AI in Crystallography: Mathematics exposes Artificial Illusions in material design

Tatiana Kurlina, University College London, tatiana.kurlina.25@ucl.ac.uk

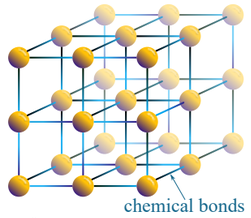
Introduction

Crystallography is the study of solid crystalline structures; one of the goals is to design new useful materials. AI has been generating a large amount of simulated crystals. It is important to analyse new creations to see which are the same and which are different.

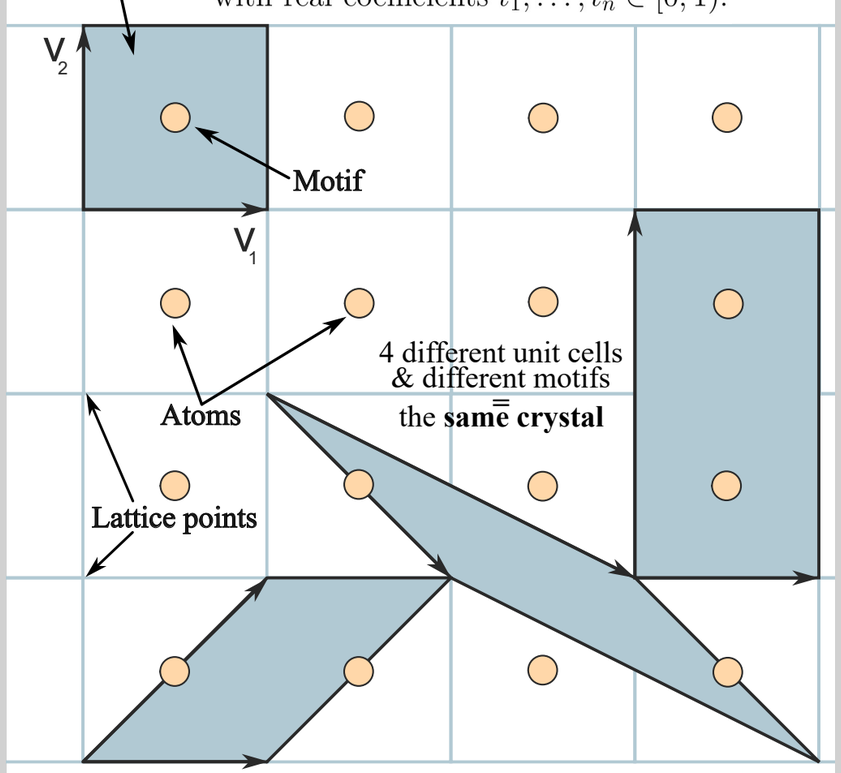
What is a crystal?

A crystal is a periodic arrangement of atoms.
To mathematically contruct a crystal, pick:

1) independent vectors $\{v_1, ..., v_n\}$ of $\mathbb{R}^n$
2) a motif (a set of points: atoms)



Unit Cell: The parallelepiped $U(v_1, ..., v_n)$ consisting of all linear combinations $\sum_{i=1}^{n} t_i v_i$ with real coefficients $t_1, \ldots, t_n \in [0, 1)$.



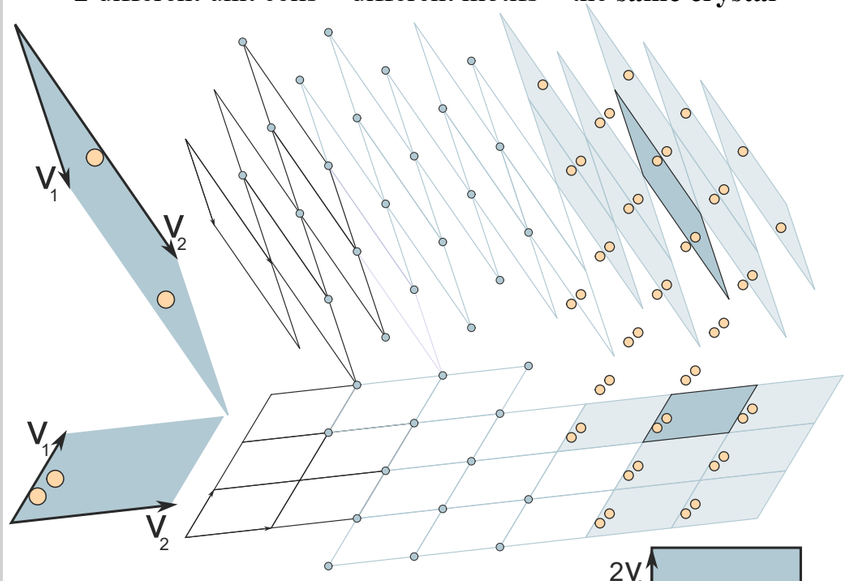
Motif: $M \subset U(v_1, \ldots, v_n)$ is a finite set M of points.

Lattice: $\Lambda(v_1, \ldots, v_n)$ consists of all linear combinations $\sum_{i=1}^{n} c_i v_i$ with integer coefficients $c_1, \ldots, c_n \in \mathbb{Z}$.

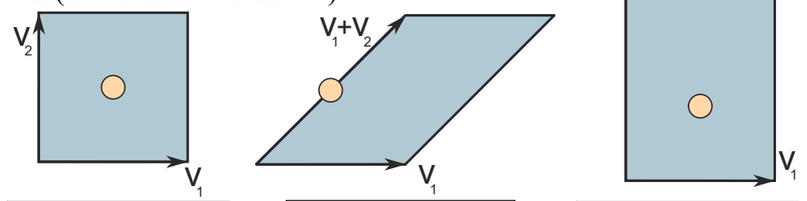
Periodic Point Set: $S = M + \Lambda(v_1, \ldots, v_n)$ is the set of points $p + v$ for all $p \in M$ and $v \in \Lambda$.

Which crystals are the same?

2 different unit cells + different motifs = the **same crystal**



Databases store crystals in **CIF**s, Crystallographic Information Files containing the unit cell and motif (fractional coordinates)



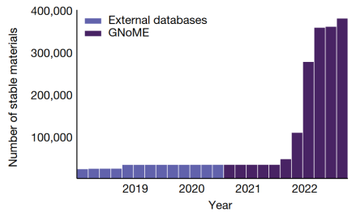

```
cell_length_a  1.0
cell_length_b  1.0
cell_angle    90.0
loop_
_atom_site_fract_x
_atom_site_fract_y
  (0.5 , 0.5)
```

```
cell_length_a   1.0
cell_length_b 1.414
cell_angle     45.0
loop_
_atom_site_fract_x
_atom_site_fract_y
  (0.0 , 0.5)
```

```
cell_length_a  1.0
cell_length_b  2.0
cell_angle    90.0
loop_
_atom_site_fract_x
_atom_site_fract_y
   (0.5 , 0.25)
   (0.5 , 0.75)
```

Artificial Illusions

Google's Deep Mind's AI tool GNoME claims to have discovered 2.2 million new crystals: "equivalent to nearly 800 years' worth of knowledge".

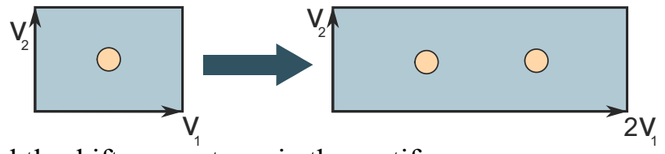


However, there is no guarantee that the the crystals they found are different or in fact real.

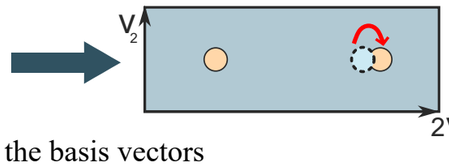
E.g. a group of 9 CIFs contain equal numbers but the chemical composition differs by only 1 or 2 atoms.

Designing your own Illusions

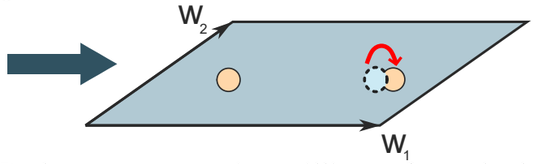
1) extend the unit cell



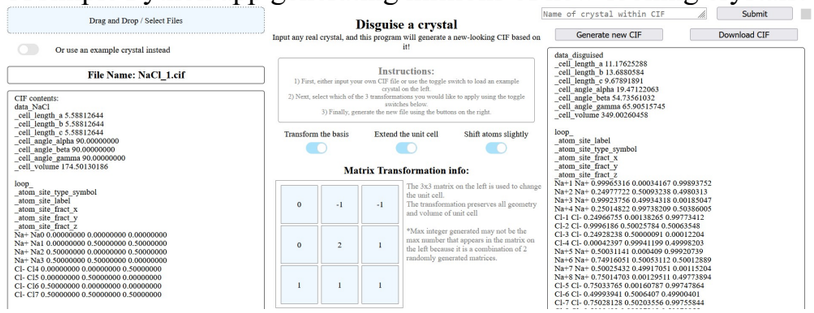
2) slightly shift some atoms in the motif



3) change the basis vectors



A simple Python app generating millions of new-looking crystals:



A **periodic structure** is an *equivalence class* of periodic point sets $S \subset \mathbb{R}^n$ under rigid motion

which can be efficiently recognised by **complete invariants** I (such that $I(S) = I(Q)$ if and only if $S \cong Q$) with Lipschitz continuous metrics.

Conclusion

A rigorous geometric approach to crystal design suggests that the AI is often more Illusions than Intelligence. Equivalence classes, invariants and continous metrics are required to compare crystals.

References

Anosova et al. The importance of definitions in crystallography. IUCrJ, v.11 (4), p.453-463, 2024, doi:10.1107/S2052252524004056

Merchant et al. Scaling deep learning for materials discovery. Nature 624, 80-85 (2023) https://doi.org/10.1038/s41586-023-06735-9

https://deepmind.google/blog/millions-of-new-materials-discovered-with-deep-learning/