of root invariants under deformation). The obtuse superbase  $\mathbf{v}_1 = (1, 0), \mathbf{v}_2(t) = (-t, 2), \mathbf{v}_0(t) = (t - 1, -2)$  continuously deforms for  $t \in [0, 1]$  in Fig. 8.3. The basis of  $\mathbf{v}_1, \mathbf{v}_2(t)$  is reduced (non-acute) for  $t \in [0, \frac{1}{2}]$  and at the critical moment  $t = \frac{1}{2}$  changes to its mirror image  $\mathbf{v}_1, \mathbf{v}_0(t)$ , which remains reduced for  $t \in [\frac{1}{2}, 1]$ . The obtuse superbase of unordered vectors  $\{\mathbf{v}_1, \mathbf{v}_2(t), \mathbf{v}_0(t)\}$  keeps changing continuously because  $\mathbf{v}_2(t), \mathbf{v}_0(t)$  only swap their places at  $t = \frac{1}{2}$ .

The discontinuity of the obtuse superbases in the above deformation emerges at  $t = 1$  when the final superbase of  $(1, 0), (-1, 2), (0, -2)$  becomes a mirror image of the initial superbase of  $(1, 0), (0, 2), (-1, -2)$ , not related by rigid motion, though

both (unordered) superbases at  $t = 0, 1$  generate the same lattice with the rectangular cell  $1 \times 2$ . The root invariants are  $r_{12} = \sqrt{t}$ ,  $r_{01} = \sqrt{1-t}$ ,  $r_{02} = \sqrt{4-t+t^2}$ . Since  $4-t+t^2 \geq \frac{15}{4} \geq \max\{t, 1-t\}$  for  $t \in [0, 1]$ , the root invariant can be written as

$$\text{RI}(\Lambda(t)) = \begin{cases} (\sqrt{t}, \sqrt{1-t}, \sqrt{4-t+t^2}) & \text{for } t \in [0, \frac{1}{2}], \\ (\sqrt{1-t}, \sqrt{t}, \sqrt{4-t+t^2}) & \text{for } t \in [\frac{1}{2}, 1]. \end{cases}$$



**Fig. 8.16** **Left:** graphs of root products in  $\text{RI}(\Lambda(t))$ , see Example 8.8.2. **Middle:** graphs of the components in  $\text{PI}(\Lambda(t))$ . **Right:** the continuous path of  $\text{PI}(\Lambda(t))$  in the quotient square  $QS$ .

By Definition 8.4.4 the size is  $|\Lambda(t)| = r_{12} + r_{01} + r_{02} = \sqrt{t} + \sqrt{1-t} + \sqrt{4-t+t^2}$ . The projected invariant is  $\text{PI}(\Lambda(t)) = (x(t), y(t))$ , see Fig. 8.16, where

$$x(t) = \frac{\sqrt{4-t+t^2} - \max\{\sqrt{t}, \sqrt{1-t}\}}{\sqrt{t} + \sqrt{1-t} + \sqrt{4-t+t^2}}, \quad y(t) = \frac{3 \min\{\sqrt{t}, \sqrt{1-t}\}}{\sqrt{t} + \sqrt{1-t} + \sqrt{4-t+t^2}}.$$

If  $t = 0$  or  $t = 1$ , then  $\text{RI}(\Lambda(t)) = (0, 1, 2)$ ,  $|\Lambda(t)| = 3$ ,  $\text{PI}(\Lambda(t)) = (\frac{1}{3}, 0)$ . If  $t = \frac{1}{2}$ , then  $\sqrt{t} = \sqrt{1-t} = \frac{\sqrt{2}}{2}$ ,  $\sqrt{4-t+t^2} = \frac{\sqrt{15}}{2}$ ,  $|\Lambda(\frac{1}{2})| = \sqrt{2} + \frac{\sqrt{15}}{2}$ . So

$$\text{RI}\left(\Lambda\left(\frac{1}{2}\right)\right) = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, \frac{\sqrt{15}}{2}\right), \quad \text{PI}\left(\Lambda\left(\frac{1}{2}\right)\right) = \left(\frac{\sqrt{15} - \sqrt{2}}{\sqrt{15} + 2\sqrt{2}}, \frac{3\sqrt{2}}{\sqrt{15} + 2\sqrt{2}}\right).$$

The last point is approximately  $(0.37, 0.63)$  in the diagonal  $x + y = 1$  of  $QS$ . Under the symmetry  $t \leftrightarrow 1-t$ , all the functions above remain invariant and  $\Lambda(t)$  changes its sign. Both paths  $\text{RI}(\Lambda(t))$  and  $\text{PI}^o(\Lambda(t)) \in QS$  are continuous everywhere, while the obtuse superbasis is discontinuous (under rigid motion) at  $t = 0, 1$ .  $\blacklozenge$

Theorem 8.8.3 below proves the inverse continuity of  $\text{RIS} \rightarrow \text{SBI}$  and a weaker claim in the oriented case, saying that we can choose an obtuse superbasis  $B'$  of a perturbed lattice arbitrarily close to a given superbasis  $B$  of an original lattice.

**Theorem 8.8.3** (continuity of LIS  $\rightarrow$  SBI, [17, Theorem 7.7]). **(a)** For  $q \in [1, +\infty]$ , let lattices  $\Lambda, \Lambda'$  in  $\mathbb{R}^2$  satisfy  $\text{RM}_q(\Lambda, \Lambda') \leq \delta$ . For any obtuse superbase  $B$  of  $\Lambda$ , there is an obtuse superbase  $B'$  of  $\Lambda'$  such that  $\text{SIM}_\infty(B, B') \leq \text{SRM}_\infty(B, B') \rightarrow 0$  as  $\delta \rightarrow 0$ .

**(b)** The bijection  $\text{LIS}(\mathbb{R}^2) \rightarrow \text{SBI}(\mathbb{R}^2)$  is continuous in the metrics  $\text{RM}_q, \text{SIM}_\infty$ .  $\text{LRS}(\mathbb{R}^2) \rightarrow \text{SBR}(\mathbb{R}^2)$  is continuous in  $\text{RM}_q^\circ, \text{SIM}_\infty^\circ$  at non-rectangular lattices.

**(c)** The above conclusions hold for lattices under dilation and homothety.  $\blacksquare$

Corollary 8.8.4 shows that Theorem 8.8.3(b) is the strongest possible continuity in the oriented case. In  $\mathbb{R}^3$ , a similar discontinuity around high-symmetry lattices will be much harder to resolve for continuous invariants even under isometry [16].

**Corollary 8.8.4** (partial discontinuity of  $\text{LRS} \rightarrow \text{SBR}$ , [17, Cor. 7.9]). The bijection  $\text{LRS} \rightarrow \text{SBR}$  is discontinuous in the metrics  $\text{RM}_\infty, \text{SIM}_\infty^\circ$  at all rectangular lattices.  $\blacksquare$

Corollary 8.8.4 should be positively interpreted in the sense that we need to study lattices under rigid motion by their complete oriented root invariants in the continuous space  $\text{LRS}(\mathbb{R}^2)$  rather than in terms of reduced bases (or, equivalently, obtuse superbases due to Proposition 8.3.8b), which are inevitably discontinuous.

Proposition 8.8.5 shows that all  $G$ -chiral distances  $\text{RC}[G] : \text{LIS}(\mathbb{R}^2) \rightarrow \mathbb{R}$  and  $\text{PC}[G] : \text{LHS}(\mathbb{R}^2) \rightarrow \mathbb{R}$  are continuous in any metrics  $\text{RM}, \text{PM}$  from Definition 8.6.1.

**Proposition 8.8.5** (continuous chiral distances, [17, Proposition 7.10]). For a crystallographic point group  $G$  and lattices  $\Lambda_1, \Lambda_2$  in  $\mathbb{R}^2$ , we have

$$|\text{RC}[G](\Lambda_1) - \text{RC}[G](\Lambda_2)| \leq \text{RM}(\Lambda_1, \Lambda_2),$$

$$|\text{PC}[G](\Lambda_1) - \text{PC}[G](\Lambda_2)| \leq \text{PM}(\Lambda_1, \Lambda_2)$$

for any metrics  $\text{RM}$  and  $\text{PM}$ .  $\blacksquare$

Now Remark 8.8.6 summarises a wide range of rich mathematical structures that can be considered on the lattice spaces in addition to continuous metrics.

**Remark 8.8.6** (linear structure, scalar product on lattices). Since the triangular cone  $\text{TC}$  in Fig. 8.9 is convex, we can consider any convex linear combination of root invariants  $t\text{RI}(\Lambda_1) + (1-t)\text{RI}(\Lambda_2) \in \text{TC}$ ,  $t \in [0, 1]$ . The resulting root invariant determines (an isometry class of) the new lattice that can be denoted by  $t\Lambda_1 + (1-t)\Lambda_2$ . The average of the square and hexagonal lattices with  $\text{RI}(\Lambda_4) = (0, 1, 1), \text{RI}(\Lambda_6) = (1, 1, 1)$  has  $\text{RI} = (\frac{1}{2}, 1, 1)$ . The new lattice  $\frac{1}{2}(\Lambda_4 + \Lambda_6)$  is centred rectangular and has the basis  $\mathbf{v}_1 = (\sqrt{\frac{3}{2}}, 0)$  and  $\mathbf{v}_2 = (-\frac{1}{9}\sqrt{\frac{3}{2}}, \frac{4}{9}\sqrt{\frac{15}{2}})$ . We can define similar sums in  $\text{LHS}(\mathbb{R}^2)$  due to the convexity of the triangle  $\text{QT}$ . The usual scalar product of vectors in  $\mathbb{R}^3$  defines the positive product of root invariants:  $\text{RI}(\Lambda_4) \cdot \text{RI}(\Lambda_6) = (0, 1, 1) \cdot (1, 1, 1) = 2$ .  $\blacklozenge$

In conclusion, Problem 8.2.2 was resolved by the new invariants  $\text{RI}, \text{RI}^\circ, \text{PI}, \text{PI}^\circ$  classifying all 2D lattices under four equivalences, see a summary in Table 8.5.

8.2.2(a) Completeness of invariants: Theorem 8.4.2 and Corollary 8.4.5.

8.2.2(b) Reconstruction: Lemma 8.4.1 and Proposition 8.5.3 with Example 8.5.4.

8.2.2(c-e) Continuous metrics: Definitions 8.6.1 and 8.6.3, Theorems 8.8.1 and 8.8.3.

8.2.2(f,g) Realisability and Euclidean embeddability: Corollary 8.5.8.

8.2.2(h) Computability of metrics: Propositions 8.6.5, 8.6.6 and Examples 8.6.2, 8.7.6.

**Table 8.5** A summary of classifications of all lattices  $\Lambda \subset \mathbb{R}^2$  under four equivalence relations.

equivalence	complete invariant	configuration space	continuous metric	visual results
isometry	root invariant	$LIS(\mathbb{R}^2) \leftrightarrow TC$	root metric	Theorem 8.4.2
	$RI(\Lambda)$	triangular cone	RM	Fig. 8.9 (left)
rigid	oriented	$LRS(\mathbb{R}^2) \leftrightarrow DC$	oriented	Theorem 8.4.2
motion	invariant $RI^o(\Lambda)$	doubled cone	metric $RM^o$	Fig. 8.11 (right)
homothety	projected	$LHS(\mathbb{R}^2) \leftrightarrow QT$	projected	Corollary 8.4.5
	invariant $PI(\Lambda)$	quotient triangle	metric PM	Fig. 8.10 (left)
dilation	oriented	$LDS(\mathbb{R}^2) \leftrightarrow QS$	oriented	Corollary 8.4.5
	projected	quotient	projected	Fig. 8.10 (right)
	invariant $PI^o(\Lambda)$	square	metric $PM^o$	Fig. 8.12

The key contributions of this chapter are the easily computable metrics in Definitions 8.6.1, 8.6.3, which led to continuous real-valued deviations of lattices from their higher symmetry neighbours. The chiral distances in Definition 8.7.1 continuously extend the binary chirality by explicit formulae in Propositions 8.7.4 and 8.7.5.

The discontinuity of basis reductions in [23, Theorem 15] was proved with a simple metric on lattice bases without isometry. When we consider obtuse superbases under isometry, the continuity holds in Theorem 8.8.3 under isometry for all lattices and under rigid motion for non-rectangular lattices. For rigid motion, when orientation is preserved, Corollary 8.8.4 proves discontinuity at any rectangular lattice in  $\mathbb{R}^2$ .

It was important to clarify the above discontinuity of bases in Corollary 8.8.4 since the 3-dimensional case is much harder to resolve even under isometry [16].

The structures in Remark 8.8.6 help treat lattices as vectors in a meaningful way (independent of a basis), for example, as inputs or outputs in machine learning algorithms. Paper [6] visualises for the first time millions of 2D lattices extracted from real crystals in the Cambridge Structural Database (CSD), see the Python code at [https://github.com/MattB-242/Lattice\\_Invariance](https://github.com/MattB-242/Lattice_Invariance).

Lattice invariants can be used as a first ultra-fast step to find (near-)duplicates of a potentially new material in all existing experimental datasets. The forthcoming work

[16, 5] extends the isometry classification of Theorem 8.4.2 to  $\mathbb{R}^3$ . The next chapter will introduce more advanced distance-based invariants of general periodic point sets.

## References

1. Relations between norms, 2021.
2. Mois I Aroyo, JM Perez-Mato, Danel Orobengoa, EMRE Tasci, Gemma de la Flor, and Asel Kirov. Crystallography online: Bilbao crystallographic server. *Bulg. Chem. Commun.*, 43(2):183–197, 2011.
3. Mois Ilia Aroyo and H Wondratschek. *International Tables for Crystallography*. Wiley Online Library, 2013.
4. B.N.Delone, N. Padurov, and A. Aleksandrov. *Mathematical Foundations of Structural Analysis of Crystals*. State Technical-Theoretical Press, USSR, 1934.
5. Matthew Bright, Andrew I Cooper, and Vitaliy Kurlin. A complete and continuous map of the lattice isometry space for all 3-dimensional lattices. *arXiv:2109.11538*, 2021.
6. Matthew Bright et al. Geographic-style maps for 2D lattices. *Acta Cryst A*, 79:1–13, 2023.
7. Matthew J Bright, Andrew I Cooper, and Vitaliy A Kurlin. Continuous chiral distances for 2-dimensional lattices. *Chirality*, 35:920–936, 2023.
8. J Conway and N Sloane. Low-dimensional lattices. VI. Voronoi reduction of three-dimensional lattices. *Proceedings Royal Society A*, 436(1896):55–68, 1992.
9. Joseph Louis De Lagrange. Recherches d’arithmétique. *Nouveaux Mémoires de l’Académie de Berlin*, 1773.
10. B. N. Delone. Geometry of positive quadratic forms. part ii (in russian). *Uspekhi Matematicheskikh Nauk*, (4):102–164, 1938.
11. BN Delone, RV Galiulin, and MI Shtogrin. On the Bravais types of lattices. *Journal of Soviet Mathematics*, 4(1):79–156, 1975.
12. Peter Engel, Louis Michel, and Marjorie Sénéchal. Lattice geometry. Technical Report IHES-P-2004-45, 2004.
13. B Gruber. Reduced cells based on extremal principles. *Acta Cryst A*, 45(1):123–131, 1989.
14. Gareth A Jones and David Singerman. *Complex functions: an algebraic and geometric viewpoint*. Cambridge University press, 1987.
15. Jürgen Jost. *Compact Riemann surfaces: an introduction to contemporary mathematics*. Springer Science & Business Media, 2013.
16. V Kurlin. A complete isometry classification of 3D lattices. *arxiv:2201.10543*, 2022.
17. Vitaliy Kurlin. Mathematics of 2D lattices. *Found. Comp. Mathematics*, 24:805–863, 2024.
18. Stephen Lawton and Robert Jacobson. The reduced cell and its crystallographic applications. Technical report, Ames Lab, Iowa State University, 1965.
19. Hermann Minkowski. Ueber die positiven quadratischen formen und über kettenbruchähnliche algorithmen. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, (107):278–297, 1891.
20. P. Niggli. *Krystallographische und strukturtheoretische Grundbegriffe*, volume 1. Akademische verlagsgesellschaft mbh, 1928.
21. Eduard Selling. Ueber die binären und ternären quadratischen formen. *Journal für die reine und angewandte Mathematik*, 77:143–229, 1874.
22. Georges Voronoi. Nouvelles applications des paramètres continus à la théorie des formes quadratiques. *J. Reine Angew. Math.*, (133):97–178, 1908.
23. Daniel Widdowson, Marco M Mosca, Angeles Pulido, Andrew I Cooper, and Vitaliy Kurlin. Average minimum distances of periodic point sets - foundational invariants for mapping all periodic crystals. *MATCH Commun. Math. Comput. Chem.*, 87:529–559, 2022.
24. B. Zhilinskii. *Introduction to lattice geometry through group action*. EDP sciences, 2016.

## Chapter 9

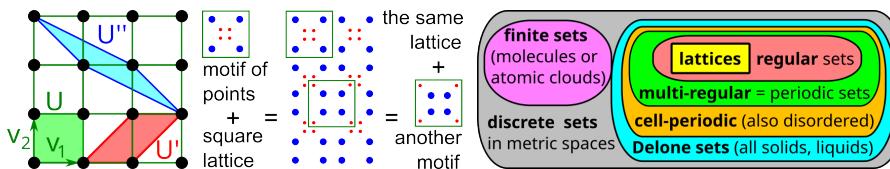
# Density functions of periodic sets of points in $\mathbb{R}^n$ and intervals in $\mathbb{R}$

**Abstract** This chapter adapts the general geo-mapping problem to periodic sets of points under isometry in high dimensions, motivated by periodic crystals in dimension 3. We introduce density functions, which extend the scalar point density to Lipschitz continuous isometry invariants depending on a variable radius of balls centred at given points. These functions can be efficiently computed at discrete radii in low dimensions and are generically complete for periodic point sets under isometry in  $\mathbb{R}^3$ . In dimension  $n = 1$ , the density functions are analytically computable for periodic sets of intervals.

### 9.1 Periodic sets in $\mathbb{R}^n$ and ambiguity of cell-based representations

This section follows paper [14] and adapts Geo-Mapping Problem 1.4.5 to periodic point sets in  $\mathbb{R}^n$ . These sets extend lattices from Definition 8.1.1, as formalised below.

**Definition 9.1.1** (motif, periodic point set in  $\mathbb{R}^n$ ). Let vectors  $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$  form a basis of  $\mathbb{R}^n$ , define the lattice  $\Lambda = \{ \sum_{i=1}^n c_i \mathbf{v}_i \mid c_1, \dots, c_n \in \mathbb{Z} \}$  and the unit cell  $U = \{ \sum_{i=1}^n x_i \mathbf{v}_i \mid x_1, \dots, x_n \in [0, 1) \} \subset \mathbb{R}^n$ . For any finite set of points (called a *motif*)  $M \subset U$ , the sum  $S = M + \Lambda = \{ \mathbf{p} + \mathbf{v} \mid \mathbf{p} \in M, \mathbf{v} \in \Lambda \}$  is a *periodic point set*.  $\blacktriangle$

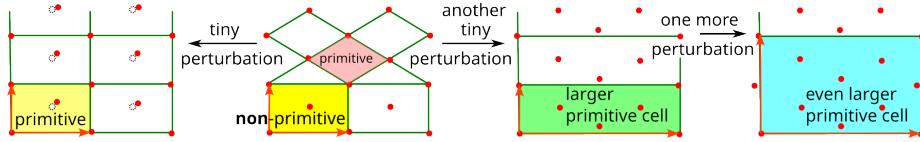


**Fig. 9.1** **Left:** a lattice can be defined by many primitive bases. **Middle:** a periodic set can be defined by different pairs (basis, motif). **Right:** a hierarchy of discrete sets, which model all crystalline materials (periodic crystals) and amorphous solids with points at atomic centers, see Definition 9.1.1.

Any unit cell  $U$  includes only a partial boundary: we exclude the points with any coefficient  $t_i = 1$ ,  $i = 1, \dots, l$ , for convenience. Then  $\mathbb{R}^n$  is tiled by the shifted cells  $\{U + \mathbf{v} \mid \mathbf{v} \in \Lambda\}$  without overlaps. Any lattice is an example of a periodic set with one point in a motif. Any periodic point set  $S = M + \Lambda$  can be considered a finite union  $\bigcup_{p \in M} (\mathbf{p} + \Lambda)$  of lattices whose origins are shifted to all  $p \in M = S \cap U$ .

If we double a unit cell in one direction, e.g. by taking the basis  $2\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ , the doubled motif  $M \cup (M + \mathbf{v}_1)$  with the sublattice on the new basis defines the original periodic point set  $S = M + \Lambda$ . A basis and its cell  $U$  of  $S$  are called *primitive* if  $S \cap U$  has the smallest size among all unit cells  $U$  of  $S$ . Fig. 9.1 (left) shows a square lattice in  $\mathbb{R}^2$ , which (as any lattice) can be generated by infinitely many primitive bases. Even if we fix a basis, Fig. 9.1 (middle) shows that different motifs in the same primitive cell  $U$  define equivalent periodic sets, which differ only by translation.

Finite and periodic point sets represent molecules and periodic crystals at the atomic scale by considering zero-sized points at all atomic centers. Chemical bonds can be modelled by straight-line edges between atomic centers. However, even the strongest covalent bonds within a molecule depend on various thresholds for distances and angles. So these bonds are not real sticks and only abstractly represent inter-atomic interactions, while atomic nuclei are real. We model all materials at the fundamental level of atoms.



**Fig. 9.2** Almost any noise arbitrarily scales up a primitive unit cell of any periodic point set.

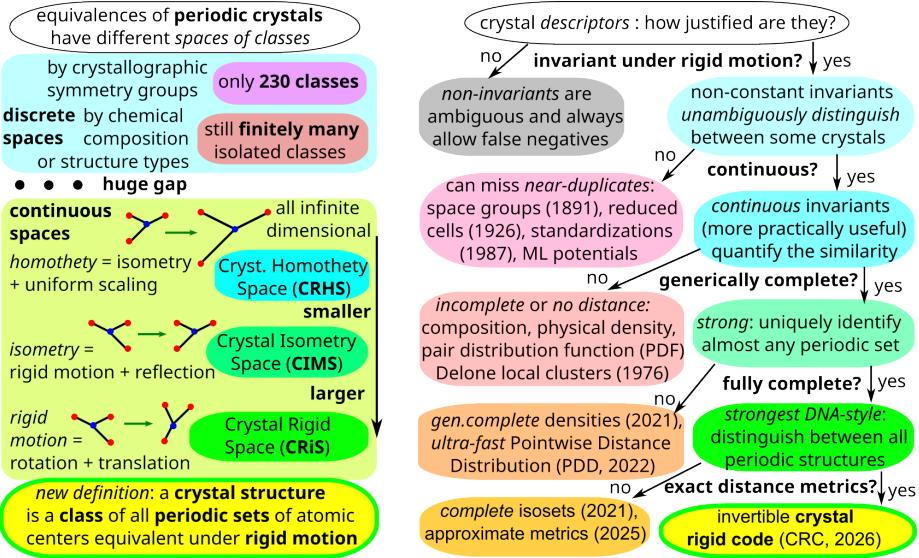
Fig. 9.2 illustrates for a hexagonal lattice of red points that the discontinuity of cell-based representations substantially worsens for periodic point sets in comparison with lattices, which have only one point in a motif. If we arbitrarily extend a given primitive unit cell  $U$  of  $S$  to a larger cell  $U'$ , almost any perturbation of one atom in  $U'$  (and all its periodic copies obtained by translations along the edges of  $U'$ ) makes  $U'$  primitive.

Definition 9.1.2 extends moduli spaces of lattices in Definition 8.2.3 to periodic point sets with up to  $m$  points in their motifs. One physically justified metric on all these spaces is the bottleneck distance  $BD$  from Example 1.3.3(b), which quantifies atomic vibrations as a maximum deviation of all points from their original positions.

However,  $BD$  is defined by minimising over bijections between infinite periodic sets and also over infinitely many equivalences, such as isometries. Hence, periodic sets need a simpler distance metric that should be efficiently computable. Nonetheless, the bottleneck distance allows us to define the concept of a generic set in crystal spaces below. We use the word *crystal* instead of the *periodic point set* to keep all names short.

**Definition 9.1.2** (moduli spaces of periodic point sets and generic subspaces). In all cases below, we consider all periodic point sets  $S \subset \mathbb{R}^n$  with motifs of up to  $m$  points.

- (a) The *Crystal Rigid Space* CRIS( $\mathbb{R}^n; m$ ): periodic point sets under rigid motion.
- (b) The *Crystal Isometry Space* CIMS( $\mathbb{R}^n; m$ ): periodic point sets under isometry.
- (c) The *Crystal Dilation Space* CRDS( $\mathbb{R}^n; m$ ): periodic point sets under dilation.
- (d) The *Crystal Homothety Space* CRHS( $\mathbb{R}^n; m$ ): periodic point sets under homothety.
- (e) For any space  $X$  above, a subspace  $Y \subset X$  is *dense* (or *generic*) if, for any  $\varepsilon > 0$ , any periodic point set  $Q$  representing a class in  $X$  can be obtained from some  $S \subset \mathbb{R}^n$  representing a class in  $Y$  by perturbing any point of  $S$  up to Euclidean distance  $\varepsilon$ .  $\blacktriangle$



**Fig. 9.3 Left:** crystals were previously studied by equivalences based on discrete features, such as compositions and symmetries, but are determined as rigid structures that live in much larger continuous spaces, see Definition 9.1.2. **Right:** the recently developed hierarchy of continuous invariants [16, 15, 4] gives rise to a crystal code that is theoretically complete, computable in polynomial time, and invertible back to a periodic structure, uniquely under rigid motion in  $\mathbb{R}^n$ .

Problem 9.1.3 extends Problem 7.1.5 to generic periodic point sets under isometry in  $\mathbb{R}^n$ . The stronger equivalence of rigid motion will be considered in Chapter 11.

**Problem 9.1.3** (isometry invariants of periodic point sets in  $\mathbb{R}^n$ ). Design an invariant  $I$  on the Crystal Isometry Space CIMS( $\mathbb{R}^n; m$ ) satisfying the following conditions.

- (a) *Generic completeness:* let  $S, Q$  be any generic sets whose isometry classes are in a dense subspace  $Y \subset \text{CIMS}(\mathbb{R}^n; m)$ , then  $S, Q$  are isometric if and only if  $I(S) = I(Q)$ .
- (b) *Metric:* there is a distance  $d$  on the Crystal Isometry Space CIMS( $\mathbb{R}^n; m$ ) satisfying all metric axioms in Definition 1.3.1(a).

**(c) Continuity:** there is a constant  $\lambda > 0$ , such that, for all sufficiently small  $\varepsilon > 0$ , if a periodic point set  $Q$  is obtained by perturbing every point of a periodic point set  $S \subset \mathbb{R}^n$  up to Euclidean distance  $\varepsilon$ , then  $d(I(S), I(Q)) \leq \lambda\varepsilon$ .

**(d) Computability:** the invariant  $I$ , a reconstruction of  $S \subset \mathbb{R}^n$  from  $I(S)$ , and the metric  $d(I(S), I(Q))$  can be computed in times that depend polynomially on the dimension  $n$  and the maximum motif size of periodic point sets  $S, Q$ .  $\star$

Condition 9.1.3(a) includes only generic completeness, which will be tackled in this and the next chapter, while the full completeness will be resolved in Chapter 11.

## 9.2 Continuity and generic completeness of density functions in $\mathbb{R}^3$

This section defines density functions, which almost solve Problem 9.1.3 in  $\mathbb{R}^3$ , though polynomial-time algorithms in 9.1.3(d) will be only approximate. Chapter 10 will introduce newer isometry invariants and metrics with exact polynomial-time algorithms. For any  $t \geq 0$  and  $p \in \mathbb{R}^n$ , let  $\bar{B}(p; t) \subset \mathbb{R}^n$  be the closed ball of radius  $t$  centred at  $p$ .

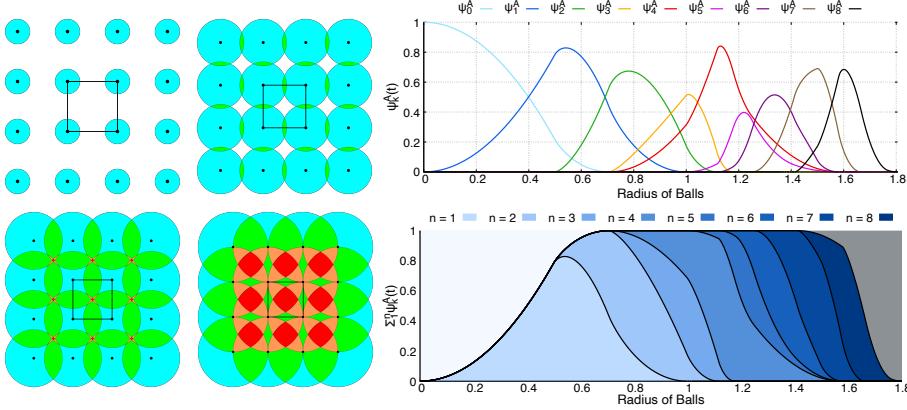
**Definition 9.2.1** (density functions of a periodic point set in  $\mathbb{R}^n$ ). Let a periodic set  $S = \Lambda + M \subset \mathbb{R}^n$  have a unit cell  $U$  of volume  $\text{vol}[U]$ . For any integer  $k \geq 0$ , let  $U_k(t)$  be the region within the cell  $U$  covered by exactly  $k$  closed balls  $\bar{B}(p; t)$  with a radius  $t \geq 0$  and centres at all points of  $S$ . The  $k$ -th *density function* is  $\psi_k[S](t) = \frac{\text{vol}[U_k(t)]}{\text{vol}[U]}$ . The *density fingerprint* is the infinite sequence  $\Psi[S] = \{\psi_k[S](t)\}_{k=0}^{+\infty}$ .  $\blacktriangle$

Notice that closed balls  $\bar{B}(p; t)$  are considered for all points  $p \in S$ , not restricted to the motif  $M = S \cap U$ . The 0-th density  $\psi_0[S](t)$  measures the subset of  $U$  that is not covered by any balls  $\bar{B}(p; t)$  for  $p \in S$ , i.e. the fractional volume subset of all points  $q \in U$  that are more than  $t$  away from all points of  $S$ . For  $k \geq 1$ ,  $\psi_k[S](t)$  measures the fractional volume of all  $k$ -fold intersections  $\bigcap_{p_1, \dots, p_k \in S} \bar{B}(p_i; t)$  within  $U$ .

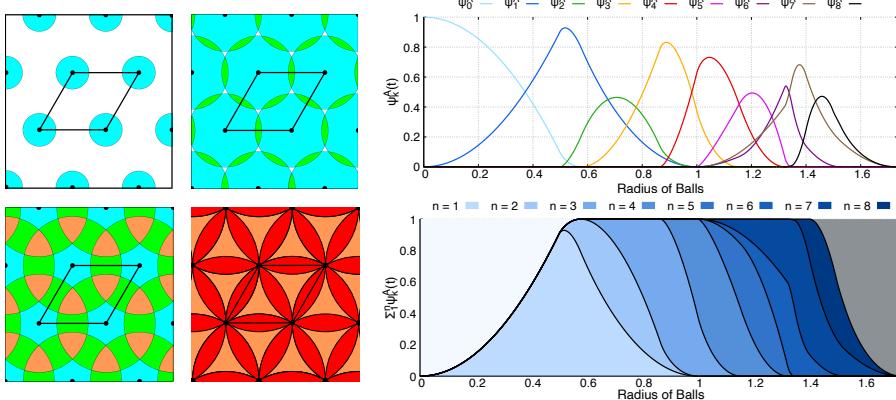
Since each density function  $\psi_k[S](t)$  depends on  $t \in [0, +\infty)$ , our computations in dimensions  $n = 2, 3$  use uniformly sampled radii  $t$ . For any fixed radius  $t$ , the density function  $\psi_k[S](t)$  will be efficiently computed in section 9.3. In dimension  $n = 1$ , all density functions will be analytically computed, also for more general periodic sequences of intervals. Fig. 9.4 and 9.5 illustrate the *densigrams* that combine several density functions in one diagram for the square and hexagonal lattices in  $\mathbb{R}^2$ .

Any density function  $\psi_k[S](t)$  can also be interpreted as the probability that a random point  $p \in \mathbb{R}^n$  is at a maximum distance  $t$  to exactly  $k$  points of  $S$ . However,  $\psi_k[S](t)$  is not a probability density function, so a  $\psi_k[S](t)$  can be called the *k-fold intersection function*, but it is now a bit late to change the name.

Since any isometry preserves distances and hence volumes of areas in  $\mathbb{R}^n$ , any density function is invariant under isometry and independent of a unit cell choice, see [5, Lemma 1]. We describe the key results from [5]. Definition 9.2.2 introduces metrics on density functions to prove their Lipschitz continuity in Theorem 9.2.4.



**Fig. 9.4** Illustration of Definition 9.2.1 for the square lattice  $\Lambda_4$ . **Left:** subregions  $U_k(t)$  are covered by  $k$  disks for the radii  $t = 0.25, 0.55, 0.75, 1$ . **Right:** the nine density functions are above the corresponding *densigram* of accumulated functions  $\sum_{i=1}^k \psi_i(\Lambda_4; t)$  [5, Fig. 2].



**Fig. 9.5** illustration of Definition 9.2.1 for the hexagonal lattice  $\Lambda_6$ . **Left:** subregions  $U_k(t)$  are covered by  $k$  disks for the radii  $t = 0.25, 0.55, 0.75, 1$ . **Right:** the nine density functions are above the corresponding *densigram* of accumulated functions  $\sum_{i=1}^k \psi_i(\Lambda_6; t)$  [5, Fig. 2].

**Definition 9.2.2** (metrics on density functions). For any  $k \geq 1$  and periodic point sets  $S, Q \subset \mathbb{R}^n$ , define the *max metrics* between their density functions as

$$|\psi_k[S] - \psi_k[Q]|_\infty = \sup_{t \geq 0} |\psi_k[S](t) - \psi_k[Q](t)|$$

and between their fingerprints  $d_\infty(\Psi[S], \Psi[Q]) = \sup_{k \geq 0} \frac{|\psi_k[S] - \psi_k[Q]|_\infty}{(\sqrt[3]{k+1})^2}$ .  $\blacktriangle$

In practice, the max distance  $|\psi_k[S] - \psi_k[Q]|_\infty$  is approximated for uniformly sampled radii  $t$ , while  $d_\infty(\Psi[S], \Psi[Q])$  is approximated by considering  $k = 1, \dots, 8$ .

For any fixed index  $k \geq 0$ , the density function  $\psi_k[S](t)$  eventually vanishes for large  $t$  because the whole unit cell  $U$  of  $S$  includes only higher  $l$ -fold intersections of ball  $B(p; t)$  for  $l > k$ . The extra factor  $\frac{1}{(\sqrt[3]{k+1})^2}$  reduces the apparently growing sensitivity of the density functions  $\psi_k[S](t)$  to perturbations of points.

**Definition 9.2.3** (radii  $r(S)$  and  $R(S)$ ). Let  $S \subset \mathbb{R}^n$  be any periodic point set.

(a) The *packing radius*  $r(S)$  is the maximum  $r$  such that the open balls  $B(p; r)$  are disjoint for all  $p \in S$ , or  $r(S)$  is the minimum half-distance between any points of  $S$ .

(b) The *covering radius*  $R(S)$  is the minimum radius  $R$  such that the union of closed balls  $\bar{B}(p; R)$  for all  $p \in S$  covers  $\mathbb{R}^n$ . Alternatively,  $R(S)$  is the maximum distance from any point  $q \in \mathbb{R}^n$  to its nearest neighbour in  $S$ .  $\blacktriangle$

In terms of density functions,  $r(S)$  is the maximum radius  $t \geq 0$  such that  $\psi_1[S](t) = 0$ , while  $R(S)$  is the maximum radius  $R$  such that  $\psi_1[S](t) = 0$  for all  $t \in [0, R]$ .

**Theorem 9.2.4** (continuity of density functions, [5, Theorem 1]). Let  $Q \subset \mathbb{R}^3$  be a periodic point set obtained from another periodic point set  $S \subset \mathbb{R}^3$  by perturbing any point of  $S$  up to Euclidean distance  $\varepsilon$  such that  $0 \leq \varepsilon < r = \min\{r(S), r(Q)\}$ . Set

$$R = \max\{R(S), R(Q)\} \text{ and } \lambda = 13 \frac{R^2}{r^3}. \text{ Then } d_\infty(\Psi[S], \Psi[Q]) \leq \lambda \varepsilon. \blacksquare$$

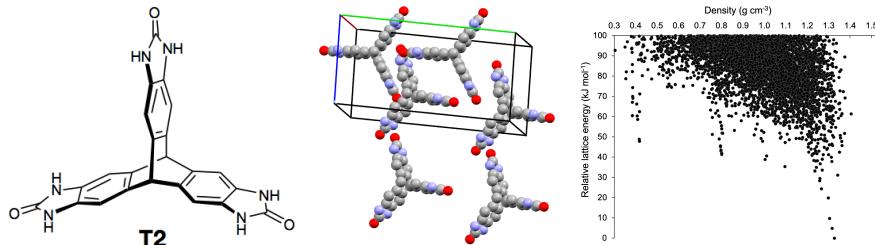
[5, section 5.1] describes technical conditions defining *generic* periodic point sets  $S \subset \mathbb{R}^3$ , which satisfy Theorem 9.2.5.

**Theorem 9.2.5** (generic completeness of density functions, [5, Theorem 1]). If any generic periodic point sets  $S, Q \subset \mathbb{R}^3$  are not isometric ( $S \neq Q$ ), then  $\Psi[S] \neq \Psi[Q]$ .  $\blacksquare$

Now we describe the first practical impact of density functions. Crystal Structure Prediction (CSP) aims to predict whether a selected molecule can be crystallised into a functional material, i.e. a crystal with useful functions or properties. In theory, CSP seeks to answer the question of whether copies of a molecule can be arranged in such a way that the resulting crystal is thermodynamically stable as well as useful. CSP aims to answer this question purely computationally to streamline the trial-and-error.

Our colleagues at Liverpool's Materials Innovation Factory [10] computationally predicted that the T2 molecule in Fig. 9.6 can be crystallised into a new nanoporous material for gas storage. As part of this process, they also identified four other structures of interest. Following the CSP predictions, they synthesised 5 families of T2-crystals in the laboratory by varying parameters like temperature and pressure, calling them T2- $\alpha$ , T2- $\beta$ , ..., T2- $\epsilon$ . One of them, T2- $\gamma$ , indeed had the desired property of having only half the physical density of the previously known structure T2- $\alpha$ .

Structures of the synthesised crystals were reconstructed using X-ray powder diffraction yielding Crystallographic Information Files (CIFs), each containing a unit cell and a motif of points representing atoms. These resulting CIFs were then compared with the results of the simulations, by filtering by physical density and then by using the



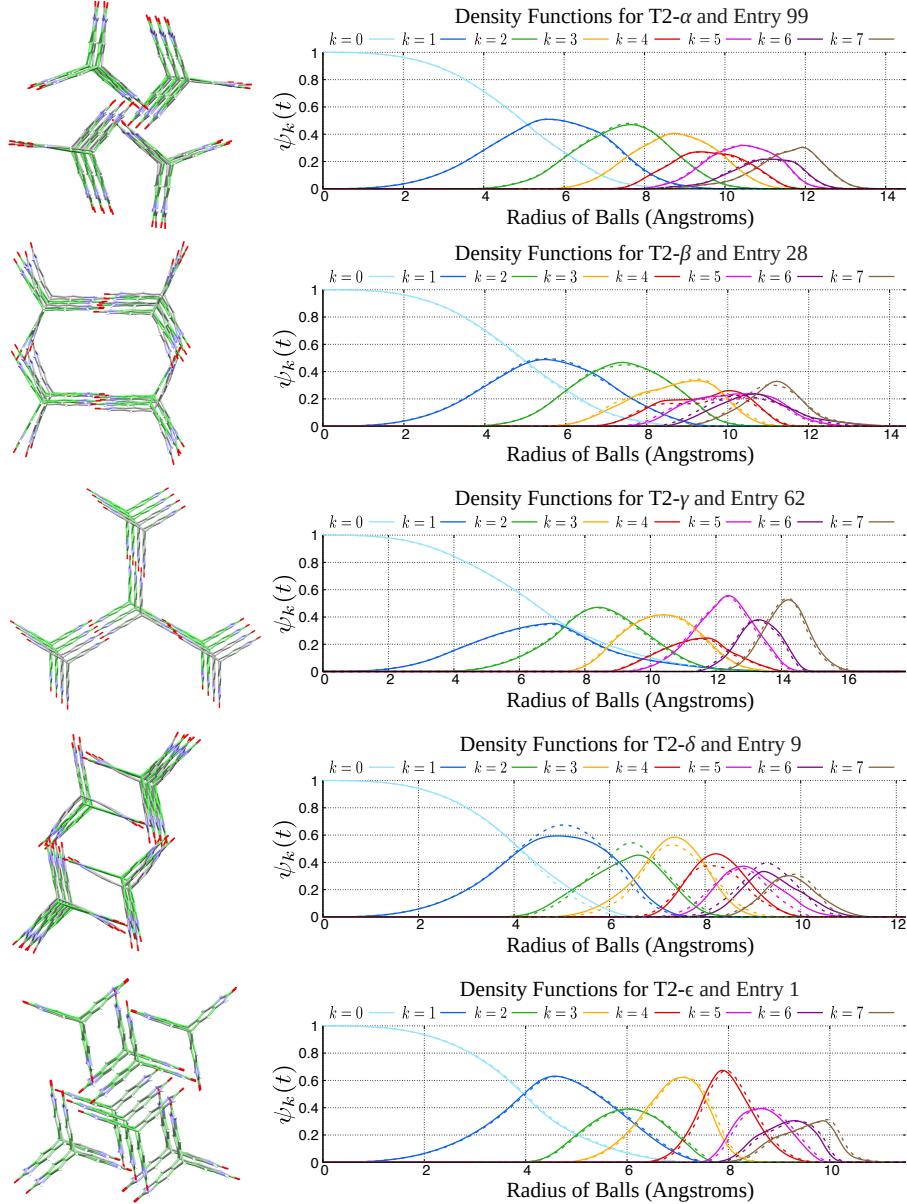
**Fig. 9.6** **Left:** a T2 molecule. **Middle:** the T2- $\delta$  crystal with highlighted unit cell. **Right:** a plot of 5679 simulated T2 crystals [10, Fig. 2d], each represented by two coordinates: the physical density (atomic mass within a unit cell divided by the unit cell volume) and energy (determining the crystal's thermodynamic stability). Structures at the bottom of the 'downward spikes' are likely to be stable.

COMPACK algorithm—which compares only finite portions of infinite crystals — or by looking at visualisations of structures. The experimental crystals were considered close to some predictions and deposited into the Cambridge Structural Database (CSD).

$ \psi_k[S] - \psi_k[Q] _\infty$	$k = 0$	1	2	3	4	5	6	7
T2- $\alpha$ vs entry 99	0.0042	0.0092	0.0125	0.0056	0.0099	0.0088	0.0127	0.0099
T2- $\beta$ vs entry 28	0.0157	0.0156	0.0159	0.0224	0.0334	0.0396	0.0357	0.0454
T2- $\gamma$ vs entry 62	0.0020	0.0080	0.0128	0.0155	0.0153	0.0250	0.0296	0.0391
T2- $\delta$ vs entry 09	0.0610	0.0884	0.1267	0.0676	0.0915	0.0801	0.0733	0.0388
T2- $\epsilon$ vs entry 01	0.0132	0.0152	0.0207	0.0571	0.0514	0.0431	0.0468	0.0550
T2- $\beta'$ vs entry 09	0.2981	0.2631	0.3718	0.3747	0.2563	0.2360	0.3161	0.3232

**Table 9.1** **First five rows:** the  $L_\infty$ -distances between the first eight pairs of corresponding density functions of physically synthesised T2 crystals (T2- $\alpha$ , T2- $\beta$ , etc.) and the simulated structures that had predicted them from the CSP output dataset (entry XX). **Last row:** the suspiciously larger numbers revealed the mix-up of the files T2- $\delta$  and T2- $\beta'$  and thus led to depositing the initially omitted Crystallographic Information File of the T2- $\delta$  crystal into the Cambridge Structural Database.

At a later time, we used our density functions to verify the past matchings between the synthesised crystals T2- $\alpha$  to T2- $\epsilon$  and their closest predictions numbered by 99, 28, 62, 09, 01 in the T2 dataset of all 5679 simulated crystals. We did so by computing, for each of the five matches, the distance between the density functions of the synthesised and the simulated crystal. As one is the prediction of the other, we expected to see small distances. And for four of the five structures this was true: T2- $\gamma$ , for example, always has an  $L_\infty$ -distance of less than 0.04 over the first eight pairs of corresponding density functions; see Table 9.1. However, when we came to check the distances between density



**Fig. 9.7** **Left:** experimental T2 crystals (curved grey molecules) and their simulated versions (straight green molecules) are overlaid. **Right:** the density functions of the periodic sets of molecular centres of the experimental T2 crystals (solid curves) vs. simulated crystals (dashed curves), produced by [12].

functions of T2- $\delta$  with its predicted structure, we were surprised to see large distances (the final row of Table 9.1). It turned out that a mix-up of files had happened, and what was uploaded to the CSD as T2- $\delta$  was in fact a version of the T2- $\beta$  polymorph. The

density fingerprint revealed this error, which was verified by chemists upon a visual inspection, and it is because of this that T2- $\delta$  was subsequently correctly deposited.

Plots of the density functions of correctly matched synthesised and simulated structures can be seen in Figure 9.7. As another application, we expect that the fingerprint will be used to simplify the large output data sets produced by CSP by comparing simulated structures with each other, thus speeding up what is currently a slow process.

### 9.3 Computing density functions via degree- $k$ Voronoi domains

This section follows papers [5, 11] with minor updates. To compute any density function  $\psi_k[S](t)$ , we introduce the *degree- $k$  Voronoi domains*, also called the  *$k$ -th Brillouin zone*, which generalise the Voronoi domain  $\bar{V}(\Lambda)$  of a lattice from Definition 8.1.3(a) and differ from the *order- $k$  Voronoi domain* [6] defined for a  $k$ -point subset  $A \subset S \subset \mathbb{R}^n$  and as the set of all points for whom the points in  $A$  are the closest  $k$  points in  $S$ .

**Definition 9.3.1** (index- $k$  Voronoi domain  $V_k(S; p)$  and degree- $k$  Voronoi domain  $Z_k(S; p)$ ). **(a)** For a finite or a periodic point set  $S \subset \mathbb{R}^n$  and a point  $p \in S$ , the  *$k$ -th Voronoi domain*  $V_k(C; p)$  is the closure of the set of all points  $q \in \mathbb{R}^n$  such that  $p$  is among the  $k$  nearest points of  $S$  to  $q$ .

**(b)** For any periodic point set  $S \subset \mathbb{R}^n$  and  $p \in S$ , the *degree- $k$  Voronoi domain* is the difference between successive closed index- $k$  Voronoi domains, i.e.  $Z_k(C; p) = V_k(C; p) - V_{k-1}(C; p)$  for  $k \geq 1$ , where we set  $V_0(C; p) = \emptyset$ .  $\blacktriangle$

If  $k = 1$ ,  $V_1(S; p) = Z_1(S; p)$  is the classical Voronoi domain for a point  $p \in S$ .

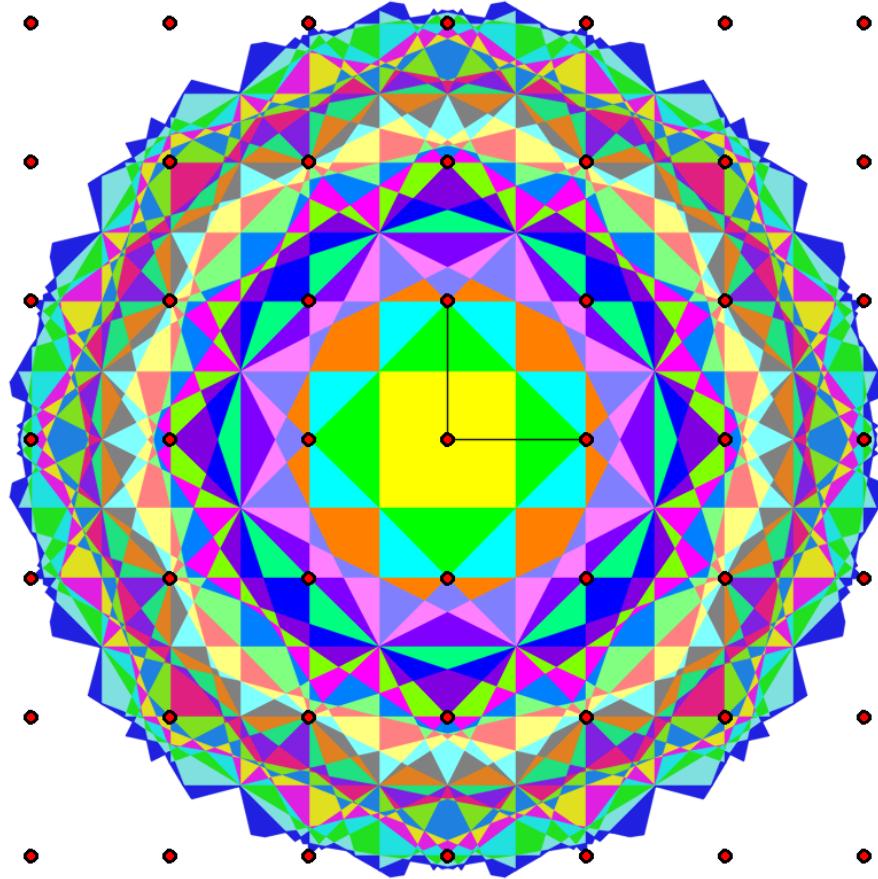
The index- $k$  Voronoi domain  $V_k(C; p) \subset \mathbb{R}^n$  is defined as a closed set above to cover all cases where  $p$  has equal distances to several neighbours, so a  $k$ -th neighbour of  $p$  may not be unique. Unlike order- $k$  Voronoi domains, which tile  $\mathbb{R}^n$  [7], index- $k$  Voronoi domains form a nested sequence. Any  $V_k(C; p)$  is *star-convex*, which means it contains all line segments connecting  $\partial V_k(C; p)$  to  $p$ . Indeed, if  $p \in C$  is among the  $k$  nearest to  $q \in \partial V_k(C; p)$ , then any intermediate point in the line segment  $[p, q]$  has  $p$  among its  $k$  nearest neighbours of  $C$ .

Fig. 9.8 and 9.9 are outputs of Philip Smith's software [13] and show degree- $k$  Voronoi domains for the square and hexagonal lattices for  $k = 1, \dots, 30$ . In each of these images, any degree- $k$  Voronoi domain is the union of polygons of the same colour, and has the origin as its  $k$ -th nearest neighbour among all lattice points.

Theorem 9.3.2 implies that Fig. 9.8 and 9.9 have the same total area of each colour.

**Theorem 9.3.2** (volumes of degree- $k$  Voronoi domains, [11, Theorem 7]). For any periodic point set  $S \subset \mathbb{R}^n$  with a motif  $M$ , the sum of the volumes of the degree- $k$  Voronoi domains  $Z_k(S; p)$  over all motif points  $p \in M$  is independent of  $k$ .  $\blacksquare$

Theorem 9.3.3 computes any density function  $\psi_k[S]$  via degree- $k$  Voronoi domains of  $S$ . Let  $\text{vol}[C]$  denote the volume of any compact region  $C \subset \mathbb{R}^n$ .



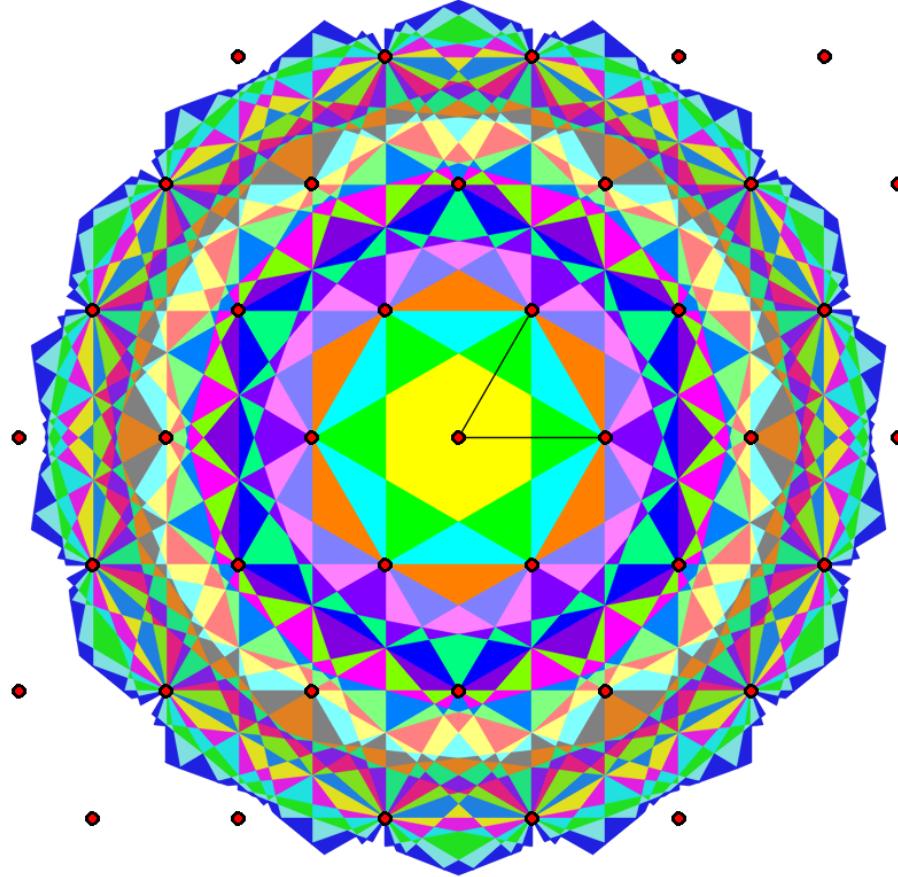
**Fig. 9.8** Degree- $k$  Voronoi domains  $Z_k(\Lambda_4; 0)$  of the square lattice  $\Lambda_4$  of red points for  $k = 1, \dots, 30$ .

**Theorem 9.3.3** (formula for density functions, [5, Theorem 2]). Let  $S = \Lambda + M \subset \mathbb{R}^n$  be a periodic point set with a motif  $M \subseteq U$  in a unit cell  $U$  of a lattice  $\Lambda \subset \mathbb{R}^n$ . Fix any integer  $k \geq 1$ . Then  $\psi_k[S](t) = \frac{1}{\text{vol}[U]} \sum_{p \in M} \text{vol}[V_k(S; p) \cap \bar{B}(p; t)]$ .  $\blacksquare$

Lemma 9.3.4 can be considered a definition of a Minkowski-reduced basis, which is needed to estimate the time for degree- $k$  Voronoi domains in Theorem 9.3.5.

**Lemma 9.3.4** ([9, Lemma 2.2.1]). A basis  $\mathbf{v}_1, \dots, \mathbf{v}_n$  of a lattice  $\Lambda \subset \mathbb{R}^n$  is *Minkowski-reduced* if and only if, for any  $i = 1, \dots, n$  and  $c_1, \dots, c_n \in \mathbb{Z}$  such that  $c_1, \dots, c_n$  have no common integer factor  $c > 1$ , the inequality  $|\sum_{j=1}^n c_j \mathbf{v}_j| \geq |\mathbf{v}_i|$  holds.  $\blacksquare$

**Theorem 9.3.5** (time of degree- $k$  Voronoi domains, [11, Theorem 7]). For  $n = 2, 3$ , let a periodic point set  $S = \Lambda + M \subset \mathbb{R}^n$  have a motif  $M$  of  $m$  points and a lattice  $\Lambda$



**Fig. 9.9** Degree- $k$  Voronoi domains  $Z_k(\Lambda_6; 0)$  of the hexagonal lattice  $\Lambda_6$  for  $k = 1, \dots, 30$ .

with a Minkowski-reduced basis. For any point  $p \in S$ , the time to compute all degree- $i$  Voronoi domains  $Z_i(S; p)$  for  $i = 1, \dots, k$  is  $O(m^n(4k)^{n^2}(n \log(4k) + \log m))$ . ■

## 9.4 A description of density functions of periodic sequences

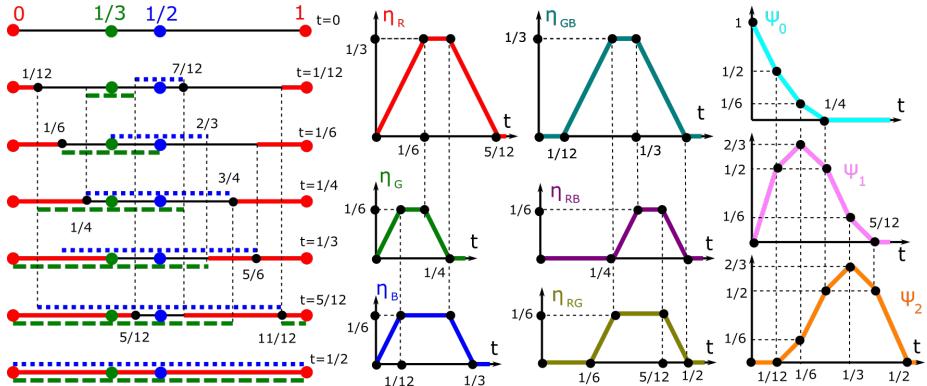
All quoted results in this section have detailed proofs in [2].

For convenience, scale any periodic sequence to period 1 so that  $S = \{p_1, \dots, p_m\} + \mathbb{Z}$ . Since the expanding balls in  $\mathbb{R}$  are growing intervals, volumes of their intersections linearly change in the variable radius  $t$ . Hence, any density function  $\psi_k(t)$  is piecewise linear and uniquely determined by *corner* points  $(a_j, b_j)$  where the gradient changes.

Examples 9.4.1 and 9.4.3 explain how the density functions  $\psi_k(t)$  are computed for the periodic sequence  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$ , see all graphs in Fig. 9.10.

**Example 9.4.1** (0-th density  $\psi_0(t)$  for  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$ ). By Definition 9.2.1  $\psi_0(t)$  is the fractional length within the period interval  $[0, 1]$  not covered by the intervals of radius  $t$  (length  $2t$ ), which are the red intervals  $[0, t] \cup [1-t, 1]$ , green dashed interval  $[\frac{1}{3}-t, \frac{1}{3}+t]$  and blue dotted interval  $[\frac{1}{2}-t, \frac{1}{2}+t]$ . The graph of  $\psi_0(t)$  starts from the point  $(0, 1)$  at  $t = 0$ . Then  $\psi_0(t)$  linearly drops to the point  $(\frac{1}{12}, \frac{1}{2})$  at  $t = \frac{1}{12}$  when a half of the interval  $[0, 1]$  remains uncovered.

The next linear piece of  $\psi_0(t)$  continues to the point  $(\frac{1}{6}, \frac{1}{6})$  at  $t = \frac{1}{6}$  when only  $[\frac{2}{3}, \frac{5}{6}]$  is uncovered. The graph of  $\psi_0(t)$  finally returns to the  $t$ -axis at the point  $(\frac{1}{4}, 0)$  and remains there for  $t \geq \frac{1}{4}$ . The piecewise linear behaviour of  $\psi_0(t)$  can be briefly described via the *corner points*  $(0, 1), (\frac{1}{12}, \frac{1}{2}), (\frac{1}{6}, \frac{1}{6}), (\frac{1}{4}, 0)$ .  $\blacklozenge$



**Fig. 9.10** **Left:** the periodic sequence  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$  with points of three colors. The growing intervals around the red point  $0 \equiv 1 \pmod{1}$ , green point  $\frac{1}{3}$ , blue point  $\frac{1}{2}$  have the same color for various radii  $t$ . **Right:** the trapezium functions  $\eta$  from Example 9.4.3.

Theorem 9.4.2 extends Example 9.4.1 to any periodic sequence  $S$  and implies that  $\psi_0(t)$  is uniquely determined by the ordered distances within a unit cell of  $S$ .

**Theorem 9.4.2** (description of  $\psi_0$ , [2, Theorem 5]). For any periodic sequence  $S = \{p_1, \dots, p_m\} + \mathbb{Z}$  with motif points  $0 \leq p_1 < \dots < p_m < 1$ , set  $d_i = p_{i+1} - p_i \in (0, 1)$ , where  $i = 1, \dots, m$  and  $p_{m+1} = p_1 + 1$ . Put the distances in the increasing order  $d_{[1]} \leq d_{[2]} \leq \dots \leq d_{[m]}$ . Then the 0-th density function  $\psi_0$  is piecewise linear with the following (unordered) corners:  $(0, 1)$  and  $(\frac{1}{2}d_{[i]}, 1 - \sum_{j=1}^{i-1} d_{[j]} - (m-i+1)d_{[i]})$  for  $i = 1, \dots, m$ , so the last corner is  $(\frac{1}{2}d_{[m]}, 0)$ . If any corner points are repeated, e.g. if  $d_{[i-1]} = d_{[i]}$ , these corners are collapsed into one.  $\blacksquare$

Theorem 9.4.2 for the sequence  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$  gives the ordered distances  $d_{[1]} = \frac{1}{6} < d_{[2]} = \frac{1}{3} < d_{[3]} = \frac{1}{2}$ , which determine the corner points  $(0, 1), (\frac{1}{12}, \frac{1}{2}), (\frac{1}{6}, \frac{1}{6}), (\frac{1}{4}, 0)$  of the density function  $\psi_0(t)$  in Fig. 9.10, see Example 9.4.1.

For any periodic sequence with  $m$  points in a unit cell, by Theorem 9.4.2, any 0th density function  $\psi_0(t)$  is uniquely determined by the (unordered) set of lengths of intervals between successive points. Hence, we can reorder these intervals without changing  $\psi_0(t)$ . For instance, the periodic sequence  $Q = \{0, \frac{1}{2}, \frac{2}{3}\} + \mathbb{Z}$  has the same set of interval lengths  $d_{[1]} = \frac{1}{6}, d_{[2]} = \frac{1}{3}, d_{[3]} = \frac{1}{2}$  as the periodic sequence  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$  in Example 9.4.1.

The above sequences  $S, Q$  are related by the mirror reflection  $t \mapsto 1 - t$ . One can easily construct many non-isometric sequences with  $\psi_0[S](t) = \psi_0[Q](t)$ . For any  $1 \leq i \leq m-3$ , the sequences  $S_{m,i} = \{0, 2, 3, \dots, i+2, i+4, i+5, \dots, m+2\} + (m+2)\mathbb{Z}$  have the same interval lengths  $d_{[1]} = \dots = d_{[m-2]} = 1, d_{[m-1]} = d_{[m]} = 2$  but are not related by isometry (translations and reflections in  $\mathbb{R}$ ) because the intervals of length 2 are separated by  $i-1$  intervals of length 1 in  $S_{m,i}$ .

Corollary 9.4.9 will prove that the 1st density function  $\psi_1[S](t)$  uniquely determines a periodic sequence  $S \subset \mathbb{R}$  in general position up to isometry of  $\mathbb{R}$ .

**Example 9.4.3** (functions  $\psi_k(t)$  for  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$ ). The 1st density function  $\psi_1(t)$  can be obtained as a sum of the three *trapezium* functions  $\eta_R, \eta_G, \eta_B$ , each measuring the length of a region covered by a single interval (of one color). The red intervals  $[0, t] \cup [1-t, 1]$  grow until  $t = \frac{1}{6}$  when they touch the green interval  $[\frac{1}{6}, \frac{1}{2}]$ . So the length  $\eta_R(t)$  of this interval linearly grows from the origin  $(0, 0)$  to the corner point  $(\frac{1}{6}, \frac{1}{3})$ . For  $t \in [\frac{1}{6}, \frac{1}{4}]$ , the left red interval is shrinking at the same rate due to the overlapping green interval, while the right red interval continues to grow until  $t = \frac{1}{4}$ , when it touches the blue interval  $[\frac{1}{4}, \frac{3}{4}]$ . Hence the graph of  $\eta_R(t)$  remains constant up to the corner point  $(\frac{1}{4}, \frac{1}{3})$ . After that  $\eta_R(t)$  linearly returns to the  $t$ -axis at  $t = \frac{5}{12}$ . Hence the trapezium function  $\eta_R$  has the piecewise linear graph through the corner points  $(0, 0), (\frac{1}{6}, \frac{1}{3}), (\frac{1}{4}, \frac{1}{3}), (\frac{5}{12}, 0)$ .

The 2nd function  $\psi_2(t)$  is the sum of the *trapezium* functions  $\eta_{GB}, \eta_{RG}, \eta_{RB}$ , each measuring the length of a double intersection. For the green interval  $[\frac{1}{3}-t, \frac{1}{3}+t]$  and the blue interval  $[\frac{1}{2}-t, \frac{1}{2}+t]$ , the graph of the trapezium function  $\eta_{GB}(t)$  is piecewise linear and starts at the point  $(\frac{1}{12}, 0)$ , where the intervals touch. The green-blue intersection interval  $[\frac{1}{2}-t, \frac{1}{3}+t]$  grows until  $t = \frac{1}{4}$ , when  $[\frac{1}{4}, \frac{7}{12}]$  touches the red interval on the left. At the same time  $\eta_{GB}(t)$  is linearly growing to the point  $(\frac{1}{4}, \frac{1}{3})$ . For  $t \in [\frac{1}{4}, \frac{1}{3}]$ , the green-blue intersection interval becomes shorter on the left, but grows at the same rate on the right until  $[\frac{1}{3}, \frac{2}{3}]$  touches the red interval  $[\frac{2}{3}, 1]$ . Then  $\eta_{GB}(t)$  remains constant up to the point  $(\frac{1}{3}, \frac{1}{3})$ . For  $t \in [\frac{1}{3}, \frac{1}{2}]$  the green-blue intersection interval is shortening from both sides. The graph of  $\eta_{GB}(t)$  returns to the  $t$ -axis at  $(\frac{1}{2}, 0)$ , see Fig. 9.10.  $\blacklozenge$

Theorem 9.4.4 extends Example 9.4.3 and proves that any  $\psi_k(t)$  is a sum of trapezium functions whose corners are explicitly described. We consider any index  $i = 1, \dots, m$  (of a point  $p_i$  or a distance  $d_i$ ) modulo  $m$  so that  $m+1 \equiv 1 \pmod{m}$ .

**Theorem 9.4.4** (description of  $\psi_k$  for  $k > 0$ , [2, Theorem 7]). For any periodic sequence  $S = \{p_1, \dots, p_m\} + \mathbb{Z}$  with points  $0 \leq p_1 < \dots < p_m < 1$  in a motif, set  $d_i = p_{i+1} - p_i \in (0, 1)$ , where  $i = 1, \dots, m$  and  $p_{m+1} = p_1 + 1$ . Any interval  $[p_i - t, p_i + t]$  is projected to  $[0, 1]$  modulo  $\mathbb{Z}$ . For  $1 \leq k \leq m$ , the density function  $\psi_k(t)$  is the sum of  $m$  trapezium functions  $\eta_{k,i}$  with the corner points  $(\frac{s}{2}, 0), (\frac{d_{i-1}+s}{2}, d), (\frac{s+d_{i+k-1}}{2}, d), (\frac{d_{i-1}+s+d_{i+k-1}}{2}, 0)$ , where  $d = \min\{d_{i-1}, d_{i+k-1}\}$ ,  $s = \sum_{j=i}^{i+k-2} d_j$ ,  $i = 2, \dots, m + 1$ . If  $k = 1$ , then  $s = 0$  is the empty sum. So  $\psi_k(t)$  is determined by the unordered set of triples  $(d_{i-1}, s, d_{i+k-1})$  whose first and last entries are swappable. ■

In Example 9.4.3 for  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$ , we have  $d_1 = \frac{1}{3}$ ,  $d_2 = \frac{1}{6}$ ,  $d_3 = \frac{1}{2} = d_0$ . For  $k = 2$ ,  $i = 2$ , we get  $d_{i-1} = d_1 = \frac{1}{3}$ ,  $d_{i+k-1} = d_3 = \frac{1}{2}$ , i.e.  $d = \min\{d_1, d_3\} = \frac{1}{3}$ ,  $s = d_2 = \frac{1}{6}$ . Then  $\eta_{22} = \eta_{GB}$  has the corner points  $(\frac{1}{12}, 0), (\frac{1}{4}, \frac{1}{3}), (\frac{1}{3}, \frac{1}{3}), (\frac{1}{2}, 0)$ .

**Theorem 9.4.5** (symmetries of  $\psi_k(t)$ , [2, Theorem 8]). For any periodic sequence  $S \subset \mathbb{R}$  with a unit cell  $[0, 1]$ , we have the periodicity  $\psi_{k+m}(t + \frac{1}{2}) = \psi_k(t)$  for  $k \geq 0$ ,  $t \geq 0$ , and the symmetry  $\psi_{m-k}(\frac{1}{2} - t) = \psi_k(t)$  for  $k = 0, \dots, [\frac{m}{2}]$ , and  $t \in [0, \frac{1}{2}]$ . ■

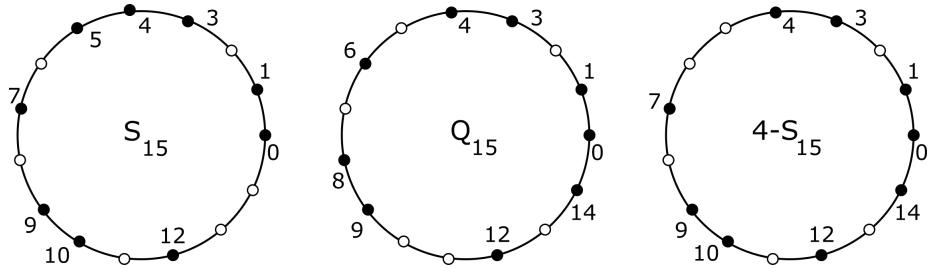
**Corollary 9.4.6** (time of  $\psi_k(t)$  for periodic sequences of points, [2, Corollary 9]). Let  $S, Q \subset \mathbb{R}$  be periodic sequences with at most  $m$  points in motifs. For  $k \geq 1$ , one can draw the graph of the  $k$ -th density function  $\psi_k[S]$  in time  $O(m^2)$ . One can check in time  $O(m^3)$  if the full density fingerprints coincide:  $\Psi[S] = \Psi[Q]$ . ■

To illustrate Corollary 9.4.6, Example 9.4.7 will justify that the periodic sequences  $S_{15}$  and  $Q_{15}$  in Fig. 9.11 have identical density fingerprints  $\Psi[S_{15}] = \Psi[Q_{15}]$ .

**Example 9.4.7** (periodic sequences  $S_{15}, Q_{15} \subset \mathbb{R}$ ). [16, Appendix B] discusses homometric periodic sets that can be distinguished by the invariant AMD (Average Minimum Distances) and not by inter-point distance distributions. The periodic sequences

$$S_{15} = \{0, 1, 3, 4, 5, 7, 9, 10, 12\} + 15\mathbb{Z}, \quad Q_{15} = \{0, 1, 3, 4, 6, 8, 9, 12, 14\} + 15\mathbb{Z}$$

have period 15 and unit cell  $[0, 15]$  shown as a circle in Fig. 9.11.



**Fig. 9.11** Circular versions of the periodic sets  $S_{15}, Q_{15}$ . Distances are along round arcs.

These periodic sequences [8] are obtained as Minkowski sums  $S_{15} = U + V + 15\mathbb{Z}$  and  $Q_{15} = U - V + 15\mathbb{Z}$  for  $U = \{0, 4, 9\}$  and  $V = \{0, 1, 3\}$ . The last picture in Fig. 9.11

shows the periodic set  $4 - S_{15}$  isometric to  $S_{15}$ . Now the difference between  $Q_{15}$  and  $4 - S_{15}$  is better visible: points  $0, 1, 3, 4, 5, 12, 14$  are common, but points  $6, 8, 9 \in Q_{15}$  are shifted to  $7, 9, 10$  in the circular set  $4 - S_{15}$ .

To avoid fractions, we keep the unit cell  $[0, 15]$  of the sequences  $S_{15}, Q_{15}$  without scaling it down to  $[0, 1]$  because all quantities in Theorem 9.4.4 can be scaled up by factor 15. To conclude that  $\psi_0[S_{15}] = \psi_0[Q_{15}]$ , by Theorem 9.4.2 we check that  $S_{15}, Q_{15}$  have the same set of the ordered distances  $d_{[i]}$  between successive points, which is shown in identical rows 3 of Tables 9.2 and 9.3.

**Table 9.2** **Row 1:** points  $p_i$  from  $S_{15}$  in Fig. 9.11. **Row 2:** the distances  $d_i$  between successive points of  $S_{15}$ . **Row 3:** the distances  $d_{[i]}$  are in the increasing order. **Row 4:** the unordered set of these pairs determines the density function  $\psi_1$  by Theorem 9.4.4. **Row 5:** the pairs are lexicographically ordered for comparison with row 5 in Table 9.3. **Rows 6,8,10:** the unordered sets of these triples determine the density functions  $\psi_2, \psi_3, \psi_4$  by Theorem 9.4.4 for  $k = 2, 3, 4$ . **Rows 7,9,11:** the triples from rows 6,8,10 are ordered for easier comparison with corresponding rows 7,9,11 in Table 9.3, see Example 9.4.7.

$p_i$	0	1	3	4	5	7	9	10	12
$d_i = p_{i+1} - p_i$	1	2	1	1	2	2	1	2	3
ordered $d_{[i]}$	1	1	1	1	2	2	2	2	3
$(d_{i-1}, d_i)$	(3,1)	(1,2)	(2,1)	(1,1)	(1,2)	(2,2)	(2,1)	(1,2)	(2,3)
order $(d_{i-1}, d_i)$	(1,1)	(1,2)	(1,2)	(1,2)	(1,2)	(1,2)	(1,3)	(2,2)	(2,3)
$(d_{i-1}, \mathbf{d}_i, d_{i+1})$	(3,1,2)	(1,2,1)	(2,1,1)	(1,1,2)	(1,2,2)	(2,2,1)	(2,1,2)	(1,2,3)	(2,3,1)
order $(d_{i-1}, \mathbf{d}_i, d_{i+1})$	(1,1,2)	(1,1,2)	(2,1,2)	(2,1,3)	(1,2,1)	(1,2,2)	(1,2,2)	(1,2,3)	(1,3,2)
$(d_{i-1}, \mathbf{s}, d_{i+2})$	(3,3,1)	(1,3,1)	(2,2,2)	(1,3,2)	(1,4,1)	(2,3,2)	(2,3,3)	(1,5,1)	(2,4,2)
order $(d_{i-1}, \mathbf{s}, d_{i+2})$	(2,2,2)	(1,3,1)	(1,3,2)	(1,3,3)	(2,3,2)	(2,3,3)	(1,4,1)	(2,4,2)	(1,5,1)
$(d_{i-1}, \mathbf{s}, d_{i+3})$	(3,4,1)	(1,4,2)	(2,4,2)	(1,5,1)	(1,5,2)	(2,5,3)	(2,6,1)	(1,6,2)	(2,6,1)
order $(d_{i-1}, \mathbf{s}, d_{i+3})$	(1,4,2)	(1,4,3)	(2,4,2)	(1,5,1)	(1,5,2)	(2,5,3)	(1,6,2)	(1,6,2)	(1,6,2)

To conclude that  $\psi_1[S_{15}] = \psi_1[Q_{15}]$  by Theorem 9.4.4, we check that  $S_{15}, Q_{15}$  have the same set of unordered pairs  $(d_{i-1}, d_i)$  of distances between successive points. Indeed, Tables 9.2 and 9.3 have identical rows 5, where pairs are *lexicographically* ordered for comparison:  $(a, b) < (c, d)$  if  $a < b$  or  $a = b$  and  $c < d$ .

To conclude that  $\psi_k[S_{15}] = \psi_k[Q_{15}]$  for  $k = 2, 3, 4$ , we compare the triples  $(d_{i-1}, \mathbf{s}, d_{i+k-1})$  from Theorem 9.4.4 for  $S_{15}, Q_{15}$ . For  $k = 2$  and  $k = 3$ , Tables 9.2 and 9.3 have identical rows 7 and 9, where the triples are ordered for easier comparison as follows. If needed, we swap  $d_{i-1}, d_{i+k-1}$  to make sure that the first entry is not larger

**Table 9.3** **Row 1:** points  $p_i$  from  $Q_{15}$  in Fig. 9.11. **Row 2:** the distances  $d_i$  between successive points of  $Q_{15}$ . **Row 3:** the distances  $d_{[i]}$  are in the increasing order. **Row 4:** the unordered set of these pairs determines the density function  $\psi_1$  by Theorem 9.4.4b. **Row 5:** the pairs are lexicographically ordered for comparison with row 5 in Table 9.2. **Rows 6,8,10:** the unordered sets of these triples determine the density functions  $\psi_2, \psi_3, \psi_4$  by Theorem 9.4.4 for  $k = 2, 3, 4$ . **Rows 7,9,11:** the triples from rows 6,8,10 are ordered for comparison with corresponding rows 7,9,11 in Table 9.2, see Example 9.4.7.

$p_i$	0	1	3	4	6	8	9	12	14
$d_i = p_{i+1} - p_i$	1	2	1	2	2	1	3	2	1
ordered $d_{[i]}$	1	1	1	1	2	2	2	2	3
$(d_{i-1}, d_i)$	(1,1)	(1,2)	(2,1)	(1,2)	(2,2)	(2,1)	(1,3)	(3,2)	(2,1)
ordered $(d_{i-1}, d_i)$	(1,1)	(1,2)	(1,2)	(1,2)	(1,2)	(1,2)	(1,3)	(2,2)	(2,3)
$(d_{i-1}, \mathbf{d}_i, d_{i+1})$	(1,1,2)	(1,2,1)	(2,1,2)	(1,2,2)	(2,2,1)	(2,1,3)	(1,3,2)	(3,2,1)	(2,1,1)
order $(d_{i-1}, \mathbf{d}_i, d_{i+1})$	(1,1,2)	(1,1,2)	(2,1,2)	(2,1,3)	(1,2,1)	(1,2,2)	(1,2,2)	(1,2,3)	(1,3,2)
$(d_{i-1}, \mathbf{s}, d_{i+2})$	(1,3,1)	(1,3,2)	(2,3,2)	(1,4,1)	(2,3,3)	(2,4,2)	(1,5,1)	(3,3,1)	(2,2,2)
order $(d_{i-1}, \mathbf{s}, d_{i+2})$	(2,2,2)	(1,3,1)	(1,3,2)	(1,3,3)	(2,3,2)	(2,3,3)	(1,4,1)	(2,4,2)	(1,5,1)
$(d_{i-1}, \mathbf{s}, d_{i+3})$	(1,4,2)	(1,5,2)	(2,5,1)	(1,5,3)	(2,6,2)	(2,6,1)	(1,6,1)	(3,4,2)	(2,4,1)
order $(d_{i-1}, \mathbf{s}, d_{i+3})$	(1,4,2)	(1,4,2)	(2,4,3)	(1,5,2)	(1,5,2)	(1,5,3)	(1,6,1)	(1,6,2)	(2,6,2)

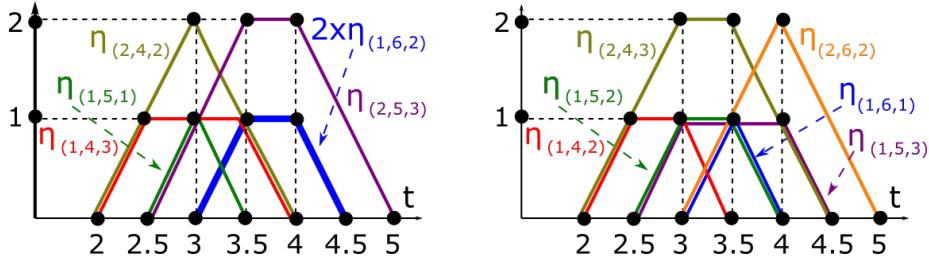
than the last. Then we order by the middle bold number  $\mathbf{s}$ . Finally, we lexicographically order the triples with the same middle value  $s$ .

Final rows 11 of Tables 9.2 and 9.3 look different for  $k = 4$ . More exactly, the rows share three triples  $(1,4,2), (1,5,2), (1,6,4)$ , but the remaining six triples differ. However, the density function  $\psi_4$  is the *sum* of nine trapezium functions. Fig. 9.12 shows that these sums are equal for  $S_{15}, Q_{15}$ . Then the sequences  $S_{15}, Q_{15}$  have identical density functions  $\psi_k$  for  $k = 0, 1, 2, 3, 4$ , hence for all  $k$  by the symmetry and periodicity from Theorem 9.4.5. Fig. 9.13 shows the density functions  $\psi_k$  for  $k = 0, 1, \dots, 9$ .  $\blacklozenge$

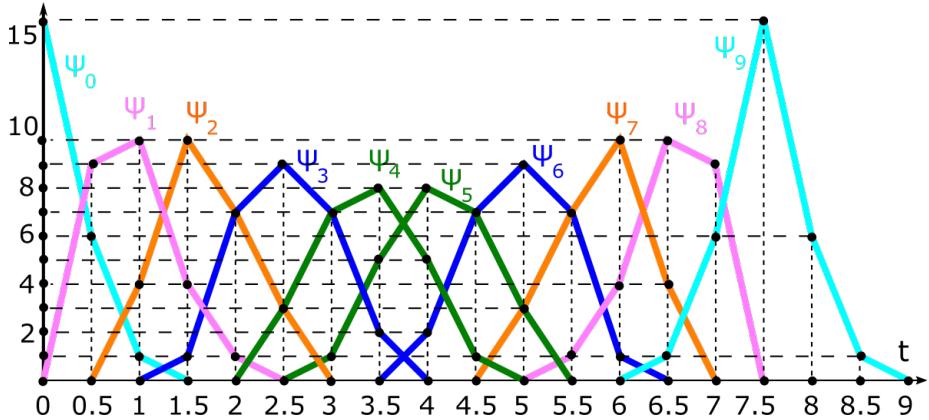
Recall that all indices  $i$  of distances  $d_i$  are considered modulo  $m$ .

**Corollary 9.4.8** ( $k$ -th density  $\rho_k$ , [2, Corollary 11]). For any periodic sequence  $S = \{p_1, \dots, p_m\} + \mathbb{Z}$  with inter-point distances  $d_i = p_{i+1} - p_i$ , where  $i = 1, \dots, m$  and  $p_{m+1} = p_1 + 1$ , the  $k$ -th density  $\rho_k[S] = \int_{-\infty}^{+\infty} \psi_k(t) dt$  defined as the area under the graph of  $\psi_k(t)$  over  $\mathbb{R}$  equals  $\rho_k[S] = \frac{1}{2} \sum_{i=1}^m d_{i-1} d_{i+k-1}$  for any  $k > 0$  and  $\rho_0[S] = \frac{1}{4} \sum_{i=1}^m d_i^2$ .  $\blacksquare$

For  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$ , Corollary 9.4.8 gives  $\rho_0 = \frac{7}{72}$ ,  $\rho_1 = \rho_2 = \frac{11}{12^2}$  as in Fig. 9.10.



**Fig. 9.12** The 4th-density function  $\psi_4[S_{15}]$  includes the six trapezium functions on the left, which are replaced by other six trapezium functions in  $\psi_4[Q_{15}]$  on the right, compare the last rows of Tables 9.2 and 9.3. However, the sums of these six functions are equal, which can be checked at critical radii: both sums of six functions have  $\eta(2.5) = 2$ ,  $\eta(3) = 5$ ,  $\eta(3.5) = 6$ ,  $\eta(4) = 4$ ,  $\eta(4.5) = 1$ . Hence, the sequences  $S_{15}, Q_{15}$  in Fig. 9.11 have identical density functions  $\psi_k$  for all  $k \geq 0$ , see Example 9.4.7.



**Fig. 9.13** The periodic sequences  $S_{15}, Q_{15}$  in Fig. 9.11 have identical density functions  $\psi_k(t)$  for all  $k \geq 0$ . Both axes are scaled by factor 15. Theorem 9.4.5 implies the symmetry  $\psi_k(\frac{15}{2} - t) = \psi_{9-k}(t)$ ,  $t \in [0, \frac{15}{2}]$ , and periodicity  $\psi_9(t + \frac{15}{2}) = \psi_0(t)$ ,  $t \geq 0$ .

**Corollary 9.4.9** (generic completeness of  $\psi_1$ , [2, Corollary 12]). Let  $S \subset \mathbb{R}$  be a sequence with period 1 and  $m$  points  $0 \leq p_1 < \dots < p_m < 1$ . The sequence  $S$  is called *generic* if  $d_i = p_{i+1} - p_i$  are distinct, where  $i = 1, \dots, m$  and  $p_{m+1} = p_1 + 1$ . Then any generic  $S$  can be reconstructed from the 1st density function  $\psi_1[S](t)$  up to isometry in  $\mathbb{R}$ . Hence  $\psi_1(t)$  is a complete isometry invariant for all generic  $S$ . ■

## 9.5 Density functions of periodic sequences of intervals in $\mathbb{R}$

This section follows [3, sections 2-5] by extending density functions to periodic sets of points with radii, including periodic sequences of disjoint intervals in  $\mathbb{R}$ .

**Definition 9.5.1** (density functions for periodic sets of points with radii). Let a periodic set  $S = \Lambda + M \subset \mathbb{R}^n$  have a unit cell  $U$ . For every point  $p \in M$ , fix a radius  $r(p) \geq 0$ . For any integer  $k \geq 0$ , let  $U_k(t)$  be the region within the cell  $U$  covered by exactly  $k$  closed balls  $\bar{B}(p; r(p) + t)$  for  $t \geq 0$  and all points  $p \in M$  and their translations by  $\Lambda$ . The  $k$ -th *density* function  $\psi_k[S](t) = \text{vol}[U_k(t)]/\text{vol}[U]$  is the fractional volume of the  $k$ -fold intersections of these balls within  $U$ . ■

In Definition 9.5.1, the balls are growing at all points of  $S$ , because centers  $p \in M$  are translated by all lattice vectors  $v \in \Lambda$ . The initially different radii  $r_i$  are motivated by real lengths of continuous events in periodic time series for  $n = 1$  and also by atomic radii of different chemical elements for  $n = 3$ . Another (possibly, non-linear) growth of radii lead to more complicated density functions.

The density  $\psi_k[S](t)$  can be interpreted as the probability that a random (uniformly chosen in  $U$ ) point  $q$  is at a maximum distance  $t$  to exactly  $k$  balls with initial radii  $r(p)$  and all centers  $p \in S$ . For  $k = 0$ , the 0-th density  $\psi_0[S](t)$  measures the fractional volume of the empty space not covered by any expanding balls  $\bar{B}(p; r(p) + t)$ .

For  $k = 1$  and small  $t > 0$  while all equal-sized balls  $\bar{B}(p; t)$  remain disjoint, the 1st density  $\psi_1[S](t)$  increases proportionally to  $t^n$  but later reaches a maximum and eventually drops back to 0 when all points of  $\mathbb{R}^n$  are covered by at least two balls.

The original densities helped find a missing crystal in the Cambridge Structural Database, which was accidentally confused with a slight perturbation (measured at a different temperature) of another crystal (polymorph) with the same chemical composition, see [5, section 7].

The new weighted case with radii  $r(p) \geq 0$  in Definition 9.5.1 is even more practically important due to different Van der Waals radii, which are individually defined for all chemical elements.

The key advantage of density functions over other isometry invariants of periodic crystals (such as symmetries or conventional representations based on a geometry of a minimal cell) is their continuity under perturbations. The only limitation is the infinite size of densities  $\psi_k(t)$  due to the unbounded parameters: integer index  $k \geq 0$  and continuous radius  $t \geq 0$ .

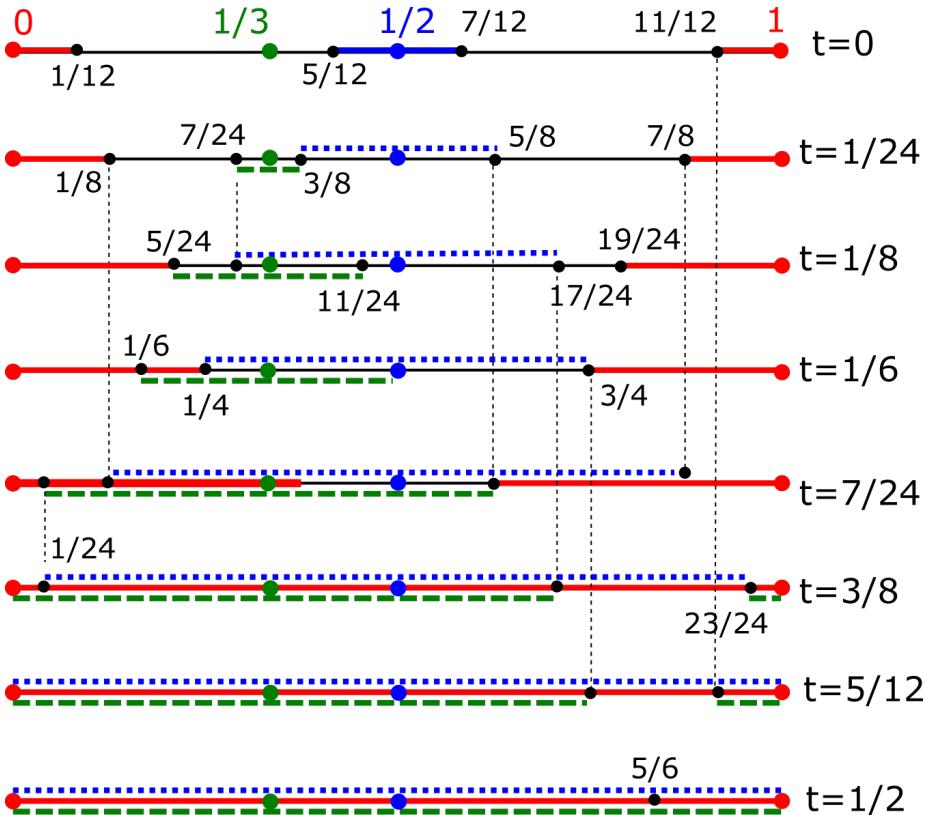
Theorem 9.5.3 will explicitly describing the 0-th density function  $\psi_0[S](t)$  for any periodic sequence  $S \subset \mathbb{R}$  of intervals. All intervals are considered closed and called *disjoint* if their open interiors (not endpoints) have no common points.

For convenience, scale any periodic sequence  $S$  to period 1 so that  $S$  is given by points  $0 \leq p_1 < \dots < p_m < 1$  with radii  $r_1, \dots, r_m$ , respectively. Since the expanding balls in  $\mathbb{R}$  are growing intervals, volumes of their intersections linearly change with respect to the variable radius  $t$ . Hence any density function  $\psi_k(t)$  is piecewise linear and uniquely determined by *corner* points  $(a_j, b_j)$  where the gradient of  $\psi_k(t)$  changes. To illustrate Theorem 9.5.3, we consider Example 9.5.2 for the simple sequence  $S$ .

**Example 9.5.2** (0-th density function  $\psi_0$ ). Let the periodic sequence  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$  have three points  $p_1 = 0, p_2 = \frac{1}{3}, p_3 = \frac{1}{2}$  of radii  $r_1 = \frac{1}{12}, r_2 = 0, r_3 = \frac{1}{12}$ , respectively. Fig. 9.14 shows each point  $p_i$  and its growing interval

$$L_i(t) = [(p_i - r_i) - t, (p_i + r_i) + t] \text{ of the length } 2r_i + 2t$$

for  $i = 1, 2, 3$  in its own color: red, green, blue.

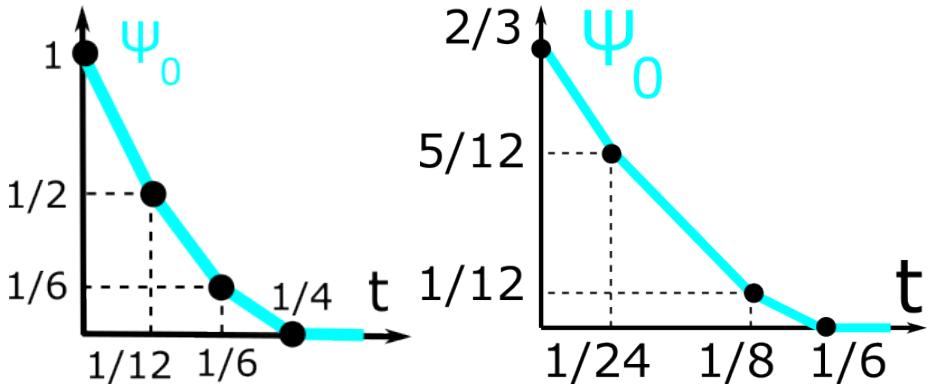


**Fig. 9.14** The sequence  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$  has the points of weights  $\frac{1}{12}, 0, \frac{1}{12}$ , respectively. The intervals around the red point  $0 \equiv 1 \pmod{1}$ , green point  $\frac{1}{3}$ , blue point  $\frac{1}{2}$  have the same color for various radii  $t$ , see Examples 9.5.2, 9.5.5, 9.5.8.

By Definition 9.5.1, each density function  $\psi_k[S](t)$  measures a fractional length covered by exactly  $k$  intervals within the unit cell  $[0, 1]$ . It is convenient to periodically map the endpoints of each growing interval to the unit cell  $[0, 1]$ .

For instance, the interval  $[-\frac{1}{12} - t, \frac{1}{12} + t]$  of the point  $p_1 = 0 \equiv 1 \pmod{1}$  maps to the red intervals  $[0, \frac{1}{12} + t] \cup [\frac{11}{12} - t, 1]$  shown by solid red lines in Fig. 9.14. The same image shows the green interval  $[\frac{1}{3} - t, \frac{1}{3} + t]$  by dashed lines and the blue interval  $[\frac{5}{12} - t, \frac{7}{12} + t]$  by dotted lines.

At the moment  $t = 0$ , since the starting intervals are disjoint, they cover the length  $l = 2(\frac{1}{12} + 0 + \frac{1}{12}) = \frac{1}{3}$ . The non-covered part of  $[0, 1]$  has length  $1 - \frac{1}{3} = \frac{2}{3}$ . So the graph of  $\psi_0(t)$  at  $t = 0$  starts from the point  $(0, \frac{2}{3})$ , see Fig. 9.15 (right).



**Fig. 9.15** **Left:** the 0-th density function  $\psi_0(t)$  for the 1-period sequence  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$  with radii 0. **Right:** the 0-th density  $\psi_0(t)$  for the 1-period sequence  $S$  whose points  $0, \frac{1}{3}, \frac{1}{2}$  have radii  $\frac{1}{12}, 0, \frac{1}{12}$ , respectively, see Example 9.5.2.

At the first critical moment  $t = \frac{1}{24}$  when the green and blue intervals collide at  $p = \frac{3}{8}$ , only the intervals  $[\frac{1}{8}, \frac{7}{24}] \cup [\frac{5}{8}, \frac{7}{8}]$  of total length  $\frac{5}{12}$  remain uncovered. Hence  $\psi_0(t)$  linearly drops to the point  $(\frac{1}{12}, \frac{5}{12})$ . At the next critical moment  $t = \frac{1}{8}$  when the red and green intervals collide at  $p = \frac{5}{24}$ , only the interval  $[\frac{17}{24}, \frac{19}{24}]$  of length  $\frac{1}{12}$  remain uncovered, so  $\psi_0(t)$  continues to  $(\frac{1}{8}, \frac{1}{12})$ .

The graph of  $\psi_0(t)$  finally returns to the  $t$ -axis at the point  $(\frac{1}{6}, 0)$  and remains there for  $t \geq \frac{1}{6}$ . The piecewise linear behaviour of  $\psi_0(t)$  can be described by specifying the corner points in Fig. 9.15:  $(0, \frac{2}{3}), (\frac{1}{24}, \frac{5}{12}), (\frac{1}{8}, \frac{1}{12}), (\frac{1}{6}, 0)$ . ◆

Theorem 9.5.3 extends Example 9.5.2 to any periodic sequence  $S$  and implies that the 0-th density function  $\psi_0(t)$  is uniquely determined by the ordered gap lengths between successive intervals.

**Theorem 9.5.3** (description of  $\psi_0$ , [3, Theorem 3.2]). Let a periodic sequence  $S = \{p_1, \dots, p_m\} + \mathbb{Z}$  consist of disjoint intervals with centers  $0 \leq p_1 < \dots < p_m < 1$  and radii  $r_1, \dots, r_m \geq 0$ . Consider the total length  $l = 2 \sum_{i=1}^m r_i$  and gaps between successive intervals  $g_i = (p_i - r_i) - (p_{i-1} + r_{i-1})$ , where  $i = 1, \dots, m$  and  $p_0 = p_m - 1$ ,  $r_0 = r_m$ . Put the gaps in increasing order:  $g_{[1]} \leq g_{[2]} \leq \dots \leq g_{[m]}$ . Then the 0-th density  $\psi_0[S](t)$  is piecewise linear with the following (unordered) corner points:  $(0, 1 - l)$  and  $(\frac{g_{[i]}}{2}, 1 - l - \sum_{j=1}^{i-1} g_{[j]} - (m - i + 1)g_{[i]})$  for  $i = 1, \dots, m$ , so the last corner is  $(\frac{g_{[m]}}{2}, 0)$ .

If any corners are repeated, e.g. if  $g_{[i-1]} = g_{[i]}$ , these corners are collapsed into one. ■

Example 9.5.4 applies Theorem 9.5.3 to get  $\psi_0$  for the sequence  $S$  in Example 9.5.2.

**Example 9.5.4** (using Theorem 9.5.3). The sequence  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$  in Example 9.5.2 with points  $p_1 = 0, p_2 = \frac{1}{3}, p_3 = \frac{1}{2}$  of radii  $r_1 = \frac{1}{12}, r_2 = 0, r_3 = \frac{1}{12}$ , respectively, has  $l = 2(r_1 + r_2 + r_3) = \frac{1}{3}$  and the initial gaps between successive intervals

$$g_1 = p_1 - r_1 - p_3 - r_3 = (1 - \frac{1}{12}) - (\frac{1}{2} + \frac{1}{12}) = \frac{1}{3},$$

$$g_2 = p_2 - r_2 - p_1 - r_1 = (\frac{1}{3} - 0) - (0 + \frac{1}{12}) = \frac{1}{4},$$

$$g_3 = p_3 - r_3 - p_2 - r_2 = (\frac{1}{2} - \frac{1}{12}) - (\frac{1}{3} + 0) = \frac{1}{12}.$$

Order the gaps:  $g_{[1]} = \frac{1}{12} < g_{[2]} = \frac{1}{4} < g_{[3]} = \frac{1}{3}$ . Then

$$1 - l = 1 - \frac{1}{3} = \frac{2}{3},$$

$$1 - l - 3g_{[1]} = \frac{2}{3} - \frac{3}{12} = \frac{5}{12},$$

$$1 - l - g_{[1]} - 2g_{[2]} = \frac{2}{3} - \frac{1}{12} - \frac{2}{4} = \frac{1}{12},$$

$$1 - l - g_{[1]} - g_{[2]} - g_{[3]} = \frac{2}{3} - \frac{1}{12} - \frac{1}{4} - \frac{1}{3} = 0.$$

By Theorem 9.5.3  $\psi_0(t)$  has the corner points

$$(0, 1 - l) = (0, \frac{2}{3}),$$

$$(\frac{1}{2}g_{[1]}, 1 - l - 3g_{[1]}) = (\frac{1}{24}, \frac{5}{12}),$$

$$(\frac{1}{2}g_{[2]}, 1 - l - g_{[1]} - 2g_{[2]}) = (\frac{1}{8}, \frac{1}{12}),$$

$$(\frac{1}{2}g_{[3]}, 1 - l - g_{[1]} - g_{[2]} - g_{[3]}) = (\frac{1}{6}, 0).$$

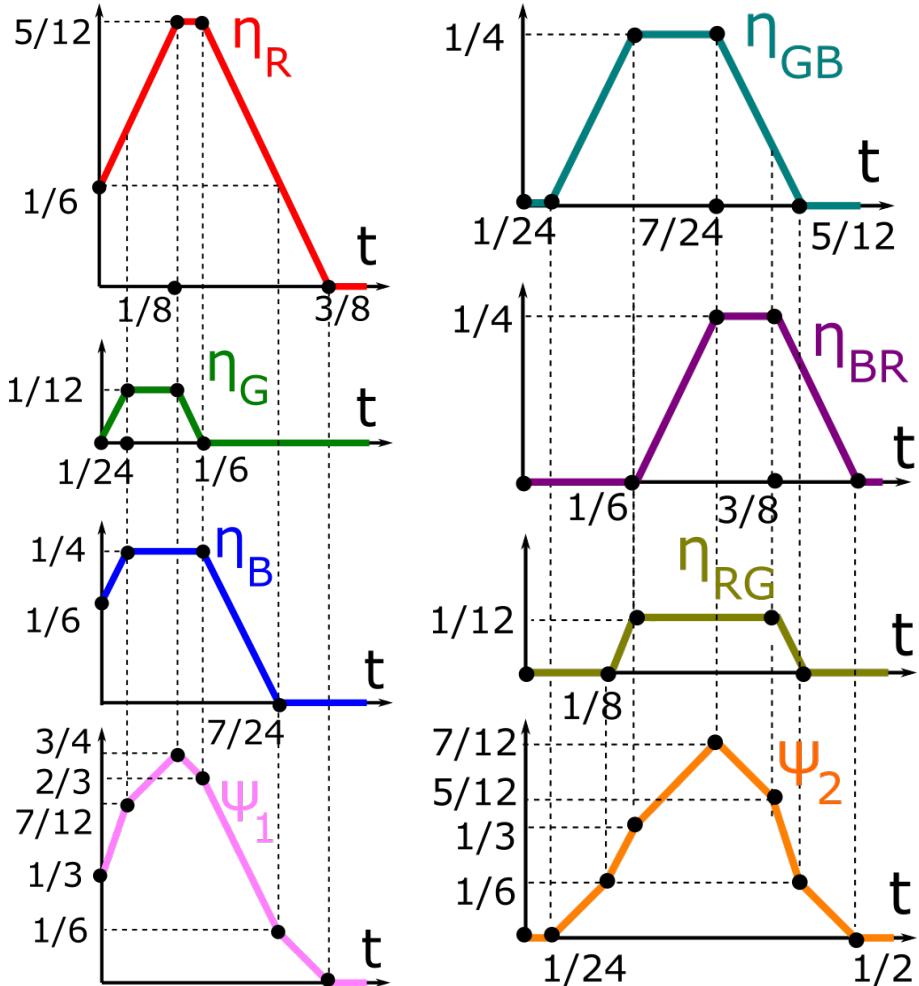
See the graph of the 0-th density  $\psi_0(t)$  in Fig. 9.15. ◆

By Theorem 9.5.3 any 0-th density function  $\psi_0(t)$  is uniquely determined by the (unordered) set of gap lengths between successive intervals. Hence we can re-order these intervals without changing  $\psi_0(t)$ . For instance, the periodic sequence  $Q = \{0, \frac{1}{2}, \frac{2}{3}\} + \mathbb{Z}$  with points  $0, \frac{1}{2}, \frac{2}{3}$  of weights  $\frac{1}{12}, \frac{1}{12}, 0$  has the same set ordered gaps  $g_{[1]} = \frac{1}{12}, d_{[2]} = \frac{1}{3}, d_{[3]} = \frac{1}{2}$  as the periodic sequence  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$  in Example 9.5.2.

The above sequences  $S, Q$  are related by the mirror reflection  $t \mapsto 1 - t$ . One can easily construct many non-isometric sequences with  $\psi_0[S](t) = \psi_0[Q](t)$ . For any  $1 \leq i \leq m - 3$ , the sequences  $S_{m,i} = \{0, 2, 3, \dots, i+2, i+4, i+5, \dots, m+2\} + (m+2)\mathbb{Z}$  have the same interval lengths  $d_{[1]} = \dots = d_{[m-2]} = 1, d_{[m-1]} = d_{[m]} = 2$  but are not

related by isometry (translations and reflections in  $\mathbb{R}$ ) because the intervals of length 2 are separated by  $i - 1$  intervals of length 1 in  $S_{m,i}$ .

Theorem 9.5.6 will explicitly describe the 1st density function  $\psi_1[S](t)$  for any periodic sequence  $S$  of disjoint intervals. To illustrate Theorem 9.5.6, Example 9.5.5 finds  $\psi_1[S]$  for the sequence  $S$  from Example 9.5.2.



**Fig. 9.16** **Left:** the trapezium functions  $\eta_R$ ,  $\eta_G$ ,  $\eta_B$  and the 1st density function  $\psi_1(t)$  for the 1-period sequence  $S$  whose points  $0, \frac{1}{3}, \frac{1}{2}$  have radii  $\frac{1}{12}, 0, \frac{1}{12}$ , see Example 9.5.5. **Right:** The trapezium functions  $\eta_{GB}$ ,  $\eta_{BR}$ ,  $\eta_{RG}$  and the 2nd density function  $\psi_2(t)$  for the 1-period sequence  $S$  whose points  $0, \frac{1}{3}, \frac{1}{2}$  have radii  $\frac{1}{12}, 0, \frac{1}{12}$ , see Example 9.5.8.

**Example 9.5.5** ( $\psi_1$  for the sequence  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$ ). The 1st density function  $\psi_1(t)$  can be obtained as a sum of the three *trapezium* functions  $\eta_R, \eta_G, \eta_B$ , each measuring the length of a region covered by a single interval of one color, see Fig. 9.14.

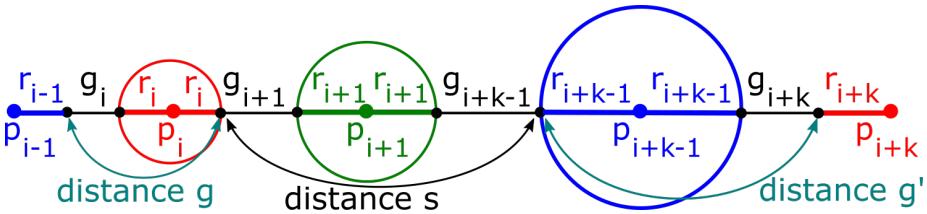
At the initial moment  $t = 0$ , the red intervals  $[0, \frac{1}{12}] \cup [\frac{11}{12}, 1]$  have the total length  $\eta_R(0) = \frac{1}{6}$ . These red intervals  $[0, \frac{1}{12} + t] \cup [\frac{11}{12} - t, 1]$  for  $t \in [0, \frac{1}{8}]$  grow until they touch the green interval  $[\frac{7}{24}, \frac{3}{8}]$  and have the total length  $\eta_R(\frac{1}{8}) = \frac{1}{6} + \frac{2}{8} = \frac{5}{12}$  in the second picture of Fig. 9.14. So the graph of the red length  $\eta_R(t)$  linearly grows with gradient 2 from the point  $(0, \frac{1}{6})$  to the corner point  $(\frac{1}{8}, \frac{5}{12})$ .

For  $t \in [\frac{1}{8}, \frac{1}{6}]$ , the left red interval is shrinking at the same rate (due to the overlapping green interval) as the right red interval continues to grow until  $t = \frac{1}{6}$ , when it touches the blue interval  $[\frac{1}{4}, \frac{3}{4}]$ . Hence the graph of  $\eta_R(t)$  remains constant for  $t \in [\frac{1}{8}, \frac{1}{6}]$  up to the corner point  $(\frac{1}{6}, \frac{5}{12})$ . After that, the graph of  $\eta_R(t)$  linearly decreases (with gradient  $-2$ ) until all red intervals are fully covered by the green and blue intervals at moment  $t = \frac{3}{8}$ , see the 6th picture in Fig. 9.14.

Hence, the trapezium function  $\eta_R$  has the piecewise linear graph through the corner points  $(0, \frac{1}{6}), (\frac{1}{8}, \frac{5}{12}), (\frac{1}{6}, \frac{5}{12}), (\frac{3}{8}, 0)$ . After that,  $\eta_R(t) = 0$  remains constant for  $t \geq \frac{3}{8}$ . Fig. 9.16 shows the graphs of  $\eta_R, \eta_G, \eta_B$  and  $\psi_1 = \eta_R + \eta_G + \eta_B$ .  $\blacklozenge$

Theorem 9.5.6 extends Example 9.5.5 and proves that any  $\psi_1(t)$  is a sum of trapezium functions whose corners are explicitly described. We consider any index  $i = 1, \dots, m$  (of a point  $p_i$  or a gap  $g_i$ ) modulo  $m$  so that  $m + 1 \equiv 1 \pmod{m}$ .

**Theorem 9.5.6** (description of  $\psi_1$ , [3, Theorem 4.2]). Let a periodic sequence  $S = \{p_1, \dots, p_m\} + \mathbb{Z}$  consist of disjoint intervals with centers  $0 \leq p_1 < \dots < p_m < 1$  and radii  $r_1, \dots, r_m \geq 0$ , respectively. Consider the gaps  $g_i = (p_i - r_i) - (p_{i-1} + r_{i-1})$ , between successive intervals, where  $i = 1, \dots, m$  and  $p_0 = p_m - 1, r_0 = r_m$ . Then the 1st density  $\psi_1(t)$  is the sum of  $m$  trapezium functions  $\eta_i, i = 1, \dots, m$ , with the corners  $(0, 2r_i), (\frac{g_i}{2}, g + 2r_i), (\frac{g_{i+1}}{2}, g + 2r_i), (\frac{g_i + g_{i+1}}{2} + r_i, 0)$ , where  $g = \min\{g_i, g_{i+1}\}$ . Hence,  $\psi_1(t)$  is determined by the unordered set of unordered pairs  $(g_i, g_{i+1}), i = 1, \dots, m$ .  $\blacksquare$



**Fig. 9.17** The distances  $g, s, g'$  between line intervals in Theorems 9.5.6 and 9.5.9, shown for  $k = 3$ .

Example 9.5.7 applies Theorem 9.5.6 to get  $\psi_1$  for the sequence  $S$  in Example 9.5.5.

**Example 9.5.7** (using Theorem 9.5.6 for  $\psi_1$ ). The sequence  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$  in Example 9.5.5 with points  $p_1 = 0, p_2 = \frac{1}{3}, p_3 = \frac{1}{2}$  of radii  $r_1 = \frac{1}{12}, r_2 = 0, r_3 = \frac{1}{12}$ ,

respectively, has the initial gaps between successive intervals  $g_1 = \frac{1}{3}$ ,  $g_2 = \frac{1}{4}$ ,  $g_3 = \frac{1}{12}$ , see all the computations in Example 9.5.4.

**Case (R).** In Theorem 9.5.6 for the trapezium function  $\eta_R = \eta_1$  measuring the fractional length covered only by the red interval, we set  $i = 1$ . Then  $r_i = \frac{1}{12}$ ,  $g_i = \frac{1}{3}$  and  $g_{i+1} = \frac{1}{4}$ ,

$$\frac{g_i+g_{i+1}}{2} + r_i = \frac{1}{2}(\frac{1}{3} + \frac{1}{4}) + \frac{1}{12} = \frac{3}{8},$$

$$g = \min\{g_i, g_{i+1}\} = \frac{1}{4}, \quad g + 2r_i = \frac{1}{4} + \frac{2}{12} = \frac{5}{12}.$$

Then  $\eta_R = \eta_1$  has the following corner points:

$$(0, 2r_i) = (0, \frac{1}{6}), \quad (\frac{g_i}{2}, g + 2r_i) = (\frac{1}{6}, \frac{5}{12}),$$

$$(\frac{g_{i+1}}{2}, g + 2r_i) = (\frac{1}{8}, \frac{5}{12}),$$

$$(\frac{g_i+g_{i+1}}{2} + r_i, 0) = (\frac{3}{8}, 0),$$

where the two middle corners are accidentally swapped due to  $g_i > g_{i+1}$  but they define the same trapezium function as in the first picture of Fig. 9.16.

**Case (G).** In Theorem 9.5.6 for the trapezium function  $\eta_G = \eta_2$  measuring the fractional length covered only by the green interval, we set  $i = 2$ . Then  $r_i = 0$ ,  $g_i = \frac{1}{4}$  and  $g_{i+1} = \frac{1}{12}$ ,  $\frac{g_i+g_{i+1}}{2} + r_i = \frac{1}{2}(\frac{1}{4} + \frac{1}{12}) + 0 = \frac{1}{6}$ ,  $g = \min\{g_i, g_{i+1}\} = \frac{1}{12}$ ,  $g + 2r_i = \frac{1}{12} + 0 = \frac{1}{12}$ .

Then  $\eta_G = \eta_2$  has the following corner points exactly as shown in the second picture of Fig. 9.16 (left):

$$(0, 2r_i) = (0, 0), \quad (\frac{g_i}{2}, g + 2r_i) = (\frac{1}{8}, \frac{1}{12}),$$

$$(\frac{g_{i+1}}{2}, g + 2r_i) = (\frac{1}{24}, \frac{5}{12}), \quad (\frac{g_i+g_{i+1}}{2} + r_i, 0) = (\frac{1}{6}, 0).$$

**Case (B).** In Theorem 9.5.6 for the trapezium function  $\eta_B = \eta_3$  measuring the fractional length covered only by the blue interval, we set  $i = 3$ . Then  $r_i = \frac{1}{12}$ ,  $g_i = \frac{1}{12}$  and  $g_{i+1} = \frac{1}{3}$ ,

$$\frac{g_i+g_{i+1}}{2} + r_i = \frac{1}{2}(\frac{1}{12} + \frac{1}{3}) + \frac{1}{12} = \frac{7}{24}, \quad g = \min\{g_i, g_{i+1}\} = \frac{1}{12}, \quad g + 2r_i = \frac{1}{12} + \frac{2}{12} = \frac{1}{4}.$$

Then  $\eta_B = \eta_3$  has the following corner points:

$$(0, 2r_i) = (0, \frac{1}{6}), \quad (\frac{g_i}{2}, g + 2r_i) = (\frac{1}{24}, \frac{1}{4}),$$

$$(\frac{g_{i+1}}{2}, g + 2r_i) = (\frac{1}{6}, \frac{1}{4}), \quad (\frac{g_i+g_{i+1}}{2} + r_i, 0) = (\frac{7}{24}, 0)$$

exactly as shown in the third picture of Fig. 9.16.  $\blacklozenge$

Theorem 9.5.9 describing the  $k$ -th density function  $\psi_k[S](t)$  for any  $k \geq 2$  and a periodic sequence  $S$  of disjoint intervals. To illustrate Theorem 9.5.9, Example 9.5.8 computes  $\psi_2[S]$  for the periodic sequence  $S$  from Example 9.5.2.

**Example 9.5.8** ( $\psi_2$  for the sequence  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$ ). The density  $\psi_2(t)$  can be found as the sum of the *trapezium* functions  $\eta_{GB}, \eta_{BR}, \eta_{RG}$ , each measuring the length of a double intersection, see Fig. 9.14. For the green interval  $[\frac{1}{3} - t, \frac{1}{3} + t]$  and the blue interval  $[\frac{5}{12} - t, \frac{7}{12} + t]$ , the graph of the function  $\eta_{GB}(t)$  is piecewise linear and starts at the point  $(\frac{1}{24}, 0)$  because these intervals touch at  $t = \frac{1}{24}$ .

The green-blue intersection  $[\frac{5}{12} - t, \frac{1}{3} + t]$  grows until  $t = \frac{1}{6}$ , when the resulting interval  $[\frac{1}{4}, \frac{1}{2}]$  touches the red interval on the left. At the same time, the graph of  $\eta_{GB}(t)$  is linearly growing (with gradient 2) to the corner  $(\frac{1}{6}, \frac{1}{4})$ , see Fig. 9.16.

For  $t \in [\frac{1}{6}, \frac{7}{24}]$ , the green-blue intersection interval becomes shorter on the left, but grows at the same rate on the right until  $t = \frac{7}{24}$  when  $[\frac{1}{8}, \frac{5}{8}]$  touches the red interval  $[\frac{5}{8}, 1]$  on the right, see the 5th picture in Fig. 9.14. So the graph of  $\eta_{GB}(t)$  remains constant up to the point  $(\frac{7}{24}, \frac{1}{4})$ .

For  $t \in [\frac{7}{24}, \frac{5}{12}]$  the green-blue intersection interval is shortening from both sides. So the graph of  $\eta_{GB}(t)$  linearly decreases (with gradient -2) and returns to the  $t$ -axis at the corner  $(\frac{5}{12}, 0)$ , then remains constant  $\eta_{GB}(t) = 0$  for  $t \geq \frac{5}{12}$ . Fig. 9.16 shows all trapezium functions for double intersections and  $\psi_2 = \eta_{GB} + \eta_{BR} + \eta_{RG}$ .  $\blacklozenge$

**Theorem 9.5.9** (description of  $\psi_k$  for  $k \geq 2$ , [3, Theorem 5.2]). Let a periodic sequence  $S = \{p_1, \dots, p_m\} + \mathbb{Z}$  consist of disjoint intervals with centers  $0 \leq p_1 < \dots < p_m < 1$  and radii  $r_1, \dots, r_m \geq 0$ , respectively. Consider the gaps  $g_i = (p_i - r_i) - (p_{i-1} + r_{i-1})$  between the successive intervals of  $S$ , where  $i = 1, \dots, m$  and  $p_0 = p_m - 1, r_0 = r_m$ .

For  $k \geq 2$ , the density function  $\psi_k(t)$  equals the sum of  $m$  *trapezium* functions  $\eta_{k,i}(t)$ ,  $i = 1, \dots, m$ , each having the corner points  $(\frac{s}{2}, 0), (\frac{g+s}{2}, g), (\frac{s+g'}{2}, g), (\frac{g+s+g'}{2}, 0)$ , where  $g, g'$  are the minimum and maximum values in the pair  $\{g_i + 2r_i, g_{i+k} + 2r_{i+k-1}\}$ , and  $s = \sum_{j=i+1}^{i+k-1} g_j + 2 \sum_{j=i+1}^{i+k-2} r_j$ . For  $k = 2$ , we have  $s = g_{i+1}$ . Hence,  $\psi_k(t)$  is determined by the unordered set of the ordered tuples  $(g, s, g')$ ,  $i = 1, \dots, m$ .  $\blacksquare$

Example 9.5.10 applies Theorem 9.5.9 to get  $\psi_2$  for the sequence  $S$  in Example 9.5.2.

**Example 9.5.10** (using Theorem 9.5.9 for  $\psi_2$ ). The sequence  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$  in Example 9.5.5 with points  $p_1 = 0, p_2 = \frac{1}{3}, p_3 = \frac{1}{2}$  of radii  $r_1 = \frac{1}{12}, r_2 = 0, r_3 = \frac{1}{12}$ , respectively, has the initial gaps  $g_1 = \frac{1}{3}, g_2 = \frac{1}{4}, g_3 = \frac{1}{12}$ , see Example 9.5.4.

In Theorem 9.5.9, the 2nd density function  $\psi_2[S](t)$  is expressed as a sum of the trapezium functions computed via their corners below.

**Case (GB).** For the function  $\eta_{GB}$  measuring the double intersections of the green and blue intervals centered at  $p_2 = p_i$  and  $p_3 = p_{i+k-1}$ , we set  $k = 2$  and  $i = 2$ . Then we have the radii  $r_i = 0$  and  $r_{i+1} = \frac{1}{12}$ , the gaps  $g_i = \frac{1}{4}, g_{i+1} = \frac{1}{12}, g_{i+2} = \frac{1}{3}$ , and the sum

$s = g_{i+1} = \frac{1}{12}$ . The pair  $\{g_i + 2r_i, g_{i+2} + 2r_{i+1}\} = \{\frac{1}{4} + 0, \frac{1}{3} + \frac{2}{12}\}$  has the minimum value  $g = \frac{1}{4}$  and maximum value  $g' = \frac{1}{2}$ . Then  $\eta_{2,2}[S](t) = \eta_{GB}$  has the following corners as in the top picture of Fig. 9.16 (right):

$$\begin{aligned} (\frac{s}{2}, 0) &= (\frac{1}{24}, 0), \\ (\frac{g+s}{2}, g) &= (\frac{1}{2}(\frac{1}{4} + \frac{1}{12}), \frac{1}{4}) = (\frac{1}{6}, \frac{1}{4}), \\ (\frac{s+g'}{2}, g) &= (\frac{1}{2}(\frac{1}{12} + \frac{1}{2}), \frac{1}{4}) = (\frac{7}{24}, \frac{1}{4}), \\ (\frac{g+s+g'}{2}, 0) &= (\frac{1}{2}(\frac{1}{4} + \frac{1}{12} + \frac{1}{2}), 0) = (\frac{5}{12}, 0). \end{aligned}$$

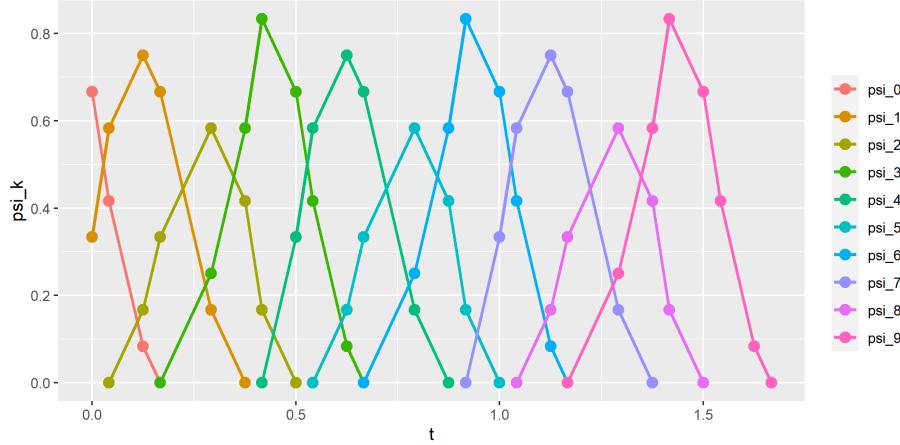
**Case (BR).** For the trapezium function  $\eta_{BR}$  measuring the double intersections of the blue and red intervals centered at  $p_3 = p_i$  and  $p_1 = p_{i+k-1}$ , we set  $k = 2$  and  $i = 3$ . Then we have the radii  $r_i = \frac{1}{12} = r_{i+1}$ , the gaps  $g_i = \frac{1}{12}$ ,  $g_{i+1} = \frac{1}{3}$ ,  $g_{i+2} = \frac{1}{4}$ , and  $s = g_{i+1} = \frac{1}{3}$ . The pair  $\{g_i + 2r_i, g_{i+2} + 2r_{i+1}\} = \{\frac{1}{12} + \frac{2}{12}, \frac{1}{4} + \frac{2}{12}\}$  has the minimum  $g = \frac{1}{4}$  and maximum  $g' = \frac{5}{12}$ . Then  $\eta_{2,3}[S](t) = \eta_{BR}$  has the following corners as expected in the second picture of Fig. 9.16 (right):

$$\begin{aligned} (\frac{s}{2}, 0) &= (\frac{1}{6}, 0), \\ (\frac{g+s}{2}, g) &= (\frac{1}{2}(\frac{1}{4} + \frac{1}{3}), \frac{1}{4}) = (\frac{7}{24}, \frac{1}{4}), \\ (\frac{s+g'}{2}, g) &= (\frac{1}{2}(\frac{1}{3} + \frac{5}{12}), \frac{1}{4}) = (\frac{3}{8}, \frac{1}{4}), \\ (\frac{g+s+g'}{2}, 0) &= (\frac{1}{2}(\frac{1}{4} + \frac{1}{3} + \frac{5}{12}), 0) = (\frac{1}{2}, 0). \end{aligned}$$

**Case (RG).** For the trapezium function  $\eta_{RG}$  measuring the double intersections of the red and green intervals centered at  $p_1 = p_i$  and  $p_2 = p_{i+k-1}$ , we set  $k = 2$  and  $i = 1$ . Then we have the radii  $r_i = \frac{1}{12}$  and  $r_{i+1} = 0$ , the gaps  $g_i = \frac{1}{3}$ ,  $g_{i+1} = \frac{1}{4}$ ,  $g_{i+2} = \frac{1}{12}$ , and  $s = g_{i+1} = \frac{1}{4}$ . The pair  $\{g_i + 2r_i, g_{i+2} + 2r_{i+1}\} = \{\frac{1}{3} + \frac{2}{12}, \frac{1}{12} + 0\}$  has the minimum  $g = \frac{1}{12}$  and maximum  $g' = \frac{1}{2}$ . Then  $\eta_{2,1}[S](t) = \eta_{RG}$  has the following corners:

$$\begin{aligned} (\frac{s}{2}, 0) &= (\frac{1}{8}, 0), \\ (\frac{g+s}{2}, g) &= (\frac{1}{2}(\frac{1}{12} + \frac{1}{4}), \frac{1}{12}) = (\frac{1}{6}, \frac{1}{12}), \\ (\frac{s+g'}{2}, g) &= (\frac{1}{2}(\frac{1}{4} + \frac{1}{2}), \frac{1}{12}) = (\frac{3}{8}, \frac{1}{12}), \\ (\frac{g+s+g'}{2}, 0) &= (\frac{1}{2}(\frac{1}{12} + \frac{1}{4} + \frac{1}{2}), 0) = (\frac{5}{12}, 0). \end{aligned}$$

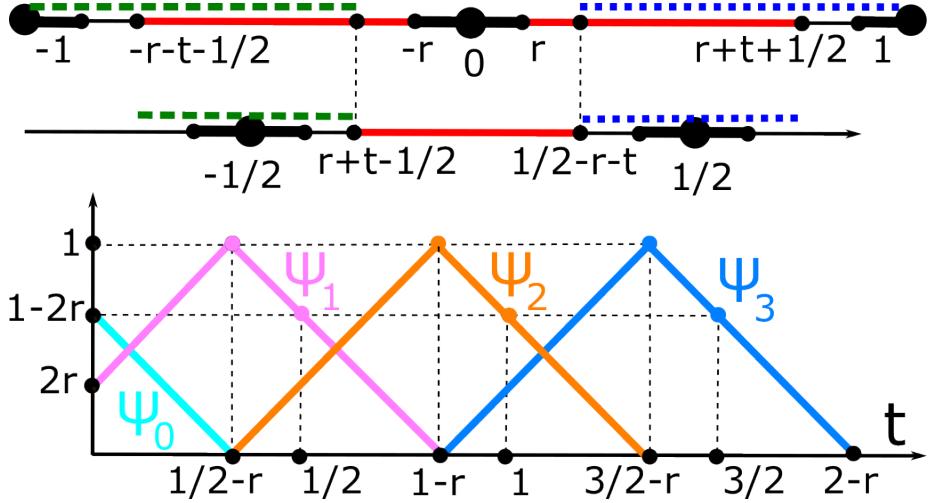
as expected in the third picture of Fig. 9.16 (right). ◆



**Fig. 9.18** The densities  $\psi_k$ ,  $k = 0, \dots, 9$  for the 1-period sequence  $S$  whose points  $0, \frac{1}{3}, \frac{1}{2}$  have radii  $\frac{1}{12}, 0, \frac{1}{12}$ , respectively. The densities  $\psi_0, \psi_1, \psi_2$  are described in Examples 9.5.2, 9.5.5, 9.5.8 and determine all other densities by periodicity in Theorem 9.6.2.

## 9.6 Properties of density functions of periodic sequences of intervals

All results in this section have detailed proofs in [3, sections 6]. Now we study the periodicity of the sequence  $\{\psi_k\}$  with respect to the index  $k \geq 0$  in Theorem 9.6.2, which was a bit unexpected from Definition 9.5.1. We start with the simpler example for the familiar 3-point sequence in Fig. 9.14.



**Fig. 9.19** **Top:** Example 9.6.3 illustrates the proof of Theorem 9.6.2 for  $m = 1$ . **Bottom:** the density functions  $\psi_k$  of  $S = \mathbb{Z}$  whose points have a radius  $0 < r < \frac{1}{4}$  satisfy the periodicity  $\psi_{k+1}(t + \frac{1}{2}) = \psi_k(t)$  for any  $k \geq 0$  and  $t \geq 0$ .

**Example 9.6.1** (periodicity of  $\psi_k$  in the index  $k$ ). Let the periodic sequence  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$  have three points  $p_1 = 0, p_2 = \frac{1}{3}, p_3 = \frac{1}{2}$  of radii  $r_1 = \frac{1}{12}, r_2 = 0, r_3 = \frac{1}{12}$ , respectively. The initial intervals  $L_1(0) = [-\frac{1}{12}, \frac{1}{12}], L_2(0) = [\frac{1}{3}, \frac{1}{3}], L_3(0) = [\frac{5}{12}, \frac{7}{12}]$  have the 0-fold intersection measured by  $\psi_0(0) = \frac{2}{3}$  and the 1-fold intersection measured by  $\psi_1(0) = \frac{1}{3}$ , see Fig. 9.15 and 9.16.

By the time  $t = \frac{1}{2}$  the initial intervals will grow to  $L_1(\frac{1}{2}) = [-\frac{7}{12}, \frac{7}{12}], L_2(\frac{1}{2}) = [-\frac{1}{6}, \frac{5}{6}], L_3(\frac{1}{2}) = [-\frac{1}{12}, \frac{13}{12}]$ . The grown intervals at the radius  $t = \frac{1}{2}$  have the 3-fold intersection  $[-\frac{1}{12}, \frac{7}{12}]$  of the length  $\psi_3(\frac{1}{2}) = \frac{2}{3}$ , which coincides with  $\psi_0(0) = \frac{2}{3}$ .

With the extra interval  $L_4(\frac{1}{2}) = [\frac{5}{12}, \frac{19}{12}]$  centered at  $p_4 = 1$ , the 4-fold intersection is  $L_1 \cap L_2 \cap L_3 \cap L_4 = [\frac{5}{12}, \frac{7}{12}]$ . With the extra interval  $L_5(\frac{1}{2}) = [\frac{5}{6}, \frac{11}{6}]$  centered at  $p_5 = \frac{4}{3}$ , the 4-fold intersection  $L_2 \cap L_3 \cap L_4 \cap L_5$  is the single point  $\frac{5}{6}$ . With the extra interval  $L_6(\frac{1}{2}) = [\frac{11}{12}, \frac{13}{12}]$  centered at  $p_6 = \frac{3}{2}$ , the 4-fold intersection is  $L_3 \cap L_4 \cap L_5 \cap L_6 = [\frac{11}{12}, \frac{13}{12}]$ . Hence the total length of the 4-fold intersection at  $t = \frac{1}{2}$  is  $\psi_4(\frac{1}{2}) = \frac{1}{3}$ , which coincides with  $\psi_1(0) = \frac{1}{3}$ .

For the larger  $t = 1$ , the six grown intervals

$$L_1(1) = [-\frac{13}{12}, \frac{13}{12}], L_2(1) = [-\frac{2}{3}, \frac{4}{3}],$$

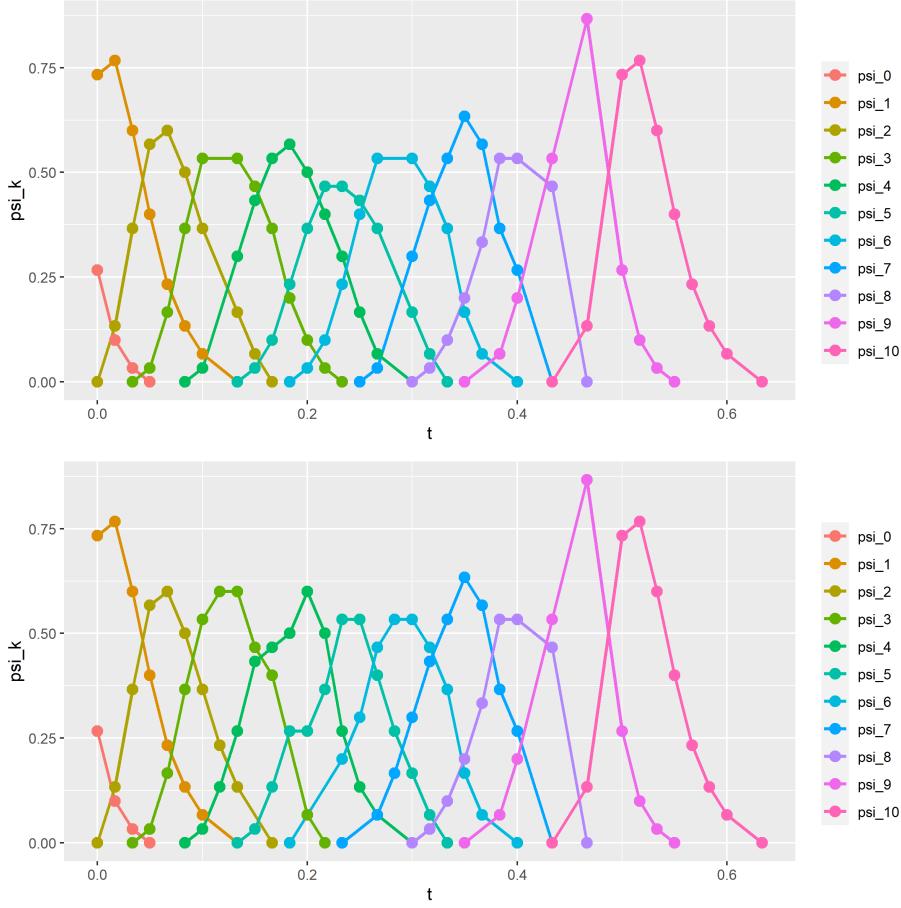
$$L_3(1) = [-\frac{7}{12}, \frac{19}{12}], L_4(1) = [-\frac{1}{12}, \frac{25}{12}],$$

$$L_5(1) = [\frac{1}{3}, \frac{7}{3}], \quad L_6(1) = [\frac{5}{12}, \frac{31}{12}]$$

have the 6-fold intersection  $[\frac{5}{12}, \frac{13}{12}]$  of length  $\psi_6(1) = \frac{2}{3}$  coinciding with  $\psi_0(0) = \psi_3(\frac{1}{2}) = \frac{2}{3}$ .  $\blacklozenge$

Corollary 9.6.2 says that the coincidences in Example 9.6.1 are not accidental. The periodicity of  $\psi_k$  with respect to  $k$  is illustrated by Fig. 9.18.

**Theorem 9.6.2** (periodicity of  $\psi_k$  in the index  $k$ , [3, Theorem 6.2]). The density functions  $\psi_k[S]$  of a periodic sequence  $S = \{p_1, \dots, p_m\} + \mathbb{Z}$  consisting of disjoint intervals with centers  $0 \leq p_1 < \dots < p_m < 1$  and radii  $r_1, \dots, r_m \geq 0$ , respectively, satisfy the *periodicity*  $\psi_{k+m}(t + \frac{1}{2}) = \psi_k(t)$  for any  $k \geq 0$  and  $t \geq 0$ .  $\blacksquare$



**Fig. 9.20** The densities  $\psi_k$ ,  $k = 0, \dots, 10$ , distinguish (already for  $k \geq 2$ ) the sequences (scaled down by period 15)  $S_{15} = \{0, 1, 3, 4, 5, 7, 9, 10, 12\} + 15\mathbb{Z}$  (**top**) and  $Q_{15} = \{0, 1, 3, 4, 6, 8, 9, 12, 14\} + 15\mathbb{Z}$  (**bottom**), where the radius  $r_i$  of any point is the half-distance to its closest neighbour, see Example 9.4.7.

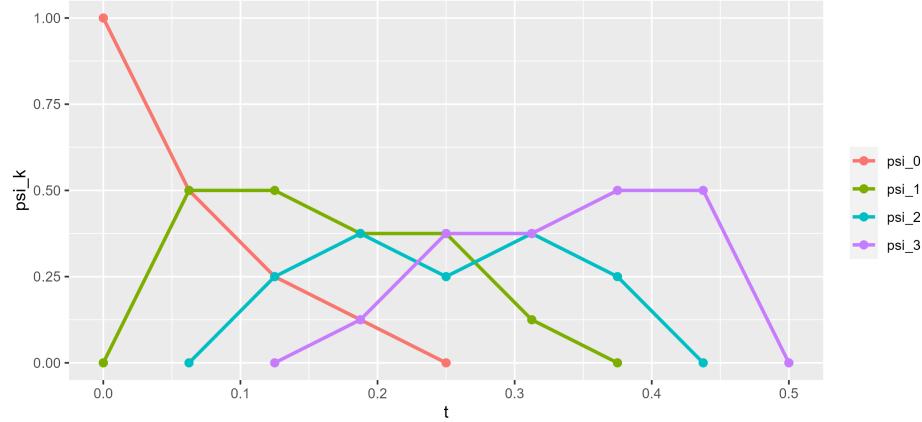
**Example 9.6.3** (Theorem 9.6.2 for  $m = 1$  in Fig. 9.19). Let a 1-period sequence  $S$  have one point  $p_1 = 0$  of a radius  $0 < r < \frac{1}{2}$ . The grown interval  $[-r - t - \frac{1}{2}, r + t + \frac{1}{2}]$  around 0 has the 1-fold intersection  $I = [r + t - \frac{1}{2}, \frac{1}{2} - r - t]$  centered at  $p = 0$  and not covered by the adjacent intervals centered at  $\pm 1$ , so  $\psi_1(t + \frac{1}{2}) = 1 - 2(t + r)$ .

After collapsing  $[-\frac{1}{2}, \frac{1}{2}]$  to 0, which is excluded from  $S$ , the periodic sequence has new points  $\pm \frac{1}{2}$  of the smaller radius  $r + t$ . The new shorter intervals have the same endpoints  $-\frac{1}{2} + (r + t)$  and  $\frac{1}{2} - (r + t)$  around  $p = 0$ . Now  $I = [r + t - \frac{1}{2}, \frac{1}{2} - r - t]$  is not covered by any shorter intervals, so the get the same length of the 0-fold intersection:  $\psi_0(t) = 1 - 2(t + r)$ .  $\blacklozenge$

The symmetry  $\psi_{m-k}(\frac{1}{2} - t) = \psi_k(t)$  for  $k = 0, \dots, [\frac{m}{2}]$ , and  $t \in [0, \frac{1}{2}]$  from [2, Theorem 8] no longer holds for points with different radii. For example,  $\psi_1(t) \neq \psi_2(\frac{1}{2} - t)$  for the periodic sequence  $S = \{0, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$ , see Fig. 9.16. If all points have the same radius  $r$ , [2, Theorem 8] implies the symmetry after replacing  $t$  by  $t + 2r$ .

Example 9.6.4 justified that all density functions cannot distinguish the non-isometric sequences  $S_{15} = \{0, 1, 3, 4, 5, 7, 9, 10, 12\} + 15\mathbb{Z}$  and  $Q_{15} = \{0, 1, 3, 4, 6, 8, 9, 12, 14\} + 15\mathbb{Z}$  of points with zero radii. Example 9.6.4 shows that the densities for sequences with non-zero radii are strictly stronger and distinguish the sequences  $S_{15} \not\cong Q_{15}$ .

**Example 9.6.4** ( $\psi_k$  for  $S_{15}, Q_{15}$  with radii). For any point  $p$  in a periodic sequence  $S \subset \mathbb{R}$ , define its *neighbour* radius as the half-distance to a closest neighbour of  $p$  within the sequence  $S$ . This choice of radii respects the isometry in the sense that periodic sequences  $S, Q$  with zero-sized radii are isometric if and only if  $S, Q$  with neighbour radii are isometric. Fig. 9.20 shows that the densities  $\psi_k$  for  $k \geq 2$  distinguish the non-isometric sequences  $S_{15}$  and  $Q_{15}$  scaled down by factor 15 to the unit cell  $[0, 1]$ , see Example 9.4.7.  $\blacklozenge$



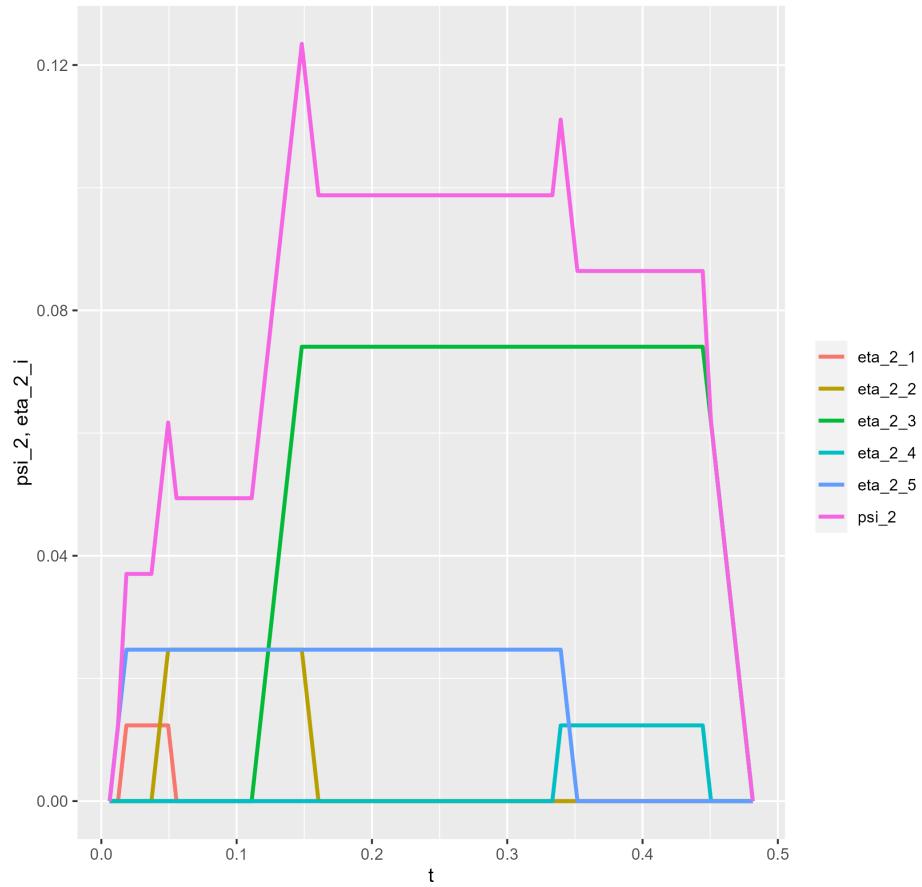
**Fig. 9.21** For the periodic sequence  $S = \{0, \frac{1}{8}, \frac{1}{4}, \frac{3}{4}\} + \mathbb{Z}$  whose all points have radii 0, the 2nd density  $\psi_2[S](t)$  has the local minimum at  $t = \frac{1}{4}$  between two local maxima.

**Corollary 9.6.5** (time of  $\psi_k(t)$ , [3, Corollary 6.5]). Let  $S, Q \subset \mathbb{R}$  be periodic sequences with at most  $m$  motif points. For  $k \geq 1$ , one can draw the graph of the  $k$ -th density function  $\psi_k[S]$  in time  $O(m^2)$ . One can check in time  $O(m^3)$  if  $\Psi[S] = \Psi[Q]$ . ■

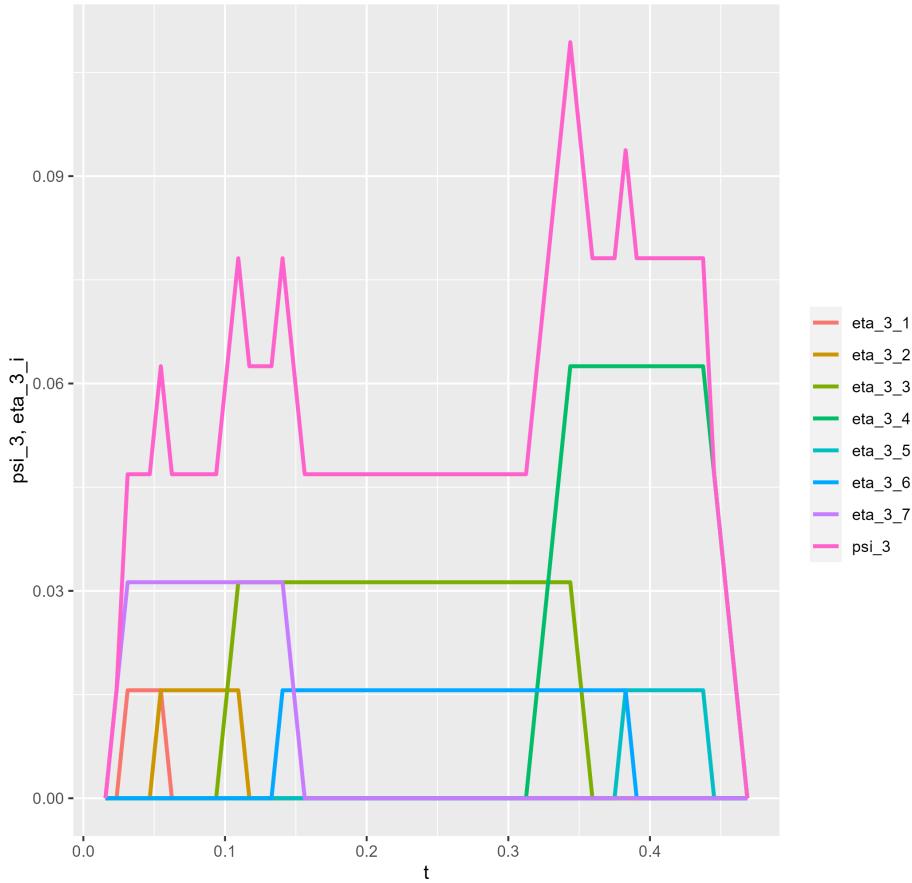
All previous examples show densities with a single local maximum. However, the new R code [1] helped us discover the opposite examples.

**Example 9.6.6** (densities with multiple maxima). Fig. 9.21 shows a simple 4-point sequence  $S$  whose 2nd density  $\psi_2[S]$  has two local maxima. Figs. 9.22 and 9.23 show complicated sequences whose density functions have more than two maxima. Fig. 9.24 shows that two local maxima are more common than one maximum. ♦

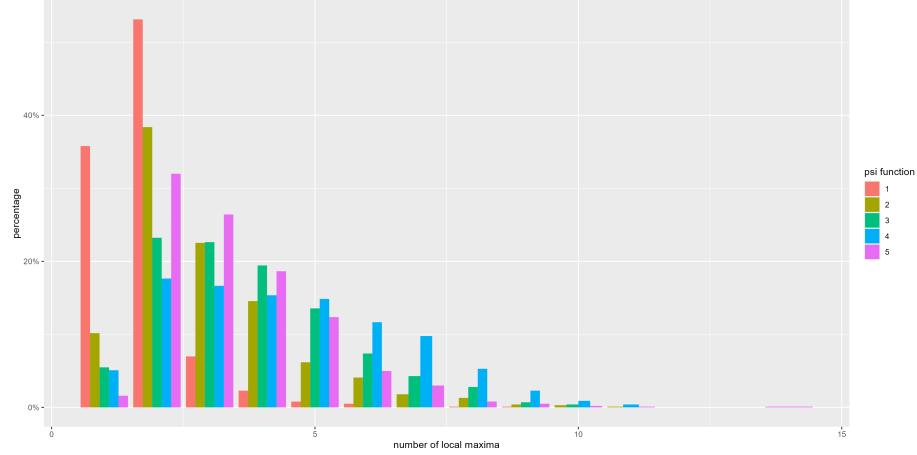
**Problem 9.6.7** (density functions for 2D lattices). Find an analytic description of all density functions  $\{\psi_k[\Lambda]\}_{k=0}^{+\infty}$  for any lattices  $\Lambda \subset \mathbb{R}^2$  in terms of the root invariant  $\text{RI}(\Lambda)$  from Definition 8.3.1. ★



**Fig. 9.22** For the sequence  $S = \left\{0, \frac{1}{81}, \frac{1}{27}, \frac{1}{9}, \frac{1}{3}\right\} + \mathbb{Z}$  whose all points have radii 0,  $\psi_2[S]$  equal to the sum of the shown five trapezium functions has three maxima.



**Fig. 9.23** For the sequence  $S = \left\{0, \frac{1}{64}, \frac{1}{16}, \frac{1}{8}, \frac{1}{4}, \frac{3}{4}\right\} + \mathbb{Z}$  whose all points have radii 0,  $\psi_3[S]$  has 5 local maxima.



**Fig. 9.24** Percentages of cases when the density functions  $\psi_k(t)$ ,  $k = 1, \dots, 5$  (shown in five different colors) have one or multiple local maxima for 1000 sequences of 10 points with zero radii, which are uniformly sampled in the interval  $[0, 1]$ .

## References

1. Olga Anosova. R code for density functions of periodic sequences, 2023.
2. Olga Anosova and Vitaliy Kurlin. Density functions of periodic sequences. In *Lecture Notes in Computer Science (Proceedings of DGMM)*, volume 13493, pages 395–408, 2022.
3. Olga Anosova and Vitaliy Kurlin. Density functions of periodic sequences of continuous events. *Journal of Mathematical Imaging and Vision*, 65:689–701, 2023.
4. Olga Anosova, Daniel Widdowson, and Vitaliy Kurlin. Recognition of near-duplicate periodic patterns by continuous metrics with approximation guarantees. *Pattern Recognition*, 171:112108, 2026.
5. 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.
6. Herbert Edelsbrunner and Georg Osang. A simple algorithm for higher-order delaunay mosaics and alpha shapes. *Algorithmica*, 85(1):277–295, 2023.
7. Herbert Edelsbrunner and Raimund Seidel. Voronoi diagrams and arrangements. *Discrete & Computational Geometry*, 1(1):25–44, 1986.
8. Grünbaum and Moore. The use of higher-order invariants in the determination of generalized Patterson cyclotomic sets. *Acta Cryst A*, 51:310–323, 1995.
9. Phong Q Nguyen and Damien Stehlé. Low-dimensional lattice basis reduction revisited. *ACM Transactions on algorithms (TALG)*, 5(4):1–48, 2009.
10. A Pulido et al. Functional materials discovery using energy–structure maps. *Nature*, 543:657–664, 2017.
11. Phil Smith and Vitaliy Kurlin. A practical algorithm for degree-k voronoi domains of three-dimensional periodic point sets. In *Lecture Notes in Computer Science (Proceedings of ISVC)*, volume 13599, pages 377–391, 2022.
12. Philip Smith. Density functions of periodic point sets in 3d. GitHub Repository, 2020.
13. Philip Smith. Voronoi zones of periodic point sets in c++. GitHub Repository, 2021.
14. Daniel Widdowson and Vitaliy Kurlin. Resolving the data ambiguity for periodic crystals. *Advances in Neural Information Processing Systems*, 35:24625–24638, 2022.

15. Daniel Widdowson and Vitaliy Kurlin. Pointwise distance distributions for detecting near-duplicates in large materials databases. *SIAM Journal on Applied Mathematics* ([doi:10.1137/25M1736657](https://doi.org/10.1137/25M1736657), [arxiv:2108.04798](https://arxiv.org/abs/2108.04798)), 2026.
16. Daniel Widdowson, Marco M Mosca, Angeles Pulido, Andrew I Cooper, and Vitaliy Kurlin. Average minimum distances of periodic point sets - foundational invariants for mapping all periodic crystals. *MATCH Commun. Math. Comput. Chem.*, 87:529–559, 2022.



## Chapter 10

# Pointwise Distance Distributions of periodic point sets in $\mathbb{R}^n$

**Abstract** This chapter extends the Pointwise Distance Distribution (PDD) from the case of finite clouds of unordered points to arbitrary periodic point sets. We prove that the PDD is Lipschitz continuous and generically complete for periodic point sets under isometry in  $\mathbb{R}^n$ . The PDD is computable in a near-linear asymptotic time of key input sizes and detects numerous near-duplicates among about 2 million crystals in major materials databases within two hours on a modest desktop computer.

### 10.1 Pointwise Distance Distributions for lattices and $l$ -periodic sets

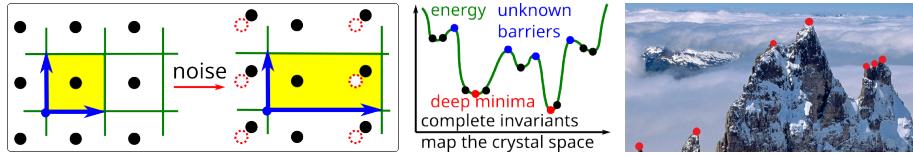
This section follows papers [31, 28, 29] with minor updates. Definition 10.1.1 extends 1-periodic point sets from Definition 7.1.1 and fully periodic sets in  $\mathbb{R}^n$  from Definition 9.1.1 to the more general sets that are periodic in  $l$  directions for  $1 \leq l \leq n$ .

**Definition 10.1.1** ( $l$ -periodic point set in  $\mathbb{R}^n$ ). Let vectors  $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$  form a basis of  $\mathbb{R}^n$ , define the lattice  $\Lambda = \{ \sum_{i=1}^l c_i \mathbf{v}_i \mid c_1, \dots, c_l \in \mathbb{Z} \}$  Fix any  $1 \leq l \leq n$ . The *unit cell* defined by  $\mathbf{v}_1, \dots, \mathbf{v}_n$  is  $U = \{ \sum_{i=1}^n x_i \mathbf{v}_i \mid x_1, \dots, x_l \in [0, 1), x_{l+1}, \dots, x_n \in \mathbb{R} \} \subset \mathbb{R}^n$ .

If  $l = n$ , then  $U$  is an  $n$ -dimensional parallelepiped. If  $l < n$ , then  $U$  is an infinite slab over an  $l$ -dimensional parallelepiped on  $\mathbf{v}_1, \dots, \mathbf{v}_l$ . For any finite *motif* of points  $M \subset U$ , the sum  $S = M + \Lambda = \{ \mathbf{p} + \mathbf{v} \mid \mathbf{p} \in M, \mathbf{v} \in \Lambda \}$  is an  $l$ -periodic point set.  $\blacktriangle$

Fig. 10.1 (left) illustrates the discontinuity of cell-based representations under noise. Definition 9.1.2 of generic periodic sets and crystal spaces under equivalences extends to  $l$ -periodic point sets with motifs of  $m$  points in  $\mathbb{R}^n$ . We use similar notations and add  $l$  as an extra parameter. This chapter studies the *Crystal Isometry Space*  $\text{CIMS}(\mathbb{R}^n; m, l)$  of  $l$ -periodic sets with motifs of up to  $m$  points under isometry in  $\mathbb{R}^n$ .

Problem 10.1.2 extends Problem 9.1.3 to  $l$ -periodic point sets and add the extra condition of realisability in 10.1.2(b). This realisability condition was unrealistic for



**Fig. 10.1** **Left:** the symmetry group and a reduced cell discontinuously change under tiny noise. **Middle:** the most stable crystals are deep minima surrounded by high energy barriers in the crystal space. **Right:** energy landscapes of crystals show optimised structures as isolated peaks of height =  $-\text{energy}$ . To see beyond the ‘fog’, we need an explicit *geomap* parametrised by geocodes in Problem 10.1.2.

density functions from the previous chapter, at least in dimensions  $n \geq 2$ , but will be achieved for generic periodic sets with  $l = n$  by Pointwise Distance Distributions.

**Problem 10.1.2** (isometry invariants of  $l$ -periodic point sets in  $\mathbb{R}^n$ ). Design an invariant  $I$  on the Crystal Isometry Space  $\text{CIMS}(\mathbb{R}^n; m, l)$  satisfying the following conditions.

- (a) *Generic completeness:* let  $S, Q$  be any generic sets whose isometry classes are in a dense subspace of  $\text{CIMS}(\mathbb{R}^n; m, l)$ , then  $S, Q$  are isometric if and only if  $I(S) = I(Q)$ .
- (b) *Reconstruction:* any generic periodic point set  $S \subset \mathbb{R}^n$  is reconstructable from its invariant  $I(S)$ , a lattice  $\Lambda$  of  $S$  and the motif size  $m$ , uniquely under isometry in  $\mathbb{R}^n$ .
- (c) *Metric:* there is a distance  $d$  on the Crystal Isometry Space  $\text{CIMS}(\mathbb{R}^n; m, l)$  satisfying all metric axioms in Definition 1.3.1(a).

(d) *Continuity:* there is a constant  $\lambda > 0$ , such that, for all sufficiently small  $\varepsilon > 0$ , if a periodic point set  $Q$  is obtained by perturbing every point of a periodic point set  $S \subset \mathbb{R}^n$  up to Euclidean distance  $\varepsilon$ , then  $d(I(S), I(Q)) \leq \lambda\varepsilon$ .

(e) *Computability:* for fixed  $1 \leq l \leq n$ , the invariant  $I(S)$ , a reconstruction of  $S \subset \mathbb{R}^n$  from  $I(S)$ , and the metric  $d(I(S), I(Q))$  can be computed in times that depend polynomially on the maximum motif size of  $l$ -periodic point sets  $S, Q$ . ★

Definition 10.1.3 introduces our main invariant PDD and its average AMD.

**Definition 10.1.3** (PDD( $S; k$ ) and AMD( $S; k$ ) of any  $l$ -periodic set  $S$ ). Let  $M = \{p_1, \dots, p_m\}$  be a motif of any  $l$ -periodic point set  $S \subset \mathbb{R}^n$ . Fix an integer  $k \geq 1$ . For every point  $p \in M$ , let  $d_1(p) \leq \dots \leq d_k(p)$  be the distances from  $p$  to its  $k$  nearest neighbours within the full set  $S$  (not restricted to  $M$ ). The matrix  $D(S; k)$  has  $m$  rows consisting of the distances  $d_1(p_i), \dots, d_k(p_i)$  for  $i = 1, \dots, m$ . If any  $l \geq 2$  rows coincide, we collapse them into a single row with the weight  $l/m$ . The resulting unordered set (written as a matrix) of maximum  $m$  rows and  $k+1$  columns, including the extra column of weights, is the *Pointwise Distance Distribution* PDD( $S; k$ ). The *Average Minimum Distance* AMD $_i$  is the weighted average of the  $i$ -th column in PDD( $S; k$ ) for each  $i = 1, \dots, k$ . Let AMD( $S; k$ ) denote the vector  $(\text{AMD}_1, \dots, \text{AMD}_k)$ . ▲

Theorem 10.1.4 shows that PDD and hence AMD are independent of a motif  $M \subset S$ , so there is no need to include a motif  $M$  in the notation PDD( $S; k$ ).

**Theorem 10.1.4** (isometry invariance of PDD, [31, Theorem 3.3(b)]). For any  $l$ -periodic point set  $S \subset \mathbb{R}^n$ , where  $1 \leq l \leq n$ , PDD( $S; k$ ) and AMD( $S; k$ ) are invariants of  $S$  (independent of a motif  $M \subset S$ ) under isometry of  $\mathbb{R}^n$  for  $k \geq 1$ . ■

## 10.2 Asymptotic and time of Pointwise Distance Distributions

This section analyses the behaviour of  $\text{PDD}(S; k)$  as the index of neighbours  $k \rightarrow +\infty$ . Definition 10.2.1 will help describe this asymptotic in Theorem 10.2.2.

**Definition 10.2.1** (Point Packing Coefficient PPC of a cell-periodic set  $S$ ). **(a)** For  $1 \leq l \leq n$  and a basis  $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$ , consider the *lattice*  $\Lambda = \{ \sum_{i=1}^l c_i \mathbf{v}_i \mid c_1, \dots, c_l \in \mathbb{Z} \}$

and the unit cell  $U = \{ \sum_{i=1}^n x_i \mathbf{v}_i \mid x_1, \dots, x_l \in [0, 1), x_{l+1}, \dots, x_n \in \mathbb{R} \}$ . A set  $S \subset \mathbb{R}^n$  is *cell-periodic* if  $S$  has a fixed number  $m$  points in every shifted cell  $U + \mathbf{v}$  for all  $\mathbf{v} \in \Lambda$ .

**(b)** If  $l < n$ , let  $R^l \subset \mathbb{R}^n$  be the subspace spanned by  $\mathbf{v}_1, \dots, \mathbf{v}_l$ , then  $U$  is an infinite slab based on the  $l$ -dimensional parallelepiped of volume  $\text{vol}[U \cap R^l]$ . The volume of the unit ball in  $\mathbb{R}^l$  is  $V_l = \frac{\pi^{l/2}}{\Gamma(\frac{l}{2} + 1)}$ , where Euler's Gamma function is  $\Gamma(m) = (m-1)!$  and  $\Gamma(\frac{m}{2} + 1) = \sqrt{\pi}(m - \frac{1}{2})(m - \frac{3}{2}) \cdots \frac{1}{2}$  for any integer  $m \geq 1$ . Define the *Point Packing Coefficient* of the cell-periodic set  $S$  as  $\text{PPC}(S) = \sqrt[l]{\frac{\text{vol}[U \cap R^l]}{mV_l}}$ .

Any  $l$ -periodic set is cell-periodic, but all cell-periodic sets form a wider collection of Delone sets and model disordered solid materials that can have an underlying lattice with atoms at different positions in periodically translated cells  $U + \mathbf{v}$ , see Fig. 9.1.

**Theorem 10.2.2** (asymptotic of  $\text{PDD}(S; k)$  as  $k \rightarrow +\infty$ , [31, Theorem 3.7]). For a point  $p$  in a cell-periodic set  $S \subset \mathbb{R}^n$ , let  $d_k(S; p)$  be the distance from  $p$  to its  $k$ -th nearest neighbour in  $S$ . Then  $\lim_{k \rightarrow +\infty} \frac{d_k(S; p)}{\sqrt[k]{k}} = \text{PPC}(S)$  and  $\lim_{k \rightarrow +\infty} \frac{\text{AMD}_k(S)}{\sqrt[k]{k}} = \text{PPC}(S)$ . ■

By Theorem 10.2.2,  $\text{AMD}_k(S)$  and all distances in the last column of  $\text{PDD}(S; k)$  asymptotically approach  $\text{PPC}(S)\sqrt[k]{k}$  as  $k \rightarrow +\infty$  and hence are mainly determined by  $\text{PPC}(S)$  for large  $k$ . That is why the most descriptive information is contained in  $\text{PDD}(S; k)$  for smaller values of  $k$ , e.g. we use  $k = 100$  atomic neighbours in most experiments on crystals. To neutralise the asymptotic growth, we subtract and also normalise by the term  $\text{PPC}(S)\sqrt[k]{k}$  to get simpler invariants under uniform scaling.

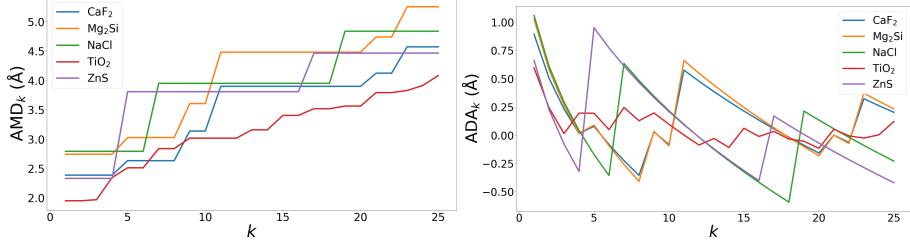
**Definition 10.2.3** (simplified invariants ADA, PDA, AND, PND). Let  $S \subset \mathbb{R}^n$  be any  $l$ -periodic set with an underlying lattice generated by  $l$  vectors.

**(a)** The *Average Deviation from Asymptotic* is  $\text{ADA}_k(S) = \text{AMD}_k(S) - \text{PPC}(S)\sqrt[k]{k}$ ,  $k \geq 1$ . To get the *Pointwise Deviation from Asymptotic*  $\text{PDA}(S; k)$  from  $\text{PDD}(S; k)$  subtracting  $\text{PPC}(S)\sqrt[j]{j}$  from each distance in a row  $i$  and a column  $j$  for  $i \geq 1 \leq j \leq k$ .

**(b)** The *Average Normalised Deviation* is  $\text{AND}_k(S) = \text{ADA}_k(S)/(\text{PPC}(S)\sqrt[k]{k})$ ,  $k \geq 1$ . The *Pointwise Normalised Deviation*  $\text{PND}(S; k)$  obtained from  $\text{PDA}(S; k)$  by dividing every element in a row  $i$  and a column  $j$  by  $\text{PPC}(S)\sqrt[j]{j}$  for  $i \geq 1 \leq j \leq k$ . ▲

The invariants  $\text{AMD}_k$  and  $\text{ADA}_k$  form vectors of length  $k$ , e.g. set  $\text{AMD}(S; k) = (\text{AMD}_1(S), \dots, \text{AMD}_k(S))$  and  $\text{ADA}(S; k) = (\text{ADA}_1(S), \dots, \text{ADA}_k(S))$ . These vectors can be compared by many metrics. The metric  $L_\infty(u, v) = \max_{i=1, \dots, k} |\mathbf{u}_i - \mathbf{v}_i|$  for

**Fig. 10.2** The average invariants  $\text{AMD}_k$  and  $\text{ADA}_k$  from Definition 10.2.3 for  $k = 1, \dots, 25$  and five simple crystals from the Materials Project.



any vectors  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^k$  preserves the intuition of atomic displacements in the following sense. If  $S$  is obtained from  $Q$  by perturbing every point up to a small  $\varepsilon$ , then  $L_\infty(\text{AMD}(S; k), \text{AMD}(Q; k)) \leq 2\varepsilon$  by [32, Theorem 9]. Other distances such as Euclidean can be considered but will accumulate a larger deviation depending on  $k$ .

All invariants above and metrics on them are measured in the same units as original coordinates, i.e. in Angstroms for crystals given by Crystallographic Information Files (CIFs). The Point Packing Coefficient  $\text{PPC}(S)$  was defined as the cube root of the cell volume per atom (of the same radius  $1\text{\AA}$ ) and can be interpreted as an average radius of balls ‘packed’ in a unit cell. So  $\text{PPC}(S)$  is roughly inversely proportional to the physical density but they are exactly related only when materials have the same average atomic mass (total mass of atoms in a unit cell divided by the cell volume).

While  $\text{AMD}_k(S)$  monotonically increases in  $k$ , the invariants  $\text{ADA}_k(S)$  can be positive or negative as deviations around the asymptotic  $\text{PPC}(S) \sqrt[3]{k}$ . Fig. 10.3 reveals geometric differences between the mainly organic databases CSD and Crystallography Open Database (COD) [15] versus the more inorganic collections ICSD and MP.

The first average of  $\text{ADA}_1 \in [-0.25, -0.17]$  in the top images of Fig. 10.3 can be explained by the presence of many hydrogen atoms, which have distances smaller than  $\text{PPC}(S)$  to their first neighbour in most organic materials. Indeed, hydrogens are usually bonded at distances less than  $1.2\text{\AA}$ , while  $\text{PPC}(S)$  is often larger than  $1.2\text{\AA}$  because most chemical elements have van der Waals radii above  $1.2\text{\AA}$  [6].

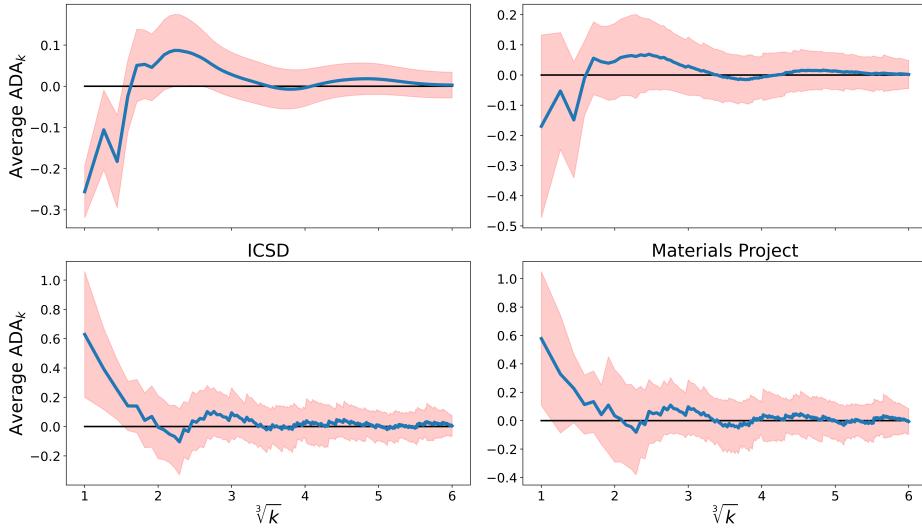
For inorganic materials, metal atoms or ions have relatively large distances to their neighbours, so the average  $\text{ADA}_1$  is in  $[0.58, 0.62]$  in the bottom images of Fig. 10.3.

For all types of materials in Fig. 10.3, the value of  $\text{ADA}_k$  experimentally converges to 0 on average, so there is no need to substantially increase  $k$  because the important structural information emerges for smaller indices  $k$  of neighbours.

If we increase  $k$ , the matrix  $\text{PDD}(S; k)$  and hence the vector  $\text{ADA}(S; k)$  become longer by including distance data to further neighbours but all initial values remain the same. Hence we consider  $k$  not as a parameter that changes the output but as a degree of approximation similarly to the number of decimal places on a calculator.

The experimental convergence  $\text{ADA}_k \rightarrow 0$  as  $k \rightarrow +\infty$  in Fig. 10.3 justifies computing the distance  $L_\infty$  between  $\text{ADA}$  vectors up to a reasonable  $k$ . We use  $k = 100$

**Fig. 10.3** The averages of  $\text{ADA}_k$  and standard deviations (1 sigma shaded) vs  $\sqrt[3]{k}$  for four databases.



because all  $\text{ADA}_k$  for  $k > 100$  are close to 0 (the range of 1 sigma between  $\pm 0.2 \text{ \AA}$ ) in Fig. 10.3.

**Corollary 10.2.4** (invariance of AND, PND under uniform scaling, [31, Corollary 3.9]). For any  $l$ -periodic set  $S \subset \mathbb{R}^n$ ,  $\text{AND}_k(S)$  and  $\text{PND}(S; k)$  in Definition 10.2.3 are invariant under isometry and uniform scaling for  $k \geq 1$ . Also,  $\text{AND}_k(S) \rightarrow 0$  as  $k \rightarrow +\infty$ . ■

The following conjecture is justified by Example 10.2.6 and Fig 10.3.

**Conjecture 10.2.5** (asymptotic of  $\text{ADA}(S; k)$  as  $k \rightarrow +\infty$ ). For any periodic point set  $S \subset \mathbb{R}^n$ , we have  $\text{ADA}_k(S) \rightarrow 0$  as  $k \rightarrow +\infty$ . ★

**Example 10.2.6** (asymptotic  $\text{ADA}_k(S) \rightarrow 0$  as  $k \rightarrow +\infty$  for the cubic lattice  $S = \mathbb{Z}^n$ ). The survey [17] describes progress on the generalised Gauss circle problem expressing the number of points from the cubic lattice  $\mathbb{Z}^n$  within a ball of a radius  $r$  as  $k = V_n r^n - O(r^{\alpha_n + \varepsilon})$  for any  $\varepsilon > 0$ , where  $\alpha_n < n - 1$  for  $n \geq 2$ , e.g.  $\alpha_2 \leq \frac{2}{3}$ ,  $\alpha_3 \leq \frac{3}{2}$ , and  $\alpha_n \leq n - 2$  for any  $n \geq 4$ . The cubic lattice  $S = \mathbb{Z}^n$  has  $\text{PPC}(\mathbb{Z}^n) = 1/\sqrt[n]{V_n}$ . Let  $d_k$  denote the distance from the origin 0 to its  $k$ -th neighbour in  $\mathbb{Z}^n$ . Then

$$k = V_n d_k^n - O(d_k^{\alpha_n + \varepsilon}), \text{ so } d_k = \sqrt[n]{\frac{k + O(d_k^{\alpha_n + \varepsilon})}{V_n}} = \text{PPC}(\mathbb{Z}^n) \sqrt[n]{k + O(d_k^{\alpha_n + \varepsilon})},$$

$$\frac{\text{ADA}_k(\mathbb{Z}^n)}{\text{PPC}(\mathbb{Z}^n)} = \frac{d_k}{\text{PPC}(\mathbb{Z}^n)} - \sqrt[n]{k + O(d_k^{\alpha_n + \varepsilon})} = \frac{O(d_k^{\alpha_n + \varepsilon})}{P_n(\sqrt[n]{k + O(d_k^{\alpha_n + \varepsilon})}, \sqrt[n]{k})},$$

where  $P_n$  is a homogeneous polynomial of degree  $n-1$ , e.g.  $P_2(x, y) = x+y$ ,  $P_3(x, y) = x^2 + xy + y^2$ . Since the numerator has the power  $\alpha_n < n-1$  of  $d_k = O(\sqrt[n]{k})$  for  $n \geq 2$ , the final expression of  $\frac{\text{ADA}_k(\mathbb{Z}^n)}{\text{PPC}(\mathbb{Z}^n)}$  and hence  $\text{ADA}_k(\mathbb{Z}^n)$  have limit 0 as  $k \rightarrow +\infty$ .  $\blacklozenge$

**Theorem 10.2.7** (time of PDD, [31, Theorem 3.10]). Let  $S \subset \mathbb{R}^n$  be any  $l$ -periodic set with a minimum inter-point distance  $d_{\min}$  and a unit cell  $U = P \times R^{n-l}$ , where  $P$  is a parallelepiped in the  $l$ -dimensional subspace  $R^l$  with the orthogonal subspace  $R^{n-l}$  in  $\mathbb{R}^n$ . Consider the *width*  $w = \sup_{u, v \in P} |\mathbf{u} - \mathbf{v}|$  and the *height*  $h$  equal to the maximum distance between points in the orthogonal projection of  $S$  to  $R^{n-l}$ . If the motif  $M = S \cap U$  consists of  $m$  points, then  $\text{PDD}(S; k)$  can be computed for any  $k \geq 1$  in time

$$O(km(2^{4n} \log k + \log m) + 2^{12n}m \log^2 k + (2^{8n}/l)k \log k + 2^{8n}a^l bk),$$

where  $a = 1 + \frac{2.5w + 2h}{\text{PPC}(S)}$  and  $b = \log(2\text{PPC}(S) + 3w + 5h) - \log d_{\min}$ . The complexity of  $\text{AMD}(S; k)$  and invariants  $\text{PDA}(S; k)$ ,  $\text{PND}(S; k)$  from Definition 10.2.3 is the same as for  $\text{PDD}(S; k)$ , because the extra computations can be done in time  $O(km)$ .  $\blacksquare$

The worst-case estimate in Theorem 10.2.7 is conservative due to the upper bound  $2^n$  for the expansion constants  $c_{\min}, c$  from [13, Definition 1.4]. We conjecture that this upper bound can be reduced to  $2^l$  for any  $l$ -periodic point set  $S \subset \mathbb{R}^n$ .

For any fixed dimensions  $l \leq n$ , if we ignore the parameters  $a, b, d_{\min}$ , and  $\text{PPC}(S)$ , then the complexity in Theorem 10.2.7 becomes  $O(km(\log k + \log m))$ , which is near-linear in both  $k, m$ . For the most practical dimensions  $l = n = 3$ , experiments in section 10.5 will report running times in minutes on a modest desktop computer for about 2 million real crystals from the world's largest materials databases.

### 10.3 Lipschitz continuous metrics and local novelty distances

This section proves the Lipschitz continuity of the vectorial invariants  $\text{AMD}$ ,  $\text{ADA}$ ,  $\text{AND}$ , matrix invariants  $\text{PDD}$ ,  $\text{PDA}$ ,  $\text{PND}$ , and their averages in Theorem 10.3.4.

The widely used Mercury software visually compares periodic structures [10] by minimizing the Root Mean Square Deviation (RMSD) of atomic positions from up to a given number  $m$  (15 by default) of closest molecules in two structures. This comparison depends on many parameters (maximum number of matched molecules, thresholds for matched distances and angles), fails the triangle inequality in metric axioms, see Definition 1.3.1, and is too slow for pairwise comparisons. In fact, this RMSD tries to measure a maximum displacement of atoms only by using finite subsets of crystals.

For full infinite crystals, the theoretically better alternative is the *bottleneck distance*  $\text{BD}(S, Q)$  equal to the maximum Euclidean distance needed to perturb every point  $p \in S$  to its unique match in  $Q$ , see Example 1.3.3(b).

Example 10.3.1 shows that the periodic sequences have  $d_B = +\infty$  for any  $\delta > 0$ .

**Example 10.3.1** (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  $p \in S$ , the image of  $2N + 1$  points  $S \cap [-N, N]$  under this bijection  $g$  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  $\frac{N+d_B}{1+\delta} < N$  or  $d_B < \delta N$ . We get a contradiction by choosing a large  $N > \frac{d_B}{\delta}$ .  $\blacklozenge$

If  $S, Q$  are lattices of equal density (equal unit cell volume), they have a finite bottleneck distance  $d_B$  by [11, Theorem 1(iii)]. If we consider only periodic point sets  $S, Q \subset \mathbb{R}^n$  with the same density (or unit cells of the same volume),  $d_B(S, Q)$  becomes a well-defined *wobbling* distance [7], which is discontinuous under perturbations below.

**Example 10.3.2** (discontinuous wobbling). 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  $\Lambda$ . We prove that  $d_B(\Lambda, \mathbb{Z}^2) \geq \frac{1}{3}$  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 the whole square lattice  $\mathbb{Z}^2$  is mapped to the single point represented by the corners of the square  $[0, 1]^2$ . The perturbed lattice  $\Lambda$  maps to the sequence of points  $\{k\varepsilon \pmod{1}\}_{k=0}^{+\infty} \times \{0, 1\}$  in the horizontal edges. If  $d_B(\Lambda, \mathbb{Z}^2) = r < \frac{1}{3}$ , then all above points should be covered by the closed disks of the radius  $r$  centred at the corners of  $[0, 1]^2$ . For  $0 < \varepsilon < \frac{1}{3} - r$ , we can find  $k \in \mathbb{Z}$  so that  $k\varepsilon$  is strictly between  $r, 1 - r$ , hence not covered by these disks, so  $d_B(\Lambda, \mathbb{Z}^2) \geq \frac{1}{3}$ .  $\blacklozenge$

We will use the Earth Mover's Distance (EMD) from Definition 3.5.4, which is well-defined for any normalised distributions of different sizes and makes sense for any matrix invariant  $I(S)$  that is an unordered collection of row vectors  $\mathbf{R}_i(S)$  with weights  $w_i(S) \in (0, 1]$  satisfying  $\sum_{i=1}^{m(S)} w_i(S) = 1$ . Each row  $\mathbf{R}_i(S)$  should have a size independent of  $i$ , e.g. the number  $k$  of neighbours in  $\text{PDD}(S; k)$ . For any vectors  $\mathbf{R}_i = (r_{i1}, \dots, r_{ik})$  and  $\mathbf{R}_j = (r_{j1}, \dots, r_{jk})$ , we will use the Minkowski and Chebyshev distances from Example 1.3.3(b):  $L_q(\mathbf{R}_i, \mathbf{R}_j) = \left( \sum_{l=1}^k |r_{il} - r_{jl}|^q \right)^{1/q}$ ,  $L_\infty(\mathbf{R}_i, \mathbf{R}_j) = \max_{l=1, \dots, k} |r_{il} - r_{jl}|$ .

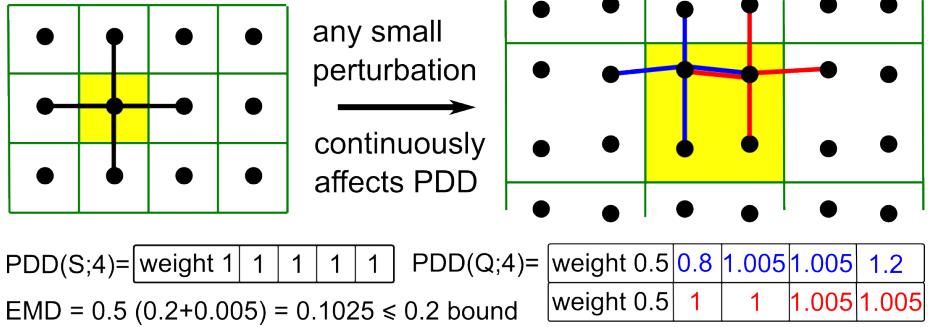
**Example 10.3.3** (EMD under noise). We illustrate EMD for perturbations that scale up a unit cell as in Fig. 10.1 (left). The integer sequence  $\mathbb{Z}$  has  $\text{PDD}(\mathbb{Z}; 2) = (1; 1, 1)$ , a single row of weight 1 and unit distances to 2 neighbours. The periodic sequence  $\mathbb{Z}_\varepsilon = \{0, 1 + \varepsilon, 2 - \varepsilon\} + 3\mathbb{Z}$  is obtained from  $\mathbb{Z}$  by  $\varepsilon$ -perturbations of points 1, 2 and

all their translates with period 3. Then  $\text{PDD}(\mathbb{Z}_\varepsilon; 2) = \begin{pmatrix} 1/3 & 1 + \varepsilon & 1 + \varepsilon \\ 2/3 & 1 - 2\varepsilon & 1 + \varepsilon \end{pmatrix}$ , where the

2nd row represents the shifted points  $1 + \varepsilon, 2 - \varepsilon$ . After splitting  $\text{PDD}(\mathbb{Z}; 2) = (1; 1, 1)$  into two identical rows of weights  $\frac{1}{3}, \frac{2}{3}$  and using  $L_\infty$  on vectors of two distances, a difference between PDDs can be defined as the weighted average  $\frac{1}{3}\varepsilon + \frac{2}{3}2\varepsilon = \frac{2}{3}\varepsilon$ .  $\blacklozenge$

Fig. 10.4 illustrates the continuity of  $\text{PDD}(S; 4)$  under a perturbation of a square lattice  $S \subset \mathbb{R}^2$ , which scales up an initial cell by a factor of 4.

**Fig. 10.4** The Pointwise Distance Distribution  $\text{PDD}(S; 4)$  of a unit square lattice  $S$  changes continuously under noise, shown here in the case when a 1-point motif extends to a motif of 4-points that form an isosceles trapezium with parallel sides of lengths 0.8 (red) and 1.2 (blue).



The Lipschitz continuity of invariants in EMD will use bounded perturbations of points up to Euclidean distance  $\varepsilon$  in  $\mathbb{R}^n$ . Recall that the *packing radius*  $r(S)$  is the minimum half-distance between any points of  $S$ , see Definition 9.2.3(a).

**Theorem 10.3.4** (Lipschitz continuity of PDA and PND, [31, Theorem 4.2]). Let  $S, Q \subset \mathbb{R}^n$  be  $l$ -periodic point sets such that  $Q$  is obtained from  $S$  by perturbing every point of  $S$  up to Euclidean distance  $\varepsilon$ . Fix any  $q \in [1, +\infty]$  and an integer  $k \geq 1$ . Interpret  $\sqrt[k]{k}$  as 1 in the limit case  $q = +\infty$ . If  $\min\{r(S), r(Q)\} > \varepsilon$ , then  $\text{PPC}(S) = \text{PPC}(Q)$ ,

(a)  $\text{EMD}_q(\text{PDA}(S; k), \text{PDA}(Q; k)) \leq 2\varepsilon \sqrt[k]{k}$ , and

(b)  $\text{EMD}_q(\text{PND}(S; k), \text{PND}(Q; k)) \leq \frac{2\varepsilon \sqrt[k]{k}}{\text{PPC}(S)}$ . ■

All columns of PDD, PDA, PND are ordered by the index  $k$  of neighbours. Though their rows are unordered (as points of a motif  $M$ ), all such matrices (even with different numbers of rows) can be compared by Earth Mover's Distance, or by any other metrics on weighted distributions, see Definition 3.5.4. We can simplify any PDD into a fixed-size matrix, which can be flattened into a vector, while keeping the continuity and almost all invariant data. Any distribution of  $m$  unordered values can be reconstructed from its  $m$  moments in Definition 5.1.4. When all weights  $w_i$  are rational, as in our case, the distribution can be expanded to equal-weighted values  $a_1, \dots, a_m$ . The  $m$  moments can recover all  $a_1, \dots, a_m$  as roots of a degree  $m$  polynomial whose coefficients are expressed via the  $m$  moments [22], e.g. any  $a, b \in \mathbb{R}$  can be found from  $a + b, a^2 + b^2$  as the roots of  $x^2 - (a + b)x + ab$ , where  $ab = \frac{1}{2}((a + b)^2 - (a^2 + b^2))$ .

For  $t = 1$ , the  $1 \times k$  matrix  $\mu^{(1)}[\text{PDD}(S; k)]$  appeared in Definition 10.1.3 as the vector  $\text{AMD}(S; k) = (\text{AMD}_1, \dots, \text{AMD}_k)$  of column averages. All rows and columns of  $\mu^{(t)}[I(S)]$  are ordered, but this matrix is a bit weaker than  $I(S)$  because each column can be reconstructed from its moments (for a large enough  $t$ ) only up to permutation. We can flatten any matrix  $\mu^{(t)}[I(S)]$  to a vector for machine learning.

**Theorem 10.3.5** (lower bounds, [31, Theorem 4.4]). For any  $l$ -periodic sets  $S, Q \subset \mathbb{R}^n$ ,

- (a)  $\text{EMD}_q(\text{PDD}(S; k), \text{PDD}(Q; k)) \geq L_q(\text{AMD}(S; k), \text{AMD}(Q; k))$ ;
- (b)  $\text{EMD}_q(\text{PDA}(S; k), \text{PDA}(Q; k)) \geq L_q(\text{ADA}(S; k), \text{ADA}(Q; k))$ ;
- (c)  $\text{EMD}_q(\text{PND}(S; k), \text{PND}(Q; k)) \geq L_q(\text{AND}(S; k), \text{AND}(Q; k))$  for  $q, k \geq 1$ .  $\blacksquare$

## 10.4 Generic completeness of PDDs for periodic point sets in $\mathbb{R}^n$

While the generic completeness of the PDD was easy for any finite clouds in  $\mathbb{R}^n$ , this section extends Theorem 4.5.4 to the much harder periodic case in Theorem 10.4.5.

For a periodic point set  $S \subset \mathbb{R}^n$ , the generic completeness of PDD is not straightforward, because infinitely many distances between points of  $S$  are repeated due to periodicity. We introduce a few auxiliary concepts for distance-generic periodic sets in Definition 10.4.3. For any point  $p$  in a lattice  $\Lambda \subset \mathbb{R}^n$ , the *open Voronoi domain*

$$V(\Lambda; p) = \{q \in \mathbb{R}^n \text{ such that } |q - p| < |q - p'| \text{ for any } p' \in \Lambda - p\}$$

is the neighbourhood of all points  $q \in \mathbb{R}^n$  that are strictly closer to  $p$  than to all other points  $p'$  of the lattice  $\Lambda$ . Definition 8.1.3(a) used the closed version  $\bar{V}(\Lambda; p)$ .

Open Voronoi domains  $V(\Lambda; p)$  of different points  $p \in \Lambda$  are disjoint translation copies of each other and their closures tile  $\mathbb{R}^n$ , so  $\bigcup_{p \in \Lambda} \bar{V}(\Lambda; p) = \mathbb{R}^n$ . For example, for a generic lattice  $\Lambda \subset \mathbb{R}^2$ , the domain  $V(\Lambda; p)$  is a centrally symmetric hexagon.

Points  $p, p' \in \Lambda$  are *Voronoi neighbours* if their Voronoi domains share a boundary point, so  $\bar{V}(\Lambda; p) \cap \bar{V}(\Lambda, p') \neq \emptyset$ . Below we always assume that any lattice  $\Lambda$  is shifted to contain the origin 0, also any periodic point set  $S = \Lambda + M$  has a point at 0.

**Definition 10.4.1** (neighbour set  $N(\Lambda)$  and base distances). For any lattice  $\Lambda \subset \mathbb{R}^n$ , the *neighbour set* of the origin 0 is  $N(\Lambda) = \Lambda \cap \bar{B}(0; r) \setminus \{0\}$  for a minimum radius  $r$  such that  $N(\Lambda)$  is not contained in any affine  $(n-1)$ -dimensional subspace of  $\mathbb{R}^n$ , and  $N(\Lambda)$  includes all  $n+1$  nearest neighbours (within  $\Lambda$ ) of any point  $q \in V(\Lambda; 0)$ .

Consider all sets of unordered points  $p_1, \dots, p_n \in N(\Lambda)$  that are *linearly independent*, i.e. the vectors  $\mathbf{p}_1, \dots, \mathbf{p}_n$  form a linear basis of  $\mathbb{R}^n$ . For any point  $q \in V(\Lambda; 0)$ , a lexicographically smallest list of distances  $d_1(q) \leq \dots \leq d_n(q)$  from  $q$  to a set of linearly independent points  $p_1, \dots, p_n \in N(\Lambda)$  is the list of *base distances* of  $q$ .  $\blacktriangle$

The linear independence of vectors  $\mathbf{p}_1, \dots, \mathbf{p}_n$  in Definition 10.4.1 guarantees that any point  $q$  is uniquely determined in  $\mathbb{R}^n$  by its distances  $|q|, d_1(q), \dots, d_n(q)$  to  $n+1$  neighbours  $0, p_1, \dots, p_n$ , which are not in the same  $(n-1)$ -dimensional subspace.

**Example 10.4.2** (neighbour sets). The vector  $(2, 0), (0, 1)$  generate the rectangular lattice  $\Lambda \subset \mathbb{R}^2$ . The Voronoi domain  $V(\Lambda; 0)$  is the rectangle  $(-1, 1) \times (-0.5, 0.5)$ . The neighbour set  $N(\Lambda) \subset \Lambda$  includes the 3rd neighbours  $(0, \pm 2)$  of the points  $(0, \pm 0.4) \in$

$V(\Lambda; 0)$ . Indeed, if in Definition 10.4.1  $\Lambda$  has a radius  $r < 2$ , then  $\Lambda \cap \bar{B}(0; r) \setminus \{0\} = \{(0, \pm 1)\}$  is in the 1-dimensional subspace (y-axis) of  $\mathbb{R}^2$ . For  $q = (0, 0.4)$ , considering all pairs  $(\mathbf{p}_1, \mathbf{p}_2)$  that generate  $\mathbb{R}^2$  among the four possibilities  $((0, \pm 1), (\pm 2, 0))$ , we find the base distances  $d_1(q) = 0.6 < d_2(q) = \sqrt{0.4^2 + 2^2} \approx 2.04$  for the 2nd and 3rd lattice neighbours  $p_1 = (0, 1)$  and  $p_2 = (\pm 2, 0)$  of  $q$ , respectively.  $\blacklozenge$

**Definition 10.4.3** (a distance-generic set). A periodic point set  $S = M + \Lambda \subset \mathbb{R}^n$  with the origin  $0 \in \Lambda \subset S$  is called *distance-generic* if the following conditions hold.

- (a) For any points  $p, q \in S \cap V(\Lambda; 0)$ , the vectors  $\mathbf{p}, \mathbf{q}$  are not orthogonal.
- (b) For vectors  $\mathbf{u}, \mathbf{v}$  between any two pairs of points in  $S$ , if  $|\mathbf{u}| = l|\mathbf{v}| \leq 2R(\Lambda)$  for  $l = 1, 2$ , then  $\mathbf{u} = \pm l\mathbf{v}$  and  $\mathbf{v} \in \Lambda$ .
- (c) For any point  $q \in S \cap V(\Lambda; 0)$ , let  $d_0 = |q|$  be its distance to the closest neighbour  $p_0 = 0$  in  $\Lambda$ . Take any linearly independent points  $p_1, \dots, p_n \in N(\Lambda)$  and any distances  $d_1 \leq \dots \leq d_n$  from  $q$  to some points in  $S \cap \bar{B}(0; 2R(\Lambda))$ . The  $n+1$  spheres  $\partial B(p_i; d_i)$  can meet at a single point of  $S \cap V(\Lambda; 0)$  only if  $d_1 \leq \dots \leq d_n$  are the base distances of  $q$  and only for two tuples  $p_1, \dots, p_n \in N(\Lambda)$  related by  $\mathbf{v} \mapsto -\mathbf{v}$ .  $\blacktriangle$

Condition 10.4.3(b) means that all inter-point distances are distinct apart from necessary exceptions due to periodicity. Since any periodic set  $S = M + \Lambda \subset \mathbb{R}^n$  is invariant under translations along all vectors of  $\Lambda$ , condition 10.4.3(b) for  $|\mathbf{v}| \leq 2R(\Lambda)$  can be checked only for vectors from all points of  $S$  in the original Voronoi domain  $V(\Lambda; 0)$  to all points in the domain  $3V(\Lambda; 0)$  extended by factor 3. Condition 10.4.3(b) implies that  $S$  has no points on the boundary  $\partial V(\Lambda; 0)$ , because any such point is equidistant to points  $0, v \in \Lambda$  and hence should belong to  $\Lambda$ . Let a *lattice distance* be the Euclidean distance from any  $p \in M = S \cap V(\Lambda; 0)$  to its lattice translate  $p + \mathbf{v}$  for all  $\mathbf{v} \in \Lambda$ . Condition 10.4.3(a) guarantees that only a lattice distance  $d$  appears together with  $2d$  (and possibly with higher multiples) in a row of  $\text{PDD}(S; k)$ . Any such  $d$  and its multiples are repeated twice in every row, because  $\Lambda$  is centrally symmetric.

**Lemma 10.4.4** (almost any periodic set is distance-generic, [31, Lemma 5.7]). Let  $S = M + \Lambda \subset \mathbb{R}^n$  be any periodic point set. For any  $\varepsilon > 0$ , one can perturb coordinates of a basis of  $\Lambda$  and of points from  $M$  up to  $\varepsilon$  such that the resulting perturbation  $S'$  of  $S$  is a distance-generic periodic point set in the sense of Definition 10.4.3.  $\blacksquare$

The size  $m$  of a motif  $M$  is an isometry invariant because any isometry maps  $N$  to another hose motif of the same size. In dimensions  $n = 2, 3$ , any lattice  $\Lambda$  can be reconstructed from its complete isometry invariants [20, 19]. Theorem 10.4.5 reconstructs a periodic point set  $S = M + \Lambda \subset \mathbb{R}^n$  in any dimension  $n \geq 2$  from the invariant  $I(S)$  consisting of  $m$ ,  $\text{PDD}(S; k)$ , and (complete invariants of) a lattice  $\Lambda$  to satisfy condition 10.1.2(a) for distance-generic periodic sets  $S \subset \mathbb{R}^n$ . Recall that the *packing radius*  $R(\Lambda)$  is the smallest radius  $R$  such that  $\bigcup_{p \in \Lambda} \bar{B}(p; R) = \mathbb{R}^n$ , see Definition 9.2.3(b).

**Theorem 10.4.5** (generic completeness of PDD for periodic sets, [31, Theorem 5.8]). Let  $S = M + \Lambda \subset \mathbb{R}^n$  be any distance-generic periodic set. For any  $k$  such that all distances in the last column of  $\text{PDD}(S; k)$  are larger than  $2R(\Lambda)$ , the set  $S$  can be

reconstructed from  $\Lambda$ ,  $\text{PDD}(S; k)$  and the size  $m$  of a motif of  $S$ , uniquely under isometry in  $\mathbb{R}^n$ .  $\blacksquare$

## 10.5 Detecting near-duplicates in the world's largest databases

This section reports thousands of previously unknown (near-)duplicates in the world's largest databases [26, 15, 33, 18, 24]. The sizes in Table 10.1 below are the numbers of all periodic crystals (with no disorder and full geometric data) in September 2024 (total number is 1,847,462, see Table 10.5 and all experimental details in [31, appendix SM1]).

**Table 10.1** Links and versions of the world's largest materials databases, see their sizes in Table 10.5.

database and web address	version
CSD: Cambridge Structural Database, <a href="http://ccdc.cam.ac.uk">http://ccdc.cam.ac.uk</a>	version 6.00
COD: Crystallography Open Database, <a href="http://crystallography.net/cod">crystallography.net/cod</a>	July 30, 2024
ICSD: Inorganic Crystal Structures, <a href="http://icsd.products.fiz-karlsruhe.de">icsd.products.fiz-karlsruhe.de</a>	Feb 25, 2025
MP: Materials Project, <a href="http://next-gen.materialsproject.org">http://next-gen.materialsproject.org</a>	v2023.11.1
GNoME: <a href="https://github.com/google-deepmind/materials_discovery">github.com/google-deepmind/materials_discovery</a>	Nov 29, 2023

We first used the vector  $\text{ADA}(S; 100)$  to find nearest neighbours across all databases by  $k$ -d trees [14] up to  $L_\infty \leq 0.01\text{\AA}$ . Since the smallest inter-atomic distances are about  $1\text{\AA} = 10^{-10}\text{m}$ , atomic displacements up to  $0.01\text{\AA}$  are considered experimental noise. For the closest pairs found by  $\text{ADA}(S; 100)$ , the stronger  $\text{PDA}(S; 100)$  can have only equal or larger  $\text{EMD} \geq L_\infty$  by Theorem 10.3.5. The CSD, COD, ICSD should contain experimental structures. MP is obtained from ICSD by extra optimisation.

Table 10.2 shows that the well-curated 60-year-old CSD has 0.9% near-duplicate crystals, while more than a third of the ICSD consists of near-duplicates that are geometrically almost identical so that all atoms can be matched by an average perturbation up to  $0.01\text{\AA}$ . Table 1 in [2, section 6] reported many thousands of exact duplicates, where chemical elements were replaced while keeping all coordinates fixed. These replacements are physically impossible without more substantial perturbations. Five journals are investigating integrity [9].

The bold numbers in Table 10.2 count near-duplicates, and their percentages within each database, which should be filtered out, else the ground truth data becomes skewed. Table 10.3 confirms that cell-based comparisons miss near-duplicates as in Fig. 10.1.

**Table 10.2** Count and percentage of all ideal periodic crystals in each database (left) found to have a near-duplicate in other databases (top) by the distance  $\text{EMD} \leq 0.01\text{\AA}$  on matrices  $\text{PDA}(S; 100)$ .

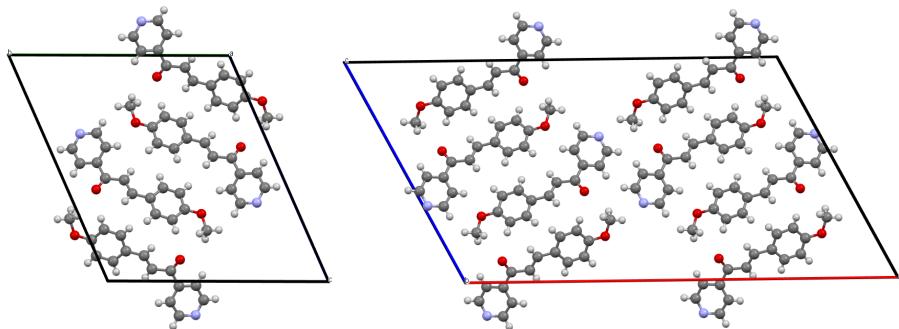
duplicates in databases	CSD		COD		ICSD		MP		GNoME	
	count	%	count	%	count	%	count	%	count	%
CSD	8343	0.92	283000	31.19	26506	2.92	33	0.00	1	0.00
COD	286663	80.18	19568	5.47	47065	13.16	5231	1.46	2705	0.76
ICSD	26853	15.78	69948	41.10	51085	30.01	27194	15.98	15449	9.08
MP	73	0.05	11986	7.82	15188	9.91	19177	12.51	10681	6.97
GNoME	2	0.00	1800	0.47	2614	0.68	3401	0.88	82859	21.53

**Table 10.3** Near-duplicates from Table 10.2 whose unit cells differ by  $0.01\text{\AA}$ . Unit cells are compared by the Chebyshev metric  $L_\infty$  between vectors of corresponding lengths of 3 edges and 3 face diagonals.

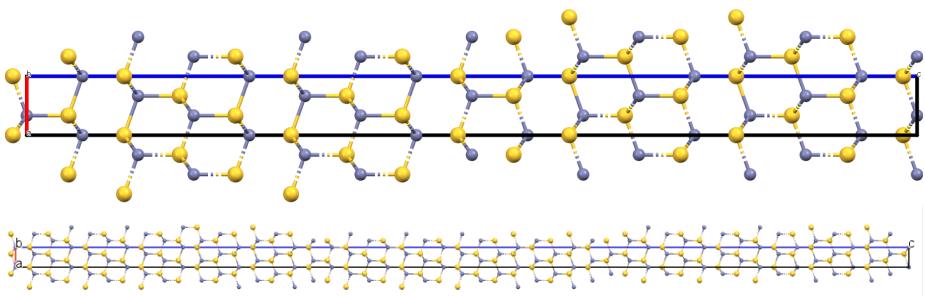
duplicates in databases	CSD		COD		ICSD		MP		GNoME	
	count	%	count	%	count	%	count	%	count	%
CSD	776	0.09	419	0.05	210	0.02	29	0.00	1	0.00
COD	472	0.13	7263	2.03	8629	2.41	5059	1.42	2684	0.75
ICSD	462	0.27	28863	16.96	42946	25.23	26554	15.60	15360	9.02
MP	70	0.05	11790	7.69	14915	9.73	18582	12.13	10608	6.92
GNoME	2	0.00	1786	0.46	2590	0.67	3346	0.87	60248	15.65

Figures 10.5, 10.6, 10.7, 10.8, 10.9, show near-duplicates with very different cells, which were counted in Table 10.3. In the past, the (near-)duplicates were impossible to detect at scale, because the traditional comparison through iterative alignment of 15 (by default) molecules by the COMPACK algorithm [10] is too slow for all-vs-all comparisons. Tables 10.4 and SM6 compare the running times: **minutes** of  $\text{PDA}(S; 100)$  vs **years** of RMSD, extrapolated for the same machine from the median time 117 milliseconds (582 ms on average) for 500 random pairs in the CSD. On the same 500 pairs,  $\text{PDA}(S; 100)$  for two crystals and EMD together took only 7.48 ms on average. All experiments were done on a modest desktop computer (AMD Ryzen 5 5600X 6-core, 32GB RAM).

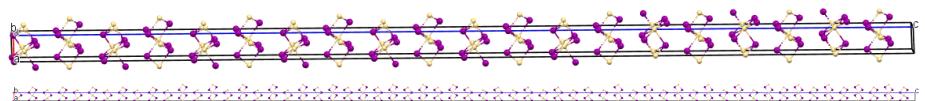
Table 10.6 compares the proven properties of past and new descriptors. All invariants based on cut-off atomic environments, such as MACE [5], discontinuously change under almost any perturbation that arbitrarily scales up a primitive cell, as in Fig. 10.1 (left).



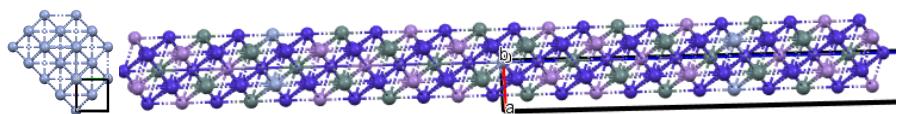
**Fig. 10.5** In the CSD, near-duplicates PUBTEM (left) and PUBTEM01 (right) have a very small EMD = 0.00038 Å on invariants PDA( $S$ ; 100), though their unit cells are rather different.



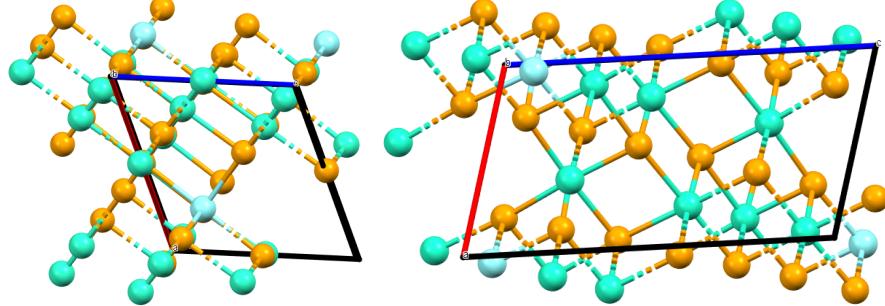
**Fig. 10.6** In the COD, near-duplicates 2310812 (top) and 2310813 (bottom) have a very small EMD = 0.0008 Å on invariants PDA( $S$ ; 100), though their unit cells differ by a factor of about 3.



**Fig. 10.7** In the ICSD, near-duplicates 42291 (top) and 42302 (bottom) have a very small EMD = 0.0024 Å on invariants PDA( $S$ ; 100), though their unit cells differ by a factor of about 3.



**Fig. 10.8** In the Materials Project, near-duplicate entries mp-90 (left) and mp-1221808 (right) have a very small EMD = 0.0087 Å on invariants PDA( $S$ ; 100), though their unit cells substantially differ.



**Fig. 10.9** In the GNoME, near-duplicates 4cb3b6ed9f (left) and 776c1b7570 (right) in the GNoME have EMD = 0.0079 Å on invariants PDA( $S$ ; 100), though their unit cells are very different.

**Table 10.4** Times to compute PDA( $S$ ; 100) and find all near-duplicates in Table 10.2 with EMD  $\leq$  0.01 Å across all major databases (seconds in the last 4 columns), compare with years in Table 10.5.

database	PDA, min:sec	EMD, min:sec	CSD	COD	ICSD	MP	GNoME
CSD	60:44	12:21	125.5	498.1	77.0	19.1	20.6
COD	30:16	16:29	524.5	122.0	235.1	79.6	27.0
ICSD	5:57	22:04	80.5	239.3	515.8	414.9	73.5
MP	1:40	13:31	28.2	82.9	413.8	222.8	63.0
GNoME	4:07	18:59	29.0	26.7	74.5	64.5	943.7

**Table 10.5** These times for all comparisons by COMPACK [10] are extrapolated from the median time of 117 ms on 500 random pairs from the CSD on the same machine, which completed Table 10.2 of near-duplicates across all the major databases within 2 hours.

database	periodic crystals	unordered pairs	COMPACK time, sec	years
CSD	907,246	411,547,198,635	$4.81 \times 10^{10}$	1526
COD	357,510	63,906,521,295	$7.48 \times 10^9$	237
ICSD	170,206	14,484,956,115	$1.69 \times 10^9$	53
MP	153,235	11,740,405,995	$1.37 \times 10^9$	43
GNoME	384,938	74,088,439,453	$8.67 \times 10^9$	274

The PDD remains continuous by taking into account only distances to neighbours rather than indices or relative positions of neighbours, which are discontinuous at cut-

off boundaries. Another exception is the complete isoset invariant in the next chapter.

Descriptor	Invariant	Continuity	Complete	Reconstruction	Time
primitive cell	✗	✗	✗	✗	✓
reduced cell	✓	✗	✗	✗	✓
space group	✓	✗	✗	✗	✓
PDF [27]	✓	✓	✗	✗	✓*
MACE [5]	✓	✗	✓*	✗	✓*
<hr/>					
densities [12]	✓	✓	✓*	✗	✓*
AMD [32]	✓	✓	✗	✗	✓
PDD [28]	✓	✓	✓*	✓*	✓
isosets [1, 3]	✓	✓	✓	✓	✓*

**Table 10.6** Comparison of crystal descriptors in the context of Problem 10.1.2. ✓\* in the ‘Computable’ column indicates that only an approximate algorithm exists for distances, and ✓\* in the ‘Complete’ and ‘Reconstruction’ columns means that the condition holds only for generic periodic sets as in 10.1.2(a).

## 10.6 Structural novelty of crystals and navigating materials space

This section leverages the strength of PDAs to quickly and continuously quantify the novelty of any periodic crystal relative to a given dataset in Definition 10.6.1.

**Definition 10.6.1** (Local Novelty Distance LND( $S; D$ )). Let  $D$  be a finite dataset of periodic point sets. Fix an integer  $k \geq 1$ . For any periodic point set  $S$ , the *Local Novelty Distance*  $\text{LND}(S; D) = \min_{Q \in D} \text{EMD}(\text{PDA}(S; k), \text{PDA}(Q; k))$  is the shortest  $L_\infty$ -based EMD distance from  $S$  to its nearest neighbour  $Q$  in the given crystal dataset  $D$ . ▲

If  $S$  is already contained in the dataset  $D$ , then  $\text{LND}(S; D) = 0$ , so  $S$  cannot be considered novel. Conversely, if  $\text{LND}(S; D) = 0$  then  $S$  highly likely belongs to  $S$ , because PDD( $S; 100$ ) distinguished all non-duplicate periodic crystals in the CSD.

$\text{LND}(S; D)$  is based on PDAs instead of PDDs because distances to  $k$ -th neighbours in  $\text{PDD}(S; k)$  asymptotically increase as  $\text{PPC}(S) \sqrt[3]{k}$  by Theorem 10.2.2. If crystals  $S, Q$  have  $\text{PPC}(S) \neq \text{PPC}(Q)$ , the Chebyshev distance  $L_\infty$  between rows of PDDs equals the largest absolute difference of  $i$ -th distances, which likely happens for  $i = k$ .

Hence, subtracting  $\text{PPC}(S)\sqrt[3]{k}$  in Definition 10.2.3 makes any metric on PDAs more informative than on PDDs. If a newly synthesised periodic crystal  $S$  is a near-duplicate of some known  $Q \in D$ , then  $\text{LND}(S; D)$  is small as justified below.

**Theorem 10.6.2** (continuity of  $\text{LND}$ , [29, Theorem 6]). For periodic point sets  $S, Q \subset \mathbb{R}^3$ , if  $S$  is obtained from  $Q$  in a dataset  $D$  by perturbing every point of  $Q$  up to  $\varepsilon < r(Q)$ , then  $\text{LND}(S; D) \leq 2\varepsilon$ . To get  $S$  from a periodic point set  $Q \in D$  with  $\text{LND}(S; D) < 2r(Q)$ , a point of  $Q$  should be perturbed by at least  $0.5\text{LND}(S; D)$ . ■

[29, section 3] describes how the 43 materials reported by Berkeley's A-lab [25] can be automatically positioned relative to the ICSD and the MP within the full Crystal Isometry Space  $\text{CIMS}(\mathbb{R}^3) = \bigcup_{m \geq 1} \text{CIMS}(\mathbb{R}^3; m)$  in seconds, see Table 10.7.

Stage	ICSD (s)	MP (s)
Binary search on ADA( $S; 100$ ) in the full database	3.023	2.450
PDA( $Q; 100$ ) for 100 neighbours $Q$ of $S$ found by ADA	5.272	5.990
EMD on PDAs for 100 neighbours $Q$ found by ADA	0.535	0.742
Elemental Mover's Distance (ElMD) for 100 neighbours	9.534	9.737

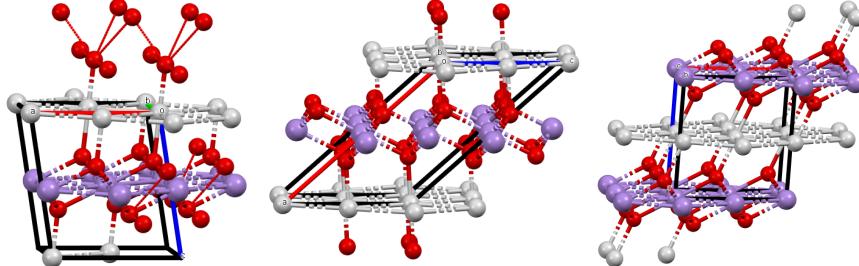
**Table 10.7** Time (seconds) to complete each stage of the process of finding nearest neighbours in the ICSD and Materials Project for 43 A-lab crystals [25] on a modest desktop computer. The binary search used 6-cores for multiprocessing.

Two A-lab crystals were found to already exist in the ICSD with the same composition:  $\text{KNaP}_6(\text{PbO}_3)_8$  matched ICSD 182501 reported in 2011 [4], and  $\text{MnAgO}_2$  matched ICSD 670065 reported as a hypothetical structure in 2015 [8]. In particular,  $\text{MnAgO}_2$  was one of three crystals that the later rebuttal said was synthesized successfully [21], and they go on to state that the material was first reported in 2021 [16] (ICSD 139006), after the snapshot used to train the GNoME, and so was not included in the original training data and could be considered a success. Our findings show this crystal did in fact exist in the ICSD prior to the 2021 snapshot. The pre-existing version of this crystal was not found by [21] using a unit cell search because the unit cell of ICSD 670065 significantly differs from that of the A-lab version or ICSD 139006, with the former listing its space group as  $A\ 2/m$  and the latter two having space group  $C\ 2/m$ , see Fig. 10.10. Such cell-based search can always miss near-duplicates as in Fig. 10.1 (left), while continuous invariants independent of a unit cell find near-duplicates despite disagreement on a space group, which breaks down under almost any noise.

Aside from the two structures above, all other A-lab crystals were found to have a geometric near-duplicate in the ICSD with a different composition. Many of these near-duplicates involve the substitution of only one atom, replacing a disordered site with a fully ordered one or adjusting the occupancy ratios of atoms at a site.

These structural analogues of A-lab's claimed crystals are not surprising, as the GNoME [23] used atomic substitution on existing crystals to generate potential new

**Fig. 10.10** **Left:**  $\text{MnAgO}_2$  synthesized by A-lab. **Middle:** ICSD entry 670065 with the same composition and  $\text{EMD} = 0.097\text{\AA}$  found by PDA( $S; 100$ ) in [29, Table 2], though its unit cell is very different from the cell of  $\text{MnAgO}_2$ . **Right:** another ICSD entry 139006 from 2021 matched by [21] and found by unit cell search, but is more distant from  $\text{MnAgO}_2$  by  $\text{EMD} = 0.368\text{\AA}$  on invariants PDA( $S; 100$ ).



ones without substantially changing the atomic geometry. The fact that pre-existing structures in the ICSD were missed by the later rebuttal [21] suggests that a more robust method is needed for comparing structures in the aid of materials discovery.

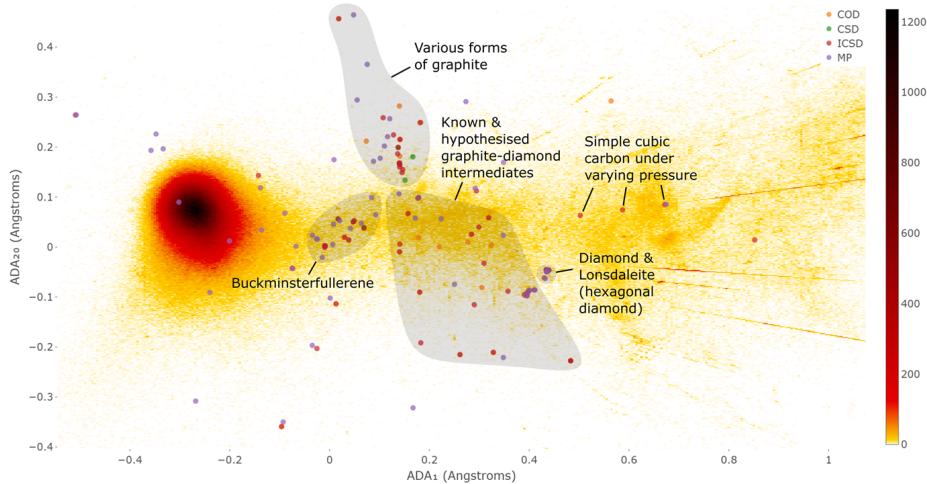
In conclusion, crystals were classified for hundreds of years almost exclusively by discrete tools such as space groups or by using reduced cells, which are unique in theory. Fig. 10.1 (left) showed that any known crystal can be disguised by changing a unit cell, shifting atoms a bit, changing chemical elements, and then claiming them as ‘new’.

Artificial near-duplicates threaten the integrity of experimental databases [9], which are skewed by previously undetectable near-duplicates. These challenges motivated the stronger question (*if different, by how much?*) that was formalised by Lipschitz continuity in condition 10.1.2(d) and inspired the research leading to this book.

Our future paper [30] extends the PDD to stronger higher-order invariants. Since the ultra-fast PDD distinguished all non-duplicate crystals among all experimental materials in the periodic case, these invariants already parametrise the known ‘universe’ of all existing crystals as ‘shiny stars’, while all not yet discovered crystals remain hidden in empty spots on the same map. Fig. 10.11 shows an example for two invariant coordinates.

## References

1. O Anosova and V Kurlin. An isometry classification of periodic point sets. In *Proceedings of Discrete Geometry and Mathematical Morphology*, pages 229–241, 2021.
2. Olga Anosova, Vitaliy Kurlin, and Marjorie Senechal. The importance of definitions in crystallography. *International Union of Crystallography Journal*, 11:453–463, 2024.
3. Olga Anosova, Daniel Widdowson, and Vitaliy Kurlin. Recognition of near-duplicate periodic patterns by continuous metrics with approximation guarantees. *Pattern Recognition*, 171:112108, 2026.
4. M Azrou et al. Rietveld refinements and vibrational spectroscopic studies of  $\text{Na}_{1-x}\text{K}_x\text{Pb}_4(\text{PO}_4)_3$  lacunar apatites ( $0 \leq x \leq 1$ ). *Journal of Physics and Chemistry of Solids*, 72(11):1199–1205, 2011.



**Fig. 10.11** All carbon allotropes (materials of pure carbon) on the heat map of four major databases.

5. Ilyes Batatia, David P Kovacs, Gregor Simm, Christoph Ortner, and Gábor Csányi. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields. *Advances in Neural Information Processing Systems*, 35:11423–11436, 2022.
6. Stepan Batsanov. Van der Waals radii of elements. *Inorganic mat.*, 37:871–885, 2001.
7. Hans-Georg Carstens et al. Geometrical bijections in discrete lattices. *Combinatorics, Probability and Computing*, 8(1-2):109–129, 1999.
8. T Cerqueira et al. Identification of novel Cu, Ag, and Au ternary oxides from global structural prediction. *Chemistry of Materials*, 27(13):4562–4573, 2015.
9. Dalmeet Singh Chawla. Crystallography databases hunt for fraudulent structures. *ACS Central Science*, 9:1853–1855, 2023.
10. J. Chisholm and S. Motherwell. Compack: a program for identifying crystal structure similarity using distances. *J. Applied Crystal.*, 38:228–231, 2005.
11. Michel Duneau and Christophe Oguey. Bounded interpolations between lattices. *Journal of Physics A: Mathematical and General*, 24(2):461, 1991.
12. 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.
13. Yury Elkin and Vitaliy Kurlin. A new near-linear time algorithm for k-nearest neighbor search using a compressed cover tree. In *Int. Conf. Machine Learning*, pages 9267–9311, 2023.
14. Fabian Gieseke, Justin Heinermann, Cosmin Oancea, and Christian Igel. Buffer kd trees: processing massive nearest neighbor queries on GPUs. In *Intern. Conf. Machine Learning*, pages 172–180, 2014.
15. Saulius Gražulis et al. Crystallography open database—an open-access collection of crystal structures. *J Appl. Crystallography*, 42(4):726–729, 2009.
16. Sean D Griesemer, Logan Ward, and Chris Wolverton. High-throughput crystal structure solution using prototypes. *Physical Review Materials*, 5(10):105003, 2021.
17. A Ivic, E Krätzel, M Kühleitner, and WG Nowak. Lattice points in large regions and related arithmetic functions: recent developments in a very classic topic. *Publications of the Scientific Society at the Johann Wolfgang Goethe University*, pages 89–128, 2006.
18. Anubhav Jain et al. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL materials*, 1(1), 2013.
19. V Kurlin. A complete isometry classification of 3D lattices. *arxiv:2201.10543*, 2022.
20. Vitaliy Kurlin. Mathematics of 2D lattices. *Found. Comp. Mathematics*, 24:805–863, 2024.

21. Josh Leeman, Yuhua Liu, Joseph Stiles, Scott B Lee, Prajna Bhatt, Leslie M Schoop, and Robert G Palgrave. Challenges in high-throughput inorganic materials prediction and autonomous synthesis. *PRX Energy*, 3(1):011002, 2024.
22. Ian Grant Macdonald. *Symmetric functions and Hall polynomials*. Oxford University Press, 1998.
23. Amil Merchant, Simon Batzner, Samuel S Schoenholz, Muratahan Aykol, Gowon Cheon, and Ekin Dogus Cubuk. Scaling deep learning for materials discovery. *Nature*, pages 80–85, 2023.
24. Amil Merchant and Ekin Dogus Cubuk. Millions of new materials discovered with deep learning. <https://deepmind.google/discover/blog/millions-of-new-materials-discovered-with-deep-learning/>, 2023.
25. Mark Peplow. Robot chemist sparks row with claim it created new materials. <https://www.nature.com/articles/d41586-023-03956-w>, 2023.
26. Robin Taylor and Peter A Wood. A million crystal structures: The whole is greater than the sum of its parts. *Chemical reviews*, 119(16):9427–9477, 2019.
27. Maxwell W Terban and Simon JL Billinge. Structural analysis of molecular materials using the pair distribution function. *Chemical Reviews*, 122(1):1208–1272, 2021.
28. Daniel Widdowson and Vitaliy Kurlin. Resolving the data ambiguity for periodic crystals. *Advances in Neural Information Processing Systems*, 35:24625–24638, 2022.
29. Daniel Widdowson and Vitaliy Kurlin. Geographic-style maps with a local novelty distance help navigate in the materials space. *Scientific Reports*, 15:27588, 2025.
30. Daniel Widdowson and Vitaliy Kurlin. Higher-order, generically complete, continuous, and polynomial-time isometry invariants of periodic sets. *arXiv:2509.15088*, 2025.
31. Daniel Widdowson and Vitaliy Kurlin. Pointwise distance distributions for detecting near-duplicates in large materials databases. *SIAM Journal on Applied Mathematics* (doi:10.1137/25M1736657, arxiv:2108.04798), 2026.
32. Daniel Widdowson, Marco M Mosca, Angeles Pulido, Andrew I Cooper, and Vitaliy Kurlin. Average minimum distances of periodic point sets - foundational invariants for mapping all periodic crystals. *MATCH Commun. Math. Comput. Chem.*, 87:529–559, 2022.
33. Dejan Zagorac, H Müller, S Ruehl, J Zagorac, and Silke Rehme. Recent developments in the inorganic crystal structure database. *J Applied Crystallography*, 52(5):918–925, 2019.



## Chapter 11

# Complete and continuous isosets of all periodic point sets in $\mathbb{R}^n$

**Abstract** This chapter adapts the general Geo-Mapping Problem to periodic point sets under rigid motion in  $\mathbb{R}^n$ . We introduce a complete isoset invariant with a Lipschitz continuous metric. For a fixed dimension  $n$ , the isoset is computable in polynomial time of the input size, while the distance metric is approximated with a constant factor in polynomial time. The isoset distinguished all known homometric crystals that have identical diffraction patterns and detected several pairs of exact (but unlikely) mirror images in the Cambridge Structural Database of experimental materials.

## 11.1 Geo-mapping for periodic point sets under rigid motion in $\mathbb{R}^n$

This chapter follows papers [1, 3]. In comparison with Problem 10.1.2 in the previous chapter, Problem 11.1.1 covers all (not only generic) periodic point sets and asks for completeness under rigid motion, which distinguishes mirror images.

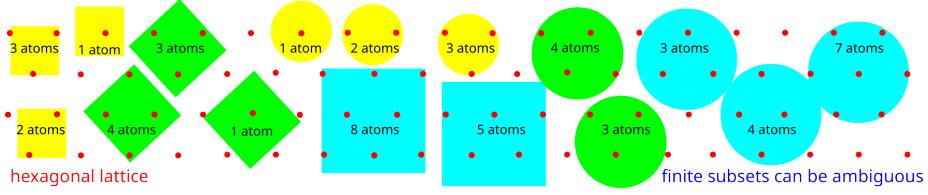
**Problem 11.1.1** (invariants of periodic point sets under rigid motion in  $\mathbb{R}^n$ ). Design an invariant  $I$  on the Crystal Rigid Space  $\text{CRIS}(\mathbb{R}^n; m)$  satisfying the conditions below.

- (a) *Completeness*: any periodic point sets  $S, Q \subset \mathbb{R}^n$  are related by rigid motion in  $\mathbb{R}^n$  if and only if  $I(S) = I(Q)$ .
- (b) *Reconstruction*: any periodic point set  $S \subset \mathbb{R}^n$  is reconstructable from its invariant value  $I(S)$ , uniquely under rigid motion in  $\mathbb{R}^n$ .
- (c) *Metric*: there is a distance  $d$  on the Crystal Rigid Space  $\text{CRIS}(\mathbb{R}^n; m)$  satisfying all metric axioms in Definition 1.3.1(a).
- (d) *Continuity*: there is a constant  $\lambda > 0$ , such that, for all sufficiently small  $\varepsilon > 0$  and periodic point set  $S, Q \subset \mathbb{R}^n$  if  $Q$  is obtained by perturbing every point of  $S$  up to Euclidean distance  $\varepsilon$ , then  $d(I(S), I(Q)) \leq \lambda\varepsilon$ .
- (e) *Computability*: for a fixed dimension  $n$ , the invariant  $I(S)$  of any periodic point set  $S$  and a reconstruction of  $S \subset \mathbb{R}^n$  from  $I(S)$  can be computed in times that depend polynomially on the motif size  $m$  of  $S$ , while the metric  $d(I(S), I(Q))$  can be approximated in polynomial time of the maximum motif size of  $S, Q$ . ★

One remaining limitation in Problem 11.1.1 is an approximate algorithm for a Lipschitz continuous metric, which will be improved to an exact one in future work.

Fig. 11.1 visualises the practical consequences of discontinuous cell-based representations. Starting from any periodic crystal with a unit cell of (say)  $m$  atoms, one can extend this cell to get an arbitrarily large motif of  $2m$  or  $3m$  or  $sm$  atoms for any integer scale  $s \geq 2$ . If  $s$  is not prime, then any extension can be done in different geometric ways. For example, an extension of unit cell  $U(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \subset \mathbb{R}^3$  by factor  $s = 8$ , can be by factor 8 in the direction of  $\mathbf{v}_i$  for  $i = 1, 2, 3$ , or by factor 4 in the direction of  $\mathbf{v}_1$  and by factor 2 in the direction of  $\mathbf{v}_2$  or  $\mathbf{v}_3$ , or by factor in each of the directions  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ .

any periodic crystal  has infinitely many **near-duplicates** with **larger motifs**                                    <img



**Fig. 11.2** Comparing periodic crystals by finite subsets is hard to justify because any periodic point set has many non-isometric subsets of different numbers of points within a box or a ball of a fixed size.

## 11.2 Isotree of local clusters of points in a periodic set in $\mathbb{R}^n$

This section defines the complete invariant [1] based on local clusters and their symmetry groups, which were previously studied in [5, 6].

**Definition 11.2.1** (*global* clusters and  $m$ -*regular* periodic sets). **(a)** For any point  $p$  in a periodic set  $S \subset \mathbb{R}^n$ , the *global cluster* is  $C(S, p) = \{\mathbf{q} - \mathbf{p} \mid q \in S\}$ . For any  $p, q \in \mathbb{R}^n$ , let the set  $O(\mathbb{R}^n; p, q)$  consist of all isometries of  $\mathbb{R}^n$  that map  $p$  to  $q$ .

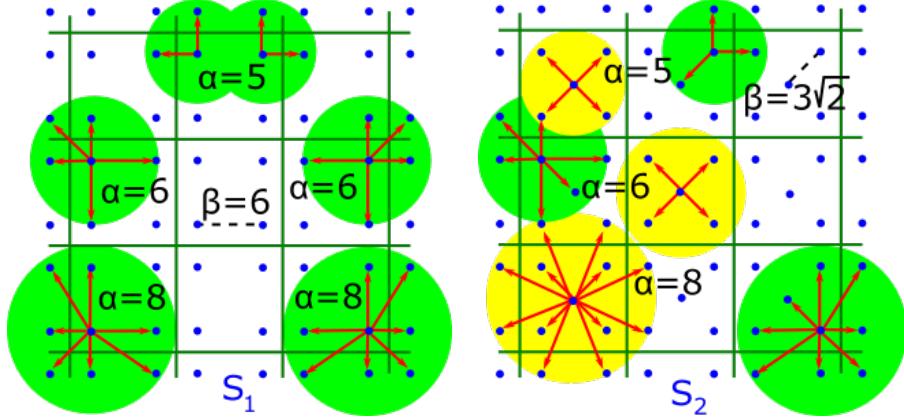
**(b)** Global clusters  $C(S, p)$  and  $C(S, q)$  are called *isometric* if there is  $f \in O(\mathbb{R}^n; p, q)$  such that  $f(S) = S$ . A periodic point set  $S \subset \mathbb{R}^n$  is called  $m$ -*regular* if all global clusters of  $S$  form exactly  $m \geq 1$  isometry classes.  $\blacktriangle$

For any point  $p \in S$ , its global cluster is a view of  $S$  from the position of  $p$ . We view all astronomical stars in the universe  $S$  from our Earth at  $p$ . Any lattice is 1-regular since all its global clusters are related by translations. Though global clusters  $C(S, p), C(S, q)$  at any different points  $p, q \in S$  contain the same set  $S$ , they may not match under the translation shifting  $p$  to  $q$ . The global clusters are infinite, hence distinguishing them under isometry is not easier than the original periodic sets. However, the  $m$ -regularity of a periodic set can be checked in terms of local  $\alpha$ -clusters below.

**Definition 11.2.2** (*local  $\alpha$ -clusters*  $C(S, p; \alpha)$  and *symmetry groups*  $\text{Sym}(S, p; \alpha)$ ). For a point  $p$  in a periodic point set  $S \subset \mathbb{R}^n$  and any  $\alpha \geq 0$ , the local  $\alpha$ -cluster  $C(S, p; \alpha)$  is the set of all vectors  $\mathbf{q} - \mathbf{p}$  such that  $q \in S$  and  $|\mathbf{q} - \mathbf{p}| \leq \alpha$ . Let the group  $O(\mathbb{R}^n; p)$  consist of all isometries that fix  $p$ . If  $p = 0$  is the origin,  $O(\mathbb{R}^n; 0)$  is the usual orthogonal group. The *symmetry group*  $\text{Sym}(S, p; \alpha)$  consists of all isometries  $f \in O(\mathbb{R}^n; p)$  that map  $C(S, p; \alpha)$  to itself so that  $f(p) = p$ .  $\blacktriangle$

Fig. 11.3 (left) shows the 1-regular periodic set  $S_1 \subset \mathbb{R}^2$  whose all points (close to vertices of square cells) have isometric global clusters related by translations and rotations through  $90^\circ, 180^\circ, 270^\circ$ . The 2-regular periodic set  $S_2$  has extra points at the centers of all square cells. The local  $\alpha$ -clusters around these centers are not isometric to  $\alpha$ -clusters around the points close to cell vertices for any  $\alpha \geq 3\sqrt{2}$ .

The 1-regular periodic point set  $S_1$  in Fig. 11.3 for any  $p \in S_1$  has the symmetry group  $\text{Sym}(S_1, p; \alpha) = O(\mathbb{R}^2)$  for  $\alpha \in [0, 4)$ . Then  $\text{Sym}(S_1, p; \alpha)$  stabilizes as  $\mathbb{Z}^2$  with one reflection for  $\alpha \geq 4$  as soon as  $C(S_1, p; \alpha)$  includes one more point.



**Fig. 11.3** **Left:** in  $\mathbb{R}^2$ , the periodic point set  $S_1$  has the square unit cell  $[0, 10]^2$  containing the four points  $(2, 2), (2, 8), (8, 2), (8, 8)$ , so  $S_1$  isn't a lattice, but is 1-regular by Definition 11.2.1, and  $\beta(S_1) = 6$ . All local  $\alpha$ -clusters of  $S_1$  are isometric, shown by red arrows for  $\alpha = 5, 6, 8$ , see Definition 11.2.2. **Right:**  $S_2$  has the extra point  $(5, 5)$  in the center of the cell  $[0, 10]^2$  and is 2-regular with  $\beta(S_2) = 3\sqrt{2}$ , so  $S_2$  has green and yellow isometry types of  $\alpha$ -clusters.

For any periodic set  $S$ , if  $\alpha$  is smaller than the minimum distance between all points of  $S$ , then any  $\alpha$ -cluster  $C(S, p; \alpha)$  is one point  $\{p\}$ . Its symmetry group consists of all isometries fixing the centre  $p$ , so  $\text{Sym}(S, p; \alpha) = O(\mathbb{R}^n; p)$ . When  $\alpha$  is increasing, the  $\alpha$ -clusters  $C(S, p; \alpha)$  become larger and there can be fewer (not more) isometries  $f \in O(\mathbb{R}^n; p)$  that bijectively map  $C(S, p; \alpha)$  to itself.

So the group  $\text{Sym}(S, p; \alpha)$  can become smaller (not larger) and eventually stabilises (stops changing), which is formalised in Definition 11.3.1. This stabilization uses the bridge length extending the idea of a longest edge in a Minimum Spanning Tree to a periodic set  $S$ , which does not easily reduce to the finite case [10].

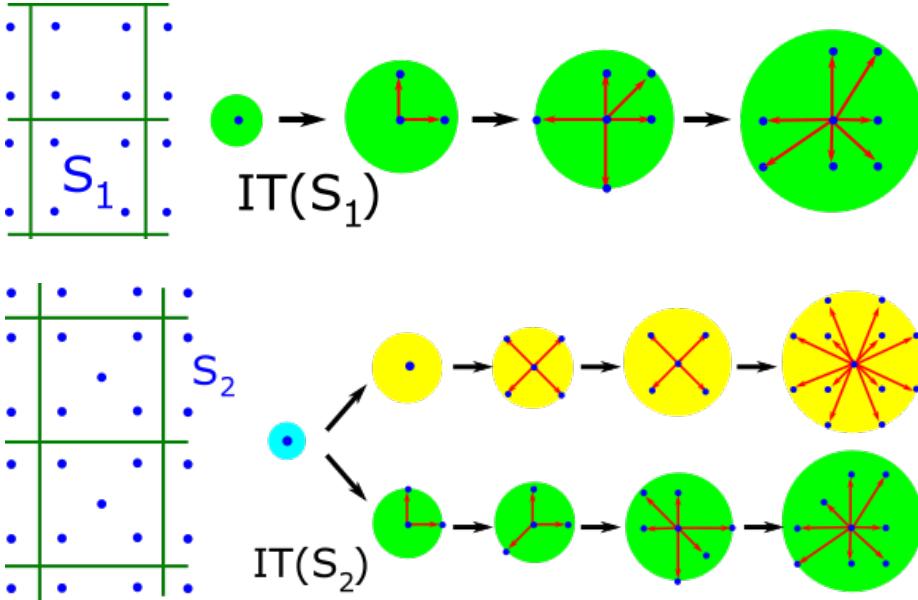
**Definition 11.2.3 (bridge length  $\beta(S)$ ).** For a periodic point set  $S \subset \mathbb{R}^n$ , the *bridge length* is a minimum distance  $\beta(S) > 0$  such that any  $p, q \in S$  can be connected by a sequence of points  $p_0 = p, p_1, \dots, p_k = q$  such that any two successive points  $p_{i-1}, p_i$  are close so that  $|\mathbf{p}_{i-1} - \mathbf{p}_i| \leq \beta(S)$  for  $i = 1, \dots, k$ .  $\blacktriangle$

The seminal result in [6, Theorem 1.3] described how a family of clusters determines a periodic point set under isometry. These results motivated the *isotree*, *stable radius*, and *isoset* in Definitions 11.2.4, 11.3.1, 11.3.5, respectively, leading to the isometry classification of periodic point sets via isosets in Theorem 11.3.7.

The *isotree* in Definition 11.2.4 is inspired by a clustering dendrogram because points of  $S$  split into isometry classes of  $\alpha$ -clusters at variable radii  $\alpha$ , not at a fixed  $\alpha$ .

Fig. 11.4 illustrates the isotrees of the periodic sets  $S_1, S_2$  in Fig. 11.3, as defined below.

**Definition 11.2.4 (isotree  $\text{IT}(S)$  of  $\alpha$ -partitions).** **(a)** Fix a periodic point set  $S \subset \mathbb{R}^n$ . Points  $p, q \in S$  are  $\alpha$ -equivalent if their  $\alpha$ -clusters  $C(S, p; \alpha)$  and  $C(S, q; \alpha)$  can



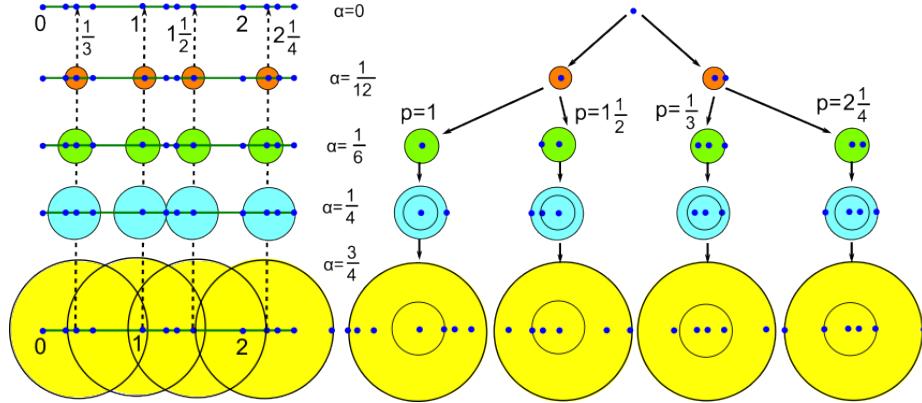
**Fig. 11.4** **Top:** the isotree  $IT(S_1)$  from Definition 11.2.4 of the 1-regular set  $S_1$  in Fig. 11.3 for any  $\alpha \geq 0$  has one isometry class of  $\alpha$ -clusters under rotations. **Bottom:** the isotree  $IT(S_2)$  of the 2-regular set  $S_2$  in Fig. 11.3 stabilizes with two non-isometric classes of  $\alpha$ -clusters for  $\alpha \geq 4$ .

be related by an isometry that matches their centres. The *isometry class*  $[C(S, p; \alpha)]$  consists of all  $\alpha$ -clusters isometric to  $C(S, p; \alpha)$ . The  $\alpha$ -partition  $P(S; \alpha)$  is the splitting of  $S$  into  $\alpha$ -equivalence classes of points. Call a value  $\alpha$  *singular* if  $P(S; \alpha) \neq P(S; \alpha - \varepsilon)$  for any small enough  $\varepsilon > 0$ .

**(b)** Represent each  $\alpha$ -equivalence class by a vertex of the *isotree*  $IT(S)$ . The top vertex of  $IT(S)$  represents the 0-equivalence class coinciding with  $S$ . For any successive singular values  $\alpha < \alpha'$ , connect the vertices representing any classes  $A \in P(S; \alpha)$  and  $A' \in P(S; \alpha')$  such that  $A' \subset A$  by an edge of the length  $\alpha' - \alpha$  in  $IT(S)$ .  $\blacktriangle$

For any periodic point set  $S \subset \mathbb{R}^n$ , the root vertex of  $IT(S)$  at  $\alpha = 0$  is the single class  $S$ , because any 0-cluster  $C(S, p; 0)$  of a point  $p \in S$  consists only of its centre  $p$ . When  $\alpha$  is increasing,  $\alpha$ -clusters  $C(S, p; \alpha)$  include more points and hence may not be isometric. In other words, any  $\alpha$ -equivalence class from  $P(S; \alpha)$  may split into two or more classes, which cannot merge at any larger  $\alpha'$ . Branched vertices of  $IT(S)$  correspond to the values of  $\alpha$  when an  $\alpha$ -equivalence class is split into subclasses for  $\alpha'$  slightly larger than  $\alpha$ . So the number  $|P(S; \alpha)|$  of  $\alpha$ -equivalence is non-decreasing in  $\alpha$ , see Fig. 11.5. The isotree  $IT(S)$  is continuously parametrised by  $\alpha \geq 0$  and is visualised as a tree of  $\alpha$ -equivalence classes in Fig. 11.5, due to Lemma 11.2.5.

Our proofs of Lemmas 11.2.5, 11.2.6, 11.2.7, and 11.3.2 were peer-reviewed but were published only in appendices of [3] online, so we include them for completeness.



**Fig. 11.5 Left:** the 1-dimensional set  $S_4 = \{0, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}\} + \mathbb{Z}$  has four points in the unit cell  $[0, 1]$  and is 4-regular by Definition 11.2.1. **Right:** the colored disks show  $\alpha$ -clusters in the line  $\mathbb{R}$  with radii  $\alpha = 0, \frac{1}{12}, \frac{1}{6}, \frac{1}{4}, \frac{3}{4}$  and represent points in the isoset tree  $IT(S_4)$  from Definition 11.2.4.

**Lemma 11.2.5** (properties of  $\alpha$ -partitions and isosets, [3, Lemma 4.1]). The isoset tree  $IT(S)$  from Definition 11.2.4 has the following properties.

- (a) for  $\alpha = 0$ , the  $\alpha$ -partition  $P(S; 0)$  consists of one class.
- (b) if  $\alpha < \alpha'$ , then  $\text{Sym}(S, p; \alpha') \subseteq \text{Sym}(S, p; \alpha)$  for any point  $p \in S$ .
- (c) If  $\alpha < \alpha'$ , the  $\alpha'$ -partition  $P(S; \alpha')$  refines  $P(S; \alpha)$ , i.e. any  $\alpha'$ -equivalence class from  $P(S; \alpha')$  is included into an  $\alpha$ -equivalence class from the partition  $P(S; \alpha)$ . So the cluster count  $|P(S; \alpha)|$  is a non-strictly increasing integer-valued function of  $\alpha$ , i.e.  $|\text{Iso}(S; \alpha)| \leq |\text{Iso}(S; \alpha')|$  for  $\alpha < \alpha'$ .  $\blacksquare$

*Proof.* (a) Let  $\alpha \geq 0$  be smaller than the minimum distance  $2r(S)$  between any points of  $S$ . Then any cluster  $C(S, p; \alpha)$  is the single-point set  $\{p\}$ . All these 1-point clusters are isometric to each other. So  $|P(S; \alpha)| = 1$  for  $\alpha < 2r(S)$ .

(b) For any  $p \in S$ , the inclusion of clusters  $C(S, p; \alpha) \subseteq C(S, p; \alpha')$  implies that any isometry  $f \in O(\mathbb{R}^n; p)$  that isometrically maps the larger cluster  $C(S, p; \alpha')$  to itself also maps the smaller cluster  $C(S, p; \alpha)$  to itself. Hence any element of  $\text{Sym}(S, p; \alpha') \subseteq O(\mathbb{R}^n; p)$  belongs to  $\text{Sym}(S, p; \alpha)$ .

(c) If points  $p, q \in S$  are  $\alpha'$ -equivalent at the larger radius  $\alpha'$ , i.e. the clusters  $C(S, p; \alpha')$  and  $C(S, q; \alpha')$  are related by an isometry from  $O(\mathbb{R}^n; p, q)$ , then  $p, q$  are  $\alpha$ -equivalent at the smaller radius  $\alpha$ . Hence any  $\alpha'$ -equivalence class of points in  $S$  is a subset of an  $\alpha$ -equivalence class in  $S$ .  $\square$

The  $\alpha$ -clusters of the periodic sequence  $S_4 \subset \mathbb{R}$  in Fig. 11.5 are intervals in  $\mathbb{R}$ , shown as disks for better visibility. In Fig. 11.5, the initial 1-point class persists until  $\alpha = \frac{1}{12}$ , when all points  $p \in S_4$  are split into two classes: one represented by 1-point cluster  $\{p\}$  for  $p \in \{0, \frac{1}{2}\} + \mathbb{Z}$ , and another represented by 2-point clusters  $\{p, p + \frac{1}{12}\}$ ,  $p \in \{\frac{1}{4}, \frac{1}{3}\} + \mathbb{Z}$ .

The sequence  $S_4$  has four  $\alpha$ -equivalence classes for any radius  $\alpha \geq \frac{1}{6}$ . For any point  $p \in \mathbb{Z} \subset S_4$ , the symmetry group  $\text{Sym}(S_4, p; \alpha) = \mathbb{Z}_2$  is generated by the reflection in  $p$  for  $\alpha \in [0, \frac{1}{4})$ . For all  $p \in S_4$ , the symmetry group  $\text{Sym}(S_4, p; \alpha)$  is trivial for  $\alpha \geq \frac{1}{4}$ .

Lemmas 11.2.6 and 11.2.7 are key steps towards a complete classification of periodic point sets under isometry and rigid motion in Theorem 11.3.7.

**Lemma 11.2.6** (local extension). Let  $S, Q \subset \mathbb{R}^n$  be periodic point sets and  $\text{Sym}(S, p; \alpha - \beta) = \text{Sym}(S, p; \alpha)$  for some point  $p \in S$  and  $\alpha > \beta$ . Assume that there is an isometry  $g \in O(\mathbb{R}^n; p, q)$  such that  $g(C(S, p; \alpha)) = C(Q, q; \alpha)$ . Let  $f \in O(\mathbb{R}^n; p, q)$  be any isometry such that  $f(C(S, p; \alpha - \beta)) = C(Q, q; \alpha - \beta)$ . Then  $f$  isometrically maps the larger clusters:  $f(C(S, p; \alpha)) = C(Q, q; \alpha)$ . ■

*Proof.* The composition  $h = f^{-1} \circ g$  fixes  $p$  and isometrically maps  $C(S, p; \alpha - \beta)$  to itself, so  $h \in \text{Sym}(S, p; \alpha - \beta)$ . The condition  $\text{Sym}(S, p; \alpha - \beta) = \text{Sym}(S, p; \alpha)$  implies that  $h \in \text{Sym}(S, p; \alpha)$ , so the isometry  $h \in O(\mathbb{R}^n; p)$  isometrically maps the larger cluster  $C(S, p; \alpha)$  to itself. Then the given isometry  $f = g \circ h^{-1}$  isometrically maps  $C(S, p; \alpha)$  to  $f(C(S, p; \alpha)) = g(C(S, p; \alpha)) = C(Q, q; \alpha)$ . □

**Lemma 11.2.7** (global extension). Let periodic point sets  $S, Q \subset \mathbb{R}^n$  have a common stable radius  $\alpha$  satisfying Definition 11.3.1 for an upper bound  $\beta \geq \beta(S), \beta(Q)$ . Let  $I(S; \alpha) = I(Q; \alpha)$  and  $p \in S, q \in Q$  be any points with an isometry  $f \in O(\mathbb{R}^n; p, q)$  such that  $f(C(S, p; \alpha)) = C(Q, q; \alpha)$ . Then  $f(S) = Q$ . ■

*Proof.* To show that  $f(S) \subset Q$ , it suffices to check that the image  $f(a)$  of any point  $a \in S$  belongs to  $Q$ . By Definition 11.2.3 the points  $p, a \in S$  are connected by a sequence of points  $p = a_0, a_1, \dots, a_k = a \in S$  such that the distances  $|a_{i-1} - a_i|$  between any successive points have the upper bound  $\beta$  for  $i = 1, \dots, k$ .

We will prove that  $f(C(S, a_k; \alpha)) = C(Q, f(a_k); \alpha)$  by induction on  $k$ , where the base  $k = 0$  is given. The induction step below goes from  $i$  to  $i + 1$ .

The ball  $\bar{B}(a_i; \alpha)$  contains the smaller ball  $\bar{B}(a_{i+1}; \alpha - \beta)$  around the closely located center  $a_{i+1}$ . Indeed, since  $|a_{i+1} - a_i| \leq \beta$ , the triangle inequality for the Euclidean distance implies that any point  $a'_{i+1} \in \bar{B}(a_{i+1}; \alpha - \beta)$  with  $|a'_{i+1} - a_i| \leq \alpha - \beta$  satisfies

$$|a'_{i+1} - a_i| \leq |a'_{i+1} - a_{i+1}| + |a_{i+1} - a_i| \leq (\alpha - \beta) + \beta = \alpha, \text{ so } \bar{B}(a_{i+1}; \alpha - \beta) \subset \bar{B}(a_i; \alpha).$$

Then the inductive assumption  $f(C(S, a_i; \alpha)) = C(Q, f(a_i); \alpha)$  gives

$$\begin{aligned} f(C(S, a_{i+1}; \alpha - \beta)) &= f(C(S, a_i; \alpha)) \cap f(\bar{B}(a_{i+1}; \alpha - \beta)) = \\ &C(Q, f(a_i); \alpha) \cap \bar{B}(f(a_{i+1}); \alpha - \beta) = C(Q, f(a_{i+1}); \alpha - \beta) \end{aligned}$$

Due to  $I(S; \alpha) = I(Q; \alpha)$ , the isometry class of  $C(S, a_{i+1}; \alpha)$  equals an isometry class of  $C(Q, b_{i+1}; \alpha)$  for some point  $b_{i+1} \in Q$ , i.e. there is an isometry  $g \in O(\mathbb{R}^n; a_{i+1}, b_{i+1})$  such that  $g(C(S, a_{i+1}; \alpha)) = C(Q, b_{i+1}; \alpha)$ .

Since  $f \circ g^{-1} \in O(\mathbb{R}^n; b_{i+1})$  isometrically maps  $C(Q, b_{i+1}; \alpha - \beta)$  to  $C(Q, f(a_{i+1}); \alpha - \beta)$ , the points  $b_{i+1}, f(a_{i+1}) \in Q$  are in the same  $(\alpha - \beta)$ -equivalence class of  $Q$ .

By condition (11.3.1a), the splitting of the periodic point set  $Q \subset \mathbb{R}^n$  into  $\alpha$ -equivalence classes coincides with its splitting into  $(\alpha - \beta)$ -equivalence classes. Hence, the points  $b_{i+1}, f(a_{i+1}) \in Q$  are in the same  $\alpha$ -equivalence class of  $Q$ . Then  $C(Q, f(a_{i+1}); \alpha)$  is isometric to  $C(Q, b_{i+1}; \alpha) = g(C(S, a_{i+1}; \alpha))$ .

Now we can apply Lemma 11.2.6 for  $p = a_{i+1}, q = f(a_{i+1})$  and conclude that the given isometry  $f$ , which satisfies  $f(C(S, a_{i+1}; \alpha - \beta)) = C(Q, f(a_{i+1}); \alpha - \beta)$ , isometrically maps the larger clusters:  $f(C(S, a_{i+1}; \alpha)) = C(Q, f(a_{i+1}); \alpha)$ .

The induction step is finished. The inclusion  $f^{-1}(Q) \subset S$  is proved similarly.  $\square$

### 11.3 A complete isoset of periodic point sets under rigid motion in $\mathbb{R}^n$

This section shows that, for any periodic point set  $S \subset \mathbb{R}^n$ , the  $\alpha$ -partition of  $S$  stabilises, which allows us to form a complete invariant at a stable radius defined below.

**Definition 11.3.1** (the *minimum stable radius*  $\alpha(S)$ ). Let  $S \subset \mathbb{R}^n$  be a periodic point,  $\beta \geq \beta(S)$  be an upper bound of the bridge length  $\beta(S)$  from Definition 11.2.3. A radius  $\alpha \geq \beta$  is called *stable* if the following conditions hold:

- (a) the  $\alpha$ -partition  $P(S; \alpha)$  equals the  $(\alpha - \beta)$ -partition  $P(S; \alpha - \beta)$ ;
- (b) the groups stabilise so that  $\text{Sym}(S, p; \alpha) = \text{Sym}(S, p; \alpha - \beta)$  for any  $p \in S$ , i.e. any isometry  $f \in \text{Sym}(S, p; \alpha - \beta)$  preserves the larger cluster  $C(S, p; \alpha)$ .

A minimum value of a stable radius  $\alpha$  satisfying 11.3.1(a,b) for  $\beta = \beta(S)$  from Definition 11.2.3 is called *the minimum stable radius* and denoted by  $\alpha(S)$ .  $\blacktriangle$

Due to the upper bounds in Lemma 11.3.3(b,c), the minimum stable radius  $\alpha(S) \geq 0$  exists and is achieved because  $P(S; \alpha)$  and  $\text{Sym}(S, p; \alpha)$  are continuous on the right (unchanged when  $\alpha$  increases by a sufficiently small value).

Any  $m$ -regular periodic point set  $S \subset \mathbb{R}^n$  has at most  $m$   $\alpha$ -equivalence classes, so the isotree  $IT(S)$  stabilises with maximum  $m$  branches. Though 11.3.1(b) is stated for all points  $p \in S$  for simplicity, it suffices to check condition 11.3.1(b) for points only from a finite motif  $M$  of  $S$  due to periodicity.

**Lemma 11.3.2** (all stable radii  $\alpha \geq \alpha(S)$ ). If  $\alpha$  is a stable radius of a periodic point set  $S \subset \mathbb{R}^n$ , then so is any larger radius  $\alpha' > \alpha$ . Then all stable radii form the interval  $[\alpha(S), +\infty)$ , where  $\alpha(S)$  is the minimum stable radius of  $S$ .  $\blacksquare$

*Proof.* Due to Lemma (11.2.5bc), conditions (11.3.1ab) imply that the  $\alpha'$ -partition  $P(S; \alpha')$  and the symmetry groups  $\text{Sym}(S, p; \alpha')$  remain the same for all  $\alpha' \in [\alpha - \beta(S), \alpha]$ , where  $\beta(S)$  is the bridle length. We need to show that they remain the same for any  $\alpha' > \alpha$  and will apply Lemma 11.2.7 for  $S = Q$  and  $\beta = \beta(S)$ .

Let points  $p, q \in S$  be  $\alpha$ -equivalent, i.e. there is an isometry  $f \in O(\mathbb{R}^n; p, q)$  such that  $f(C(S, p; \alpha)) = C(S, q; \alpha)$ . By Lemma 11.2.7,  $f$  isometrically maps the full set  $S$  to itself. Then all larger  $\alpha'$ -clusters of  $p, q$  are matched by  $f$ , so  $p, q$  are  $\alpha'$ -equivalent and  $P(S; \alpha) = P(S, \alpha')$ . Similarly, any isometry  $f \in \text{Sym}(S, p; \alpha)$  by Lemma 11.2.7 for  $S = Q$  and  $p = q$ , isometrically maps the full set  $S$  to itself. Then  $\text{Sym}(S, p; \alpha')$  coincides with  $\text{Sym}(S, p; \alpha)$  for any  $\alpha' > \alpha$ .  $\square$

All stable radii of  $S$  form the interval  $[\alpha(S), +\infty)$  by Lemma 11.3.2. The periodic sequence  $S_4$  in Fig. 11.5 has  $\beta(S_4) = \frac{1}{2}$  and  $\alpha(S) = \frac{3}{4}$  since the  $\alpha$ -partition and symmetry groups  $\text{Sym}(S_4, p; \alpha)$  are stable for  $\frac{1}{4} \leq \alpha \leq \frac{3}{4}$ .

Condition 11.3.1(b) doesn't follow from condition 11.3.1(a) due to the following example. Let  $\Lambda$  be the 2D lattice with the basis  $(1, 0)$  and  $(0, \beta)$  for  $\beta > 1$ . Then  $\beta$  is the bridge length of  $\Lambda$ . Condition 11.3.1(a) is satisfied for any  $\alpha \geq 0$ , because all points of any lattice are equivalent under translations.

However, condition 11.3.1(b) fails for any  $\alpha < \beta + 1$ . Indeed, the  $\alpha$ -cluster of the origin  $(0, 0)$  contains five points  $(0, 0), (\pm 1, 0), (0, \pm \beta)$ , whose symmetries are generated by the two reflections in the axes  $x, y$ , but the  $(\alpha - \beta)$ -cluster of the origin  $(0, 0)$  consists of its centre and has the symmetry group  $O(\mathbb{R}^2)$ .

It is possible that condition 11.3.1(b) might imply 11.3.1(a), but in practice it makes sense to verify 11.3.1(b) only after checking much simpler condition 11.3.1(a). Both conditions are essentially used in the proof of Isometry Classification Theorem 11.3.7.

Conditions 11.3.1(ab) appeared in [6] with different notations  $\rho, \rho + t$ . Since many applied papers use  $\rho$  for the physical density and have many types of bond distances, we replaced  $t$  and  $\rho + t$  with the bridge length  $\beta$  and radius  $\alpha$ , respectively, as for growing  $\alpha$ -shapes in Topological Data Analysis [12].

Recall that the *covering radius*  $R(S)$  of a periodic point set  $S \subset \mathbb{R}^n$  is the minimum radius  $R > 0$  such that  $\bigcup_{p \in S} \bar{B}(S; R) = \mathbb{R}^n$ , or the largest radius of an open ball in the complement  $\mathbb{R}^n \setminus S$ . For  $m$ -regular point sets in  $\mathbb{R}^n$ , an upper bound of  $\alpha(S)$  can be extracted from [6, Theorem 1.3] whose proof motivated a stronger bound in Lemma 11.3.3(c), see comparisons in Example 11.3.4(c).

A periodic point set  $S$  is *locally antipodal* if the local cluster  $C(S, p; 2R(S))$  is centrally symmetric for any point  $p \in S$ , i.e. bijectively maps to itself under  $\mathbf{q} \mapsto 2\mathbf{p} - \mathbf{q}$ ,  $\mathbf{q} \in \mathbb{R}^n$ . The important result in [7, Theorem 1] says that all locally antipodal Delone sets, hence all periodic sets  $S$ , are globally antipodal, i.e.  $S$  is preserved under the isometry  $\mathbf{q} \mapsto 2\mathbf{p} - \mathbf{q}$  for any fixed  $p \in S$ , e.g. any lattice is antipodal.

**Lemma 11.3.3** (upper bounds for a stable radius  $\alpha(S)$ , [3, Lemma 3.6]). **(a)** Let  $S \subset \mathbb{R}^n$  be a periodic point set with a unit cell  $U$ , which has the longest edge  $b$  and longest diagonal  $d$ . Set  $r(U) = \max\{b, \frac{d}{2}\}$ . Then the bridge length  $\beta(S)$  from Definition 11.2.3 has the upper bound  $\min\{2R(S), r(U)\} \geq \beta(S)$ .

**(b)** For any antipodal periodic set  $S \subset \mathbb{R}^n$  whose covering radius is  $R(S)$ , the minimum stable radius has the upper bound  $2R(S) + \beta(S) > \alpha(S)$ .

**(c)** Let  $S \subset \mathbb{R}^n$  be any periodic point set with the bridge length  $\beta$ . For any point  $p \in S$  and a radius  $\alpha_0 \geq 2R(S)$ , the order  $|\text{Sym}(S, p; \alpha_0)|$  of the group  $\text{Sym}(S, p; \alpha_0)$  should be finite. Let  $p_1, \dots, p_m \in S$  be all points of an asymmetric unit of  $S$ . Set  $L = \left[ \sum_{i=1}^m (\log_2 |\text{Sym}(S, p_i; \alpha_0)| - \log_2 |\text{Sym}(S, p_i)|) \right]$ . Then the minimum stable radius  $\alpha(S)$  from Definition 11.3.1 has the upper bound  $\alpha_0 + (L + m)\beta \geq \alpha(S)$ . If  $\alpha_0 = 2R(S)$ , then  $(L + m + 1)2R(S) \geq \alpha(S)$ .  $\blacksquare$

The upper bound in Lemma 11.3.3(a) holds for any unit cell of  $S$ . If a cell is non-reduced and too long, its reduced form can have smaller bounds for  $\beta(S)$ .

**Example 11.3.4** (upper bounds for  $\alpha(S)$  and  $\beta(S)$ ). Let  $\Lambda(b) \subset \mathbb{R}^n$  be a lattice whose unit cell is a rectangular box with the longest edge  $b \geq 1$ .

**(a)** In Lemma 11.3.3(a), the upper bound  $b \geq \beta(S)$  is tight because  $\beta(\Lambda(b)) = b$ .

**(b)** In Lemma 11.3.3(b), the ratio  $(2R(S) + \beta(S))/\alpha(S) \geq 1$  tends to 1 as  $b \rightarrow +\infty$  for any fixed  $n$ . Indeed, a cluster  $C(\Lambda(b), 0; \alpha)$  is  $n$ -dimensional only for  $\alpha \geq b$ , so the group  $\text{Sym}(\Lambda(b), 0; \alpha)$  stabilises at  $\alpha = b$ , hence  $\alpha(S) = b + \beta(\Lambda(b)) = 2b$  is the minimum stable radius. The covering radius  $R(\Lambda(b))$  is half of the longest diagonal of the rectangular cell  $U$ . If  $b \rightarrow +\infty$  and all other sizes of  $U$  remain fixed, the ratio  $(2R(\Lambda(b)) + \beta(\Lambda(b)))/\alpha(S)$  tends to 1 for any fixed  $n$ .

**(c)** Lemma 11.3.3(c) was motivated by [6, Theorem 1.3], which implies the upper bound  $\beta(S) + 2m(n^2 + 1) \log_2(2 + R(S)/r(S)) > \alpha(S)$  for  $m$ -regular point sets. Let  $\Lambda \subset \mathbb{R}^2$  be a lattice whose unit cell is a rhombus with sides 1. Then  $m = 1, n = 2, r(\Lambda) = 0.5, \beta(\Lambda) = 1$ , and  $\alpha(\Lambda) = 2$ . If  $\Lambda$  deforms from a square lattice to a hexagonal lattice, the covering radius  $R(\Lambda)$  varies in the range  $[\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{2}}]$ . The past bound above gives the estimate  $1 + 2(2^2 + 1) \log_2(2 + \frac{2}{\sqrt{3}}) \approx 17.6 > \alpha(\Lambda) = 2$ . For any lattice  $\Lambda$  in this family, the symmetry group  $\text{Sym}(\Lambda, 0) = \text{Sym}(\Lambda, 0; 1)$  stabilises at  $\alpha_0 = 1$ .

Lemma 11.3.3(c) for  $\alpha_0 = 1$  gives  $L = \log_2(2) - \log_2(2) = 0$ , so the upper bound  $\alpha_0 + (L + m)\beta(S) \geq \alpha(S)$  is tight:  $2 \geq \alpha(\Lambda)$ . In practice, if  $L$  is large because some local clusters  $C(S; p; \alpha_0)$  have too many symmetries, one can increase the radius  $\alpha_0$  to reduce  $L$  for a better bound of  $\alpha(S)$ .  $\spadesuit$

Definition 11.3.5 reminds of the *isoset*, which was initially introduced in [1, Definition 9]. We also cover the case of rigid motion and prove Completeness Theorem 11.3.7 in the appendix in more detail than in [1, Theorem 9].

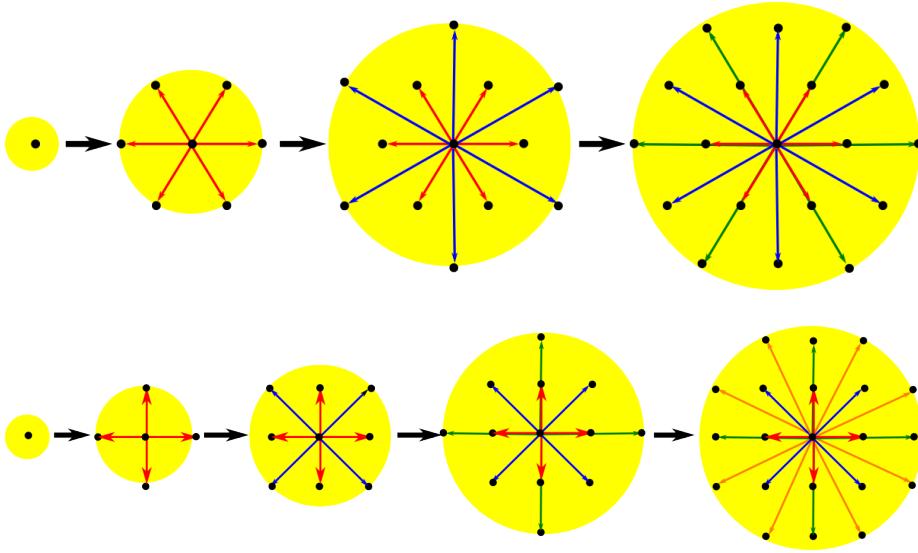
**Definition 11.3.5** (isoset  $I(S; \alpha)$  at a radius  $\alpha \geq 0$ ). Let a periodic point set  $S \subset \mathbb{R}^n$  have a motif  $M$  of  $m$  points. Split all points  $p \in M$  into  $\alpha$ -equivalence classes. Each  $\alpha$ -equivalence class of (say)  $k$  points in  $M$  can be associated with the *isometry class*  $\sigma = [C(S, p; \alpha)]$  of an  $\alpha$ -cluster centred at some  $p \in M$ . The *weight* of  $\sigma$  is  $w = k/m$ . Then the *isoset*  $I(S; \alpha)$  is the unordered set of all isometry classes  $(\sigma; w)$  with weights  $w$  for all points  $p$  in the motif  $M$ . If we replace isometry with rigid motion, we get the *oriented isoset*  $I^o(S; \alpha)$ .  $\blacktriangle$

All points  $p$  of a lattice  $\Lambda \subset \mathbb{R}^n$  from one  $\alpha$ -equivalence class for any radius  $\alpha \geq 0$  because all  $\alpha$ -clusters  $C(\Lambda, p; \alpha)$  are isometrically equivalent to each other by

translations. Hence the isoset  $I(\Lambda; \alpha)$  is one isometry class of weight 1 for  $\alpha \geq 0$ , see examples in Fig. 11.7. All isometry classes  $\sigma$  in  $I(S; \alpha)$  are in a 1-1 correspondence with all  $\alpha$ -equivalence classes in the  $\alpha$ -partition  $P(S; \alpha)$  from Definition 11.2.4.

Hence, the isoset  $I(S; \alpha)$  without weights can be viewed as a set of points in the isotree  $IT(S)$  at the radius  $\alpha$ . The size of the isoset  $I(S; \alpha)$  equals the number  $|P(S; \alpha)|$  of  $\alpha$ -equivalence classes in the  $\alpha$ -partition. Formally,  $I(S; \alpha)$  depends on  $\alpha$  because  $\alpha$ -clusters grow in  $\alpha$ . To distinguish any  $S, Q \subset \mathbb{R}^n$  under isometry, we will compare their isosets at a maximum stable radius of  $S, Q$ .

**Example 11.3.6** (isosets of simple lattices). **(a)** Any lattice  $\Lambda \subset \mathbb{R}^n$  is 1-regular by Definition 11.2.1 and can be assumed to contain the origin 0 of  $\mathbb{R}^n$ . Then the isoset  $I(\Lambda; \alpha)$  consists of a single isometry class of a cluster  $C(\Lambda, 0; \alpha)$ . So the isotree  $IT(\Lambda)$  is a linear path, which is horizontally drawn for the hexagonal and square lattices  $\Lambda_6, \Lambda_4$  in Fig. 11.6. If both  $\Lambda_6, \Lambda_4$  have a minimum inter-point distance 1, then the bridge length from Definition 11.2.3 is  $\beta = 1$ .



**Fig. 11.6** The isotree of any lattice  $\Lambda$  is  $[0, +\infty)$  is a line  $\mathbb{R}$  parametrised by the radius  $\alpha$ . **Top:** the isotree of the hexagonal lattice  $\Lambda_6$ . **Bottom:** the isotree of the square lattice  $\Lambda_4$ .

**(b)** For the hexagonal lattice  $\Lambda_6 \subset \mathbb{R}^2$ ,  $C(\Lambda_6, (0, 0); \alpha)$  includes points  $p \neq (0, 0)$  only for  $\alpha \geq 1$ . The cluster  $C(\Lambda_6, (0, 0); 1) = \{(0, 0), (\pm 1, 0), (\pm \frac{1}{2}, \pm \frac{\sqrt{3}}{2})\}$  appears in the 2nd step of Fig. 11.6 (left). The symmetry group  $\text{Sym}(\Lambda_6, (0, 0); \alpha)$  becomes the dihedral group  $D_6$  (all symmetries of a regular hexagon) for  $\alpha \geq 1$ . Hence any  $\alpha \geq \beta + 1 = 2$  is stable. The isoset  $I(\Lambda_6; 1)$  is the isometry class of the cluster  $C(\Lambda_6, (0, 0); 1)$  of six vertices of the regular hexagon and its centre.

**(c)** For the square lattice  $\Lambda_4 \subset \mathbb{R}^2$ ,  $C(\Lambda_4, (0, 0); \alpha)$  has points  $p \neq (0, 0)$  only for  $\alpha \geq 1$ .  $C(\Lambda_4, (0, 0); 2) = \{(0, 0), (\pm 1, 0), (0, \pm 1), (\pm\sqrt{2}, \pm\sqrt{2}), (\pm 2, 0), (0, \pm 2)\}$  includes the origin  $(0, 0)$  with its 12 neighbors in the 4th step of Fig. 11.6 (right). The group  $\text{Sym}(\Lambda_4, (0, 0); \alpha)$  becomes the dihedral group  $D_4$  (all symmetries of a square) for  $\alpha \geq 1$ . So any  $\alpha \geq \beta + 1 = 2$  is stable. The isoset  $I(\Lambda_4; 1)$  is the isometry class of  $C(\Lambda_4, (0, 0); 1)$  of four vertices of the square and its centre.  $\blacklozenge$

An equality  $\sigma = \xi$  between isometry classes of clusters means that some (hence any) clusters  $C(S, p; \alpha)$  and  $C(Q, q; \alpha)$  representing  $\sigma, \xi$ , respectively, are related by  $f \in O(\mathbb{R}^n; p, q)$ , which will be algorithmically tested in Corollary 11.5.3.

**Theorem 11.3.7** (isometry classification of periodic point sets, [3, Theorem 3.10]). For any periodic point sets  $S, Q \subset \mathbb{R}^n$ , let  $\alpha$  be a common stable radius satisfying Definition 11.3.1 for an upper bound  $\beta \geq \beta(S), \beta(Q)$ . Then  $S, Q$  are isometric (related by rigid motion, respectively) if and only if there is a bijection  $\varphi : I(S; \alpha) \rightarrow I(Q; \alpha)$  (between oriented isosets, respectively) that preserves all their weights.  $\blacksquare$

Theorem 11.3.7 was inspired by the seminal result in [6, Theorem 1.3] saying that, for a multi-regular point set  $X$ , “the only Delone sets  $Y$  all of whose  $\rho$ -stars are isometric to  $\rho$ -stars of  $X$  are sets globally isometric to  $X$ ”. After renaming  $\rho$ -stars as  $\alpha$ -clusters, we collected their isometry classes (with weights) into the *isoset* to rephrase [6, Theorem 1.3] as a classification of all periodic point sets by isosets.

The  $\alpha$ -equivalence and isoset in Definition 11.3.5 can be refined by labels such as chemical elements, which keeps Theorem 11.3.7 valid for labelled points.

When comparing sets from a finite database, it suffices to build their isosets only up to a common upper bound of a stable radius  $\alpha$  in Lemma 11.3.3(c).

## 11.4 Continuous metrics on isometry classes of periodic sets in $\mathbb{R}^n$

This section proves the continuity of the isoset  $I(S; \alpha)$  in Theorem 11.4.5 by using the Earth Mover’s Distance (EMD) from Definition 11.4.3.

For a point  $p \in \mathbb{R}^n$  and a radius  $\varepsilon$ , the *closed ball*  $\bar{B}(p; \varepsilon) = \{q \in \mathbb{R}^n \mid |\mathbf{q} - \mathbf{p}| \leq \varepsilon\}$  has as its the boundary  $(n - 1)$ -dimensional sphere  $\partial\bar{B}(p; \varepsilon) \subset \mathbb{R}^n$ . The  $\varepsilon$ -*offset* of any set  $C \subset \mathbb{R}^n$  is the Minkowski sum  $C + \bar{B}(0; \varepsilon) = \{\mathbf{p} + \mathbf{q} \mid \mathbf{p} \in C, \mathbf{q} \in \bar{B}(0; \varepsilon)\}$ .

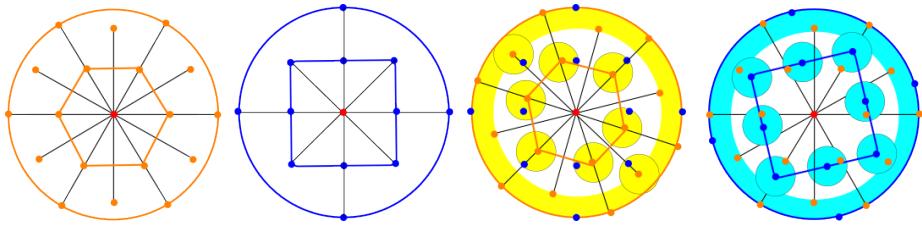
Then the directed *Hausdorff* distance from Example 1.3.3(b)  $d_H(C, D)$  is the minimum radius  $\varepsilon \geq 0$  such that  $C \subseteq D + \bar{B}(0; \varepsilon)$ . Definition 11.4.1 introduces the crucial new metric, which will be explicitly computed in Lemma 11.5.5.

**Definition 11.4.1** (boundary tolerant metric BT on isometry classes of clusters). For a radius  $\alpha$  and periodic point sets  $S, Q \subset \mathbb{R}^n$ , let clusters  $C(S, p; \alpha), C(Q, q; \alpha)$  represent isometry classes  $\sigma \in I(S; \alpha), \xi \in I(Q; \alpha)$ , respectively. The *boundary tolerant* metric  $BT(\sigma, \xi)$  is defined as the minimum  $\varepsilon \geq 0$  such that

- (11.4.1a)  $C(Q, q; \alpha - \varepsilon) \subseteq f(C(S, p; \alpha)) + \bar{B}(0; \varepsilon)$  for some  $f \in O(\mathbb{R}^n; p, q)$ , and
- (11.4.1b)  $C(S, p; \alpha - \varepsilon) \subseteq g(C(Q, q; \alpha)) + \bar{B}(0; \varepsilon)$  for some  $g \in O(\mathbb{R}^n; q, p)$ .  $\blacktriangle$

In Definition 11.4.1, if one cluster consists of only its centre, e.g.  $C(S, p; \alpha) = \{p\}$ , then the boundary tolerant metric is  $\text{BT} = \max\{|\mathbf{s} - \mathbf{q}| \mid s \in C(Q, q; \alpha)\}$ . [3, Lemma 4.2] proves that BT is independent of cluster representatives and satisfies all metric axioms from Definition 1.3.1(a).

**Example 11.4.2** (square lattice vs hexagonal). The isoset  $I(\Lambda; \alpha)$  of any lattice  $\Lambda \subset \mathbb{R}^n$  containing the origin 0 consists of a single isometry class  $[C(\Lambda, 0; \alpha)]$ , see Example 11.3.6. For the square (hexagonal) lattice with minimum inter-point distance 1 in Fig. 11.7, the cluster  $C(\Lambda, 0; \alpha)$  consists of only 0 for  $\alpha < 1$  and includes four (six) nearest neighbors of 0 for  $\alpha \geq 1$ . Hence  $\text{Sym}(\Lambda, 0; \alpha)$  stabilises as the symmetry group of the square (regular hexagon) for  $\alpha \geq 1$ . The lattices have the minimum stable radius  $\alpha(\Lambda) = 2$  and  $\beta(\Lambda) = 1$  by Example 11.3.4(c).



**Fig. 11.7** Example 11.4.2 computes the metric BT from Definition 11.4.1 for the isometry classes of the 2-clusters in the square and hexagonal lattices  $\Lambda_4, \Lambda_6$ . **1st:** the 2-cluster  $C(\Lambda_6, 0; 2)$  with its boundary circle  $\partial\bar{B}(0; 2)$ ; **2nd:** the 2-cluster  $C(\Lambda_4, 0; 2)$  with its boundary circle  $\partial\bar{B}(0; 2)$ ; **3rd:** for  $\varepsilon = \sqrt{2} - 1 \approx 0.41$ , the cluster  $C(\Lambda_6, 0; 2)$  is covered by the yellow  $\varepsilon$ -offset of  $C(\Lambda_6, 0; 2) \cup \partial\bar{B}(0; 2)$  rotated through  $15^\circ$  clockwise. **4th:**  $C(\Lambda_6, 0; 2)$  is covered by the blue  $\varepsilon$ -offset of  $C(\Lambda_4, 0; 2) \cup \partial\bar{B}(0; 2)$  rotated through  $15^\circ$  anticlockwise, so  $\text{BT} = \sqrt{2} - 1$ .

Fig. 11.7 shows the stable 2-clusters  $C(\Lambda_4, 0; 2)$  and  $C(\Lambda_6, 0; 2)$  of the square ( $\Lambda_4$ ) and hexagonal ( $\Lambda_6$ ) lattices. Without rotations, the 1st picture of Fig. 11.7 shows the directed Hausdorff distance  $d_H = \sqrt{(1 - \frac{\sqrt{3}}{2})^2 + (\frac{1}{2})^2} = \sqrt{2 - \sqrt{3}} \approx 0.52$  between clusters with the added boundary circle  $\partial B(0; 2)$ . Due to high symmetry, it suffices to consider rotations of the square vertex  $(1, 1)$  for angles  $\gamma \in [45^\circ, 60^\circ]$  because all other ranges can be isometrically mapped to this range for another vertex of the square.

We find the squared distances  $s_1(\gamma)$  and  $s_2(\gamma)$  from the vertex  $(\sqrt{2} \cos \gamma, \sqrt{2} \sin \gamma)$  rotated from  $(1, 1)$  at  $\gamma = 45^\circ$  through the angle  $\gamma - 45^\circ$  to its closest neighbors  $(\frac{1}{2}, \frac{\sqrt{3}}{2})$  and  $(\frac{3}{2}, \frac{\sqrt{3}}{2})$  in  $C(\Lambda_6, 0; 2)$ .

$$\begin{aligned} s_1(\gamma) &= \left| \left( \sqrt{2} \cos \gamma, \sqrt{2} \sin \gamma \right) - \left( \frac{1}{2}, \frac{\sqrt{3}}{2} \right) \right|^2 = \\ &= \left( \sqrt{2} \cos \gamma - \frac{1}{2} \right)^2 + \left( \sqrt{2} \sin \gamma - \frac{\sqrt{3}}{2} \right)^2 = \\ &= 3 - \sqrt{2} \cos \gamma - \sqrt{6} \sin \gamma, \end{aligned}$$

$$\frac{ds_1}{d\gamma} = \sqrt{2} \sin \gamma - \sqrt{6} \cos \gamma = 0, \tan \gamma = \sqrt{3}, \gamma = 60^\circ, s_1 = (\sqrt{2} - 1)^2$$

is minimal for the points in the line  $y = \sqrt{3}x$  at distances  $1, \sqrt{2}$  from 0.

$$\begin{aligned} s_2(\gamma) &= \left| \left( \sqrt{2} \cos \gamma, \sqrt{2} \sin \gamma \right) - \left( \frac{3}{2}, \frac{\sqrt{3}}{2} \right) \right|^2 = \\ &= \left( \sqrt{2} \cos \gamma - \frac{3}{2} \right)^2 + \left( \sqrt{2} \sin \gamma - \frac{\sqrt{3}}{2} \right)^2 = \\ &= 5 - 3\sqrt{2} \cos \gamma - \sqrt{6} \sin \gamma, \end{aligned}$$

$$\frac{ds_2}{d\gamma} = 3\sqrt{2} \sin \gamma - \sqrt{6} \cos \gamma = 0, \gamma = 30^\circ, s_2 = (\sqrt{3} - \sqrt{2})^2$$

is minimal for the points in the line  $y = \frac{x}{\sqrt{3}}$  at distances  $\sqrt{2}, \sqrt{3}$  from 0.

It might look that the second minimum is smaller. However, for the angle  $\gamma = 30^\circ$ , another vertex  $(-1, 1)$  rotated through  $\gamma - 45^\circ = -15^\circ$  has distance  $\sqrt{2} - 1$  to its closest neighbor  $(-\frac{1}{2}, \frac{\sqrt{3}}{2}) \in C(\Lambda_6, 0; 2)$ . For any angle  $\gamma \in [45^\circ, 60^\circ]$ , the second function has the minimum  $s_2(45^\circ) = 2 - \sqrt{3} = d_H^2$  in the 1st picture of Fig. 11.7.

Hence, the vertex  $(1, 1)$  has the minimum distance  $\sqrt{2} - 1 \approx 0.41 < \sqrt{2} - \sqrt{3} \approx 0.52$  in the 3rd picture of Fig. 11.7. All other points of the square cluster  $C(\Lambda_4, 0; 2)$  are even closer to their neighbors in  $C(\Lambda_6, 0; 2)$ . For example, the point  $(1, 0)$  rotated by  $15^\circ$  has the distance to  $(1, 0)$  equal to  $\sqrt{(\cos 15^\circ - 1)^2 + \sin^2 15^\circ} \approx 0.26$ .

The final picture in Fig. 11.7 confirms that all points of the hexagonal cluster  $C(\Lambda_6, 0; 2)$  are covered by the  $(\sqrt{2} - 1)$ -offset of  $C(\Lambda_4, 0; 2)$  and the boundary circle. So  $BT = \sqrt{2} - 1 \approx 0.41$ .  $\blacklozenge$

Non-isometric periodic sets  $S, Q$  can have isosets of different numbers of isometry classes. The distance between these weighted distributions of different sizes can be measured by the Earth Mover's Distance below.

**Definition 11.4.3** (Earth Mover's Distance on isosets). Let periodic point sets  $S, Q \subset \mathbb{R}^n$  have a common stable radius  $\alpha$  and isosets  $I(S; \alpha) = \{(\sigma_i, w_i)\}$  and  $I(Q; \alpha) = \{(\xi_j, v_j)\}$ , where  $i = 1, \dots, m(S)$  and  $j = 1, \dots, m(Q)$ . The *Earth Mover's Distance* is

$$EMD(I(S; \alpha), I(Q; \alpha)) = \sum_{i=1}^{m(S)} \sum_{j=1}^{m(Q)} f_{ij} BT(\sigma_i, \xi_j)$$

minimised over flows  $f_{ij} \in [0, 1]$  subject to the conditions  $\sum_{j=1}^{m(Q)} f_{ij} \leq w_i$  for  $i = 1, \dots, m(S)$ ,  $\sum_{i=1}^{m(S)} f_{ij} \leq v_j$  for  $j = 1, \dots, m(Q)$ , and  $\sum_{i=1}^{m(S)} \sum_{j=1}^{m(Q)} f_{ij} = 1$ .  $\blacktriangle$

**Example 11.4.4** (EMD for lattices with  $\text{BD} = +\infty$ ). [13, Example 2.1] showed that the lattices  $S = \mathbb{Z}$  and  $Q = (1 + \delta)\mathbb{Z}$  have the bottleneck distance  $\text{BD}(S, Q) = +\infty$  for any  $\delta > 0$ . We show that  $S, Q$  have Earth Mover's Distance  $\text{EMD} = 2\delta$  at their common stable radius  $\alpha = 2 + 2\delta$ . The bridge lengths are  $\beta(S) = 1$  and  $\beta(Q) = 1 + \delta$ . The  $\alpha$ -cluster  $C(S, 0; \alpha)$  contains non-zero points for  $\alpha \geq 1$ , e.g.  $C(S, 0; 1) = \{0, \pm 1\}$ .

The symmetry group  $\text{Sym}(S, 0; \alpha) = \mathbb{Z}_2$  includes a non-trivial reflection with respect to 0 for all  $\alpha \geq 1$ , so the stable radius of  $S$  is any  $\alpha \geq \beta + 1 = 2$ . Similarly,  $Q$  has  $\beta(Q) = 1 + \delta$  and stable radii  $\alpha \geq 2(1 + \delta)$ . The Earth Mover's Distance between  $I(S; \alpha)$  and  $I(Q; \alpha)$  at the common stable radius  $\alpha = 2 + 2\delta$  equals the metric  $\text{BT}$  between the only  $\alpha$ -clusters  $C(S, 0; \alpha) = \{0, \pm 1, \pm 2\}$  and  $C(Q, 0; \alpha) = \{0, \pm(1 + \delta), \pm 2(1 + \delta)\}$ .

By Definition 11.4.1 we look for a minimum  $\varepsilon > 0$  such that the cluster  $C(S, 0; \alpha - \varepsilon)$  is covered by  $\varepsilon$ -offsets of  $\pm(1 + \delta), \pm 2(1 + \delta)$  and vice versa. If  $\varepsilon < 2\delta$ , the points  $\pm 2 \in C(S, 0; \alpha - \varepsilon)$  cannot be  $\varepsilon$ -close to  $\pm(1 + \delta), \pm 1(1 + \delta)$ , but  $\varepsilon = 2\delta$  is large enough. The cluster  $C(Q, 0; \alpha - 2\delta) = \{0, \pm(1 + \delta)\}$  is covered by the  $2\delta$ -offset of  $C(S, 0; \alpha) = \{0, \pm 1, \pm 2\}$ , so  $\text{EMD}(I(S; \alpha), I(Q; \alpha)) = 2\delta$ .  $\blacklozenge$

For rigid motion instead of general isometry, Definition 11.4.1 of a boundary tolerant metric  $\text{BT}$  is updated to  $\text{BT}^o$  by considering only orientation-preserving isometries from  $\text{SO}(\mathbb{R}^n; p, q)$ , which also makes the continuity below valid for oriented isosets  $I^o(S; \alpha)$  under EMD using  $\text{BT}^o$  instead of  $\text{BT}$  in Definition 11.4.3.

**Theorem 11.4.5** (continuity of isosets under perturbations, [3, Theorem 4.9]). Let periodic point sets  $S, Q \subset \mathbb{R}^n$  have a bottleneck distance  $\text{BD}(S, Q) < r(Q)$ , where  $r(Q)$  is the packing radius in Definition 9.2.3(a). Then the isosets  $I(S; \alpha), I(Q; \alpha)$  are close in the Earth Mover's Distance:  $\text{EMD}(I(S; \alpha), I(Q; \alpha)) \leq 2\text{BD}(S, Q)$  for  $\alpha \geq 0$ .

■

Corollary 11.4.6(a) justifies that the EMD satisfies all metric axioms for periodic point sets that have a stable radius  $\alpha$ . Corollary 11.4.6(b) avoids this dependence on  $\alpha$  and scales any periodic point set  $S$  to the minimum stable radius  $\alpha(S) = 1$ .

**Corollary 11.4.6** (metric on periodic point sets, [3, Corollary 4.10]). **(a)** For  $\alpha > 0$ ,  $\text{EMD}(I(S; \alpha), I(Q; \alpha))$  is a metric on the space of isometry classes of all periodic point sets with a stable radius  $\alpha$  in  $\mathbb{R}^n$ .

**(b)** For a periodic point set  $S \subset \mathbb{R}^n$ , let  $S/r(S) \subset \mathbb{R}^n$  denote  $S$  after uniformly dividing all vectors by the packing radius  $r(S)$ . Then

$$|r(S) - r(Q)| + \text{EMD}(I(S/r(S); 1), I(Q/r(Q); 1))$$

is a metric on all periodic point sets.  $\blacksquare$

## 11.5 Algorithms to compute isosets and their approximate metrics

This section describes time complexities for computing the complete invariant isoset (Theorem 11.5.2), comparing isosets (Corollary 11.5.3), approximating the boundary tolerant metric  $\text{BT}$  and Earth Mover's Distance on isosets (Corollary 11.5.9).

All time estimates will use the geometric complexity  $GC(S)$  defined below.

**Definition 11.5.1** (geometric complexity  $GC$ ). Let a periodic point set  $S \subset \mathbb{R}^n$  have an asymmetric unit of  $m$  points in a cell  $U$  of volume  $\text{vol}[U]$ . Let  $L$  be the symmetry characteristic for  $\alpha_0 = 2R(S)$  in Lemma 11.3.3(c), where  $R(S)$  is the covering radius. The *geometric complexity* is  $GC(S) = \frac{(10(L + m + 2)R(S)/n)^n}{2\text{vol}[U]}$ .  $\blacktriangle$

Let  $V_n = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2} + 1)}$  be the volume of the unit ball in  $\mathbb{R}^n$ , where the Gamma function  $\Gamma$  is defined as  $\Gamma(k) = (k - 1)!$  and  $\Gamma(\frac{k}{2} + 1) = \sqrt{\pi}(k - \frac{1}{2})(k - \frac{3}{2}) \cdots \frac{1}{2}$  for any integer  $k \geq 1$ . Set  $\nu(U, \alpha, n) = \frac{(\alpha + d)^n V_n}{\text{vol}[U]}$ , where  $d = \sup_{p, q \in U} |p - q|$  is a longest diagonal of  $U$ .

The main input size of a periodic set is the number  $m$  of motif points because the length of a standard Crystallographic Information File (CIF) is linear in  $m$ .

For a fixed dimension  $n$ , the big  $O$  notation  $O(m^n)$  in all complexities means a function  $t(m)$  such that  $t(m) \leq Cm^n$  for a fixed constant  $C$  independent of  $m$ . We will include all other parameters depending on a periodic point set  $S$ .

**Theorem 11.5.2** (time of an isoset, [3, Theorem 5.3]). For any periodic point set  $S \subset \mathbb{R}^n$  given by a motif  $M$  of  $m$  points in a unit cell  $U$ , the isoset  $I(S; \alpha)$  at a stable radius  $\alpha$  can be found in time  $O(m^2 k^{\lceil n/3 \rceil} \log k)$ , where  $k = \nu m$  for  $\nu \leq GC(S)$ .  $\blacksquare$

**Corollary 11.5.3** (comparing isosets, [3, Corollary 5.4]). There is an algorithm to check if any periodic point sets  $S, Q \subset \mathbb{R}^n$  with motifs of at most  $m$  points are isometric in total time  $O(m^2 k^{\lceil n/3 \rceil} \log k)$ , where  $k = \nu m$  for  $\nu \leq \max\{GC(S), GC(Q)\}$ .  $\blacksquare$

**Definition 11.5.4** (directed distances  $d_R$  and  $d_M$ ). **(a)** For any sets  $C, D \subset \mathbb{R}^n$ , the directed *rotationally invariant* distance  $d_R(C, D) = \min_{f \in O(\mathbb{R}^n)} d_H(C, f(D))$  is minimised over all maps  $f \in O(\mathbb{R}^n; 0)$ , which fix the origin  $0 \in \mathbb{R}^n$ .

**(b)** For any finite sets  $C, D \subset \mathbb{R}^n$ , order all points  $p_1, \dots, p_k \in C$  by increasing distance to the origin  $0$ . The *radius* of  $C$  is  $R(C) = \max_{p \in C} |p|$ . Define the directed *max-min distance* as  $d_M(C, D) = \max_{i=1, \dots, k} \min\{ \alpha - |p_i|, d_R(\{p_1, \dots, p_i\}, D) \}$ .  $\blacktriangle$

If  $C' \subset C$ , then  $d_R(C', D) \leq d_R(C, D)$ . Let  $C, D \subset \bar{B}(0; \alpha)$  be finite sets including the origin  $0$ . If  $C = \{0\}$ , then  $d_R(C, D) = 0$  because  $C \subset D$ , but  $d_R(D, C) = R(D)$  is the radius of  $D$  because  $D \subset \{0\} + \bar{B}(0; \varepsilon)$  only for  $\varepsilon \geq R(D)$ .

Definition 11.5.4, Lemma 11.5.5 and hence all further results work for rigid motion by restricting all maps to the special orthogonal group  $SO(\mathbb{R}^n; 0)$ .

**Lemma 11.5.5** (max-min formula for  $d_R$ , [3, Lemma 5.6]). For any finite sets  $C, D \subset \mathbb{R}^n$ , if  $\alpha \geq R(C)$ , then  $d_R(C \cup \partial \bar{B}(0; \alpha), D \cup \partial \bar{B}(0; \alpha))$  equals  $d_M(C, D)$ .  $\blacksquare$

**Example 11.5.6** (max-min formula). Consider the subcluster  $C \subset C(\Lambda_4, 0; 2)$  of the points  $p_1 = (1, 0)$ ,  $p_2 = (1, 1)$ ,  $p_3 = (1, -1)$ ,  $p_4 = (2, 0)$  from the square lattice  $\Lambda_4$  in

Fig. 11.7. Let  $\alpha = 2$  and  $D = C(\Lambda_6, 0; 2)$  be the 2-cluster of the hexagonal lattice  $\Lambda_6$ . Then  $d_R(p_1, D) = 0$  because  $p_1$  coincides with  $(1, 0) \in D$ . Then  $d_R(\{p_1, p_2\}, D) = \sqrt{2} - 1$ , because the cloud  $D$  after the clockwise rotation through  $15^\circ$  has the points  $(\cos 15^\circ, -\sin 15^\circ)$  and  $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$  at distances  $\sqrt{(\cos 15^\circ - 1)^2 + \sin^2 15^\circ} \approx 0.26$ ,  $\sqrt{2} - 1 \approx 0.41$  to  $p_1, p_2$ , respectively. Then  $d_R(\{p_1, p_2, p_3\}, D) = \sqrt{2} - 1$  because the same rotated image of  $D$  has  $(\sqrt{\frac{3}{2}}, -\sqrt{\frac{3}{2}})$  at the distance  $\sqrt{3} - \sqrt{2} \approx 0.32$  to  $p_3$ .

For  $i = 1$ ,  $\min\{\alpha - |p_1|, d_R(p_1, D)\} = \min\{2 - 1, 0\} = 0$ . For  $i = 2, 3$ ,

$$\begin{aligned} \min\{\alpha - |p_2|, d_R(\{p_1, p_2\}, D)\} &= \\ \min\{\alpha - |p_3|, d_R(\{p_1, p_2, p_3\}, D)\} &= \\ \min\{2 - \sqrt{2}, \sqrt{2} - 1\} &= \sqrt{2} - 1. \end{aligned}$$

For  $i = 4$ ,  $\min\{\alpha - |p_4|, d_R(C, D)\} = 0$  since  $\alpha = 2 = |p_4|$ .

The maximum value is  $\sqrt{2} - 1$ , so Example 11.4.2 fits Lemma 11.5.5.  $\blacklozenge$

Lemma 11.5.7 extends [8, section 2.3] from  $n = 3$  to any dimension  $n > 1$ .

**Lemma 11.5.7** (approximating  $d_R$ , [3, Lemma 5.8]). Let a cloud  $C \subset \mathbb{R}^n$  consist of  $k = |C|$  points ordered by distances  $|p_1| \leq \dots \leq |p_k|$  from the origin and  $\langle C \rangle$  denote the number of different vectors  $\mathbf{p}/|\mathbf{p}|$  for  $p \in C$ . For each  $j = 1, \dots, k$ , consider the subcloud  $C_j = \{p_1, \dots, p_j\}$ . For any cloud  $D \subset \mathbb{R}^n$  of  $|D|$  points, all distances  $d_j = d_R(C_j, D)$  from Definition 11.5.4 for  $j = 1, \dots, k$  can be approximated by some  $d'_j$  in time  $O(|C|\langle C \rangle^{n-1}|D|)$  so that  $d_j \leq d'_j \leq \omega d_j$ ,  $\omega = 1 + \frac{1}{2}n(n-1)$ .  $\blacksquare$

The proof of Lemma 11.5.7 uses only orientation-preserving isometries from  $\text{SO}(\mathbb{R}^n, 0)$ . Hence the upper bounds from Lemma 11.5.7, Theorem 11.5.8, and Corollary 11.5.9 work for both cases of rigid motion and general isometry in  $\mathbb{R}^n$ .

**Theorem 11.5.8** (approximating BT, [3, Theorem 5.9]). Let periodic point sets  $S, Q \subset \mathbb{R}^n$  have isometry classes  $\sigma, \xi$  represented by clusters  $C, D$  of a radius  $\alpha$ , respectively. In the notations of Lemma 11.5.7,  $\text{BT}(\sigma, \xi)$  from Definition 11.4.1 can be approximated with the factor  $\omega = 1 + \frac{1}{2}n(n-1)$  in time  $O(|C|(\langle C \rangle^{n-1} + \langle D \rangle^{n-1})|D|)$ .  $\blacksquare$

**Corollary 11.5.9** (approximating EMD on isosets, [3, Corollary 5.10]). Let  $S, Q \subset \mathbb{R}^n$  be periodic point sets whose motifs have at most  $m$  points  $p$  and  $\chi$  different vectors  $\mathbf{p}/|\mathbf{p}|$ . For any  $\alpha > 0$ , the metric  $\text{EMD}(I(S; \alpha), I(Q; \alpha))$  can be approximated with the factor  $\omega = 1 + \frac{1}{2}n(n-1)$  in time  $O(\nu^2 m^4 \chi^{n-1})$ , where  $\nu \leq \max\{\text{GC}(S), \text{GC}(Q)\}$ .  $\blacksquare$

Counting directions  $\mathbf{p}/|\mathbf{p}|$  as points ( $\chi \leq m$ ), for dimension  $n = 3$ , the rough bounds for the isoset and its approximate EMD' in Theorem 11.5.2 and Corollary 11.5.9 are  $O(m^3 \log m)$  and  $O(m^6)$ , respectively. Algorithms 1-2 in the appendix describe pseudocodes for Lemma 11.5.7, Theorem 11.5.8, and Corollary 11.5.9.

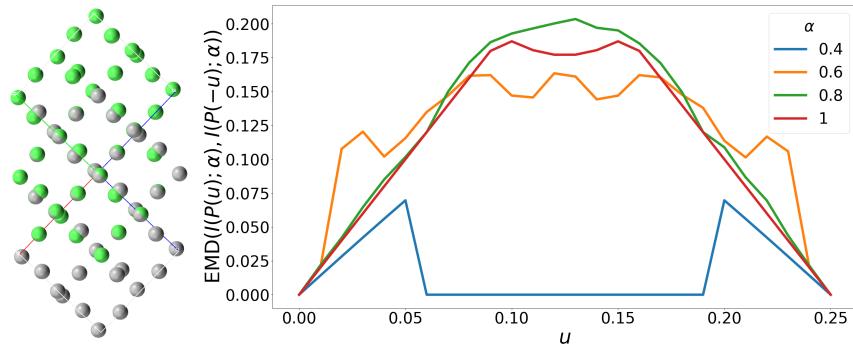
## 11.6 Comparisons of isosets with past invariants and experiments

This section justifies that the isoset can be efficiently used with the faster PDD due to a lower bound in Theorem 11.6.1.

**Theorem 11.6.1** (lower bound for EMD, [3, Theorem 6.5]). Let  $S, Q \subset \mathbb{R}^n$  be periodic sets with a common stable radius  $\alpha$ . Let  $\varepsilon = \text{EMD}(I(S; \alpha), I(Q; \alpha))$  and  $k$  be the maximum number of points of  $S, Q$  in their  $(\alpha - \varepsilon)$ -clusters. If  $\varepsilon$  is less than the half-distance between any points of  $S, Q$ , then  $\text{EMD}(\text{PDD}(S; k), \text{PDD}(Q; k)) \leq \varepsilon$ . ■

Hence, PDD( $S; k$ ) can be used for a fast filtering of distant crystals so that the isoset is computed only for near-duplicates that are hard to distinguish.

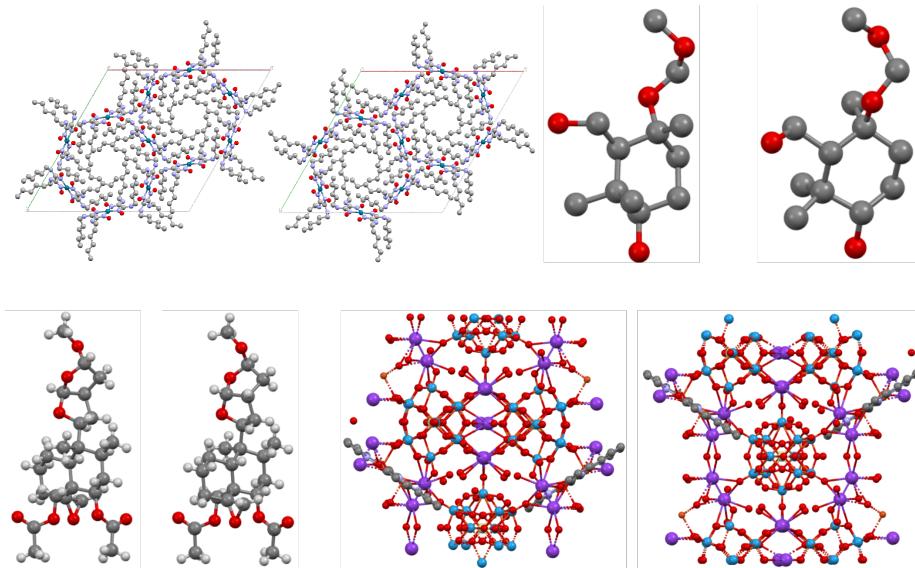
In 1930, future Nobel laureate Linus Pauling noticed the ambiguity of crystal structures obtained by diffraction [11]. Such *homometric* crystals with identical diffraction patterns were only manually distinguished until now because even the generically complete PDDs coincide for the Pauling periodic sets  $P(\pm u)$  for all  $u \in (0, 0.25)$ , see the real overlaid crystals for  $u = 0.03$  in Fig. 11.8 (left).



**Fig. 11.8** **Left:** a comparison of Pauling's homometric crystals  $P(\pm u)$  for  $u = 0.03$  [11], by COMPACK [4] aligning subsets of 48 atoms and outputs RMSD, which fails the triangle inequality. The atoms from different  $P(\pm 0.03)$  are shown in green and grey. **Right:** the pairs of  $P(\pm u)$  have  $\text{EMD}' > 0$  for all  $u \in (0, 0.25)$  and  $\alpha > 0.4$  (running time 50 ms for  $u = 0.03$  and  $\alpha = 0.5$ ).

The strongest past invariant PDD is based on distances and cannot distinguish mirror images. In the CSD, we found four pairs that have identical PDDs but are mirror images shown in Fig. 11.8 (right), distinguished by isosets with  $\alpha \geq 1.5\text{\AA}$  in Fig. 11.10 (left). For WODLOS vs XAWGAE and  $\alpha = 2$ , the total time including isosets and EMD is about 4.3 seconds. All experiments were run on CPU AMD Ryzen 5 5600X, 32GB RAM.

The limitations of the EMD metric on isosets in Definition 11.4.3 are a slower running time than for AMD, PDD and the approximate (not exact) algorithm in Corollary 11.5.9, which are outweighed by the following crucial advantages.



**Fig. 11.9** Four pairs of mirror images in the CSD are indistinguishable by all past invariants but have approximate  $\text{EMD}' > 0$  for all radii  $\alpha > 1.5\text{\AA}$  in Fig. 11.10 (left).

First, all past invariants could not distinguish infinitely many periodic sets (including all mirror images) under rigid motion, e.g. the real crystals in Fig. 11.8. The new continuous EMD fully solved Problem 11.1.1, which remained open since 1965 [9].

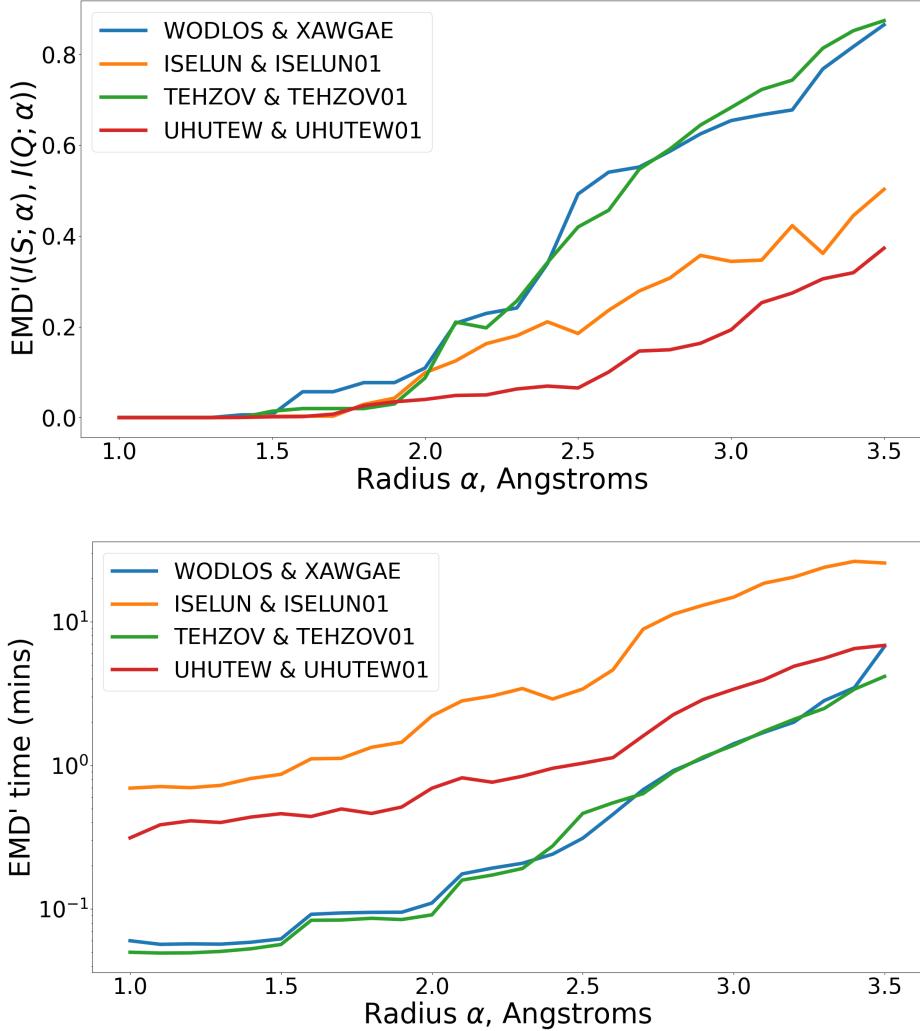
Second, because the proved error factor in the practical dimension  $n = 3$  is close to 4, any near-duplicate crystals that differ by atomic deviations of up to  $\varepsilon$  have an exact distance  $\text{EMD} \leq 2\varepsilon$  by main Theorem 11.4.5 and hence an approximate distance up to about  $8\varepsilon$  by Corollary 11.5.9.

Any crystals that can be matched under rigid motion are recognisable since our approximation of  $\text{EMD} = 0$  is also 0. Any approximate value  $\delta$  of EMD for real crystals  $S, Q$  implies that all atoms of  $S$  should be perturbed by at least  $\delta/8$  on average for a complete match with  $Q$ .

Future work can use the EMD to continuously quantify changes in material properties under perturbations of atoms and extend Problem 11.1.1 to metrics on finite or periodic sets of points under affine and projective transformations.

In conclusion, sections 11.3 and 11.4 prepared the complexity results in section 11.5: algorithms for computing and comparing isosets (Theorems 11.5.2, Corollary 11.5.3), and approximating the new boundary tolerant metric BT (Theorem 11.5.8), and EMD on isosets (Corollary 11.5.9). The proofs expressed polynomial bounds in terms of the motif size  $m = |S|$  of a periodic set  $S$  because the input size of a Crystallographic Information File is linear in  $m$ , e.g. any lattice has  $m = 1$ .

The factors depending on the dimension and geometric complexity  $\text{GC}(S)$  are inevitable due to the curse of dimensionality and the infinite nature of crystals. In practice,



**Fig. 11.10** The isosets distinguish all four pairs of mirror images given by their codes in the CSD. **Top:** approximate EMD' for different radii  $\alpha$ . **Bottom:** running times on a modest desktop.

crystal symmetries reduce a motif to a smaller asymmetric part, which usually has fewer than 20 atoms, even for large molecules in the CSD. The lower bound via faster PDD invariants in Theorem 11.6.1 justifies applying the algorithm of Corollary 11.5.9 only for a final confirmation of near-duplicates. So the isosets finalised the hierarchy of the faster but incomplete invariants.

The main novelty is the boundary-tolerant metric in Definition 11.4.1 that makes the complete invariant isoset Lipschitz continuous (Theorem 11.4.5) without extra parameters that are needed to smooth past descriptors, such as powder diffraction

patterns and atomic environments with fixed cut-off radii. Since the isoset is the only Lipschitz continuous invariant whose completeness under isometry was proved for all periodic point sets in  $\mathbb{R}^n$ , the isoset was used to confirm near-duplicates in the CSD (Table 11.6) and GNoME (Table 11.6).

**Table 11.1** The first pair consists of rigidly different mirror images from Fig. 11.10 (right). All others are geometric near-duplicates from (surprisingly) different families in the CSD, confirmed by tiny values of the EMD metric on isosets. The distance units are in *attometers*: 1 am =  $10^{-8}\text{\AA} = 10^{-18}$  meter. The run times in milliseconds (ms) depend on the cluster size (maximum number of atoms in  $\alpha$ -clusters) according to Theorem 11.5.2 and Corollary 11.5.9.

CSD id1	CSD id2	EMD, am	isosets time, ms	EMD time, ms	cluster size
WODLOS	XAWGAE	85856.22	129.619	1204.58	7
TAFQIA	VAVQIS	952.96	1690.949	321603.86	20
FIJKIU	IPEQUR	728.43	407.579	77455.47	16
JIZMIR01	JIZNAK	496.08	40.454	634.73	5
HIYVUG01	MASPIF	334.62	35.518	543.45	7
KIVXEW10	KIWCEC	125.03	22.456	32.32	5
XAYZOP	ZEMDAZ	89.47	301.217	1697.26	4
KIVXEW10	KIWCEC28	83.07	21.701	32.49	5
AFIBOH	NENCUF	31.67	126.582	1160.58	5
KIVXEW07	KIWCEC09	31.11	22.287	36.95	5
KIVXEW07	KIWCEC11	31.11	22.434	36.75	5
KIVXEW11	KIWCEC26	26.11	21.646	32.33	5
SERKIL	SERKOR	23.78	2444.885	18485.57	6
ADESAG	REWPOB	5.81	54.675	5689.27	15
GEQRAX	IFOQOL	0.05	265.4	2090.42	6
BUKYEN	UYOCES	0.03	398.129	15739.11	11
GOHYOT	VIHCEY	0.01	100.031	940.16	5
JUMCUP	QAHBOT	0.01	179.367	4234.6	5
CALMOV	CALNAI	0.01	128.437	3913.32	4
NABKOT	ZIVSEF	0.01	75.401	796.14	5
LIBGAE	VESJUY	0.01	41.535	403.87	3
AMEVEV	OLERON	0	70.172	558.78	4
SIHFIZ	TEZBUV	0	207.984	1761.39	5
XATCAA	ZAQMEN	0	60.254	394.74	4
PIDREA	XIZNOL	0	94.135	243.46	5

**Table 11.2** After excluding 3248 exact numerical duplicates from [2, Table 1], the next 25 pairs of closest near-duplicates in the GNoME database are confirmed by tiny values of EMD on isosets, see the first pair in Fig. 11.11. The distance units are *attometers*:  $1 \text{ am} = 10^{-8} \text{ Å} = 10^{-18} \text{ meter}$ . The run times are in milliseconds (ms). The cluster size is the maximum number of atoms in  $\alpha$ -clusters.

GNoME id1	GNoME id2	EMD, am	isosets time, ms	EMD time, ms	cluster size
1547d30046	ddc216e80c	1	1.659	434.362	14
b4065a4798	e78d3559e6	1.7	3.034	13.271	6
98ab164895	df1252bc44	2	1.002	419.142	14
0de9d25713	b1733941a7	2.7	1.971	49.816	6
0e79f7c053	6cf951ac6f	3	1.035	429.487	14
07ece241f0	45cacc8d45	3.2	0.618	14.374	6
a58dc74a92	c16bf63220	4.1	2.532	641.086	14
5023e3a4b8	8f7ffb4d4a	4.6	2.776	10.02	6
3198d1a3ea	35f67abe6d	5	1.031	403.398	14
6826b81efb	76ee112799	5	0.985	407.618	14
6826b81efb	e9be17f0ee	5	1.008	404.306	14
2cff5f2fa0	f470a5f6fa	5.3	0.635	169.911	17
2ce912f039	9de239ee0c	5.5	0.632	3.456	2
c9f5a7a51b	fd9f40e0e1	6	1.14	195.261	10
18078e002b	aca2a892a5	6	1.028	421.009	14
18078e002b	b9722429b1	6	1.18	445.453	14
18078e002b	b702e73db3	6	1.035	414.325	14
34b4204eee	adee17535b	6	1.017	396.855	14
506b8b5646	60d266db80	6	1.174	413.254	14
506b8b5646	ec7b789cb3	6	1.174	403.014	14
780741962f	a19688f106	6.5	1.804	870.629	15
780741962f	c6af1fc763	6.5	2.731	921.496	15
780741962f	c64c3e245c	6.5	1.792	829.979	15
b06353561c	b6d2341d32	6.6	2.387	259.359	12
ebb33e044c	ebc9a4db61	6.8	1.232	450.351	14

```

1 # generated using pymatgen
2 data_Ca4Sm3Y3Hg8
3 _symmetry_space_group_name_H-M 'P 1'
4 _cell_length_a 7.53235200
5 _cell_length_b 7.55268100
6 _cell_length_c 7.49193700
7 _cell_angle_alpha 90.0000000
8 _cell_angle_beta 90.00005341
9 _cell_angle_gamma 90.0000000
10 symmetry_Int_Tables_number 1
11 chemical_formula_structural Ca4Sm3Y3Hg8
12 chemical_formula_sum 'Ca4 Sm1 Y3 Hg8'
13 cell_volume 426.21218912
14 cell_formula_units_Z 1
15 loop_
16 _symmetry_equiv_pos_site_id
17 _symmetry_equiv_pos_as_xyz
18 1 'x, y, z'
19 loop_
20 _atom_site_type_symbol
21 _atom_site_label
22 _atom_site_symmetry_multiplicity
23 _atom_site_fract_x
24 _atom_site_fract_y
25 _atom_site_fract_z
26 _atom_site_occupancy
27 Ca Ca0 1 0.749546 0.250000 0.000123 1
28 Ca Ca1 1 0.249534 0.750000 0.499878 1
29 Ca Ca2 1 0.750454 0.250000 0.500121 1
30 Ca Ca3 1 0.250466 0.750000 0.999877 1
31 Sm Sm4 1 0.250759 0.250000 0.999999 1
32 Y Y5 1 0.749300 0.750000 0.000001 1
33 Y Y6 1 0.750715 0.750000 0.499986 1
34 Y Y7 1 0.249294 0.250000 0.500013 1
35 Hg Hg8 1 0.999372 0.500079 0.249529 1
36 Hg Hg9 1 0.999372 0.999023 0.249529 1
37 Hg Hg10 1 0.501025 0.999266 0.252226 1
38 Hg Hg11 1 0.501025 0.500734 0.252226 1
39 Hg Hg12 1 0.500607 0.500991 0.747772 1
40 Hg Hg13 1 0.500607 0.999011 0.747772 1
41 Hg Hg14 1 0.998962 0.999261 0.750475 1
42 Hg Hg15 1 0.998962 0.500739 0.750475 1

```

---

```

1 # generated using pymatgen
2 data_CdTb3SmCd8
3 _symmetry_space_group_name_H-M 'P 1'
4 _cell_length_a 7.53235200
5 _cell_length_b 7.55268100
6 _cell_length_c 7.49193700
7 _cell_angle_alpha 90.0000000
8 _cell_angle_beta 90.00005325
9 _cell_angle_gamma 90.0000000
10 symmetry_Int_Tables_number 1
11 chemical_formula_structural Ca4Tb3Sm1Cd8
12 chemical_formula_sum 'Ca4 Tb3 Sm1 Cd8'
13 cell_volume 426.21218912
14 cell_formula_units_Z 1
15 loop_
16 _symmetry_equiv_pos_site_id
17 _symmetry_equiv_pos_as_xyz
18 1 'x, y, z'
19 loop_
20 _atom_site_type_symbol
21 _atom_site_label
22 _atom_site_symmetry_multiplicity
23 _atom_site_fract_x
24 _atom_site_fract_y
25 _atom_site_fract_z
26 _atom_site_occupancy
27 Ca Ca0 1 0.749546 0.250000 0.000123 1
28 Ca Ca1 1 0.249534 0.750000 0.499878 1
29 Ca Ca2 1 0.750454 0.250000 0.500121 1
30 Ca Ca3 1 0.250466 0.750000 0.999877 1
31 Tb Tb4 1 0.749300 0.750000 0.000001 1
32 Tb Tb5 1 0.750715 0.750000 0.499986 1
33 Tb Tb6 1 0.249294 0.250000 0.500013 1
34 Sm Sm7 1 0.250759 0.250000 0.999999 1
35 Cd Cd8 1 0.999372 0.500079 0.249529 1
36 Cd Cd9 1 0.999372 0.999023 0.249529 1
37 Cd Cd10 1 0.501025 0.999266 0.252226 1
38 Cd Cd11 1 0.501025 0.500734 0.252226 1
39 Cd Cd12 1 0.500607 0.500991 0.747772 1
40 Cd Cd13 1 0.500607 0.999011 0.747772 1
41 Cd Cd14 1 0.998962 0.999261 0.750475 1
42 Cd Cd15 1 0.998962 0.500739 0.750475 1

```

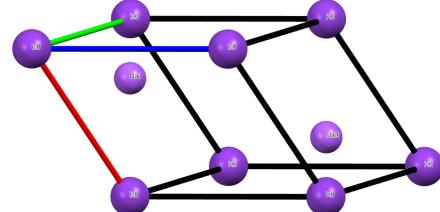
**Fig. 11.11** The GNoME crystals 1547d30046 and ddc216e80c in the first row of Table A.2 are compared as texts by <https://text-compare.com>. All differences are highlighted in blue.

Fig. 11.12 shows the most striking pair of exact duplicates in the GNoME is cdc06a1a2a and 0e2d8f26d6, whose CIFs are identical symbol by symbol in addition to two pairs of atoms at the same positions (Na1=Na2 and Na3=Na4).

```

20 _atom_site_type_symbol
21 _atom_site_label
22 _atom_site_symmetry_multiplicity
23 _atom_site_fract_x
24 _atom_site_fract_y
25 _atom_site_fract_z
26 _atom_site_occupancy
27 K K0 1 0.000000 0.000000 0.000000 1
28 Na Na1 1 0.250000 0.250000 0.250000 1
29 Na Na2 1 0.250000 0.250000 0.250000 1
30 Na Na3 1 0.749999 0.749999 0.749999 1
31 Na Na4 1 0.749999 0.749999 0.749999 1

```



**Fig. 11.12** Different entries cdc06a1a2a and 0e2d8f26d6 in the GNoME database are not only identical symbol by symbol but also contain two pairs of atoms (Na1=Na2 and Na3=Na4) at the same positions. **Left:** a screenshot from the CIF. **Right:** Mercury visualisation can show only one atom in each pair of coinciding atoms, e.g. only Na1 and not Na2 from the CIF.

## References

1. O Anosova and V Kurlin. An isometry classification of periodic point sets. In *Proceedings of Discrete Geometry and Mathematical Morphology*, pages 229–241, 2021.
2. Olga Anosova, Vitaliy Kurlin, and Marjorie Senechal. The importance of definitions in crystallography. *International Union of Crystallography Journal*, 11:453–463, 2024.
3. Olga Anosova, Daniel Widdowson, and Vitaliy Kurlin. Recognition of near-duplicate periodic patterns by continuous metrics with approximation guarantees. *Pattern Recognition*, 171:112108, 2026.
4. J. Chisholm and S. Motherwell. Compack: a program for identifying crystal structure similarity using distances. *J. Applied Crystal.*, 38:228–231, 2005.
5. Boris Nikolaevich Delone, Nikolai Petrovich Dolbilin, Mikhail Ivanovich Shtogrin, and Ravil V Galiulin. A local criterion for regularity of a system of points. In *Dokl. Akad. Nauk SSSR*, volume 227, pages 19–21, 1976.
6. N. Dolbilin, J. Lagarias, and M. Senechal. Multiregular point systems. *Discrete & Computational Geometry*, 20(4):477–498, 1998.
7. Nikolay Petrovich Dolbilin and Aleksandr Nikolaevich Magazinov. Uniqueness theorem for locally antipodal delaunay sets. *Proceedings of the Steklov Institute of Mathematics*, 294(1):215–221, 2016.
8. Michael Goodrich, Joseph S Mitchell, and Mark Orletsky. Approximate geometric pattern matching under rigid motions. *Trans. Pattern Analysis and Machine Intel.*, 21(4):371–379, 1999.
9. Stephen Lawton and Robert Jacobson. The reduced cell and its crystallographic applications. Technical report, Ames Lab, Iowa State University, 1965.
10. Jonathan McManus and Vitaliy Kurlin. Computing the bridge length: the key ingredient in a continuous isometry classification of periodic point sets. *Acta Cryst A*, 81, 2025.
11. Linus Pauling and Maple D Shappell. The crystal structure of bixbyite and the c-modification of the sesquioxides. *Zeitschrift für Kristallographie-Crystalline Materials*, 75(1):128–142, 1930.
12. Philip Smith and Vitaliy Kurlin. Generic families of finite metric spaces with identical or trivial 1-dimensional persistence. *Journal of Applied and Computational Topology*, 8:839–855, 2024.
13. Daniel Widdowson and Vitaliy Kurlin. Resolving the data ambiguity for periodic crystals. *Advances in Neural Information Processing Systems*, 35:24625–24638, 2022.



## Chapter 12

# A history of developing the most significant results and open problems

**Abstract** This chapter concludes with the most important theorems from previous chapters and a brief history of Geometric Data Science in a chronological order. The developed hierarchies of Lipschitz continuous invariants from the ultra-fast to complete ones (under rigid motion in Euclidean space  $\mathbb{R}^n$ ) allowed us to distinguish all non-duplicate objects in major databases of molecules and crystals. Experimental validation justified new concepts of geometric structures, such as a crystal structure defined as an equivalence class of periodic sets of only atomic centres without chemical elements, under rigid motion in  $\mathbb{R}^3$ . The resulting Crystal Isometry Principle uniquely identifies any real periodic material in the continuous moduli space of all periodic point sets. The book finishes by highlighting open problems and future work.

### 12.1 The most important results of Geometric Data Science by 2025

This section briefly summarises the most significant results from Chapters 1-11.

**Chapter 1** motivated and then formalised the three practical questions

*(1) Same or different? (2) If different, by how much? (3) Where do all real objects live?*

in Geo-Mapping Problem 1.4.5 for all types of geometric data, which can be studied under practically important equivalence relations, such as rigid motion in Euclidean space  $\mathbb{R}^n$ . If all conditions of Problem 1.4.5 are satisfied, the resulting moduli spaces of data objects can be explored by *geocodes*, similar to geographic coordinates on Earth.

**Chapter 2** described the Backbone Rigid Invariant (BRI) in Definition 2.3.3. Theorems 2.3.4, 2.3.7, 2.3.9 proved that the BRI is a geocode (a geographic-style invariant) on the moduli space of non-degenerate protein backbones in  $\mathbb{R}^3$ . Within a few hours on a modest desktop computer, this invariant detected thousands of exact duplicate backbones and many more near-duplicate chains in the Protein Data Bank (PDB).

**Chapter 3** presented in Theorem 3.4.4 complete polynomial-time invariants for finite clouds of unordered points under rigid motion in  $\mathbb{R}^n$ . These invariants were outper-

formed by faster invariants in later chapters, but influenced the latest developments towards a full solution of Geo-Mapping Problem 1.4.5 for unordered points in  $\mathbb{R}^n$ .

**Chapter 4** extended earlier invariants of finite clouds of unordered points (incomplete under isometry in  $\mathbb{R}^n$ ) to the stronger and continuous Pointwise Distance Distribution (PDD) in Definition 4.5.1. In addition to being generically complete under isometry in any  $\mathbb{R}^n$ , the PDD was proved to be fully complete in Theorem 4.6.1 for any 4 unordered points under isometry in  $\mathbb{R}^n$ . It seems the first faster-than-brute-force extension of the side-side-side theorem, known in Euclidean geometry for 2000+ years.

**Chapter 5** generalised the PDD to a stronger Simplexwise Distance Distribution (SDD) in Definition 5.1.3 for an arbitrary metric space. Theorems 5.3.3 and 5.3.4 proved the Lipschitz continuity and polynomial-time computability of the SDD, which is simple enough to distinguish all (infinitely many) known pairs of non-isometric clouds in  $\mathbb{R}^3$  with the same PDD through manual computations in Examples 5.2.1, 5.2.2, and 5.2.3.

**Chapter 6** improved the SDD to the faster Simplexwide Centred Distribution (SCD) in Definition 6.2.1 for clouds of unordered points in  $\mathbb{R}^n$ . Theorems 6.2.3, 6.4.4, and 6.4.5 for guaranteed completeness, Lipschitz continuity, and polynomial-time computability of the SCD for any  $n$ -dimensional cloud of unordered points under rigid motion in  $\mathbb{R}^n$ . The key ingredient was the concept of the strength of a simplex in Definition 6.3.1.

**Chapter 7** initiated a continuous approach to point sets that are periodic in one direction in a high-dimensional space  $\mathbb{R} \times \mathbb{R}^{n-1}$ . Theorem 7.2.8 improved the distance matrix to a Lipschitz continuous invariant that is complete under rigid motion, distinguishing all mirror images in  $\mathbb{R}^n$ . Theorem 7.4.3 developed a complete invariant with Lipschitz continuous and polynomial-time metrics for 1-periodic point sets in  $\mathbb{R} \times \mathbb{R}^{n-1}$ .

**Chapter 8** fully solved Geo-Mapping Problem 1.4.5 (restated as Problem 8.2.2) for all 2-dimensional lattices under four equivalences in  $\mathbb{R}^2$  in Theorem 8.4.2 and Corollary 8.4.5. The case of rigid motion remained discontinuous since the time of Lagrange [15], who classified 2D lattices in terms of quadratic forms, not distinguishing mirror images.

**Chapter 9** studied density functions, which are Lipschitz continuous isometry invariants of arbitrary periodic point sets in  $\mathbb{R}^n$ . Theorem 9.2.1 proved the generic completeness of density functions under isometry in  $\mathbb{R}^3$ . Theorems 9.5.3, 9.5.6, and 9.5.9 analytically described density functions for periodic sequences of intervals in  $\mathbb{R}$ .

**Chapter 10** extended the Pointwise Distance Distribution (PDD) to infinite point sets that are periodic in  $l$  directions. Theorems 10.3.4 and 10.4.5 proved Lipschitz continuity and generic completeness of the PDD for periodic point sets under isometry in any  $\mathbb{R}^n$ .

**Chapter 11** developed the *isoset* invariant for any periodic point sets in  $\mathbb{R}^n$ . Theorem 11.3.7 finalised the full completeness of the isoset under isometry and rigid motion in  $\mathbb{R}^n$ . Definition 11.4.3 introduced the Earth Mover's Distance (EMD) on isosets. Theorem 11.4.5 and Corollary 11.5.9 proved Lipschitz continuity and approximate polynomial-time algorithms for the EMD on isosets in  $\mathbb{R}^n$ , for a fixed dimension  $n$ .

In fact, all results in this book were developed in almost an opposite chronological order, starting with isosets from chapter 11 in early 2020 and finishing with protein invariants from chapter 2 in late 2024, which will be explained in sections 12.4-12.5. The next section defines new concepts of geometric structures motivated by applications.

## 12.2 New definitions of geometric structures and verified principles

Mathematics allows other sciences to progress by developing new concepts that formalise practical questions. It was essential to state the challenges of ambiguity and discontinuity of cell-based representations of periodic crystals in terms of complete invariants with Lipschitz continuous metrics. The resulting invariants (from the simplest and ultra-fast to slower but complete ones) opened ‘black boxes’ of databases, which keep their data in ambiguous photograph-style forms. For example, crystals are usually stored as Crystallographic Information Files (CIFs), while molecular geometries are represented by xyz files listing atomic positions in arbitrary coordinate systems.

The approach through continuous invariants revealed thousands of exact geometric duplicates and many more near-duplicates in the PDB, CSD, ICSD, and other databases. Some duplicates should have been found by dataset curators by direct comparisons of digital representations. For instance, Olga counted in [4, Table 1] thousands of symbol-by-symbol duplicate CIFs in the GNoME dataset [1]. Experimental validation on available data justified the new concept of a periodic (crystal) structure below.

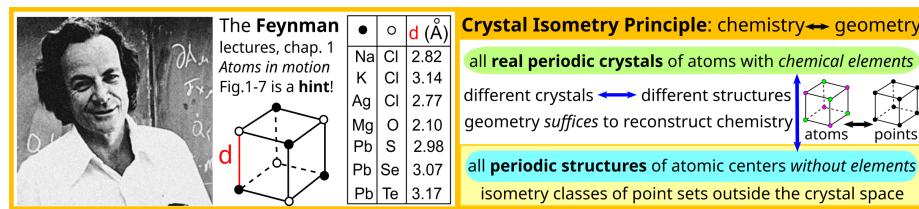
**Definition 12.2.1** ([4, Definition 6]). The *crystal structure* of a real periodic crystal  $S \subset \mathbb{R}^3$  is an equivalence class of all periodic sets of atoms that can be *exactly* matched with  $S$  atom by atom under rigid motion in  $\mathbb{R}^3$ . In general, a *periodic structure* is an equivalence class of periodic sets of unlabelled points in  $\mathbb{R}^n$  under rigid motion. ▲

The words *crystal structure* imply that an underlying object is a real (not simulated) crystal of atoms labelled by chemical elements and charges. In particular, distances between atoms in  $\mathbb{R}^3$  cannot be arbitrary. The more general concept of a periodic structures applies to any periodic sets of points in  $\mathbb{R}^n$  without physical restrictions.

In both cases, if points are shifted by the same vector within a fixed unit cell, all fractional coordinates change, but the new periodic set can be exactly matched with the original one by translation, so the periodic (crystal) structure remains the same. A cell-based representation can more dramatically change under a rotation or a cell transformation without affecting the underlying rigid structure. Definition 12.2.1 emphasises the importance of *exact* matching. Indeed, ignoring noise up to any tiny threshold  $\varepsilon > 0$  leads to the *sorites paradox* [28] and a trivial classification of all objects within a continuous space due to the transitivity axiom in Definition 1.2.1.

Fig. 12.1 (left) shows Feynman’s table distinguishing 7 cubic crystals by their single geometric invariant  $d$  that is the smallest interatomic distance. The much stronger invariants AMD, PDD, and isosets distinguished all non-duplicate periodic crystals in major materials databases. Some of the found duplicates had different chemical compositions. In the first striking example, the pair of the CSD entries HIFCAB and JEPLIA has all numbers in their CIFs identical, but one atom (Mn) is replaced with a different one (Cd), see Fig. 12.2. Since these elements have very different atomic masses 25 and 48, these coincidences of all coordinates seem physically impossible.

Indeed, any atomic replacement should change inter-atomic interactions and hence distances to neighbours, which is immediately detected by the PDD invariant. Our colleagues at the Cambridge Crystallographic Data Centre, who curate the CSD, checked



**Fig. 12.1** **Left:** Feynman's first lecture "Atoms in motion" [24] showed that 7 cubic crystals differ by the smallest interatomic distance  $d$ , which was the Eureka moment for Vitaliy in May 2021 to realise that all real periodic crystals can be distinguished as periodic sets of atomic centres without chemical elements. **Right:** our experiments on the world's largest databases of real materials [52] confirmed the physical intuition, now stated as the *Crystal Isometry Principle*: any real periodic material (under fixed ambient conditions) is uniquely determined by a precise enough atomic geometry within a common continuous space of all periodic structures, independent of their chemistry and symmetries.

that the raw diffraction data (structure factors) were also identical in this case. Several more pairs of duplicates are under investigation by five journals for data integrity.

The much more important consequence of the ability to distinguish all periodic crystals by geometry is the *Crystal Isometry Principle* (CRISP) in Fig. 12.1 (right).

A mathematical version of CRISP says that all real periodic crystals (independent of their chemistry and symmetries) live in a common moduli space  $\text{CRIS}(\mathbb{R}^3)$ , which is called the *Crystal Rigid Space*, see its 2D projections in Fig. 10.11 and [51]. Since the first invariants AMD, PDD were developed for isometry, not for rigid motion, the initial name was the *Crystal Isometry Space* [50], which is now denoted by  $\text{CIMS}(\mathbb{R}^3)$ .

Of course, not any periodic set of points can be realised by a periodic crystal. For instance, distances between neighbouring atoms are usually in small ranges, especially for fixed elements. In the geographic analogy, not every location on Earth is suitable for humans to live. However, the knowledge of a full geographic map certainly helped to find all hospitable places and not to waste time on exploring many hostile regions.

Chemistry has substantially benefited from the periodic table, as a map of all known chemical elements, though it was initially half-empty [37]. Indeed, *Mendeleev's geocode* consisting of the period and group number provided a complete invariant and guided the search for new elements. More than 150 years after Mendeleev's breakthrough in 1869, coordination chemistry [7] can progress from studying isolated shapes of local atomic environments to continuous maps parametrised by complete invariants [54].

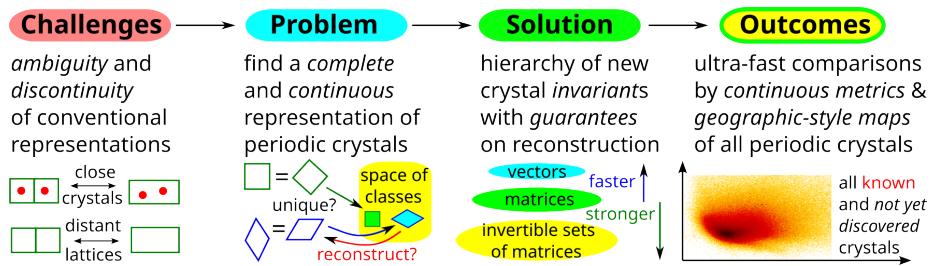
### 12.3 Geometric approaches in crystallography until 2019

This section describes a personal perspective on past developments that led by 2020 to new problems in periodic geometry and a later expansion to Geometric Data Science.

41	_symmetry_cell_setting	Monoclinic	47	_symmetry_cell_setting	Monoclinic
42	_symmetry_space_group_name_H-M	C2/c	48	_symmetry_space_group_name_H-M	C2/c
43	loop_		49		
44	_symmetry_equiv_pos_as_xyz		50	loop_	
45			51	_symmetry_equiv_pos_as_xyz	
46	'x, y, z'		52	'x, y, z'	
47	'-x, y, -z+1/2'		53	'-x, y, -z+1/2'	
48	'x+1/2, y+1/2, z'		54	'x+1/2, y+1/2, z'	
49	'-x+1/2, y+1/2, -z+1/2'		55	'-x+1/2, y+1/2, -z+1/2'	
50	'-x, -y, -z'		56	'-x, -y, -z'	
51	'x, -y, z-1/2'		57	'x, -y, z-1/2'	
52	'-x+1/2, -y+1/2, -z'		58	'-x+1/2, -y+1/2, -z'	
53	'x+1/2, -y+1/2, z-1/2'		59	'x+1/2, -y+1/2, z-1/2'	
54			60		
55			61	_cell_length_a	16.984(3)
56	_cell_length_a	16.984(3)	62	_cell_length_b	11.686(2)
57			63	_cell_length_c	10.0952(17)
58	_cell_length_b	11.686(2)	64	_cell_angle_alpha	90.00
59			65	_cell_angle_beta	109.455(11)
60	_cell_length_c	10.0952(17)	66	_cell_angle_gamma	90.00
61			67	_cell_volume	1889.2(6)
62	_cell_angle_alpha	90.00	68	_cell_formula_units_Z	4
63			69	cell_measurement_temperature	293(2)
64	_cell_angle_beta	109.455(11)	70	cell_measurement_reflns_used	2669
65			71	cell_measurement_theta_min	2.16
66	_cell_angle_gamma	90.00	72	cell_measurement_theta_max	26.43
67			73	cell_measurement_temperature	293(2)
68	_cell_volume	1889.2(6)			
69					
70	_cell_formula_units_Z	4			
71					
72	cell_measurement_reflns_used	2669			
73					
74	cell_measurement_theta_min	2.16			
75					
76	cell_measurement_theta_max	26.43			
77					
78	cell_measurement_temperature	293(2)			
171	atom_site_occupancy				
172	atom_site_symmetry_multiplicity				
173	atom_site_calc_flag				
174	atom_site_refinement_flags				
175	atom_site_disorder_assembly				
176	atom_site_disorder_group				
177	H1 Mn 0.0000 0.03972(3) 0.2500 0.02174(12) Uani d 1 2 d s .				
178	O1 Mn 0.08938(9) 0.04933(9) 0.15612(13) 0.04933(9) Uani d 1 1 .				
179	O2 Mn 0.10776(9) 0.04933(9) -0.05187(12) 0.02923(12) Uani d 1 1 .				
204	O 0 0.11381(9) 0.20935(11) 0.11112(14) 0.0433(4) Uani d 1 1 .				
205	N1 N 0.0000 0.15483(16) 0.2500 0.0256(4) Uani d 5 1 2 .				
206	N2 N 0.0000 0.76204(16) 0.2500 0.0286(5) Uani d 5 1 2 .				
207	C1 C 0.11997(10) 0.03634(13) 0.07794(16) 0.02277(3) Uani d 1 1 .				
208	C2 C 0.11997(10) 0.11387(15) 0.17277(17) 0.0294(4) Uani d 1 1 .				
209	C3 C 0.11997(10) 0.1219(2) 0.3126(2) 0.0461(5) Uani d 1 1 .				
210	C4 C 0.21673(13) 0.1219(2) 0.3126(2) 0.0461(5) Uani d 1 1 .				
211	H3 H 0.2068 0.0734 0.3784 0.051 Uiso calc R 1 1 .				
212	C4 C 0.27016(14) 0.2194(2) 0.3408(2) 0.0573(6) Uani d 1 1 .				
213	C5 C 0.26691(14) 0.2635(2) 0.2974(2) 0.0532(6) Uani d 1 1 .				
214	C6 C 0.26691(14) 0.3228(2) 0.3404(2) 0.0532(6) Uani d 1 1 .				
205	C6 C 0.02362(12) 0.21475(15) 0.12091(18) 0.03204(4) Uani d 1 1 .				
206	H6 H -0.0397 0.1748 0.0454 0.038 Uiso calc R 1 1 .				
207	C7 C 0.02529(12) 0.33268(15) 0.12561(19) 0.0347(4) Uani d 1 1 .				
208	H7 H -0.0433 0.3794 0.0396 0.042 Uiso calc R 1 1 .				
209	C8 C 0.0000 0.39518(19) 0.2500 0.0287(5) Uani d 5 1 2 .				
210	C9 C 0.0000 0.5216(2) 0.2500 0.0312(6) Uani d 5 1 2 .				
291	C10 C 0.10537(13) 0.58383(16) 0.1537(2) 0.0408(5) Uani d 1 1 .				
292	H10 H -0.1059 0.5462 0.0867 0.049 Uiso calc R 1 1 .				
293	C11 C -0.06601(16) 0.70177(16) 0.1569(2) 0.0384(5) Uani d 1 1 .				
294	H11 H -0.1026 0.7415 0.0996 0.046 Uiso calc R 1 1 .				
295					
171	atom_site_occupancy				
172	atom_site_symmetry_multiplicity				
173	atom_site_calc_flag				
174	atom_site_refinement_flags				
175	atom_site_disorder_assembly				
176	atom_site_disorder_group				
177	H1 Mn 0.0000 0.03972(3) 0.2500 0.02174(12) Uani 1 2 d s .				
178	O1 Mn 0.08938(9) 0.04933(9) 0.15615(12) 0.02773(3) Uani 1 1 d .				
179	O2 Mn 0.10778(9) 0.04933(9) -0.05187(12) 0.02923(3) Uani 1 1 d .				
204	O 0 0.11381(9) 0.20935(11) 0.11112(14) 0.0433(4) Uani 1 1 d .				
205	N1 N 0.0000 0.15483(16) 0.2500 0.0256(4) Uani 1 2 d s .				
206	N2 N 0.0000 0.76204(16) 0.2500 0.0286(5) Uani 1 2 d s .				
207	C1 C 0.11997(10) 0.03634(13) 0.07794(16) 0.02277(3) Uani 1 1 d .				
208	C2 C 0.11997(10) 0.11387(15) 0.17277(17) 0.0294(4) Uani 1 1 d .				
209	C3 C 0.11997(10) 0.1219(2) 0.3126(2) 0.0461(5) Uani 1 1 d .				
210	C4 C 0.21673(13) 0.1219(2) 0.3126(2) 0.0461(5) Uani 1 1 d .				
211	H3 H 0.2068 0.0734 0.3784 0.051 Uiso calc R 1 1 d .				
212	C4 C 0.27016(14) 0.2194(2) 0.3408(2) 0.0573(6) Uani 1 1 d .				
213	C5 C 0.26691(14) 0.2635(2) 0.2974(2) 0.0532(6) Uani 1 1 d .				
214	C6 C 0.26691(14) 0.3228(2) 0.3404(2) 0.0532(6) Uani 1 1 d .				
205	C6 C 0.02362(12) 0.21475(15) 0.12091(18) 0.03204(4) Uani 1 1 d .				
206	H6 H -0.0397 0.1748 0.0454 0.038 Uiso calc R 1 1 d .				
207	C7 C 0.02529(12) 0.33268(15) 0.12561(19) 0.0347(4) Uani 1 1 d .				
208	H7 H -0.0433 0.3794 0.0396 0.042 Uiso calc R 1 1 d .				
209	C8 C 0.0000 0.39518(19) 0.2500 0.0287(5) Uani 1 2 d s .				
210	C9 C 0.0000 0.5216(2) 0.2500 0.0312(6) Uani 1 2 d s .				
291	C10 C 0.10537(13) 0.58383(16) 0.1537(2) 0.0408(5) Uani 1 1 d .				
292	H10 H -0.1059 0.5462 0.0867 0.049 Uiso calc R 1 1 d .				
293	C11 C -0.06601(16) 0.70177(16) 0.1569(2) 0.0384(5) Uani 1 1 d .				
294	H11 H -0.1026 0.7415 0.0996 0.046 Uiso calc R 1 1 d .				
295					

**Fig. 12.2** The entries HIFCAB and JEPLIA in the CSD were reported to the CCDC in September 2021, but their public pages still have no references to each other and no acknowledgement of the underlying research [53]. The textual comparisons of the cell parameters (top images) and fractional coordinates (bottom images) are based on the CIFs that were downloaded from the CSD on 23 December 2025.

In early 2017, Vitaliy Kurlin joined the Materials Innovation Factory whose director Prof Andy Cooper FRS posed the practical challenge in (almost exactly) the following



**Fig. 12.3** This book exemplified how mathematics can advance physical sciences through rigorous definitions and theorem-based principles verified on all available experimental data, as in Fig. 12.1 (right).

words: “continuously quantify the similarity between crystal structures”. That is how we started with the periodic case of crystals, which turned out to be the hardest so far.

Crystallography emerged from empirical studies of natural minerals. The geometry of lattices goes back to Lagrange in 1771 [15], whose theory of quadratic forms can be interpreted as a reduction of lattice bases. Reduced bases of lattices in  $\mathbb{R}^3$  were explicitly described by Selling in 1874 [46, p.171-173] via 6 parameters, which are equivalent to scalar products  $\mathbf{v}_i \cdot \mathbf{v}_j$  for  $i, j \in \{0, 1, 2, 3\}$ , where  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$  from a lattice basis, and  $\mathbf{v}_0 = -(\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3)$ . This quadruple of vectors was called *Selling’s star* (or a hedgehog) by Delone [9] and later a *superbase* by Conway and Sloane [14], and finally led to a complete isometry invariant of 3D lattices [32] whose continuity is being finalised.

In parallel, discrete classifications by symmetry groups were completed for 3D lattices by Bravais in 1850 [10] and for 3D periodic crystals in 1891 by Fedorov [23] and Schoenflies [45]. In 1908, Voronoi [49] classified all 5 types of local neighbourhoods of points in 3D lattices, which are now called *Voronoi domains*. Though the combinatorial structure 4 non-generic Voronoi domains discontinuously changes under almost any perturbation of a lattice basis, their geometric characteristics, such as volume, change continuously. At this point in history, crystallography had a chance to move towards a continuous approach, but Voronoi unfortunately died at the age of 40 in 1908.

After that, X-ray applications brought to crystallography many more people who chose simpler unit cells (parallelepipeds), where atoms have easy fractional coordinates, instead of more continuous Voronoi domains. In 1928, Niggli proposed a reduced cell [39], which gradually became a canonical choice despite its discontinuity might have been clear to mathematicians already in the 19th century. In the 1920s, a small number of known crystals was easy to recognise and classify by manual measurements.

As a result, crystallography continued in the discrete direction by adding more sophisticated features such as group-subgroup relations, see more details in [6, chapter 1.1]. In 1934, another hope for continuous crystallography was born due to B.N.Delone [9], who studied the 6 geometric parameters  $p_{ij} = -\mathbf{v}_i \cdot \mathbf{v}_j$  based on Selling’s star. However, the cell-based approach was hard to abandon, so Voronoi domains were thoroughly described by using canonical parameters of unit cells [16, p.154-155].

In 1959, a distance metric between lattices was defined via a minimisation over infinitely many affine transformations in [13, section V.3.2]. Then the triangle axiom of a metric holds due to compositions of affine maps within an infinite group. If the minimisation is reduced to finitely many transformations of lattice bases, the triangle axiom and continuity were not checked and are likely to fail in boundary cases.

Feynman's lectures on physics in 1961-1964 [24] started with the most important chapter "Atoms in motion", which highlighted the key challenge right in the title.

It may be a coincidence, but experimentalists finally reported discontinuous cells in 1965 [35, p.80]. Ironically, the Cambridge Structural Database (CSD) started in the same year by representing crystals in unit cells, without realising long-term consequences of discontinuity, which had been known both theoretically and experimentally.

In 1976, Delone and his school revived a geometric approach [17] by initiating a local theory of solids via atomic environments and by modelling amorphous materials through infinite atomic sets that are now called *Delone sets*. The major result of the local theory in the periodic case was a characterisation of any periodic point set among all Delone sets in terms of sufficiently large  $\alpha$ -clusters in 1998 by Dolbilin, Lagarias, and Senechal [18], which we translated into the isoset invariant in early 2020 [2].

By 2016, several books [21, 56] thoroughly summarised an algebraic approach to lattices via group actions, but the continuous metric problem remained unaddressed.

Remembering our initial discussions in 2016-2017, chemists said that there was no noise in simulated data, which actually meant "negligible" relative to real experimental noise. As soon as the first CIF was opened in a text editor, noise in simulated crystals was clear at first glance, because Vitaliy saw nearly round numbers in outputs of many iterative algorithms when he optimised superpixel meshes for 2D images [26, 25, 33].

Mathematicians often start solving hard problems with partial cases, so we thought that periodic lattices were studied for a long time and should be simple enough.

To our surprise, we were stuck on the simplest case of 2D lattices until 2019, because the discontinuity problem was "swept under the carpet" of symmetries for centuries.

In 1992, the important paper by Conway and Sloane claimed that Delone parameters  $p_{ij}$  "vary continuously with the lattice" [14, item 1 on page 55] without any proof, which we are still finalising. About 2.5 years were spent on correcting significant misunderstandings in papers of well-known mathematicians, see details in [32].

Numerous discussions with chemists helped create the inter-disciplinary seminar MIF++ (<https://kurlin.org/MIFplusplus.php>), whose original name was MIF+CS (Materials Innovation Factory plus Computer Science), but we quickly included other subjects areas, such as Mathematics and Engineering. Though the seminar started in person, all activities moved online in March 2020, so we continue the MIF++ in a hybrid form.

By the summer of 2019, the discontinuity problem was finally "crystallised" and PhD student Marco Mosca implemented the first approximate algorithm for a continuous metric between any 3D lattices based on Voronoi domains. In August 2019, Marco and

Vitaliy came to the European Crystallographic Meeting (ECM), which was conveniently preceded by the satellite on Mathematical Crystallography, also in Vienna.

By this time we had known that the discontinuity is much more common for general periodic point sets than for lattices because any extended cell becomes primitive under almost any noise. Without knowing yet the history above, we hoped that mathematical crystallographers had already worked on this problem and were keen to learn more.

At the satellite of ECM 2019, right after Marco Mosca started to talk about discontinuity of lattice bases, chemist Martin U. Schmidt jumped to the board and drew an extended cell (parallelogram) with a single perturbed atom. We had known about the discontinuity of the underlying lattice for periodic point sets, so our first work [38] focused only on lattices to avoid this more general case. However, we understood that some chemists are aware that cell-based representations of crystals are discontinuous, so our images in Fig. 10.1 (left) are simplified versions of Martin’s drawing.

As usual at conferences, the most valuable discussions happen during breaks, so we asked several mathematical crystallographers how they compare nearly identical crystals that substantially differ by unit cells. It was an eye-opening experience when we were passionately told that crystals of different symmetries are *different and hence are not compared*. This moment showed that the research field is completely open, because some chemists saw the challenge [8] but had no tools to rigorously state the problem [43], while most mathematical crystallographers focused on discrete symmetries and missed the “gold mine” of unexplored continuous invariants of periodic crystals.

The discussions at the ECM satellite in August 2019 motivated us to rigorously state the problem of continuously comparing crystals independent of their symmetries.

## 12.4 Towards the Crystal Isometry Principle in 2019-2021

Sections 12.4-12.5 trace the milestones of Geometric Data Science in a chronological order from 2019, which substantially differs from the order of chapters in this book. This section starts with the historical origins of the Crystal Isometry Principle.

Since the satellite of ECM 2019 on mathematical crystallography invited submissions for post-proceedings, in September 2019 we finished the paper [38], which stated the equivalence problem for lattices in terms of isometry invariants and the distance problem, including the metric axioms and the new continuity condition.

After the continuity problem was understood, we extended the widely used continuous invariant, namely the scalar density of points in a crystal, to functions counting fractional volumes of  $k$ -fold intersections of balls of variable radii centred at given atoms. These invariants were first called *packing functions* and later renamed as *density functions*, see Definition 9.2.1. Philip Smith gave a talk “Packing functions are new isometry invariants of periodic crystals” at our group seminar on 10 January 2020.

Shortly after that, Philip noticed large differences between density functions of supposedly similar versions of the same T2 polymorph whose structure was determined

several times under different temperatures. It turned that some authors of [41] confused a near-duplicate version of T2- $\beta$  with a rather different polymorph T2- $\delta$ . As a result, the latter structure was accidentally missed in the original deposit of the DEBXIT family in the CSD and later got another refcode SEMDIA. This example showed that identifying crystal structures even by their creators was still an art, not yet a science.

By this time, we studied the major past work on geometric approaches to crystals, including Delone's local theory [17] and a characterisation of multiregular sets [18]. By the end of January 2020, we finished the first version of [2] introducing the complete isoset invariant from last Chapter 11. This draft even discussed an initial distance metric on isosets, which required more updates and appeared later in [5], so the first paper [2] on isosets included only a complete invariant under isometry, without continuity.

January 2020 is considered the beginning of Geometric Data Science (in the periodic case) due to the first versions of isosets and density functions, and the publication of [38], which formalised the main problems. Since both isosets and density functions are rather complicated, we started to look for easier isometry invariants in February 2020.

The simple vectorial invariant of average distances to the first  $k$  atomic neighbours was called the AMD (Average Minimum Distances), see Definition 10.1.3, though the word "minimum" could have been replaced with the more appropriate "neighbour" in this context. The initial version of [53] was ready in the summer of 2020 due to the thorough implementation by Daniel Widdowson, who was an MSc student at that time.

Though the AMD turned out to be very powerful, it was theoretically incomplete. Since density functions can be computed only at finitely many radii, we started to look for an intermediate invariant that is stronger than the AMD, but is still fast in practice.

The Pointwise Distance Distribution was a natural extension of the AMD vector to a matrix invariant. Since the Earth Mover's Distance (EMD) was already applied to isosets that are unordered distributions of isometry classes of  $\alpha$ -clusters around points, it was straightforward to adapt EMD to PDD, considered as a distribution of unordered distances. The initial version of the PDD paper [50] was completed by April 2020.

The three new invariants (density functions, isosets, and AMD) with proved theoretical guarantees and implemented software justified the significance of a new continuous classification of periodic crystals. Moreover, the MIF++ seminar had been running for 1.5 years and naturally "overflew" to a new annual meeting MACSMIN (Mathematics and Computer Science for Materials Innovation), see <https://kurlin.org/macsmi.php>.

The first MACSMIN in September 2020 required no physical space, and most participants were already familiar with zoom. At MACSMIN 2020, we introduced the new area of Periodic Geometry by formalising the isometry classification problem and continuous metric problem for all periodic sets of points in any  $\mathbb{R}^n$ . Now similar questions make sense for other objects, such as periodic graphs, surfaces, and tilings.

By May 2021, it was clear that ignoring discontinuity of cell-based representations has led to unrecognised near-duplicates among simulated and experimental crystals.

Around the same time we learned about the paper [44], whose title “Same or different - that is the question” inspired the whole area of Geometric Data Science.

Since the initial implementation of the PDD was running in minutes on thousands of crystals, Daniel Widdowson bravely moved to hundreds of thousands of experimental crystals in the world’s largest database CSD to check if there are any non-isometric crystals with the same PDD. In parallel, Vitaliy was writing a proof that the PDD is complete under isometry for all periodic point sets in general position in  $\mathbb{R}^n$ .

Unfortunately, there was no record left on the day when Vitaliy was reading first Feynman’s lecture on physics, which contains Fig. 1-7 with a table illustrating 7 cubic crystals and their smallest interatomic distances  $d$ , now re-drawn in Fig. 12.1 (left).

Here is full quote from [24, section 1.3]: “Figure 1-7 is an illustration of the three-dimensional structure of common salt, sodium chloride”, the only place where Feynman mentioned Fig.1-7. Other 6 crystals were not discussed and could have been skipped.

However, many thanks to Feynman for including all numerical distances so that they can be unconsciously compared by a mathematician to conclude that all 7 simple crystals are geometrically different even if we forget about their chemistry.

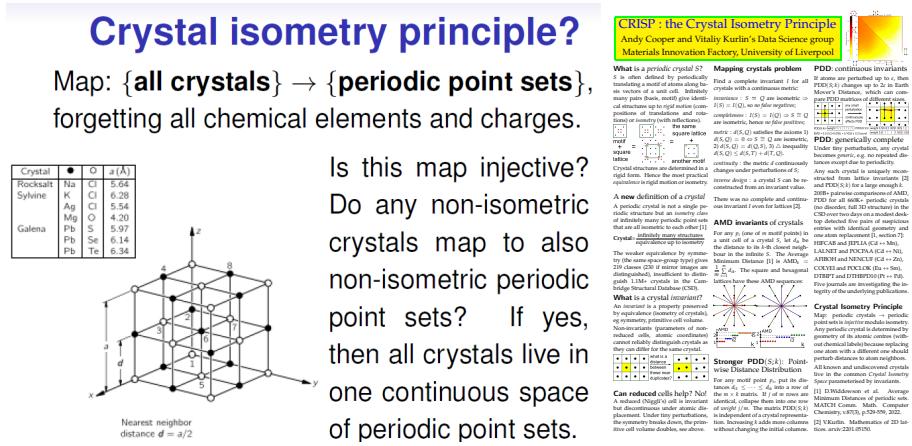
This was an Eureka moment to realise that all real crystals can also be distinguished purely geometrically, and hence their chemistry should be uniquely reconstructable from a precise enough geometry. At this moment, it was only a conjecture and later we added that any comparisons should be under the same ambient conditions, such as temperature and pressure, and only for real crystals, not for simulated or artificially generated ones, which can have arbitrary geometries. Nonetheless, the conclusion is simple and reflects our physical intuition: changing a chemical element should affect its interaction with atomic neighbours and hence perturb the crystal geometry.

This experimental observation was verified first on the CSD in September 2021 and presented at the 2nd MACSMIN on 16 September 2021 as the *Crystal Isometry Principle* (CRISP). More exactly, there were two online presentations on this day.

The first presentation at MACSMIN 2021 (<https://kurlin.org/macsmi/2021.php>) was titled “The Crystal Isometry Principle justifies a continuous map of periodic crystals”. Here is the original abstract. “More than 150 years ago Mendeleev put all known chemical elements into the same periodic table despite their differences to better understand their similarities. The recent advances in the new area of Periodic Geometry show that replacing any atom by its centre loses no information about a crystal structure, because all resulting periodic point sets remain non-isometric for all known crystals in the Cambridge Structural Database. This Crystal Isometry Principle (CRISP) justifies a map of all known and not yet discovered crystals continuously parameterised by complete isometry invariants. The IUCr congress included the related introductory talks: A unique and continuous code of all periodic crystals (16 min) and Introduction to invariant-based machine learning for periodic crystals (20 min).”

The aforementioned talks at the IUCr congress in August 2021 presented isosets [2] and the first energy predictions based on continuous AMD invariants [42].

The second 5-min presentation in the afternoon of 16 September 2021 was for the advisory board of the Leverhulme Research Centre at the MIF, where Richard Catlow FRS quickly grasped the significance, asked if the CRISP was published, and advised us to talk to crystallographers. The original slide from both talks in Fig. 12.4 (left) included a question mark in the title of CRISP, which we cautiously considered a conjecture.



**Fig. 12.4 Left:** the slide introducing the Crystal Isometry Principle (CRISP) on 16 September 2021 included original Feynman's Fig. 1-7 [24]. **Right:** the first conference poster of CRISP at BCA 2022.

When the Crystal Isometry Principle was externally presented in person after the covid pandemic in early March 2022, James Cumby proposed a potential counterexample of KBr and RbCl. These and all other simple cubic crystals, as in Fig. 12.1 (left), are geometrically determined by a single inter-atomic distance  $d$  between closest atoms.

Considering KBr and RbCl, if all atoms of K and Br are replaced with heavier atoms of Rb and lighter atoms of Cl, respectively, the distance between these atoms might remain exactly the same in theory. Indeed, we checked that the ICSD has several entries of KBr and RbCl under room temperature and ambient pressure with very similar shortest distances  $2d$  between atoms of the same chemical type, see Fig. 12.5.

All values of cube side lengths  $2d$  in Fig. 12.5 differ at least slightly, but one pair (KBr with ICSD id 53826 and RbCl with ICSD id 53829) has the same  $2d = 6.6\text{\AA}$ . However, these crystals were reported in the same paper [40] published in 1922.

Since experimental accuracy was lower 100 years ago than today, we asked our colleagues in the MIF to re-synthesise these crystals. In December 2022, Marc Little confirmed that the modern structure determination gives the shortest inter-atomic distance  $d = 3.32\text{\AA}$  for KBr and  $d = 3.29\text{\AA}$  for RbCl. Hence two decimal places in Angstroms sufficed to distinguish these very close geometric structures. Though there might be other cases of highly symmetric crystals with close geometries, we expect that accurate structure determination with modern precision (three decimal places reported for most experimental structures) will keep validating the Crystal Isometry Principle.

		Coll. Code	HMS	Struct.	Reference	Standardised Cell Parameter	Temperat	Pressure	Publ
		149928	P m -3 n	K Br3	JACS Au (2023) 3, (6) p. 1634-1641	4.8898 4.8898 4.8898 90.000 90.000 90.000	293.00	80000.0	2023
<input checked="" type="checkbox"/>	8673	F m -3 m	K Br	Luminescence (2018) 33, (5) p. 885-890	6.5960 6.5960 6.5960 90.000 90.000 90.000	293.00	0.101325	2018	
<input checked="" type="checkbox"/>	18015	F m -3 m	K Br	Zeitschrift fuer Kristallographie (1926) 63, p. 222-230	6.5980 6.5980 6.5980 90.000 90.000 90.000	293.00	0.101325	1926	
<input checked="" type="checkbox"/>	22157	F m -3 m	K Br	Annales Academiae Scientiarum Fenniae, Series A6: Physi	6.5847 6.5847 6.5847 90.000 90.000 90.000	293.00	0.101325	1969	
<input checked="" type="checkbox"/>	52243	F m -3 m	K Br	Proceedings of the Physical Society, London (1936) 48, p. 8	6.6430 6.6430 6.6430 90.000 90.000 90.000	293.00	0.101325	1936	
<input checked="" type="checkbox"/>	53826	F m -3 m	K Br	Journal of the Washington Academy of Sciences (1922) 12,	6.6000 6.6000 6.6000 90.000 90.000 90.000	293.00	0.101325	1922	
<input checked="" type="checkbox"/>	53842	F m -3 m	K Br	Zeitschrift fuer Physik (1921) 5, p. 17-26	6.5970 6.5970 6.5970 90.000 90.000 90.000	293.00	0.101325	1921	
	61556	P m -3 m	K Br	Journal of Research of the National Bureau of Standards, Sect	3.7400 3.7400 3.7400 90.000 90.000 90.000	293.00	2200.0	1964	
<input checked="" type="checkbox"/>	187220	F m -3 m	K Br	Archives of Physics Research (2012) 3, (5) p. 407-410	6.6790 6.6790 6.6790 90.000 90.000 90.000	293.00	0.101325	2012	

(1 of 1) 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

		Coll. Code	HMS	Struct.	F	Reference	Standardised Cell Parameter	Temperature	Pressure	Publ
<input checked="" type="checkbox"/>	199932	F m -3 m	Rb Cl	Acta Crystallographica (1961) 14, (4) p. 353-355	6.5989 6.5989 6.5989 90.000 90.000 90.000	293.00	0.101325	1961		
	199933	F m -3 m	Rb Cl	Acta Crystallographica (1961) 14, (4) p. 353-355	6.5916 6.5916 6.5916 90.000 90.000 90.000	301.00	0.101325	1961		
	199934	F m -3 m	Rb Cl	Acta Crystallographica (1961) 14, (4) p. 353-355	5.5972 5.5972 5.5972 90.000 90.000 90.000	322.00	0.101325	1961		
	199935	F m -3 m	Rb Cl	Acta Crystallographica (1961) 14, (4) p. 353-355	6.6060 6.6060 6.6060 90.000 90.000 90.000	359.00	0.101325	1961		
	199936	F m -3 m	Rb Cl	Acta Crystallographica (1961) 14, (4) p. 353-355	6.6110 6.6110 6.6110 90.000 90.000 90.000	375.00	0.101325	1961		
	199937	F m -3 m	Rb Cl	Acta Crystallographica (1961) 14, (4) p. 353-355	6.6296 6.6296 6.6296 90.000 90.000 90.000	447.00	0.101325	1961		
	199938	F m -3 m	Rb Cl	Acta Crystallographica (1961) 14, (4) p. 353-355	6.6332 6.6332 6.6332 90.000 90.000 90.000	466.00	0.101325	1961		
<input checked="" type="checkbox"/>	23761	F m -3 m	Rb Cl	Acta Crystallographica, Section A (1970) 26, p. 377-379	6.5810 6.5810 6.5810 90.000 90.000 90.000	293.00	0.101325	1970		
<input checked="" type="checkbox"/>	22166	F m -3 m	Rb Cl	Annales Academiae Scientiarum Fenniae, Series A6: Physica (1969)	6.5790 6.5790 6.5790 90.000 90.000 90.000	293.00	0.101325	1969		
	44620	P m -3 m	Rb Cl	Doklady Akademii Nauk SSSR (1957) 113, p. 797-798	3.8200 3.8200 3.8200 90.000 90.000 90.000	298.00	750.0	1957		
	61561	P m -3 m	Rb Cl	Journal of Research of the National Bureau of Standards, Section A: Phys	3.9100 3.9100 3.9100 90.000 90.000 90.000	293.00	500.0	1964		
<input checked="" type="checkbox"/>	53829	F m -3 m	Rb Cl	Journal of the Washington Academy of Sciences (1922) 12, p. 248-251	6.6000 6.6000 6.6000 90.000 90.000 90.000	293.00	0.101325	1922		
	61535	P m -3 m	Rb Cl	Proceedings of the National Academy of Sciences of the United States	3.8840 3.8840 3.8840 90.000 90.000 90.000	293.00	550.0	1962		
<input checked="" type="checkbox"/>	61538	F m -3 m	Rb Cl	Proceedings of the National Academy of Sciences of the United States	6.5820 6.5820 6.5820 90.000 90.000 90.000	293.00	0.101325	1962		
	61521	P m -3 m	Rb Cl	Proceedings of the Physical Society, London (1961) 77, p. 471-475	3.9700 3.9700 3.9700 90.000 90.000 90.000	193.00	0.101325	1961		
<input checked="" type="checkbox"/>	53844	F m -3 m	Rb Cl	Zeitschrift fuer Kristallographie (1925) 61, p. 463-514	6.5700 6.5700 6.5700 90.000 90.000 90.000	293.00	0.101325	1925		
<input checked="" type="checkbox"/>	18016	F m -3 m	Rb Cl	Zeitschrift fuer Kristallographie (1926) 63, p. 222-230	6.5480 6.5480 6.5480 90.000 90.000 90.000	293.00	0.101325	1926		
	26877	P m -3 m	Rb Cl	Zeitschrift fuer Physikalische Chemie, Abteilung B: Chemie der Element	3.7490 3.7490 3.7490 90.000 90.000 90.000	83.00	0.101325	1936		
	60401	P m -3 m	Rb Cl	Zhurnal Tekhnicheskoi Fiziki (1969) 39, p. 1115-1118	3.8700 3.8700 3.8700 90.000 90.000 90.000	293.00	560.0	1969		

(1 of 1) 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

**Fig. 12.5** All simple cubic crystals KBr (top) and RbCl (bottom) in the Inorganic Crystal Structure Database (ICSD). The cases under room temperature and ambient pressure are shaded.

In the geographic analogy, a complete invariant of a periodic crystal, e.g. the isoset, is similar to the latitude and longitude coordinates on Earth. If we know the geographic map precisely enough, say with a precision of 1 metre, these coordinates theoretically suffice to describe all other properties of any location, e.g. the altitude (height above the sea level), average temperature, etc. However, if we additionally include the altitude in the input, predictions of average temperature become much easier. Similarly with materials, though a precise enough atomic geometry should suffice by the Crystal Isometry Principle, using chemical elements certainly helps predict other properties.

Going back, in September 2021, we privately reported to the CCDC the first five pairs of exact duplicates in the CSD, including HIFCAB vs JEPLIA in Fig. 12.2.

Also in September 2021, we released the papers on complete isometry invariants of lattices in dimensions 2 and 3 at arxiv:2109.10885 and arxiv:2109.11538, respectively. Though we correctly guessed that these dimensions deserve separate papers, we could not anticipate yet that the paper on 2D lattices will be split into three [11, 12, 34], while a continuous metric on complete invariants of 3D lattices still needs more work [32].

The paper [53], which coined the Crystal Isometry Principle, was accepted by September 2021 and focused on simpler AMD invariants, but we mentioned the stronger PDD invariants, which appeared in arxiv:2108.04798 on 10 August 2021 and distinguished all non-duplicate crystals in the CSD more reliably. We fixed a time of Eureka's moment of the CRISP as May 2021 to commemorate Feynman's birthday in May.

## 12.5 Further history of Geometric Data Science in 2022 - 2026

This section discusses how the initial research on periodic crystals expanded to other data objects, such as finite point sets and molecules, including proteins.

September 2021 was a month of research explosions in our group also because the PDD was finally understood as an isometry invariant of a finite set of unordered points in addition to the periodic case. In the hindsight, we could have started in 2017 with the simpler finite case, especially due to our earlier work on point clouds [30, 31] in Topological Data Analysis (TDA). Then it became natural to compare the strengths of the PDD and persistent homology as isometry invariants of unordered clouds.

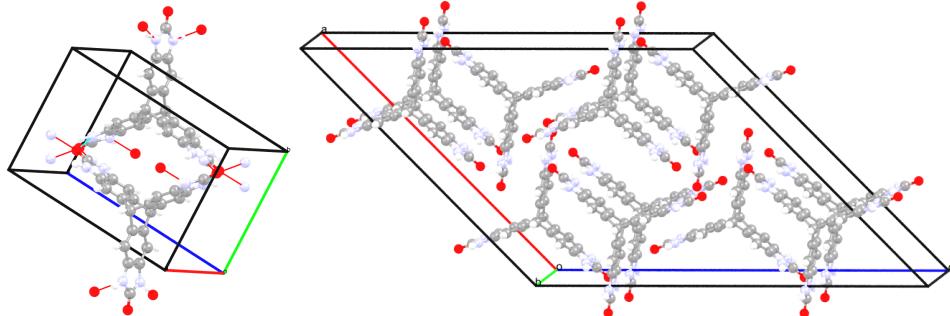
The 0-dimensional persistence of standard complexes on point clouds was known to be equivalent to a set of edge-lengths of a Minimum Spanning Tree (MST) and not distinguishing finite sets even in the line  $\mathbb{R}$ , but was extended to a stronger invariant (*mergegram*) that is Lipschitz continuous and has the same asymptotic time [19, 20].

Since Philip Smith previously computed 1-dimension persistence on many point clouds [47], he quickly checked Vitaliy's conjecture that persistence cannot distinguish generic clouds. In November 2021, Philip presented the first results at Oxford and also mentioned that the PDD is generically complete for finite and periodic point sets.

Since the PDD was clearer, easier, faster, and stronger than persistence, we offered our colleagues the collaborate beyond persistence towards a stronger invariant theory. In response, they advised us to find more applications of our invariants. Gratefully motivated by the advice from Oxford, the weaknesses of persistence became theorems in the paper [48] whose first version (arxiv:2202.00577) appeared in April 2022. The 3rd MACSMIN in September 2022 expanded Periodic Geometry to Geometric Data Science by stating the classification problem for all finite point clouds under isometry.

Returning to the periodic case and following Richard Catlow's helpful advice in September 2021, an abstract on CRISP was submitted to the British Crystallography Association (BCA) annual meeting, where the CRISP was presented as a poster (-!) in April 2022 and discussed with Tony Speck, see Fig. 12.4(right). Tony Speck is

well-known in crystallography for implementing the PLATON software, which validates CIFs. This software crushed when comparing our favourite simulated crystals in Fig. 12.6 whose cells differ by a factor of 4, though they can be matched almost perfectly.



**Fig. 12.6** Entries 14 and 15 from the database of 5679 simulated crystals reported in [41] consist of identical T2 molecules and have very different Crystallographic Information Files (with different motifs in unit cells of distinct shapes) but can be matched almost perfectly under isometry.

The first half of 2022 was occupied by three papers on 2D lattices [11, 12, 34].

After the Crystal Isometry Principle was presented at BCA 2022, Mike Glazer (Oxford) advised us to check the Protein Data Bank (PDB) for potential duplicates. This work started in October and quickly led to initial duplicates, which were confirmed Kay Diederichs already in December 2022, though our papers [3, 55] appeared later.

In the finite case, the PDD invariant was studied earlier under the name of a local distribution of distances [36]. In 2020, [22] described counter-examples to the completeness of the PDD for clouds of 5, 6, and 7 points in  $\mathbb{R}^3$ . These examples crucially helped us extend the PDD to the stronger SDD and SCD invariants. In October 2022, the SCD was proved to be complete under rigid motion in any  $\mathbb{R}^n$ . It was the first expansion from isometry to the stronger equivalence of rigid motion. Polynomial-time algorithms for SDDs, SCDs, and their potential distances were very technical but already routine due to the previous work on unordered distributions for periodic crystals. The Lipschitz continuity was much harder to guarantee because the sign of orientation, which distinguishes mirror images, discontinuously changes for degenerate clouds, while the volume of a simplex (or the convex hull of a cloud) is not Lipschitz continuous.

The first half of November 2022 was devoted to very intense attempts to finish a paper on SCDs before the submission deadline of CVPR 2023. Vitaliy tried numerous functions instead of the volume, always found counter-examples, and nearly gave up on Lipschitz continuity, but a systematic approach prevailed through partial cases. This time, Heron's formula  $S = \sqrt{p(p - a)(p - b)(p - c)}$  for the area  $S$  of a triangle

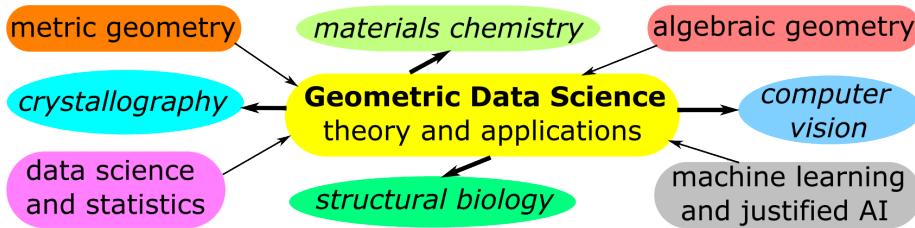
with sides  $a, b, c$  and half-perimeter  $p = \frac{a+b+c}{2}$  inspired the new strength  $\sigma = \frac{(p-a)(p-b)(p-c)}{p^2}$  and its generalisation to higher dimensions in Definition 6.3.1.

Our pure luck was to get reviewers, who grasped the significance of a complete, Lipschitz continuous and polynomial-time invariant, possibly without understanding all details of long proofs. As a result, the paper [54] was likely the most theoretical at CVPR 2023, similar to the previous results on point clouds [30] at CVPR 2014.

The complete hierarchy of point cloud invariants [54] helped us to geometrically compare all molecules as clouds of unordered atomic centres, even without covalent bonds and chemical elements. Large-scale experiments on molecular databases confirmed that a precise enough atomic geometry determines any real *molecular structure*, now defined as an equivalence class of only atomic centres under rigid motion in  $\mathbb{R}^3$ .

In the partial but important case of proteins, structural biologists often talk about a *shape* of a protein [27], though the word “shape” usually refers to an equivalence class under rigid motion composed with uniform scaling, as in shape theory [29]. However, the subject name of *structural biology* requires us to define a *protein structure*, proposed here as an equivalence class of (all chains of) a protein under rigid motion in  $\mathbb{R}^3$ .

The book included several conjectures and also postponed for future work some conditions of Geo-Mapping Problem 1.4.5, such as inverse continuity and Euclidean embeddability for finite clouds of unordered points in  $\mathbb{R}^n$ . We highlight Conjecture 4.5.5 about the completeness of the Pointwise Distance Distribution for any  $m \geq 5$  unordered points under isometry in  $\mathbb{R}^2$ , which can be accessible even to school children.



**Fig. 12.7** Geometric Data Science develops methods of metric and algebraic geometry to enrich statistics and machine learning for applications in computer vision, chemistry, and structural biology.

In conclusion, Geometric Data Science ‘connected the dots’ in practical challenges, unified all important requirements, and rigorously converted the fundamental questions (*Same or different? If different, by how much? Where do all objects live?*) into a list of verifiable conditions in Geo-Mapping Problem 1.4.5 as a guide for future developments. Fig. 12.7 illustrates theoretical sources and applications of Geometric Data Science.

## References

1. Google DeepMind’s GNoME database of 384,938 CIFs. <https://deepmind.google/discover/blog/millions-of-new-materials-discovered-with-deep-learning>, 2023.
2. O Anosova and V Kurlin. An isometry classification of periodic point sets. In *Proceedings of Discrete Geometry and Mathematical Morphology*, pages 229–241, 2021.
3. Olga Anosova et al. A complete and bi-continuous invariant of protein backbones under rigid motion. *MATCH Comm. Math. Comp. Chemistry*, 94:97–134, 2025.
4. Olga Anosova, Vitaliy Kurlin, and Marjorie Senechal. The importance of definitions in crystallography. *International Union of Crystallography Journal*, 11:453–463, 2024.
5. Olga Anosova, Daniel Widdowson, and Vitaliy Kurlin. Recognition of near-duplicate periodic patterns by continuous metrics with approximation guarantees. *Pattern Recognition*, 171:112108, 2026.
6. Mois Ilia Aroyo and H Wondratschek. *International Tables for Crystallography*. Wiley Online Library, 2013.
7. Paul V Bernhardt and Geoffrey A Lawrance. *Introduction to coordination chemistry*. John Wiley & Sons, 2025.
8. Harshad KDH Bhadeshia. *Geometry of crystals, polycrystals, and phase transformations*. CRC press, 2017.
9. B.N.Delone, N. Padurov, and A. Aleksandrov. *Mathematical Foundations of Structural Analysis of Crystals*. State Technical-Theoretical Press, USSR, 1934.
10. Auguste Bravais. Les systemes formes par des pointes distribuées régulièrement sur un plan ou dans l'espace. *J. Ecole. Polytech.*, pages 1–128, 1850.
11. Matthew Bright et al. Geographic-style maps for 2D lattices. *Acta Cryst A*, 79:1–13, 2023.
12. Matthew J Bright, Andrew I Cooper, and Vitaliy A Kurlin. Continuous chiral distances for 2-dimensional lattices. *Chirality*, 35:920–936, 2023.
13. J Cassels. *An introduction to the geometry of numbers*. Springer, 1959.
14. J Conway and N Sloane. Low-dimensional lattices. VI. Voronoi reduction of three-dimensional lattices. *Proceedings Royal Society A*, 436(1896):55–68, 1992.
15. Joseph Louis De Lagrange. Recherches d'arithmétique. *Nouveaux Mémoires de l'Académie de Berlin*, 1773.
16. BN Delone, RV Galiulin, and MI Shtogrin. On the Bravais types of lattices. *Journal of Soviet Mathematics*, 4(1):79–156, 1975.
17. Boris Nikolaevich Delone, Nikolai Petrovich Dolbilin, Mikhail Ivanovich Shtogrin, and Ravil V Galiulin. A local criterion for regularity of a system of points. In *Dokl. Akad. Nauk SSSR*, volume 227, pages 19–21, 1976.
18. N. Dolbilin, J Lagarias, and M. Senechal. Multiregular point systems. *Discrete & Computational Geometry*, 20(4):477–498, 1998.
19. Y. Elkin and V. Kurlin. The mergegram of a dendrogram and its stability. In *Proceedings of MFCS (Mathematical Foundations of Computer Science)*, pages 32:1–32:13, 2020.
20. Y. Elkin and V. Kurlin. Isometry invariant shape recognition of projectively perturbed point clouds by the mergegram extending 0d persistence. *Mathematics*, 9(17), 2021.
21. Peter Engel, Louis Michel, and Marjorie Sénéchal. Lattice geometry. Technical Report IHES-P-2004-45, 2004.
22. S Pozdnyakov et al. Incompleteness of atomic structure representations. *Phys. Rev. Lett.*, 125:166001, 2020.
23. Evgraf Stepanovich Fedorov. The symmetry of regular systems of figures. *Trans. Miner. Soc. Saint Petersburg*, 28(2):1–146, 1891.
24. Richard Feynman. *The Feynman lectures on physics: the new millennium edition*, volume 1. 2011.
25. J. Forsythe and V. Kurlin. Convex constrained meshes for superpixel segmentations of images. *Journal of Electronic Imaging*, 26(6):061609, 2017.
26. J. Forsythe, V. Kurlin, and A. Fitzgibbon. Resolution-independent superpixels based on convex constrained meshes without small angles. In *Lecture Notes in Computer Science (Proceedings of ISVC 2016)*, volume 10072, pages 223–233, 2016.

27. Elspeth F Garman and Nicola Laurieri. *Mathematics for Biosciences: From Theory to Worked Examples and Applications*. World Scientific, 2025.
28. Dominic Hyde. The sorites paradox. In *Vagueness: A guide*, pages 1–17. Springer, 2011.
29. David Kendall, Dennis Barden, Thomas Carne, and Huiling Le. *Shape and shape theory*. Wiley & Sons, 2009.
30. V. Kurlin. A fast and robust algorithm to count topologically persistent holes in noisy clouds. In *Proceedings of CVPR 2014: Computer Vision and Pattern Recognition*, pages 1458–1463. IEEE, 2014.
31. V. Kurlin. A one-dimensional homologically persistent skeleton of an unstructured point cloud in any metric space. *Computer Graphics Forum*, 34(5):253–262, 2015.
32. V Kurlin. A complete isometry classification of 3D lattices. *arxiv:2201.10543*, 2022.
33. V. Kurlin and D. Harvey. Superpixels optimized by color and shape. 10746:297–311, 2018.
34. Vitaliy Kurlin. Mathematics of 2D lattices. *Found. Comp. Mathematics*, 24:805–863, 2024.
35. Stephen Lawton and Robert Jacobson. The reduced cell and its crystallographic applications. Technical report, Ames Lab, Iowa State University, 1965.
36. Facundo Mémoli. Gromov–Wasserstein distances and the metric approach to object matching. *Foundations of Computational Mathematics*, 11(4):417–487, 2011.
37. Dmitri Ivanovich Mendeleev. *Mendeleev on the periodic law: Selected writings, 1869-1905*. Courier Corporation, 2013.
38. M Mosca and V Kurlin. Voronoi-based similarity distances between arbitrary crystal lattices. *Crystal Research and Technology*, 55(5):1900197, 2020.
39. P. Niggli. *Krystallographische und strukturtheoretische Grundbegriffe*, volume 1. Akademische Verlagsgesellschaft mbh, 1928.
40. Eugen Posnjak and Ralph WG Wyckoff. The crystal structures of the alkali halides. ii. *Journal of the Washington Academy of Sciences*, 12(10):248–251, 1922.
41. A Pulido et al. Functional materials discovery using energy–structure maps. *Nature*, 543:657–664, 2017.
42. Jakob Ropers, Marco M Mosca, Olga D Anosova, Vitaliy A Kurlin, and Andrew I Cooper. Fast predictions of lattice energies by continuous isometry invariants of crystal structures. In *International Conference on Data Analytics and Management in Data Intensive Domains*, pages 178–192, 2022.
43. Galiulin R.V. *Crystallographic Geometry (in Russian)*. Librokom, Moscow, 2009.
44. Pietro Sacchi et al. Same or different – that is the question: identification of crystal forms from crystal structure data. *Cryst Eng Comm*, 22(43):7170–7185, 2020.
45. AM Schönfli. *Kristallsysteme und Kristallstruktur Leipzig*. Teubner, 1891.
46. Eduard Selling. Ueber die binären und ternären quadratischen formen. *Journal für die reine und angewandte Mathematik*, 77:143–229, 1874.
47. Philip Smith and Vitaliy Kurlin. Skeletonisation algorithms with theoretical guarantees for unorganised point clouds with high levels of noise. *Pattern Recognition*, 115:107902, 2021.
48. Philip Smith and Vitaliy Kurlin. Generic families of finite metric spaces with identical or trivial 1-dimensional persistence. *Journal of Applied and Computational Topology*, 8:839–855, 2024.
49. Georges Voronoi. Nouvelles applications des paramètres continus à la théorie des formes quadratiques. *J. Reine Angew. Math.*, (133):97–178, 1908.
50. Daniel Widdowson and Vitaliy Kurlin. Resolving the data ambiguity for periodic crystals. *Advances in Neural Information Processing Systems*, 35:24625–24638, 2022.
51. Daniel Widdowson and Vitaliy Kurlin. Continuous invariant-based maps of the cambridge structural database. *Crystal Growth and Design*, 24:5627–5636, 2024.
52. Daniel Widdowson and Vitaliy Kurlin. Pointwise distance distributions for detecting near-duplicates in large materials databases. *SIAM Journal on Applied Mathematics* (doi:10.1137/25M1736657, arxiv:2108.04798), 2026.
53. Daniel Widdowson, Marco M Mosca, Angeles Pulido, Andrew I Cooper, and Vitaliy Kurlin. Average minimum distances of periodic point sets - foundational invariants for mapping all periodic crystals. *MATCH Commun. Math. Comput. Chem.*, 87:529–559, 2022.
54. Daniel E Widdowson and Vitaliy A Kurlin. Recognizing rigid patterns of unlabeled point clouds by complete and continuous isometry invariants with no false negatives and no false positives. In *Computer Vision and Pattern Recognition*, pages 1275–1284, 2023.

55. A Wlodawer et al. Duplicate entries in the protein data bank: how to detect and handle them. *Acta Cryst D*, 81:170–180, 2025.
56. B. Zhilinskii. *Introduction to lattice geometry through group action*. EDP sciences, 2016.

## Acronyms

ADA	Average Deviation from Asymptotic
ADD	Average Distance Distribution
AMD	Average Minimum Distances
AND	Average Normalised Distances
ASD	Average Simplexwise Distribution
ASM	Average Simplexwise Moments
BD	Bottleneck Distance
BRI	Backbone Rigid Invariant
BRIS	Backbone Rigid Invariant Space
BRAIN	Backbone Rigid Average Invariant
BT	Boundary Tolerant metric
CDM	Cyclic Distance Matrix
CDS	Cyclic Distances with Signs
CIF	Crystallographic Information File
CIS	Cloud Isometry Space
CIM	Cyclic Isometry Metric
CIMS	Crystal Isometry Space
COD	Crystallography Open Database
CR	Cyclic Rigid invariant
CRM	Cyclic Rigid Metric
CRS	Cloud Rigid Space
CRIS	Crystal Rigid Space
CRISP	Crystal Isometry Principle
CSD	Cambridge Structural Database
DIM	Dihedral Isometry Metric
DR	Dihedral Rigid invariant
DRM	Dihedral Rigid Metric
EMD	Earth Mover's Distance
GC	Geometric Complexity
GDS	Geometric Data Science
GNoME	Graph Network Materials Exploration dataset

HD	Hausdorff Distance
ICSD	Inorganic Crystal Structural Database
LAC	Linear Assignment Cost
LDS	Lattice Dilatation Space
LHS	Lattice Homothety Space
LIS	Lattice Isometry Space
LRS	Lattice Rigid Space
MSD	Measured Simplexwise Distribution
MP	Materials Project
OCD	Oriented Centred Distribution
ORD	Oriented Relative Distribution
OSD	Oriented Simplexwise Distribution
PCA	Principal Components Analysis
PCI	Principal Coordinates Invariant
PCM	Principal Coordinates Matrix
PDA	Pointwise Deviation from Asymptotic
PDD	Pointwise Distance Distribution
PND	Pointwise Normalised Distribution
PPC	Point Packing Coefficient
PI	Projected Invariant
QT	Quotient Triangle
RDD	Relative Distance Distribution
RI	Root Invariant
SBD	Superbases under Dilatation
SBH	Superbases under Homothety
SBI	Superbases under Isometry
SBR	Superbases under Rigid Motion
SCD	Simplexwise Centred Distribution
SDD	Simplexwise Distance Distribution
SDM	Superbase Dilatation Metric
SHM	Superbase Homothety Metric
SIM	Superbase Isometry Metric
SLM	Spherical Lattice Map
SM	Symmetrised Metric
SPD	Sorted Pairwise Distances
SRD	Sorted Radial Distances
SRM	Superbase Rigid Metric
TC	Triangular Cone
TDA	Topological Data Analysis
TRIN	Triangular Invariant
WMI	Weighted Matrices Invariant
WSD	Weighted Simplexwise Distribution

# Index

## Symbols

$l$ -periodic point set 186  
 $m$ -regular periodic point set 207  
0D homology 53  
1-periodic sequence 103

## A

affine dimension 24  
Average Deviation from Asymptotic 187  
Average Normalised Deviation 187

## B

backbone invariant barcode 32  
backbone invariant diagram 32  
backbone rigid average invariant 32  
Backbone Rigid Invariant 30–32  
Backbone Rigid Space 28  
backbone structure 28  
bottleneck distance 7, 8  
bridge length 208

## C

cell-periodic point set 187  
Chebyshev metric 8  
cloud 7  
Cloud Isometry Space 10  
Cloud Rigid Space 10  
cloud space 10  
complete invariant 6, 11  
conorm 118  
covariance matrix 39  
covering radius 154  
Crystal Dilation Space 151  
Crystal Homothety Space 151

Crystal Isometry Space 151  
Crystal Rigid Space 151  
crystal spaces 150  
cyclic distance matrix 105  
cyclic equivalence 104  
cyclic rigid invariant 111

## D

density fingerprint 152  
density function 152, 153, 165  
dihedral equivalence 104  
dihedral rigid invariant 111  
dilation 5, 129  
directed distance 220

## E

Earth Mover’s Distance 46, 85, 87, 218  
embedding 10, 135  
equivalence relation 3, 5

## G

general position 59  
generically complete invariant 60  
geo-mapping 121  
Geo-Mapping Problem v, 11, 16  
geocode 11, 13–15, 121  
geometric complexity 220  
graph 49

## H

Hausdorff distance 7, 8  
homothety 5, 129

## I

invariant 6, 11  
isometry 5, 127

isoset 214, 220  
isotree 208

## L

lattice 115, 121  
lattice spaces 121  
Linear Assignment Cost 45, 85  
Lipschitz continuity 9, 85, 97, 99  
local cluster 207

## M

max metric 84, 98  
measured simplexwise distribution 86  
metric 7  
metric graph 49  
metric space 7  
metric-measure space 85–87  
Minkowski metric 8  
moduli space 9  
moments of a weighted distribution 73  
motif 149

## O

obtuse superbase 118, 126, 144

## P

packing radius 154  
partition 50  
partition set 50  
periodic point set 149  
persistence diagram 53  
persistence module 52  
point cloud 49  
Point Packing Coefficient 187  
Pointwise Deviation from Asymptotic 187  
Pointwise Distance Distribution 62–64  
Pointwise Normalised Deviation 187  
polynomial-time complexity 10, 75, 85, 111  
Principal Coordinates Invariant 39, 40, 42  
Principal Coordinates Matrix 39  
principally generic cloud 39

projected invariant 127, 129

## Q

quotient square 127  
quotient triangle 127

## R

reduced basis 126  
Relative Distance Distribution 71, 84  
rigid motion 5, 127  
root invariant 123, 124

## S

Simplexwise Distance Distribution 73, 75, 85  
Simplexwise Distance Moments 73  
sorites paradox 4  
sorted pairwise distances 60, 62  
stable radius 212  
strength of a simplex 96–99  
superbase 118  
symmetrised metric 41, 42

## T

transitivity axiom 3  
triangular cone 127  
triangular invariant 29

## U

unit cell 115

## V

vonorm 118  
Voronoi domain 116, 117, 124

## W

Weighted Matrices Invariant 43, 44  
Weighted Matrix Invariant 44  
weighted simplexwise distribution 87