Geographic-style maps for 2-dimensional lattices

Matthew Bright,[a] Andrew I Cooper[a] and Vitaliy Kurlin [a]*

[a]*Materials Innovation Factory, University of Liverpool, UK.*

E-mail: vitaliy.kurlin@gmail.com

2-dimensional lattice, reduced basis, obtuse superbase, isometry, complete invariant, metric, continuity

Abstract

This paper develops geographic-style maps containing 2D lattices in all known periodic crystals parameterised by recent complete invariants. Motivated by rigid crystal structures, lattices are considered up to rigid motion and uniform scaling. The resulting space of 2D lattices is a square with identified edges or a punctured sphere. The new continuous maps show all Bravais classes as low-dimensional subspaces, visualise hundreds of thousands of real crystal lattices from the Cambridge Structural Database, and motivate the development of continuous and invariant-based crystallography.

1. Practical motivations for the problem to continuously classify lattices

This paper for mathematical crystallographers presents applications of (Kurlin, 2022*b*) written for mathematicians and computer scientists, with proofs of the invariance of map coordinates up to basis choice, and their continuity under perturbations of a basis. A lattice can be considered as a periodic crystal whose atomic motif consists of a single point. In Euclidean space $\mathbb{R}^n$, a *lattice* $\Lambda \subset \mathbb{R}^n$ consists of all integer linear combinations of basis vectors $v_1, \dots, v_n$, which span a primitive *unit cell* U of Λ.

Crystallography traditionally splits crystals into only finitely many classes, for instance by their space-group types. These discrete symmetry-based classifications were suitable for distinguishing highly symmetric crystals manually or simply by eye. Nowadays crystals are simulated and synthesised on an industrial scale. The Cambridge Structural Database (CSD) contains nearly 1.2M existing crystals (Groom *et al.*, 2016). Crystal Structure Prediction (CSP) tools generate millions of crystals even for a fixed chemical composition (Pulido *et al.*, 2017), mostly with P1 symmetry. Datasets of this size require finer classifications than by 230 crystallographic groups.

A more important reason for a continuous approach to classifying periodic structures is the inevitability of noise in data. Slight changes in initial simulated or actual crystallisation conditions mean that the same crystal can have slightly different X-ray patterns leading to close but distinct structures. Fig. 1 shows that a reduced cell cannot be used to continuously quantify a distance between general periodic sets. If we consider only lattices, a similar discontinuity of a reduced basis arises in Fig. 2.

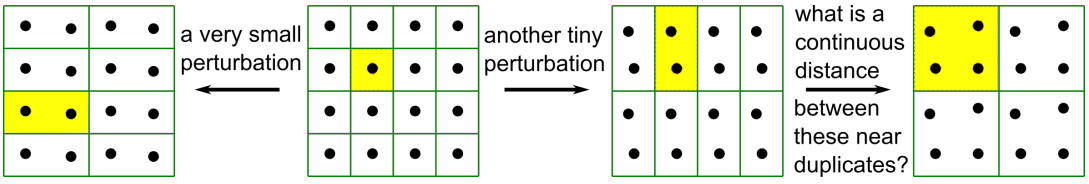

Fig. 1. For almost any perturbation of atoms, the symmetry group and any reduced cell (even its volume) discontinuously change, which justifies a continuous classification.

Consider the family of lattices with the basis $v_1 = (1, 0)$, $v_2(t) = (-t, 2)$ in Fig. 2, where the parameter t varies continuously in $[0, 1]$. Since the initial basis $v_1 = (1, 0)$, $v_2(0) = (0, 2)$ and final basis $v_1 = (1, 0)$, $v_2(1) = (-1, 2)$ define identical lattices, this continuous family of lattices is a closed loop in the space of all lattices. For $t \in [0, \frac{1}{2})$, the given basis $v_1 = (1, 0)$, $v_2(t) = (-t, 2)$ is reduced by Definition 2.1. At the critical moment $t = \frac{1}{2}$, the lattice has several primitive bases that can be chosen as reduced.

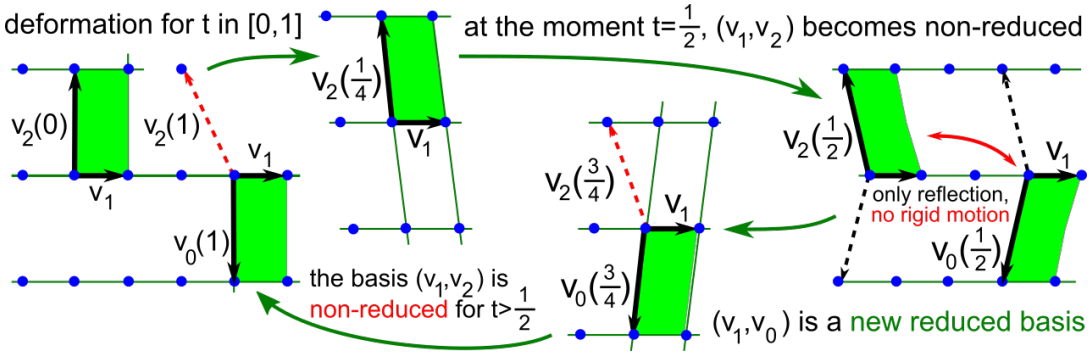

Fig. 2. The deformation of the basis $v_1 = (1, 0)$, $v_2 = (-t, 2)$ for $t \in [0, 1]$ defines a continuous loop of lattices. The basis v_1, v_2 is reduced for $t \in [0, \frac{1}{2})$ but after $t = \frac{1}{2}$ switches to a non-equivalent (up to rigid motion) reduced basis $v_1, v_0 = (t-1, -2)$.

When t passes through $\frac{1}{2}$, if we keep the angle between basis vectors continuous, the reduced basis $v_1 = (1, 0)$, $v_2(\frac{1}{2}) = (-\frac{1}{2}, 2)$ switches to $v_1 = (1, 0)$, $v_0(\frac{1}{2}) = (-\frac{1}{2}, -2)$. For any choice at $t = \frac{1}{2}$, the basis $v_1 = (1, 0), v_0(t) = (t-1, -2)$ will be a new reduced basis for $t \in (\frac{1}{2}, 1]$. The above change at $t = \frac{1}{2}$ creates the discontinuity because the given bases $v_1 = (1, 0), v_2(\frac{1}{2} - \varepsilon) = (\varepsilon - \frac{1}{2}, 2)$ and $v_1 = (1, 0), v_2(\frac{1}{2} + \varepsilon) = (-\varepsilon - \frac{1}{2}, 2)$ at $t = \frac{1}{2} \mp \varepsilon$ differ only by a small perturbation $2\varepsilon > 0$ in all coordinates but the lattices have the reduced bases v_1, $v_2(\frac{1}{2} - \varepsilon) = (\varepsilon - \frac{1}{2}, 2)$ and v_1, $v_0(\frac{1}{2} + \varepsilon) = (\varepsilon - \frac{1}{2}, -2)$, whose last coordinates differ by 4. These reduced bases cannot be made close by rigid motion because they have opposite anticlockwise angles from v_1 to the longer vector.

One way to call lattices identical (or equivalent) is to ignore deviations of lattice parameters up to a certain threshold. An equivalence gives rise to a justified classification only if this *equivalence relation* (denoted by $\sim$) satisfies the axioms:

(1) reflexivity : any lattice Λ is equivalent to itself, so $\Lambda \sim \Lambda$;

(2) symmetry : if $\Lambda \sim \Lambda'$ then $\Lambda' \sim \Lambda$;

(3) transitivity : if $\Lambda \sim \Lambda'$ and $\Lambda' \sim \Lambda''$ then $\Lambda \sim \Lambda''$.

The transitivity axiom is needed to split lattices into disjoint *equivalence classes*:

the class $[\Lambda]$ consists of all lattices equivalent to Λ, since if Λ is equivalent to Λ', which is equivalent to Λ'', all three lattices are in the same class. Past equivalences in (Lima-de Faria *et al.*, 1990) use numerical thresholds to determine a lattice class, but as Fig. 3 illustrates, all lattices can be made equivalent through sufficiently many slight perturbations up to any positive threshold due to the transitivity axiom.

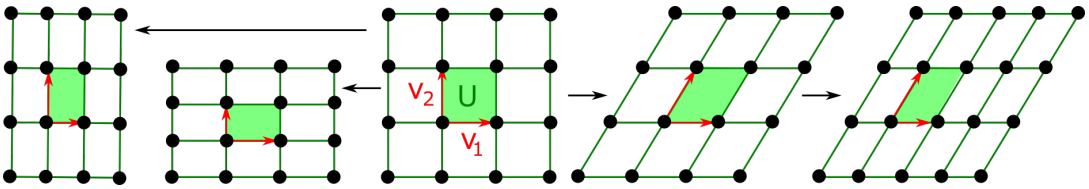

Fig. 3. All lattices continuously deform into each other if we allow any small changes.

An alternative mathematical approach classifies lattices by space groups and finer algebraic structures (Nespolo, 2008). Since crystal structures are determined as rigid forms, the most practically important equivalence of crystals and their lattices is a *rigid motion*, which in $\mathbb{R}^2$, is any composition of translations and rotations. This is the strongest possible equivalence on crystals that are indistinguishable as rigid bodies.

Slightly weaker is equivalence based on *isometry* or congruence, denoted by $\Lambda \cong \Lambda'$, which is any rigid motion composed with mirror reflections. Even if we fix an equivalence such as isometry, (Sacchi *et al.*, 2020) highlights that the key question 'same or different' remained unanswered. What is needed is the notion of an *invariant*.

Definition 1.1 (*invariants* vs *complete invariants*)**.** A descriptor I, such as a numerical vector, is called an *isometry invariant* of a lattice $\Lambda \subset \mathbb{R}^2$ if I takes the same value on all isometric lattices: if $\Lambda \cong \Lambda'$ are isometric then $I(\Lambda) = I(\Lambda')$, so I has no *false negatives*. An isometry invariant I is called *complete* (or *injective*) if the converse also holds: if $I(\Lambda) = I(\Lambda')$ then $\Lambda \cong \Lambda'$, so I distinguishes all non-isometric lattices. Hence a complete invariant I has neither false negatives, nor no false positives, see Fig. 4.

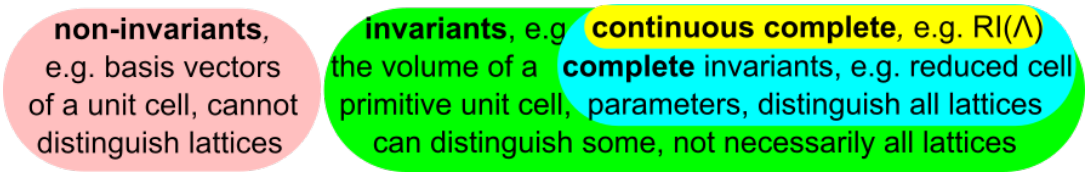

Fig. 4. The root invariant RI(Λ) from Definition 3.1 used for mapping CSD crystals in this paper is a continuous and complete isometry invariant of all 2D lattices.

In a fixed coordinate system, the basis vectors are not isometry invariants as they change under rotation, but the primitive cell area is preserved by isometry. If an invariant I takes different values on lattices Λ, Λ', these lattices are certainly not isometric, while non-invariants cannot help distinguish equivalent objects. For example, isometric lattices $\Lambda \cong \Lambda'$ can have infinitely many primitive bases. Most isometry invariants allow *false positives* that are non-isometric lattices $\Lambda \ncong \Lambda'$ with $I(\Lambda) = I(\Lambda')$. For instance, infinitely many non-isometric lattices have the same primitive cell area.

Complete invariants are the main goal of all classifications. *Continuous* invariants, which change only slightly under small perturbations of the underlying object, are even better. The dependence of pseudosymmetry on thresholds discussed in (Zwart *et al.*, 2008) can be resolved in a continuous way by finding, for any given lattice, its closest higher-symmetry neighbour through continuous invariants as in Problem 1.2.

Problem 1.2. Find a complete isometry invariant $I(\Lambda)$ of any lattice $\Lambda \subset \mathbb{R}^2$ with a metric d satisfying all necessary axioms and the new continuity condition below:

(1) *first axiom* : $d(\Lambda, \Lambda') = 0$ if and only if $\Lambda \cong \Lambda'$ are isometric;

(2) *symmetry axiom* : $d(\Lambda, \Lambda') = d(\Lambda', \Lambda)$ for any lattices $\Lambda, \Lambda' \subset \mathbb{R}^2$;

(3) *triangle axiom* : $d(\Lambda, \Lambda') + d(\Lambda', \Lambda'') \geq d(\Lambda, \Lambda'')$ for any lattices $\Lambda, \Lambda', \Lambda'' \subset \mathbb{R}^2$;

(4) *Lipschitz continuity* : there is a constant C such that, for any lattices $\Lambda, \Lambda' \subset \mathbb{R}^2$, if corresponding coordinates of basis vectors differ by at most $\varepsilon > 0$, then $d(\Lambda, \Lambda') \leq C\varepsilon$.

This paper applies a solution of Problem 1.2 from (Kurlin, 2022*b*) to visualise CSD crystals on continuous maps. Sections 2 and 3 review the related past work. Section 4 maps hundreds of thousands of CSD crystals. Section 5 explains the geographical metaphor by mapping the invariant values to a sphere, where every 2D lattice (up to rigid motion and uniform scaling) has unique latitude and longitude coordinates.

2. Overview of key concepts and past work on classifications of lattices

Crystallography traditionally uses a conventional cell to uniquely represent any periodic crystal, see Aroyo (2016, Chapter 2.1). In the simpler case of 3-dimensional lattices, the cell used is Niggli's reduced cell (Niggli, 1928). Since the current paper studies lattices in $\mathbb{R}^2$, we give the 2-dimensional version obtained from the 3-dimensional definition, which is derived as a limit of the reduction conditions for a 3-dimensional reduced basis with an orthogonal third vector v_3 whose length becomes infinite. For vectors $v_1 = (a_1, a_2)$ and $v_2 = (b_1, b_2)$ in $\mathbb{R}^2$, the determinant of the matrix $\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$ with the columns v_1, v_2 is defined as $\det(v_1, v_2) = a_1 b_2 - a_2 b_1$.

Definition 2.1 (*reduced cell*)**.** For a lattice up to isometry, a basis and its cell $U(v_1, v_2)$ are called *reduced* (non-acute) if $|v_1| \le |v_2|$ and $-\frac{1}{2}v_1^2 \le v_1 \cdot v_2 \le 0$. Up to rigid motion, the conditions are weaker: $|v_1| \le |v_2|$ and $-\frac{1}{2}v_1^2 < v_1 \cdot v_2 \le \frac{1}{2}v_1^2$, $\det(v_1, v_2) > 0$, and the new *special condition* for rigid motion : if $|v_1| = |v_2|$ then $v_1 \cdot v_2 \ge 0$. ■

The new conditions for rigid motion did not appear in Aroyo (2016, section 9.2.2) because reduced bases were considered up to isometry including reflections. Any rectangular lattice has a unique (up to rigid motion) reduced cell $a \times b$, but two 'potentially reduced' bases $v_1 = (a, 0)$ and $v_2 = (0, \pm b)$, which are not related by rigid motion for $0 < a < b$. Definition 2.1 chooses only one of these bases, namely $v_1 = (a, 0)$ and $v_2 = (0, b)$. So $\det(v_1, v_2) > 0$ defines a right-handed basis in $\mathbb{R}^2$.

Since reduced bases are easy to compute (Křivỳ & Gruber, 1976), they can be used to define the discrete metric $d(\Lambda, \Lambda')$ taking the same non-zero value (say, 1) for any non-isometric lattices $\Lambda \ncong \Lambda'$. Discontinuity of a reduced basis up to perturbations was practically demonstrated in the seminal work (Andrews *et al.*, 1980). The introduction of Edelsbrunner *et al.* (2021) said that "There is no method for choosing a unique basis for a lattice in a continuous manner. Indeed, continuity contradicts uniqueness as we can continuously deform a basis to a different basis of the same lattice", see Fig. 2 and a formal proof in Widdowson *et al.* (2022, Theorem 15). Since a reduced basis is discontinuous under perturbations, then so is any metric on these reduced bases.

L. Andrews and H. Bernstein made important advances in (Andrews & Bernstein, 1988; Andrews & Bernstein, 2014; McGill *et al.*, 2014; Andrews *et al.*, 2019*a*; Bernstein *et al.*, 2022) by analysing complicated boundary cases where cell reductions can be discontinuous. Since these advances are specialised for $\mathbb{R}^3$, we defer a detailed review of reduced bases for 3-dimensional lattices to another paper (Bright *et al.*, 2021).

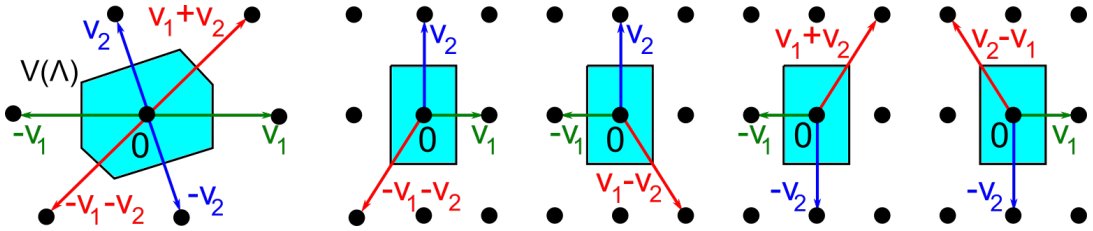

Fig. 5. **Left**: a generic 2D lattice has a hexagonal Voronoi domain with an obtuse superbase $v_1, v_2, v_0 = -v_1 - v_2$, which is unique up to permutations and central symmetry. **Other pictures**: isometric superbases for a rectangular Voronoi domain.

Another way to represent a lattice $\Lambda \subset \mathbb{R}^n$ is by its Wigner-Seitz cell (Wigner & Seitz, 1933) or Voronoi domain $V(\Lambda)$ consisting of all points $p \in \mathbb{R}^n$ that are closer to the origin $0 \in \Lambda$ than to all other points of Λ. Though $V(\Lambda)$ uniquely determines Λ up to rotations, almost any tiny perturbation of a rectangular lattice Λ converts the rectangular domain $V(\Lambda)$ into a hexagon. Hence all combinatorial invariants (numbers

of vertices or edges) of $V(\Lambda)$ are discontinuous, similarly in higher dimensions.

However, comparing Voronoi domains as geometric shapes by optimal rotation (Mosca & Kurlin, 2020) around a common centre led to two continuous metrics on lattices up to rigid motion and uniform scaling. The minimisation over infinitely many rotations was resolved only by finite sampling, so the exact computation of these metrics is still open. Similar computational difficulties remain for stronger isometry invariants of general periodic sets (Anosova & Kurlin, 2021*b*; Anosova & Kurlin, 2021*a*; Smith & Kurlin, 2022; Anosova & Kurlin, 2022*a*; Anosova & Kurlin, 2022*b*).

Another attempt to produce computable metrics was to consider distance-based invariants (Widdowson *et al.*, 2022; Widdowson & Kurlin, 2022) whose completeness was proved for generic crystals. These invariants helped establish the Crystal Isometry Principle by experimentally checking that all periodic crystals from the CSD remain non-isometric after forgetting all chemical information. This principle implies that all periodic crystals can be studied in the common Crystal Isometry Space (CRISP) whose version for 2-dimensional lattices is the Lattice Isometry Space LIS($\mathbb{R}^2$).

Though the paper by (Conway & Sloane, 1992) 30 years ago aimed for continuous invariants of 3D lattices, no formal proofs were given even for the isometry invariance. This past work for 3D lattices has been corrected and extended in (Kurlin, 2022*a*).

Kurlin (2022*b*, Proposition 3.10) proves that a reduced basis from Definition 2.1 is unique (also in the case of rigid motion) and all reduced bases are in a 1-1 correspondence with obtuse superbases, which are easier to visualise, especially for $n \leq 3$.

Definition 2.2 (*superbase*, *conorms* p_{ij})**.** For any basis $v_1, \dots, v_n$ in $\mathbb{R}^n$, the *superbase* $v_0, v_1, \dots, v_n$ from (Conway & Sloane, 1992) includes the vector $v_0 = -\sum_{i=1}^{n} v_i$. The *conorms* $p_{ij} = -v_i \cdot v_j$ are the negative scalar products of the vectors. The superbase is called *obtuse* if all $p_{ij} \geq 0$, so all angles between the vectors v_i, v_j are non-acute for distinct indices $i, j \in \{0, 1, \dots, n\}$. The obtuse superbase is *strict* if all $p_{ij} > 0$. ■

Definition 2.2 uses the conorms p_{ij} from (Conway & Sloane, 1992), which were also known as negative Selling parameters (Selling, 1874) and Delaunay parameters (Delaunay *et al.*, 1934). Lagrange (Lagrange, 1773) proved that the isometry class of any lattice $\Lambda \subset \mathbb{R}^2$ with a basis v_1, v_2 is determined by 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 $q_{11} = v_1^2$, $q_{12} = v_1 \cdot v_2$, $q_{22} = v_2^2$. The triple $(v_1^2, v_1 \cdot v_2, v_2^2)$ is also called a metric tensor of (a basis of) Λ. Any $Q(x,y)$ has a reduced (non-acute) form with $0 < q_{11} \leq q_{22}$ and $-q_{11} \leq 2q_{12} \leq 0$, which is equivalent to reducing a basis up to isometry.

The bases $v_1 = (3,0)$, $v_2^{\pm} = (-1, \pm 2)$ generate the mirror images not related by rigid motion, but define the same form $Q = 9x^2 - 6xy + 5y^2$ satisfying the reduction conditions above. So quadratic forms do not distinguish mirror images (enantiomorphs). Hence the new conditions for the rigid motion were needed in Definition 2.1.

Motivated by the non-homogeneity of the metric tensor (two squared lengths and scalar product), Delaunay proposed (Delaunay, 1937) the homogeneous parameters

$$p_{12} = -v_1 \cdot v_2 = -q_{12}, \quad p_{01} = -v_0 \cdot v_1 = q_{11} + q_{12}, \quad p_{02} = -v_0 \cdot v_2 = q_{22} + q_{12},$$

called conorms in (Conway & Sloane, 1992), see Definition 2.2. Then any permutation of superbase vectors satisfying $v_0 + v_1 + v_2 = 0$ changes p_{12}, p_{01}, p_{02} by the same permutation of indices. For example, swapping v_1, v_2 is equivalent to swapping p_{01}, p_{02}.

Delaunay's reduction (Delaunay *et al.*, 1973) proved the key existence result: any lattice in dimensions 2 and 3 has an obtuse superbase with all $p_{ij} \geq 0$. Section 3 further develops the Delaunay parameters to show in section 4 how millions of lattices from real crystals in the CSD are distributed in continuous spaces of lattices.

3. Homogeneous complete invariants of 2D lattices up to four equivalences

This section reminds the lattice classifications in Theorem 3.4 based on the recent invariants introduced in Definitions 3.1 and 3.2 from Kurlin (2022*b*, sections 3-4).

Definition 3.1 (sign(Λ) and root invariants $\mathrm{RI}, \mathrm{RI}^o$)**.** Let $B = \{v_0, v_1, v_2\}$ be any obtuse superbase of a lattice $\Lambda \subset \mathbb{R}^2$. If Λ is mirror-symmetric (achiral), set $\mathrm{sign}(\Lambda) = 0$. Otherwise v_0, v_1, v_2 have different lengths and no right angles, and hence can be ordered so that $|v_1| < |v_2| < |v_0|$. Let $\mathrm{sign}(\Lambda)$ be the sign of $\det(v_1, v_2)$ of the matrix with the columns v_1, v_2. The *root invariant* $\mathrm{RI}(\Lambda)$ is the triple of the root products $r_{ij} = \sqrt{-v_i \cdot v_j}$, which have original units of vector coordinates such as Angstroms and are ordered by their size for distinct indices $i, j \in \{0, 1, 2\}$. The *oriented root invariant* $\mathrm{RI}^o(\Lambda)$ is $\mathrm{RI}(\Lambda)$ with $\mathrm{sign}(\Lambda)$ as a superscript, which we skip if $\mathrm{sign}(\Lambda) = 0$. ■

We assume that $r_{ij} = r_{ji}$. If $|v_1| < |v_2| < |v_0|$ then $r_{12} < r_{01} < r_{02}$. If some v_i, v_j have equal lengths, then $r_{ik} = r_{jk}$ for $k \neq i, j$. Writing $\mathrm{RI}(\Lambda) = (r_{12}, r_{01}, r_{02})$ means that $|v_1| \leq |v_2| \leq |v_0|$ for a suitable indexing of obtuse superbase vectors v_0, v_1, v_2.

Kurlin (2022*b*, Lemma 3.8) proved that $\mathrm{RI}(\Lambda)$ is an isometry invariant of Λ, independent of an obtuse superbase B because an obtuse superbase of Λ is unique up to isometry, also up to rigid motion for non-rectangular lattices. This uniqueness was missed in (Conway & Sloane, 1992) and actually fails in $\mathbb{R}^3$, see (Kurlin, 2022*a*).

Definition 3.2 (projected invariants $\mathrm{PI}, \mathrm{PI}^o$)**.** The root invariants of all lattices $\Lambda \subset \mathbb{R}^2$ live in the triangular cone TC in Fig. 6. The triangular projection $\mathrm{TP} : \mathrm{TC} \to \mathrm{QT}$ divides each coordinate by the *size* $\sigma(\Lambda) = r_{12} + r_{01} + r_{02}$ and projects $\mathrm{RI}(\Lambda)$ to $(\bar{r}_{12}, \bar{r}_{01}, \bar{r}_{02})$ in the quotient triangle QT in Fig. 7. This triangle can be visualised as the isosceles right-angled triangle $\mathrm{QT} = \{x, y \geq 0,\ x + y \leq 1\} \subset \mathbb{R}^2$ parameterised by $x = \bar{r}_{02} - \bar{r}_{01}$ and $y = 3\bar{r}_{12}$ The resulting pair $\mathrm{PI}(\Lambda) = (x, y)$ is the *projected invariant*. The oriented invariant $\mathrm{PI}^o(\Lambda)$ is obtained by adding the superscript $\mathrm{sign}(\Lambda)$. ■

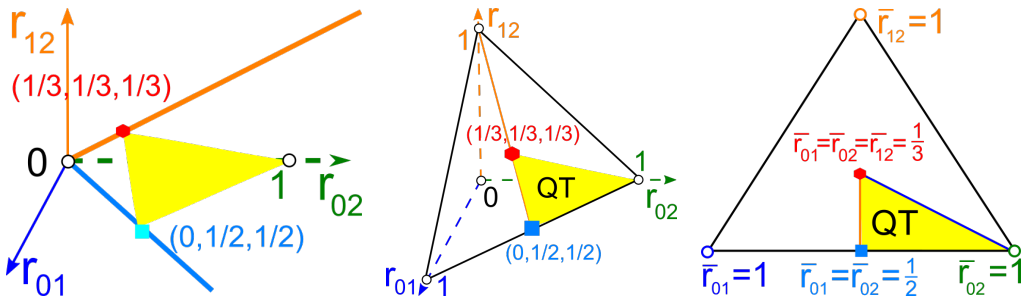

Fig. 6. **Left**: the triangular cone $\mathrm{TC} = \{(r_{12}, r_{01}, r_{02}) \in \mathbb{R}^3 \mid 0 \le r_{12} \le r_{01} \le r_{02} \ne 0\}$ is the space of all root invariants, see Definition 3.1. **Middle**: TC projects to the quotient triangle QT representing all 2D lattices up to isometry and uniform scaling. **Right**: QT is parameterised by $x = \bar{r}_{02} - \bar{r}_{01} \in [0, 1)$ and $y = 3\bar{r}_{12} \in [0, 1]$.

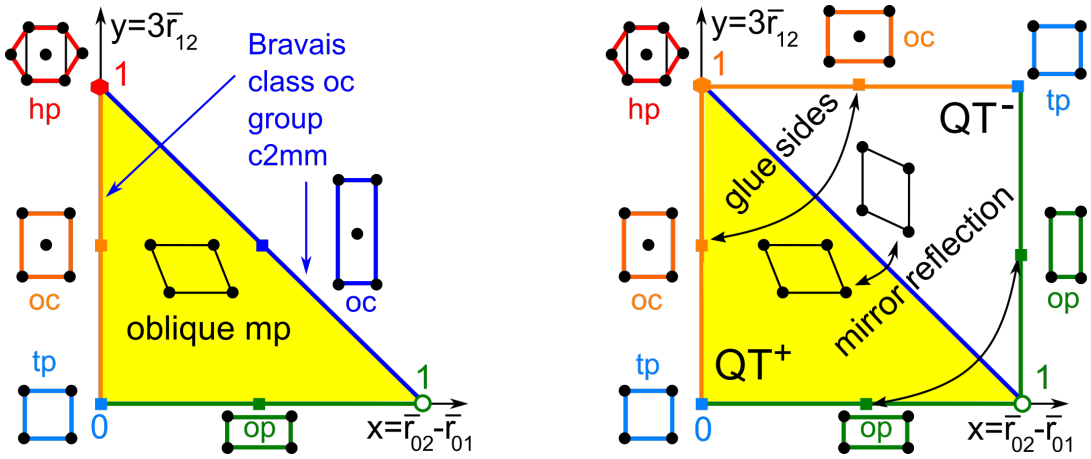

Fig. 7. **Left**: all projected invariants $\mathrm{PI}(\Lambda)$ live in the quotient triangle QT parameterised by $x = \bar{r}_{02} - \bar{r}_{01} \in [0, 1)$ and $y = 3\bar{r}_{12} \in [0, 1]$. **Right**: mirror images (enantiomorphs) of any oblique lattice are represented by a pair $(x, y) \leftrightarrow (1 - y, 1 - x)$ in the quotient square $\mathrm{QS} = \mathrm{QT}^+ \cup \mathrm{QT}^-$ symmetric in the diagonal $x + y = 1$.

All oriented projected invariants $\mathrm{PI}^o(\Lambda)$ with $\mathrm{sign}(\Lambda)$ live in a union of two quotient triangles $\mathrm{QT}^+ \cup \mathrm{QT}^-$. These triangles should be glued along the common subspace of mirror-symmetric latices (all non-oblique lattices $\Lambda \subset \mathbb{R}^2$), whose $\mathrm{PI}(\Lambda)$ belong to the boundary of QT. Fig. 7 (right) glues the hypotenuses of $\mathrm{QT}^{\pm}$ and indicates how to glue the remaining sides. We get a punctured sphere due to the excluded vertex $(1, 0)$.

Example 3.3 (subspaces of Bravais classes in QT)**. (tp)** The square lattice $\Lambda_4 \subset \mathbb{R}^2$ with a unit cell $a \times a$ has $\mathrm{RI}(\Lambda_4) = (0, a, a) \in \mathrm{TC}$ projected by TP to $(\bar{r}_{12}, \bar{r}_{01}, \bar{r}_{02}) = (0, \frac{1}{2}, \frac{1}{2})$. By Definition 3.2 the projected invariant $\mathrm{PI}(\Lambda_4) = (x, y) = (\bar{r}_{02} - \bar{r}_{01}, 3\bar{r}_{12}) = (0, 0) \in \mathrm{QT}$, see Fig. 7 (left). So the Bravais class (tp) of all square (tetragonal) lattices $\Lambda_4 \subset \mathbb{R}^2$ is represented by the bottom-left vertex $(0, 0)$ in the quotient triangle QT, identified with the top-right vertex of the quotient square QS in Fig. 7 (right).

(hp) The hexagonal lattice Λ_6 with a minimum inter-point distance a has the root invariant $\mathrm{RI}(\Lambda_6) = (\frac{a}{\sqrt{2}}, \frac{a}{\sqrt{2}}, \frac{a}{\sqrt{2}})$ projected by TP to $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. The projected invariant is $\mathrm{PI}(\Lambda_6) = (x, y) = (0, 1) \in \mathrm{QT}$, see Fig. 7 (left). The Bravais class (hp) of all hexagonal lattices $\Lambda_6 \subset \mathbb{R}^2$ is represented by the top-left vertex $(0, 1)$ in the quotient triangle QT, identified with the bottom-right vertex of the quotient square QS.

(op) Any rectangular lattice Λ with a unit cell $a \times b$ for $0 < a < b$ has the obtuse superbase $v_1 = (a, 0)$, $v_2 = (0, b)$, $v_0 = (-a, -b)$, see Fig. 8 (left). Then $\mathrm{RI}(\Lambda) = (0, a, b)$ and $\mathrm{PI}(\Lambda) = (\frac{b-a}{b+a}, 0)$ belongs to the horizontal side of QT, which represents the Bravais class (op). We approach the excluded vertex $(1, 0)$ as $b \to +\infty$.

(oc) Any centred rectangular lattice Λ with a conventional unit cell $2a \times 2b$ for $0 < a < b$ has the obtuse superbase $v_1 = (2a, 0)$, $v_2 = (-a, b)$, $v_0 = (-a, -b)$, see Fig. 8. Then $r_{01} = a\sqrt{2} = r_{02}$ and $r_{12} = \sqrt{b^2 - a^2}$. If $b \leq a\sqrt{3}$, then $\mathrm{RI}(\Lambda) = (\sqrt{b^2 - a^2}, a\sqrt{2}, a\sqrt{2})$ and $\mathrm{PI}(\Lambda) = (0, \frac{3\sqrt{b^2-a^2}}{2a\sqrt{2}+\sqrt{b^2-a^2}})$ belongs to the vertical (orange) edge of QT. This vertical edge is the shortest straight-line path between the vertices $(x, y) = (0, 0)$ representing the tetragonal and hexagonal Bravais classes, where $a = b$ and $b = a\sqrt{3}$, respectively. Hence the subspace of centred rectangular lattices for $b \leq a\sqrt{3}$ can be considered as having the symmetries of both hexagonal and square lattices. If $b > a\sqrt{3}$, then $\mathrm{RI}(\Lambda) = (a\sqrt{2}, a\sqrt{2}, \sqrt{b^2 - a^2})$ and $\mathrm{PI}(\Lambda) = (\frac{3a\sqrt{2}}{2a\sqrt{2}+\sqrt{b^2-a^2}}, \frac{\sqrt{b^2-a^2}-a\sqrt{2}}{2a\sqrt{2}+\sqrt{b^2-a^2}})$ belongs to the hypotenuse $x + y = 1$ of the triangle QT. The open vertical edge and open hypotenuse of QT represent the Bravais class *oc* of all centred rectangular lattices. ■

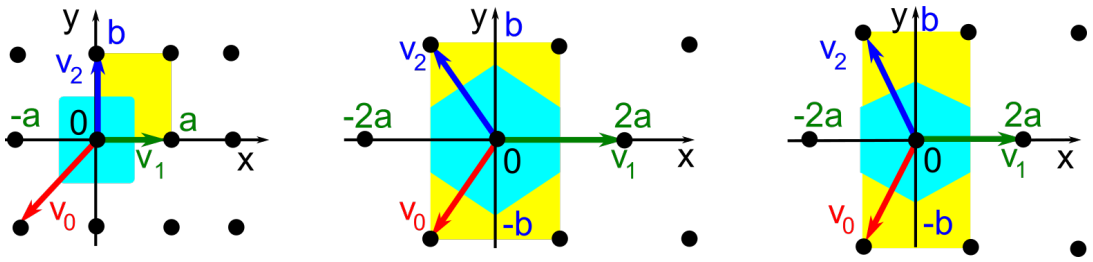

Fig. 8. **Left**: any rectangular lattice Λ with a unit cell $a \times b$ has the obtuse superbase B with $v_1 = (a, 0)$, $v_2 = (0, b)$, $v_0 = (-a, -b)$, see Example 3.3(op). Other lattices Λ have a rectangular cell $2a \times 2b$ and an obtuse superbase B with $v_1 = (2a, 0)$, $v_2 = (-a, b)$, $v_0 = (-a, -b)$. **Middle**: $\mathrm{RI}(\Lambda) = (\sqrt{b^2 - a^2}, a\sqrt{2}, a\sqrt{2})$, $a \le b \le a\sqrt{3}$. **Right**: $\mathrm{RI}(\Lambda) = (a\sqrt{2}, a\sqrt{2}, \sqrt{b^2 - a^2})$, $a\sqrt{3} \le b$, see Example 3.3(oc).

The companion paper (Kurlin, 2022*b*) proves the following classifications of 2D lattices up to four equivalences, fulfilling the invariance and completeness conditions.

Theorem 3.4 (proved in Kurlin (2022*b*, Thm 4.2, Cor 4.6))**.** For a lattice $\Lambda \subset \mathbb{R}^2$,

(a) the invariant $\mathrm{RI}(\Lambda)$ uniquely identifies Λ up to isometry,

(b) the invariant $\mathrm{RI}^o(\Lambda)$ uniquely identifies Λ up to rigid motion,

(c) the invariant $\mathrm{PI}(\Lambda)$ uniquely identifies Λ up to isometry and uniform scaling,

(d) the invariant $\mathrm{PI}^o(\Lambda)$ uniquely identifies Λ up to rigid motion and uniform scaling.

Each part in Theorem 3.4 can be rephrased as a two-directional criterion. For example, part **(a)**: any lattices $\Lambda, \Lambda' \subset \mathbb{R}^2$ are isometric if and only if $\mathrm{RI}(\Lambda) = \mathrm{RI}(\Lambda')$. The first (*only if*) direction means that if $\Lambda \cong \Lambda'$ are isometric, then $\mathrm{RI}(\Lambda) = \mathrm{RI}(\Lambda')$, so $\mathrm{RI}(\Lambda)$ is an isometry invariant taking the same value on all isometric lattices. The second (*if*) direction means that if $\mathrm{RI}(\Lambda) = \mathrm{RI}(\Lambda')$, then $\Lambda \cong \Lambda'$ are isometric.

4. Mapping millions of 2-dimensional lattices extracted from CSD crystals

For any periodic crystal from the Cambridge Structural Database (CSD), which has full geometric data of its lattice $\Lambda \subset \mathbb{R}^3$, we extract three 2-dimensional lattices

generated by three pairs $\{v_2, v_3\}$, $\{v_1, v_3\}$, $\{v_1, v_2\}$ of given basis vectors of Λ. So the CSD provides a huge collection of 2.6 million 2-dimensional lattices, which our reduction approach maps to the triangle QT in under an hour on a standard laptop.

Fig. 9 shows all resulting 2.6 million lattices in QT. Only about 55% of all lattices have Bravais classes oc, op, hp, tp. The remaining 45% of lattices are oblique, with Bravais class mp. These occupy almost the full quotient square QT, although we see a somewhat greater density close to subspaces representing higher symmetry lattices - especially around hexagonal and rectangular centered lattices.

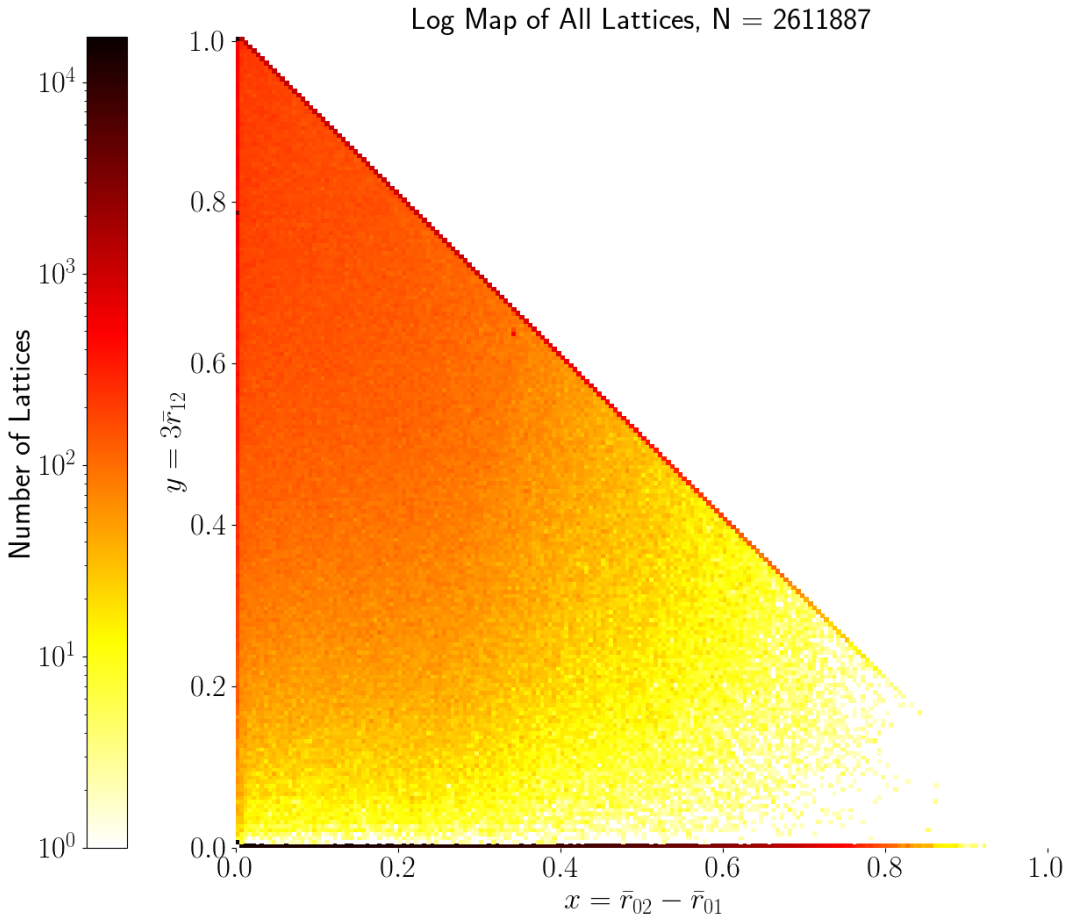

Fig. 9. The heat map in QT of all 2D lattices extracted from 870K+ CSD crystals. The colour of each pixels indicates (on the logarithmic scale) the number of lattices whose projected invariant $\mathrm{PI}(\Lambda) = (x, y) = (\bar{r}_{02} - \bar{r}_{01}, 3\bar{r}_{12})$ belongs to this pixel. The darkest pixels represent rectangular lattices on the bottom edge of QT.

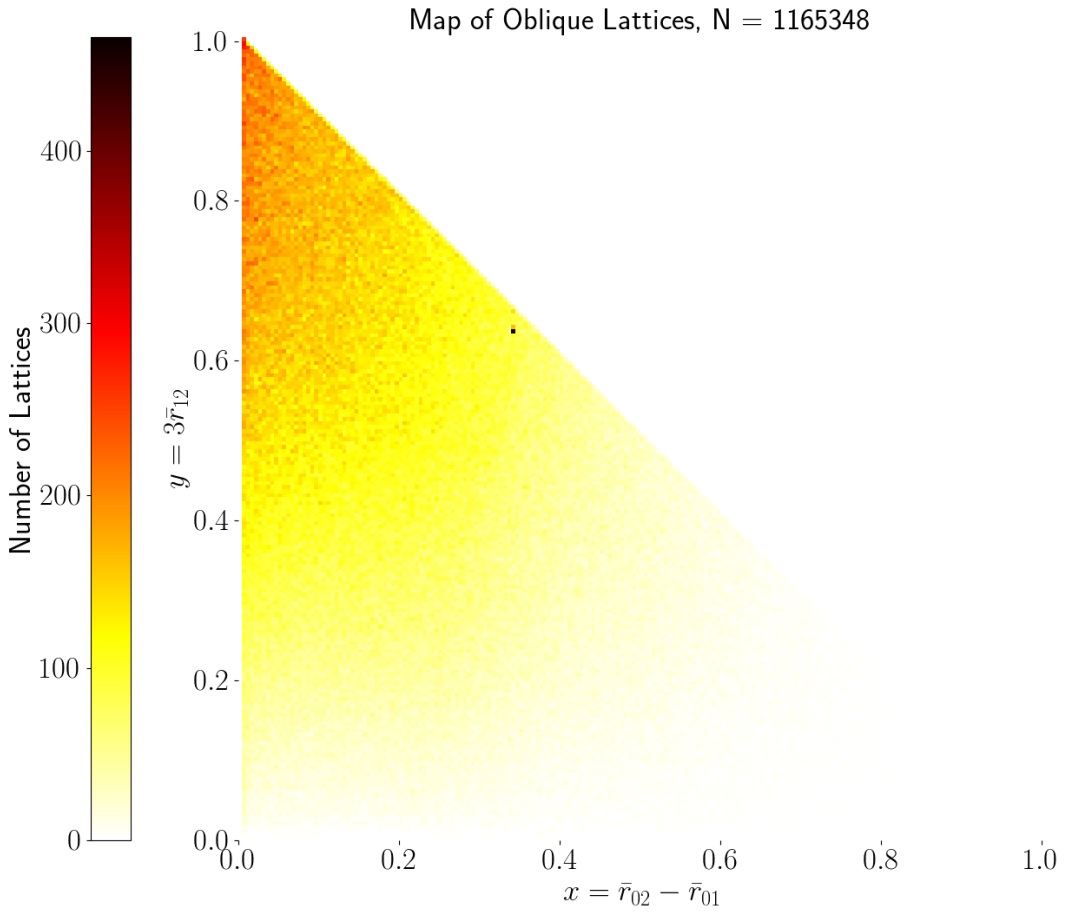

Fig. 10. The normal-scale heat map in QT of all 2D oblique lattices from CSD crystals. After removing mirror-symmetric lattices on the boundary of QT, we can better see the tendency towards hexagonal lattices at the top left corner $(0, 1) \in$ QT.

The gap of about two pixels near the horizontal edge in Fig. 9 corresponds to $\bar{r}_{12} = 0.01$. The relevant lattices have basis vectors v_1, v_2 whose angle is perturbed from 90° by less than 0.03°. The CSD has only 399 such lattices and $\bar{r}_{12} > 0.005$ for all but one of them. After removing all non-oblique lattices represented by root invariants along the boundary of QT, the map in Fig. 10 shows more clearly that all oblique lattices extracted from the CSD occupy the triangle QT without any gaps.

The heat map of rectangular lattices in Fig. 11 (left) has two high-concentration (black) pixels at $a \approx 3.5$ Angstroms arising from 386 near-identical primitive mon-

oclinic crystals of α-oxalic acid dihydrate. This molecule was used as a benchmark for the calculation of electron densities since its crystallographic properties were thoroughly documented in (Stevens & Coppens, 1980). So hundreds of publications have since generated and deposited further refinements of its structural determination.

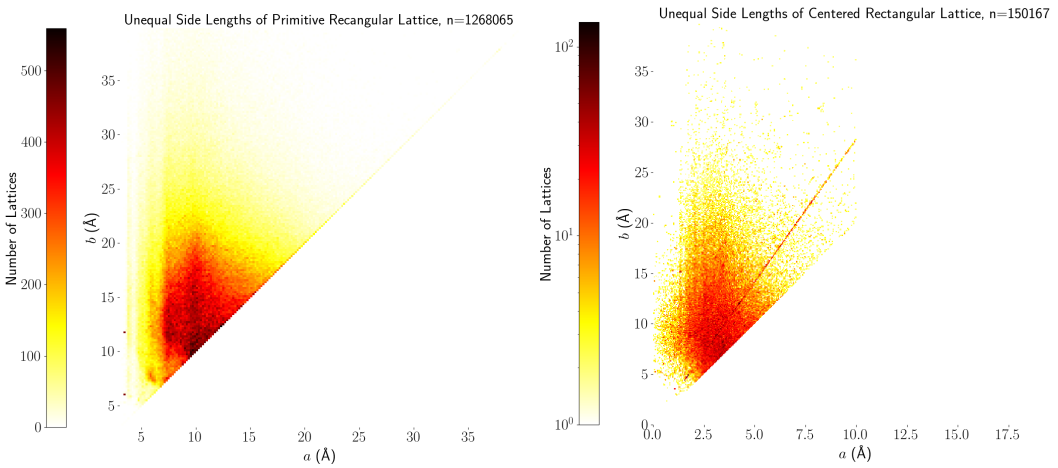

Fig. 11. Heat maps of parameters (a, b) in $\mathring{A}$. **Left**: rectangular lattices with primitive unit cells $a \times b$ in $N = 1268065$ crystals in the CSD. **Right**: centred-rectangular lattices with conventional cells $2a \times 2b$ in $N = 150167$ crystals in the CSD.

In the heat map of centered rectangular lattices in Fig. 11 (right), the most prominent feature is the hottest area in the region where the shortest side length is between $2.5\mathring{A}$ and $5\mathring{A}$. We also see a visible line $b = \sqrt{2}a$ of high-concentration pixels. This line represents 2D lattices in body-centered cubic lattices, where the ratio of side lengths is $\sqrt{2}$. This ratio was reported among preferred values for lattice length ratios in dimension 3 by de Gelder & Janner (2005). Another high-concentration pixel represents 130 structures of a standard test molecule (hexamethylenetetramine), which was frequently used in the investigation of lattice vibrations (Becka & Cruickshank, 1963).

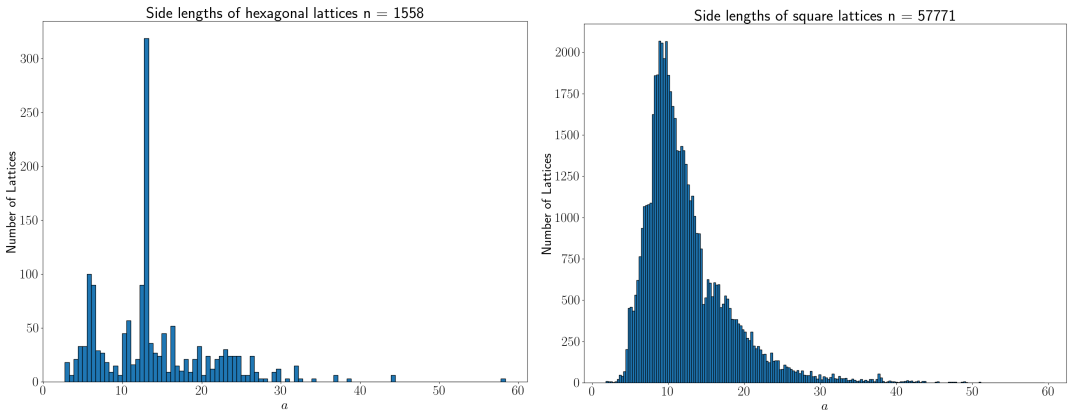

Fig. 12. The histograms of minimum inter-point distances a in Angstroms.

Hexagonal and square lattices are characterised by the inter-point distance a. Fig. 12 shows distributions and prefered values of a (in Angstroms) among CSD lattices.

5. Other complete invariants and a spherical map of 2-dimensional lattices

In comparison with other complete invariants, $\mathrm{RI}(\Lambda)$ has the advantage of homogeneity so that any permutation σ of (indices of) superbase vectors v_0, v_1, v_2 permutes the three root products accordingly: $r_{ij} \mapsto r_{\sigma(i)\sigma(j)}$. The metric tensor $\mathrm{MT} = (v_1^2, v_1 \cdot v_2, v_2^2)$ including the coefficients of the form $Q_\Lambda(x,y) = q_{11}x^2 + 2q_{12}xy + q_{22}y^2$ representing Λ is not homogeneous in the above sense. Taking square roots gives the *quadratic invariant* $\mathrm{QI}(\Lambda) = (\tau_{11}, \tau_{12}, \tau_{22}) = (\sqrt{q_{11}}, \sqrt{-q_{12}}, \sqrt{q_{22}})$ in the units of basis coordinates. The quadratic invariant $\mathrm{QI}(\Lambda)$ is complete up to isometry by Theorem 3.4(a).

In the isosceles triangle QT, continuous metrics and chiral distances have simple formulae in Kurlin (2022*b*, sections 5-6) for the coordinates $x = \bar{r}_{02} - \bar{r}_{01}$, $y = 3\bar{r}_{12}$ but can be now re-written for any coordinates on $\mathrm{LIS}(\mathbb{R}^2)$, see the earlier non-isosceles triangles in Engel *et al.* (2004, Fig. 1.2 on p. 82) and Zhilinskii (2016, Fig. 6.2).

Since the quotient square $\mathrm{QS} = \mathrm{QT}^+ \cup \mathrm{QT}^-$ with identified sides is a punctured sphere, it is natural to visualise QS as the round surface of Earth with $\mathrm{QT}^\pm$ as the

north/south hemispheres separated by the equator along their common boundary of QT represented by projected invariants PI(Λ) of all mirror-symmetric lattices Λ.

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 $90°, 45°, 45°$. The incentre of QT^+ has the coordinates (x, x), where $x = 1 - \frac{1}{\sqrt{2}} = \frac{1}{2+\sqrt{2}}$. The lattice Λ_2^+ with the projected invariant $\mathrm{PI}(\Lambda_2^+) = (x, x)$ has the basis $v_1 \approx (1.9, 0)$, $v_2 \approx (-0.18, 3.63)$ inversely designed in Kurlin (2022*b*, Example 4.10 (Λ_2)).

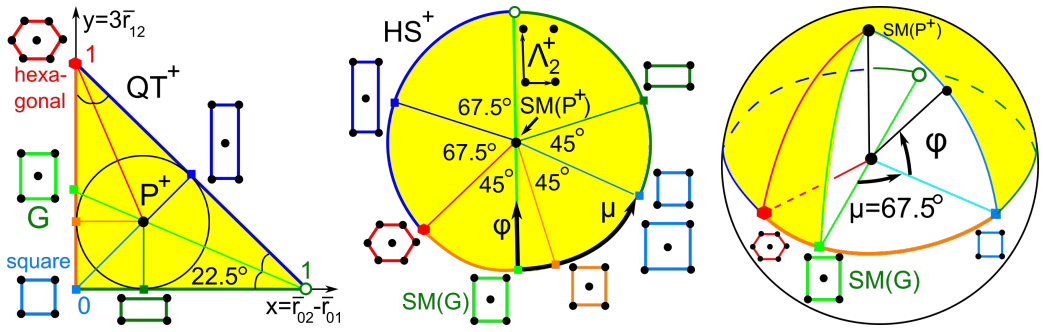

Fig. 13. **Left**: in QT^+, the Greenwich line goes from the 'empty' point (1,0) through incentre P^+ to the point $G = (0, \sqrt{2} - 1)$. **Middle**: the hemisphere HS^+ has the north pole at P^+, the equator $\partial\mathrm{QT}^+$ of mirror-symmetric lattices. **Right**: the longitude $\mu \in (-180°, +180°]$ anticlockwise measures angles from the Greenwich line, the latitude $\varphi \in [-90°, +90°]$ measures angles from the equator to the north.

Definition 5.1 (spherical map $\mathrm{SM} : \mathrm{QS} \to S^2$)**. (a)** The *spherical map* SM sends the incentre P^+ of QT to the north pole of the hemisphere HS^+ and the boundary ∂QT to the equator of HS^+, see Fig. 13 (middle). Linearly map the line segment between P^+ and any point (x, y) in the boundary ∂QT to the shortest arc connecting the north pole $\mathrm{SM}(P^+)$ to $\mathrm{SM}(x, y)$ in the equator of HS^+. Extend the *spherical map* to $\mathrm{SM} : \mathrm{QS} \to S^2$ by sending any pair of invariants $\mathrm{PI}^o(\Lambda^{\pm})$ with $\mathrm{sign}(\Lambda^{\pm}) = \pm 1$ to the northern/southern hemispheres of the 2-dimensional sphere S^2, respectively.

(b) For any lattice $\Lambda \subset \mathbb{R}^2$, the *latitude* $\varphi(\Lambda) \in [-90°, +90°]$ is the angle from the

equatorial plane EP of S^2 to the radius-vector to the point $\mathrm{SM}(\mathrm{PI}^o(\Lambda)) \in S^2$ in the upwards direction. Let $v(\Lambda)$ be the orthogonal projection of this radius-vector to EP. Define the *Greenwich* point as $G = (0, \sqrt{2}-1) \in \partial\mathrm{QT}$ in the line through P^+ and $(1,0)$. This G represents all centered rectangular lattices with a conventional unit cell $2a \times 2b$ whose ratio $r = \frac{b}{a}$ can be found from Example 3.3: $\sqrt{2}-1 = \frac{3\sqrt{b^2-a^2}}{2a\sqrt{2}+\sqrt{b^2-a^2}}$. 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 the sphere S^2 passing through the point $\mathrm{SM}(G)$ in the equator E. The longitude $\mu(\Lambda) \in (-180°, 180°]$ is the anticlockwise angle from the *Greenwich plane* through the Greenwich meridian to the vector $v(\Lambda)$ above. ■

For lattices with $\mathrm{PI}(\Lambda)$ in the straight-line segment between the excluded vertex $(1,0)$ and the incentre P^+, we choose the longitude $\mu = +180°$ rather than $-180°$. Proposition 5.2 computes $\mu(\Lambda), \varphi(\Lambda)$ via $\mathrm{PI}(\Lambda) = (x,y)$ and is proved in the appendix.

Proposition 5.2 (formulae for SM)**.** For any lattice $\Lambda \subset \mathbb{R}^2$ with $\mathrm{PI}(\Lambda) = (x,y) \in \mathrm{QT}$, if $x \neq t = 1 - \dfrac{1}{\sqrt{2}}$, then set $\psi = \arctan\dfrac{y-t}{x-t}$, otherwise $\psi = \mathrm{sign}(y-t)90°$.

(5.2a) The longitude of the lattice Λ is $\mu(\Lambda) = \begin{cases} \psi + 22.5° \text{ if } x < t, \\ \psi - 157.5° \text{ if } x \geq t, \psi \geq -22.5°, \\ \psi + 202.5° \text{ if } x \geq t, \psi \leq -22.5°. \end{cases}$

(5.2b) The latitude is $\varphi(\Lambda) = \mathrm{sign}(\Lambda) \cdot \begin{cases} \frac{x\sqrt{2}}{\sqrt{2}-1}90° \text{ if } \mu(\Lambda) \in [-45°, +67.5°], \\ \frac{y\sqrt{2}}{\sqrt{2}-1}90° \text{ if } \mu(\Lambda) \in [+67.5°, +180°], \\ \frac{1-x-y}{\sqrt{2}-1}90° \text{ if } \mu(\Lambda) \in [-180°, -45°]. \end{cases}$

The incentres $P^\pm \in \mathrm{QT}^\pm$ have $\psi = 0$ and $\varphi = \pm 90°$, respectively, μ is undefined. ■

Example 5.3 (prominent lattices)**.** Any mirror-symmetric lattice $\Lambda \subset \mathbb{R}^2$ has $\mathrm{sign}(\Lambda) = 0$, hence belongs to the equator E of S^2 and has $\varphi(\Lambda) = 0$ by (5.2b). Any square lattice Λ_4 with $\mathrm{PI}(\Lambda_4) = (0,0)$ has $\mu(\Lambda_4) = \arctan 1 + 22.5° = 67.5°$ by (5.2a). Any hexagonal lattice Λ_6 with $\mathrm{PI}(\Lambda_4) = (0,1)$ has $\mu(\Lambda_4) = \arctan\frac{1}{1-\sqrt{2}} + 22.5° = -45°$. Any

rectangular lattice Λ with $\mathrm{PI}(\Lambda) = (1 - \frac{1}{\sqrt{2}}, 0)$ has $\mu(\Lambda) = -90° + 202.5° = 112.5°$. Any centered rectangular lattice Λ with $\mathrm{PI}(\Lambda) = (\frac{1}{2}, \frac{1}{2})$ at the mid-point of the diagonal of QT has $\mu(\Lambda) = \arctan 1 - 157.5° = -112.5°$. Any *Greenwich* lattice Λ_G with $\mathrm{PI}(\Lambda_G) = G = (0, \sqrt{2} - 1)$ has $\mu(\Lambda_G) = \arctan(1 - \sqrt{2}) + 22.5° = 0$. ■

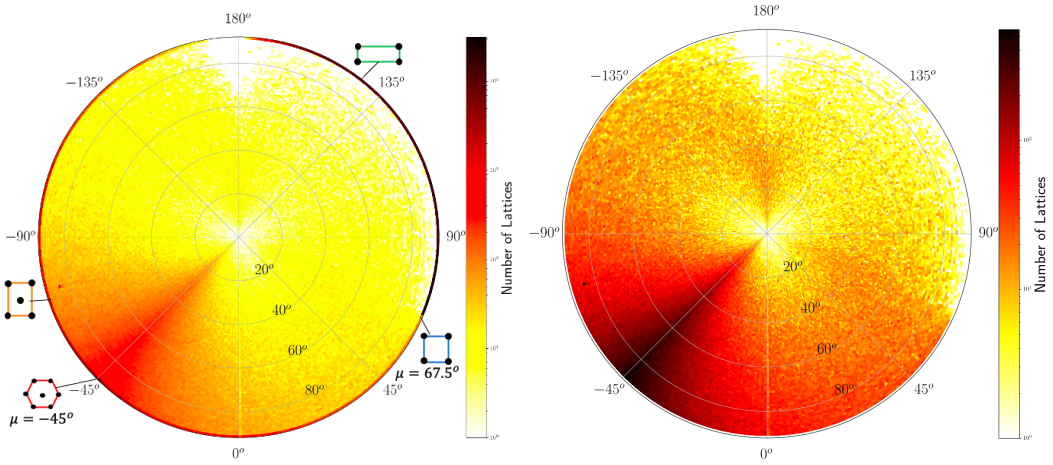

Fig. 14. The heat map of 2D lattices from CSD crystals on the northern hemisphere. The radial distance is the latitude $\phi \in [0°, 90°]$. **Left**: all $N = 2191887$ lattices with $\mathrm{sign}(\Lambda) \geq 0$, $\phi \geq 0$. **Right**: all $N = 741105$ oblique lattices with $\mathrm{sign}(\Lambda) > 0$, $\phi > 0$.

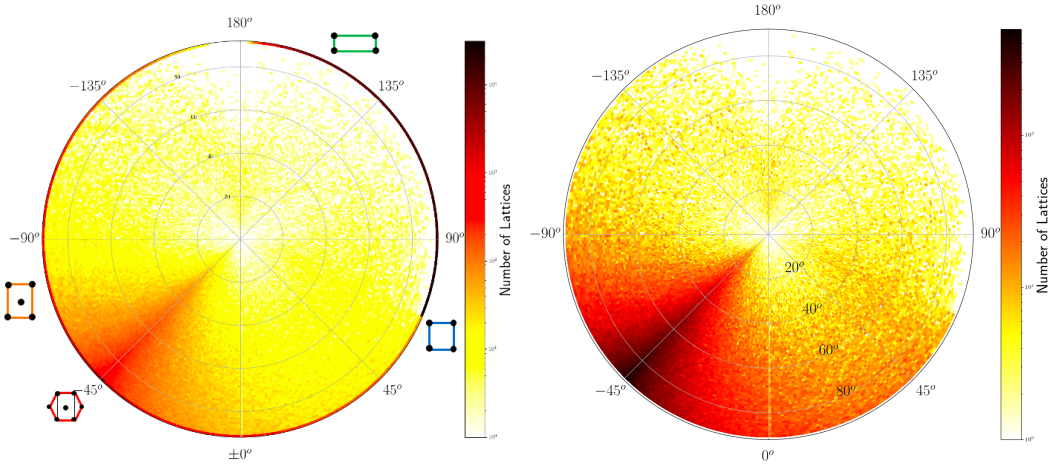

Fig. 15. The heat map of 2D lattices from CSD crystals on the northern hemisphere. The radial distance is the latitude $\phi \in [0°, 90°]$. **Left**: all $N = 1854209$ lattices with $\mathrm{sign}(\Lambda) \leq 0$, $\phi \leq 0$. **Right**: all $N = 406930$ oblique lattices with $\mathrm{sign}(\Lambda) < 0$, $\phi < 0$.

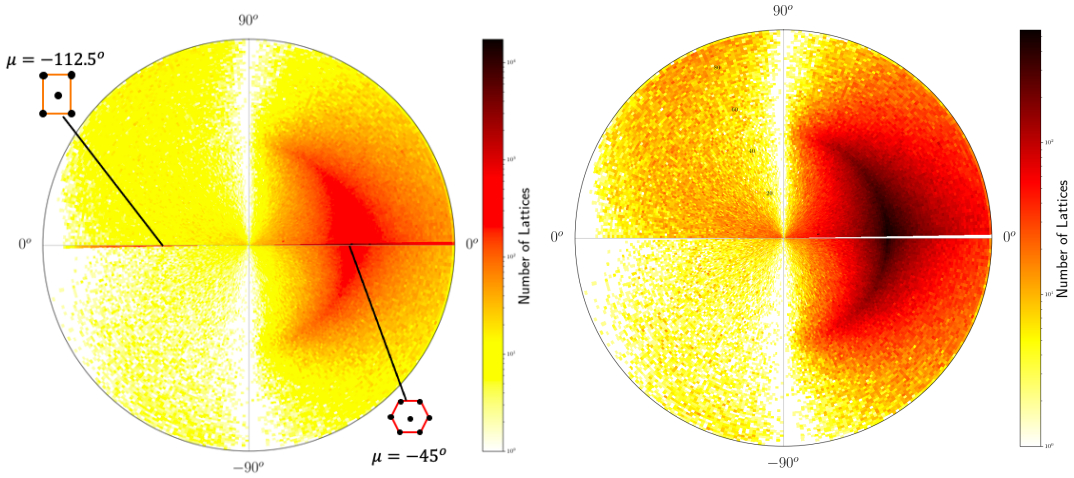

Fig. 16. The heat map of 2D lattices from CSD crystals on the western hemisphere. Angles on the circumference show the latitude $\phi \in [-90°, 90°]$. **Left**: $N = 1100580$ lattices with $\mu \in (-180°, 0°]$. The hexagonal lattice at $\mu = -45°$ and the centered rectangular lattice at $\mu = -112.5°$ are marked on the horizontal arc (western half-equator). **Right**: all $N = 932626$ oblique lattices with $\mu \in (-180°, 0°]$ and $\phi \neq 0$.

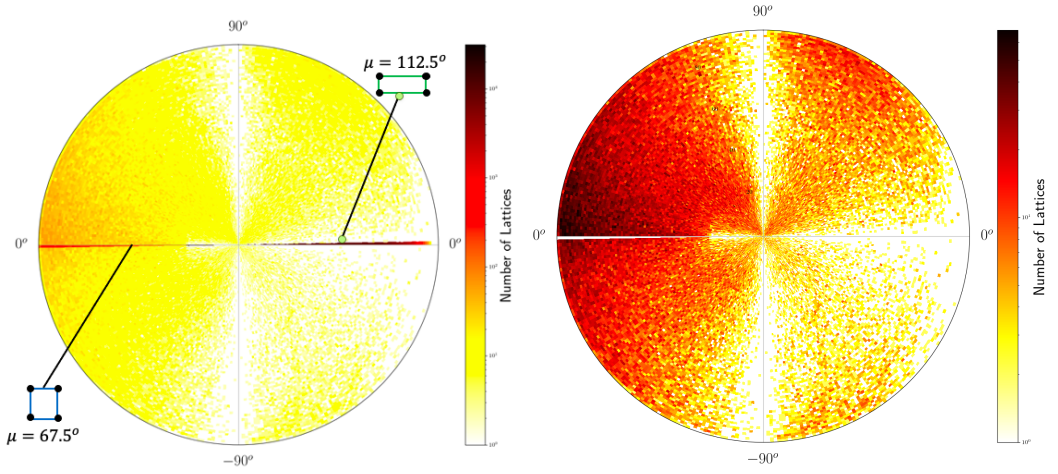

Fig. 17. The heat map of 2D lattices from CSD crystals on the eastern hemisphere. Angles on the circumference show the latitude $\phi \in [-90°, 90°]$. **Left**: all $N = 1511307$ lattices with $\mu \in [0°, 180°)$, the square lattice point at $\mu = 67.5°$ and the rectangular lattice at $\mu = 112.5°$ are marked on the horizontal arc (eastern half-equator). **Right**: all $N = 215409$ oblique lattices with $\mu \in [0°, 180°)$, $\phi \neq 0$.

The north pole represents the incentre P^+ whose pixel contains 230 lattices in Figure 10 but appears sparsely populated in Fig. 14 because this incentre pixel is

split into many 1×1 degree curved 'pixels' of a much lower concentration. The high concentration near the point representing hexagonal lattices is visible in Figures 14, 15 as dark pixels near the longitude $\mu = -45°$. Where non-oblique lattices are included, we see the high concentrations along the borders of QT, with primitive rectangular lattices appearing as a dark thick arc on the equator for $\mu \in [67.5°, 180°)$.

The heat maps show a hexagonal 'ridge' along the meridional arc at $\mu = -45°$ in Figures 14 and 15, which appears as a round arc in Figures 16 and 17. The concentration of exact square and rectangular lattices is even higher (dark pixels for the Bravais classes tp and op), but there are fewer lattices close to these classes possibly because manual or automatic adjustments are easier for angles close to 90° than to 60°.

6. Main conclusions and motivations for a continuous crystallography

The heat maps in Fig. 9-10 and 14-17 visualise for the first time 2.6 million 2-dimensional lattices in real crystals from the Cambridge Structural Database. The preprint (Bright *et al.*, 2021) extends this approach to 3-dimensional lattices, but there is a growing database of real and theoretical two dimensional lattice structures with potentially interesting properties (Mounet *et al.*, 2020) for which $2D$ lattice invariants may have direct utility. The maps indicate that lattices occur naturally in continuous distributions, and their geometry can be investigated by continuous invariant-based classification in addition to using discrete symmetry groups. The continuous approach has the added advantage of more easily spotting structures that are geometrically nearly identical, but where small variances in crystallisation conditions have led to slight structure perturbations which disrupt higher lattice symmetries. The Python code for new invariants is at https://github.com/MattB-242/Lattice_Invariance.

Using a geographic analogue, the recent invariants create complete and continuous maps for efficient navigation in the Lattice Isometry Space $\mathrm{LIS}(\mathbb{R}^2)$, which can be

zoomed in as satellite images and explored at any desirable resolution. Since each invariant is a point in a space on which various metrics can be defined, this representation leads to a continuous 'distance' between two lattices based on their separation in LIS($\mathbb{R}^2$) and a continuous measure of 'dissymmetry' as the closest distance to the subspace corresponding to lattices with higher symmetry, see (Kurlin, 2022*b*). See also symmetries of densest packings in $\mathbb{R}^2$ in (Torda *et al.*, 2022*a*; Torda *et al.*, 2022*b*).

The four non-generic Bravais classes of 2D lattices are lower-dimensional subspaces in LIS($\mathbb{R}^2$) whose separate maps in Fig. 11 and 12 have no intermediate gaps and include sparse or empty regions only for small or very large values of cell parameters.

Using a biological analogue, crystallography previously took a similar approach to the classical taxonomy, dividing lattices into an increasingly complex sequence of discrete categories based on symmetries as they divided organisms according to physical characteristics, see a comprehensive review in (Nespolo *et al.*, 2018).

The new area of *continuous crystallography* uses the geometric properties of the lattice itself to continuously classify an individual lattice in as granular a manner as we like, in a manner akin to the modern use of genetic sequences and markers to classify organisms. Indeed, since the root invariant RI(Λ) of a lattice Λ is complete, this RI(Λ) could be said to represent the DNA of Λ. Even better than the real DNA, any 2D lattice can be explicitly built up from RI(Λ), see Kurlin (2022*b*, Proposition 4.9).

The complete root invariant from Definition 3.1 extends to a 3D lattice as follows. For any 3D lattice, depending on its Voronoi domain, all obtuse superbases $\{v_i\}_{i=0}^3$ with $v_0+v_1+v_2+v_3=0$ are described in Kurlin (2022*a*, Lemmas 4.1-4.5). Any generic 3D lattices has a unique (up to isometry) obtuse superbase whose root products $r_{ij}=\sqrt{-v_i\cdot v_j}$ can be considered as labels on edges of a 3D tetrahedron or written in the matrix $\begin{pmatrix} r_{23} & r_{13} & r_{12} \\ r_{01} & r_{02} & r_{03} \end{pmatrix}$. Permutations of 4 superbase vectors induce $4!=24$

permutations of the above 6 root products. Other non-generic cases require other permutations, which were not previously considered by (Andrews *et al.*, 2019*b*), to guarantee a complete invariant of all 3D lattices in Kurlin (2022*a*, Theorem 6.3). Maps of 3D lattices extracted from CSD crystals appear in (Bright *et al.*, 2021).

Working towards a complete materials genome, Widdowson *et al.* (2022, section 7) introduced the Pointwise Distance Distribution (PDD). This PDD invariant distinguished all periodic point sets after a tiny perturbation. More than 200 billion pairwise comparisons of all 660K+ periodic crystals in the CSD over two days on a modest desktop detected five pairs of isometric duplicates, see Widdowson *et al.* (2022, section 7), where two crystals are geometrically identical to the last decimal place in all data including structure factors but one atom is replaced with a different one: Cd with Mn in the pair HIFCAB vs JEPLIA. These pairs are under investigation by five journals for data integrity. (Near-)duplicates in the CSD can be recognised only by a *continuous* invariant taking close values for close crystals. The CSD entries DEBXIT01,...,DEBXIT06 represent two polymorphs: four (near-)duplicates of T2-γ and two (near-)duplicates of T2-β reported in our past work (Pulido *et al.*, 2017). (Zhu *et al.*, 2022) predicted and synthesised new material based on PDD invariants.

Acknowledgements. We thank Mois Aroyo, Bernd Souvignier, and Nikolai Dolbilin for helpful discussions. This research was partially supported by the £3.5M EPSRC grant 'Application-driven Topological Data Analysis' (2018-2023, EP/R018472/1), and the last author's Royal Academy of Engineering Fellowship 'Data Science for Next Generation Engineering of Solid Crystalline Materials' (2021-2023, IF2122/186), and the EPSRC grant 'Inverse design of periodic crystals' (2022-2024, EP/X018474/1).

References

Andrews, L., Bernstein, H. & Pelletier, G. (1980). *Acta Cryst. A*, **36**(2), 248–252.
Andrews, L. C. & Bernstein, H. J. (1988). *Acta Cryst. A*, **44**(6), 1009–1018.
Andrews, L. C. & Bernstein, H. J. (2014). *J Applied Cryst.* **47**(1), 346–359.

Andrews, L. C., Bernstein, H. J. & Sauter, N. K. (2019*a*). *Acta Cryst. A*, **75**(1), 115–120.

Andrews, L. C., Bernstein, H. J. & Sauter, N. K. (2019*b*). *Acta Cryst. A*, **75**(3), 593–599.

Anosova, O. & Kurlin, V. (2021*a*). *arxiv:2103.02749*.

Anosova, O. & Kurlin, V. (2021*b*). In *Lecture Notes in Computer Science (Proceedings of DGMM)*, vol. 12708, pp. 229–241.

Anosova, O. & Kurlin, V. (2022*a*). *arxiv:2205.15298*.

Anosova, O. & Kurlin, V. (2022*b*). In *Lecture Notes in Computer Science (Proceedings of DGMM)*, vol. 13493.

Aroyo, M. I. (2016). *International Tables for Crystallography*. Wiley Online Library.

Becka, L. N. & Cruickshank, W. J. (1963). *Proceedings of the Royal Society A*, **273**, 435–454.

Bernstein, H. J., Andrews, L. C. & Xerri, M. (2022). *arxiv:2208.04152*.

Bright, M., Cooper, A. I. & Kurlin, V. (2021). *arxiv:2109.11538*.

Conway, J. H. & Sloane, N. J. (1992). *Proceedings of the Royal Society A*, **436**(1896), 55–68.

Delaunay, B., Galiulin, R., Dolbilin, N., Zalgaller, V. & Stogrin, K. (1973). In *Dokl. Akad. Nauk SSSR*, vol. 209, pp. 309–313.

Delaunay, B., Padurov, N. & Aleksandrov, A. (1934). *Mathematical Foundations of Structural Analysis of Crystals*. State Technical-Theoretical Press, USSR.

Delaunay, B. N. (1937). *Uspekhi Matematicheskikh Nauk*, (3), 16–62.

Edelsbrunner, H., Heiss, T., Kurlin, V., Smith, P. & Wintraecken, M. (2021). In *Proceedings of SoCG*, pp. 32:1–32:16.

Engel, P., Michel, L. & Sénéchal, M. (2004). *Lattice geometry*. Tech. Rep. IHES-P-2004-45. .

Lima-de Faria, J., Hellner, E., Liebau, F., Makovicky, E. & Parthé, E. (1990). *Acta Cryst A*, **46**(1), 1–11.

de Gelder, R. & Janner, A. (2005). *Acta Crystallographica Section B*, **61**(3), 296–303.

Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. (2016). *Acta Cryst B*, **72**(2), 171–179.

Křivỳ, I. & Gruber, B. (1976). *Acta Cryst. A*, **32**(2), 297–298.

Kurlin, V. (2022*a*). *arxiv:2201.10543*.

Kurlin, V. (2022*b*). *arxiv:2201.05150*.

Lagrange, J. L. (1773). *Nouveaux Mémoires de l'Académie de Berlin*.

McGill, K. J., Asadi, M., Karakasheva, M. T., Andrews, L. C. & Bernstein, H. J. (2014). *J Applied Cryst.* **47**(1), 360–364.

Mosca, M. & Kurlin, V. (2020). *Crystal Research and Technology*, **55**(5), 1900197.

Mounet, N., Gibertini, M. *et al.* (2020). *Materials Cloud Archive*.

Nespolo, M. (2008). *Acta Cryst A*, **64**(1), 96–111.

Nespolo, M., Aroyo, M. I. & Souvignier, B. (2018). *Journal of Applied Crystallography*, **51**(5), 1481–1491.

Niggli, P. (1928). *Krystallographische und strukturtheoretische Grundbegriffe*, vol. 1. Akademische verlagsgesellschaft mbh.

Pulido, A., Chen, L., Kaczorowski, T., Holden, D., Little, M., Chong, S., Slater, B., McMahon, D., Bonillo, B., Stackhouse, C., Stephenson, A., Kane, C., Clowes, R., Hasell, T., Cooper, A. & Day, G. (2017). *Nature*, **543**, 657–664.

Sacchi, P., Lusi, M., Cruz-Cabeza, A. J., Nauha, E. & Bernstein, J. (2020). *Cryst Eng Comm*, **22**(43), 7170–7185.

Selling, E. (1874). *Journal für die reine und angewandte Mathematik*, **77**, 143–229.

Smith, P. & Kurlin, V. (2022). In *Proceedings of International Symposium Visual Computing*.

Stevens, E. & Coppens, P. (1980). *Acta. Cryst. B*, **36**, 1864–1876.

Torda, M., Goulermas, J. Y., Kurlin, V. A. & Day, G. M. (2022*a*). *Physical Review E (arXiv:2207.08959)*.

Torda, M., Goulermas, J. Y., Púček, R. & Kurlin, V. A. (2022*b*). *arXiv:2202.11959*.

Widdowson, D. & Kurlin, V. (2022). *Advances in Neural Information Processing Systems (Proceedings of NeurIPS 2022), arxiv:2108.04798*, **35**.

Widdowson, D., Mosca, M., Pulido, A., Kurlin, V. & Cooper, A. (2022). *MATCH Communications in Mathematical and in Computer Chemistry*, **87**, 529–559.

Wigner, E. & Seitz, F. (1933). *Physical Review*, **43**(10), 804.

Zhilinskii, B. (2016). *Introduction to lattice geometry through group action*. EDP sciences.

Zhu, Q., Johal, J., Widdowson, D., Pang, Z., Li, B., Kane, C., Kurlin, V., Day, G., Little, M. & Cooper, A. I. (2022). *Journal of the American Chemical Society*, **144**, 9893–9901.

Zwart, P. H., Grosse-Kunstleve, R. W., Lebedev, A. A., Murshudov, G. N. & Adams, P. D. (2008). *Acta Cryst D*, **64**(1), 99–107.

Appendix A
A proof of Proposition 5.2, and plots of orientation-aware invariants.

Proof of Proposition 5.2. **(a)** For any point $P = (x, y) \in \mathrm{QT}$, the vector $\overrightarrow{P^+P}$ has coordinates $(x-t, y-t)$, where $P^+ = (t, t)$ is the incentre (the centre of the inscribed circle) of the quotient triangle QT and $t = 1 - \dfrac{1}{\sqrt{2}}$, see Fig. 13 (left). Recall that, for any $b \in \mathbb{R}$, the function $\arctan(b)$ outputs a unique angle $\alpha \in (-90°, 90°)$ such that $\tan(\alpha) = b$. If $x > t$, then $\psi = \arctan \dfrac{y-t}{x-t} \in (-90°, 90°)$ is the anticlockwise angle from the positive x-direction (with the origin at P^+) to the vector $\overrightarrow{P^+P}$.

For $x = t$, the limit values of arctan give $\psi = \mathrm{sign}(y-t)90°$. For $x < t$, the anticlockwise angle from the positive x-direction to $\overrightarrow{P^+P}$ is $\psi + 180°$. For example, the Greenwich vector $\overrightarrow{G}$ from the excluded vertex $(1, 0)$ to $G = (0, \sqrt{2}-1) \in \mathrm{QT}$ has the anticlockwise angle $\psi + 180° = 157.5°$ from the positive x-direction because

$$\frac{\sqrt{2}-1-t}{-t} = \frac{\sqrt{2}-1-(1-\frac{1}{\sqrt{2}})}{\frac{1}{\sqrt{2}}-1} = \frac{3-2\sqrt{2}}{1-\sqrt{2}} = 1-\sqrt{2} \text{ and } \arctan(1-\sqrt{2}) = -22.5°.$$

The anticlockwise angle from the x-axis to $\overrightarrow{P^+P}$ is $\alpha = \begin{cases} \psi \text{ if } x > t, \\ \psi + 180° \text{ if } x < t, \\ \mathrm{sign}(y-t)90° \text{ if } x = t, y \neq t. \end{cases}$

In all cases above, since the Greenwich vector $\overrightarrow{G}$ was chosen as the 0-th meridian, the anticlockwise angle from $\overrightarrow{G}$ to $\overrightarrow{P^+P}$ is the longitude $\mu = \alpha - 157.5$. For example, any centred rectangular lattice Λ with $\mathrm{PI}(\Lambda) = (x, y) = (\frac{1}{2}, \frac{1}{2})$ has $\psi = \arctan \dfrac{y-t}{x-t} =$

$\arctan 1 = 45^\circ = \alpha$ and longitude $\mu = \alpha - 157.5^\circ = -112.5^\circ$. If $\alpha - 157.5^\circ$ is outside the expected range of $\mu \in (-180^\circ, 180^\circ]$, we add or subtract 360°. Any hexagonal lattice Λ_6 with $\mathrm{PI}(\Lambda_6) = (0, 1)$ has $\psi = \arctan \dfrac{y-t}{x-t} = \arctan \dfrac{1-(1-\frac{1}{\sqrt{2}})}{\frac{1}{\sqrt{2}}-1} = \arctan \dfrac{1}{1-\sqrt{2}} = -67.5^\circ$, $\alpha = \psi + 180^\circ = 112.5^\circ$ and longitude $\mu = \alpha - 157.5^\circ = -45^\circ$. Any square lattice Λ_4 with $\mathrm{PI}(\Lambda_6) = (0, 0)$ has $\psi = \arctan \dfrac{y-t}{x-t} = \arctan 1 = 45^\circ$, $\alpha = \psi + 180^\circ = 225^\circ$ and longitude $\mu = \alpha - 157.5^\circ = 67.5^\circ$. Formula (5.2a) is split into three subcases only to guarantee the range of a longitude $\mu \in (-180^\circ, 180^\circ]$ for the anticlockwise angle $\alpha - 157.5^\circ$ from $\overrightarrow{G}$ to $\overrightarrow{P^+P}$, where α is computed above.

(b) For a fixed longitude $\mu(\Lambda)$, the projected invariant $\mathrm{PI}(\Lambda)$ varies along the line segment L at a fixed angle from the incentre P^+ to the boundary ∂QT. Formula (5.2b) is split into three subcases according to the three boundary edges of QT.

Consider the vertical edge between hexagonal and square lattices, where $\mu(\Lambda) \in [-45^\circ, 67.5^\circ]$. The latitude $\varphi(\Lambda)$ is proportional to the ratio in which the point $\mathrm{PI}(\Lambda) = (x, y)$ splits the line segment L from P^+ to the vertical edge. The endpoint $x = 0$ means that $\mathrm{SM}(\mathrm{PI}(\Lambda))$ is in the equator with $\varphi = 0$. The endpoint $x = t = 1 - \frac{1}{\sqrt{2}}$ means that $\mathrm{PI}(\Lambda) = P^+$ is in the centre whose image $\mathrm{SM}(P^+)$ is the north pole with $\varphi = 90^\circ$. The linear map between these extreme cases gives $\varphi(\Lambda) = \dfrac{x}{t} 90^\circ = \dfrac{x\sqrt{2}}{\sqrt{2}-1} 90^\circ$. The case of the horizontal edge of QT gives a similar φ after replacing x with y. The hypotenuse of QT, where $x + y = 1$, is also similar as the incentre $P^+ = (x, y) = (1 - \frac{1}{\sqrt{2}}, 1 - \frac{1}{\sqrt{2}})$ has the latitude $\varphi(\Lambda) = \dfrac{1-x-y}{\sqrt{2}-1} 90^\circ = \dfrac{1-2(1-\frac{1}{\sqrt{2}})}{\sqrt{2}-1} 90^\circ = 90^\circ$ as expected. The factor $\mathrm{sign}(\Lambda)$ in (5.2b) guarantees a symmetry of $\mathrm{SM} : \mathrm{QS} \to S^2$ in the equator. □

In the main body of this paper, we show heat maps of orientation unaware projected invariants, which clearly demonstrates the way that lattices generated from the CSD distribute through the lattice invariant space without gaps. Figure 18 shows plots of orientation aware projected invariants $PI^o(\Lambda)$ for the same dataset.

In both plots, we see an additional apparent non-smooth jump across the diagonal representing higher symmetry lattices, so that there is some apparent favouring of positive chirality among 2D lattices. This is an artefact of the interaction between vectors in the initial CSD data, and our consistently ordered selection of pairs from those vectors, and should not be read as a real physical effect. We also note that there is a much lower relative concentration, apparent from the lightness of the colour of each pixel, in the standard plot of oblique lattices. In this case the Oxalic acid structures mentioned in the main body of the paper all have consistent chirality and remain below the diagonal of the quotient square, while the lattices in any other pixel split between each half of the plot and therefore have much lower relative counts.

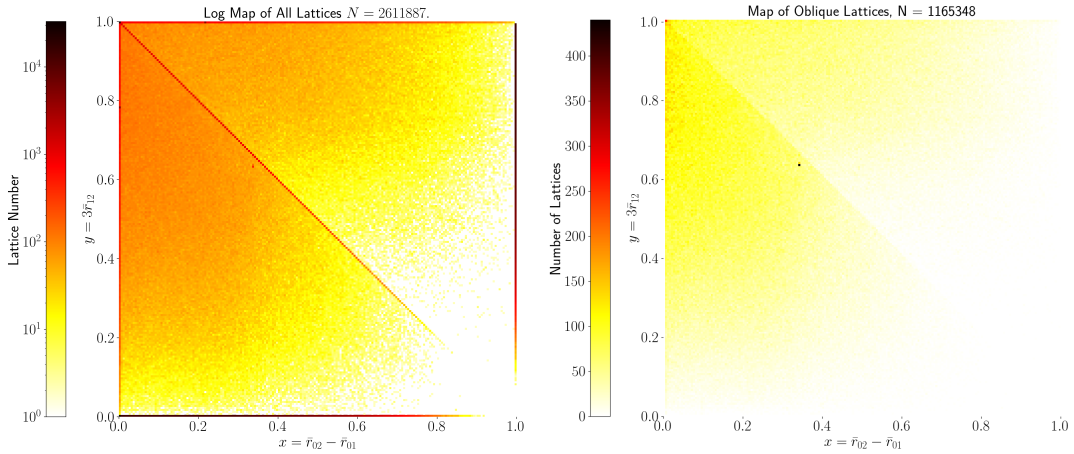

Fig. 18. Heat maps of 2D lattices derived from the CSD in the quotient square QS. Each pixel in the map represents a 0.005×0.005 interval of projected form invariant value, where each such value uniquely represents a lattice up to rigid motion only. **Left**: $N = 2611887$ lattices derived from the CSD. Projected Invariants for primitive and centered rectangular lattices are duplicated at the boundaries of the quotient square - indicative positions of non-trivially symmetric lattices are shown. **Right**: all $N = 1165348$ oblique (non-mirror-symmetric) lattices derived from the CSD.

Synopsis

Continuous invariant-based maps visualise for the first time all 2-dimensional lattices extracted from hundreds of thousands of known crystals in the Cambridge Structural Database.