CLASSIFYING STRUCTURE IN TWO-DIMENSIONAL POINT SETS VIA VORONOI TOPOLOGY

EMANUEL A. LAZAR

Abstract. Many two-dimensional systems can be studied as collections of particles in two dimensions. Understanding the way in which such sets of particles are arranged is thus a natural and important question. We suggest an approach towards classifying structure in such systems using Voronoi topology.

1. Introduction

We often encounter sets of particles in two dimensions whose "structure" we would like to analyze. What can we say about the way in which such particles are ordered? Perhaps we wish to identify defects in a single-phase system, or perhaps

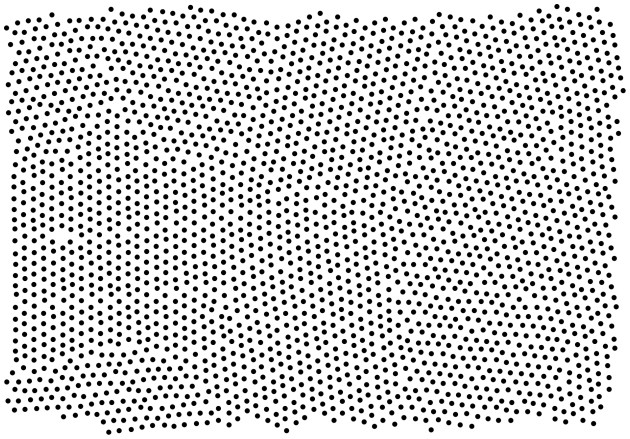

Figure 1. A two-dimensional system of particles.

we wish to identify different phases in a multi-phase system. Figure 1 illustrates a polycrystalline system with several defects, including grain boundaries, dislocations, and vacancies. Although these defects can be seen with our eyes, defining them precisely enough as to make them amenable to automated computer analysis is difficult.

Here we suggest a topological method, based on Voronoi topology, for classifying structure in two-dimensional sets of particles. This approach allows us to naturally ignore small fluctuations in atomic positions associated with thermal vibrations and small strains, without the need for quenching or temporal averaging. Many ideas suggested here can be considered as an extension of ideas first introduced in [1].

Date: May 13, 2016.

2. Voronoi Topology

One way to define the "local structure" of a particle in a two-dimensional system is by considering its number of neighbors. Sometimes this number is called the "coordination number" of a particle. Although we have a rough intuition of what it means for a particle to have "neighbors", it is not obvious how to make this term precise. One way to do this is by considering all other particles which are at most a fixed distance away. One limitation of this general approach is the need for choosing that fixed distance. How large should it be? Should it be the lattice constant? Should it be slightly larger than the lattice constant, so that particles slightly further away are included? If so, how much larger? Should this cutoff distance be system-dependent?

Alternatively, we might always consider the set of n nearest neighbors as *the* neighbors of a particle. One limitation of this approach that particle A might be a neighbor of particle B, even though particle B is not a neighbor of particle A. It seems reasonable to desire that the property of being a neighbor is reflexive.

A third approach that does not require choosing a cutoff distance, and which is also reflexive, involves using Voronoi tessellations. The **Voronoi** cell of a particle in a system of particles is the region of space closer to that particle than to any other. Constructing a Voronoi cell for every particle in a system subdivides the larger space into convex polygonal regions, each containing a single particle near its center. We can then define the number of neighbors of a particle as the number of edges of its Voronoi cell. The idea behind this definition is that every edge of a Voronoi cell indicates the presence of another, neighboring, particle in a system. This can be seen in Figure 2, which shows a central Voronoi cell with six edges;

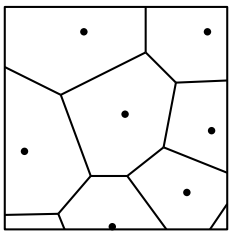

Figure 2. A central particle with a Voronoi cell with six edges; each edge is shared with a neighboring particle.

each edge is shared with a single neighboring particle. For this reason, the number of edges of a Voronoi cell can be considered as a proxy for counting the number of neighbors of a particle.

Figure 3 illustrates the same polycrystalline system shown before, this time with each particle colored according to the number of edges of its Voronoi cell. The majority of particles have six edges and are colored yellow; other Voronoi cells have five or seven edges and colored blue and blue respectively.

If we use the Voronoi cell of a particle to define its number of neighbors, then we can use this number to define the local "structure" near the particle. For example, particles whose Voronoi cells have 6 edges might be called "bulk", whereas those with more or fewer neighbors we can might label as defects. Once we have identified defects at the single-particle level, we can then identify larger-scale defects

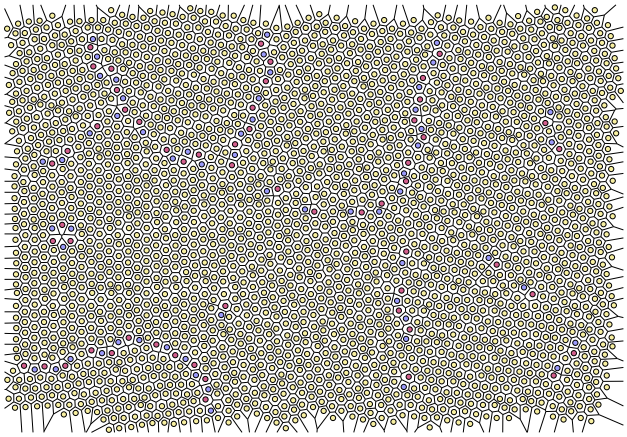

FIGURE 3. A two-dimensional system of particles. Each particle is colored by the number of edges of its Voronoi cell.

as contiguous regions of particle-defects. For example, notice a ring of 3 blue and 3 blue particles on the left side of the illustrated polycrystalline system. This ring identifies the presence of a vacancy. A long chain of alternating red and blue particles might indicate the presence of a high-angle grain boundary. And so on and so forth.

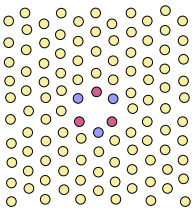

FIGURE 4. A vacancy seen as a collection of three 5-sided Voronoi cells and three 7-sided cells.

3. Extended Voronoi Topology

We sometimes find Voronoi cells with 6 edges which do not appear as part of the bulk. This motivates us to extend the simple model suggested above, in which we define the local structure of a particle as its number of Voronoi edges. Instead, we might consider the Voronoi structure not only of the immediate neighbors, but also of the second nearest neighbors. We can then define two particles to have the same local structure if the Voronoi structure around them and their neighbors is identical.

Figure 5 illustrates three central particles, each with a six-sided Voronoi cells. If we only used the number of Voronoi neighbors to classify local structure, then we would consider the three particles to have the same local structure. However, if we take into consideration the Voronoi structure around the immediate neighbors, then the three central particles each has a different local structure: the central particle

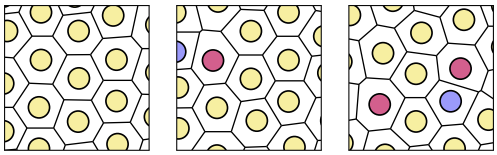

FIGURE 5. Three central particles with six-sided Voronoi cells. If we look beyond the immediate Voronoi neighbors then we can notice that the "local structure" is different near the three central particles.

in (a) is surrounded by 6 particles with hexagonal Voronoi cells; the central particle in (b) is surrounded by 5 particles with hexagonal Voronoi cells and 1 particle with a seven-sided Voronoi cell; the central particle in (c) is surrounded by six particles with different kinds of Voronoi cells.

4. Canonical codes

If we want to determine whether two particles have the same local structure, then we need a good way to record that information. Consider for a moment the arrangement of particles in Figure 5(c). The central particle has 3 neighbors with hexagonal Voronoi cells, 1 neighbor with a pentagonal Voronoi cell, and 2 neighbors with seven-sided Voronoi cells. If we only record this information, then we fail to capture information regarding the way in which those neighbors are arranged relative to one another. Instead we should record the way in which they are ordered. We could, for example, list the number of edges of the Voronoi cell of each neighbor as arranged in a clockwise fashion. For example, if we begin with the neighbor immediately below the central particle, we can write (6,7,6,6,7,5). If we begin with the neighbor immediately above the central particle we would write (6,7,5,6,7,6). Although these two sets of numbers indicate the same information, it is not immediately obvious whether the two local structures are identical by looking at them. "Local structure" should not depend on which neighbor we use first for our list and should not depend on whether we list neighbors in clockwise or counterclockwise manner.

We therefore adopt the following convention. For every arrangement of particles, we use the lexicographically smallest representation of the ordering, considered over all cyclic permutations and both orientations (clockwise and counterclockwise). That is, for any choice of N neighbors, we consider all cyclic permutations of $(n_1, n_2, \ldots n_N)$ and all cyclic permutations of $(n_N, n_{N-1}, \ldots, n_2, n_1)$, and choose the lexicographically smallest such representation. We then say that two particles have the same local structure if and only if they have the same canonical list of Voronoi neighbor edges. The examples in Figures 5(a), (b), and (c) have the canonical codes (6,6,6,6,6,6), (6,6,6,6,6,7), and (5,6,7,6,6,7) respectively.

5. Local Orders in Ideal Structures

In order for the codes to be useful in identifying structure in realistic systems, it is helpful to determine the set of structure codes associated with a particular

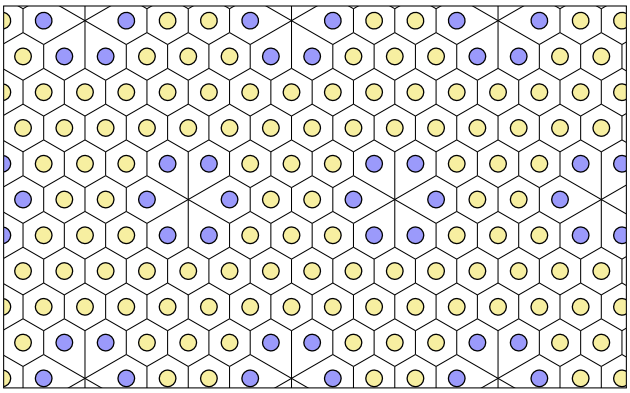

FIGURE 6. A special kind of crystal.

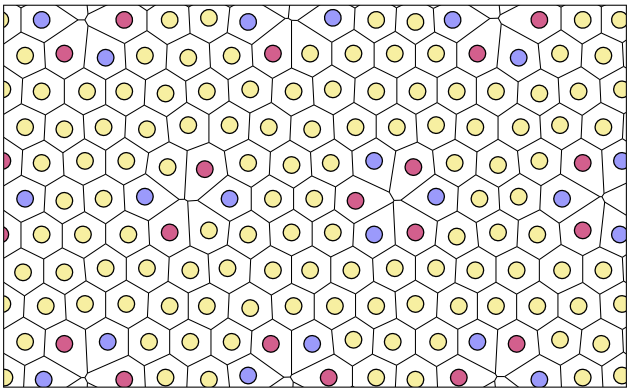

FIGURE 7. A special kind of crystal, perturbed via small thermal vibrations.

kind of phase. For example, in a perfect close-packed two-dimensional system, we should expect that the code associated with every particle is (6,6,6,6,6,6). If we find particles with other codes, then we can know that they reside nearby to defects. These codes are "stable" under perturbations of the particle positions [2].

What happens if instead of a close-packed system we consider a less trivial crystal? Figure 6 illustrates a close-packed lattice with a sublattice removed. Instead of all particles having hexagonal Voronoi cells, now many particles have pentagonal Voronoi cells. Particles in this structure can have one of four canonical codes: (6,6,6,6,6,6), (5,6,6,6,6,6), (5,5,6,6,6) and (5,5,6,6,6,6).

Unlike in the close-packed case, Voronoi cells here are highly unstable, in that small perturbations of the positions of the particles will change the structures codes. Figure 7 illustrates a "finite-temperature" version of the special structure. Notice that many of the five-sided Voronoi cells change to six- and seven-sided Voronoi cells. Along with changing the topology of the Voronoi cells, small perturbations also change the associated canonical codes. Codes that were previously (6,6,6,6,6,6) remain unchanged. However, particles initially having codes (5,6,6,6,6,6), (5,5,6,6,6), and (5,5,6,6,6,6) can change. In particular, particles with

(5,6,6,6,6,6) codes in the unperturbed structure can either remain unchanged or else change to (6,6,6,6,6,7). Particles that began as (5,5,6,6,6,6) can change to either (5,6,6,6,6,7), (5,6,6,6,6,6), or (6,6,6,6,6,7). Particles that began as (5,5,6,6,6) can resolve into many more types: (6,6,6,6,7), (6,6,6,7,7), (5,6,6,6,7,7), (5,6,6,6,6,7,6), and (5,6,6,6,5,7,7). A complete enumeration of all possible codes and changes resulting from small perturbations would take some more work; some details related to this can be found in [2].

6. Suggested Approach

Before describing a possible numerical approach, it is important to consider a theoretical problem of fundamental importance (in my opinion at least) related to identifying different "phases" in a sample. Namely, it seems to me that there are no precise definitions of the meaning "phase". For example, in Figure 6 above, every fifth particle along each direction is removed. What happens if instead it is every third, or everything seventh, or every twenty-third? Should each of those be considered different phases? What if vacancies are removed at irregular spacings in different parts of the sample? Should we classify that as yet another kind of phase, or a combination of small pieces of other phases? If we don't have a precise definition of "phases", then it makes little sense to discuss trying to identify them. I believe that this question has been, to a large degree, ignored by most people who deal with this kind of problem. So while I will suggest an approach that I believe is not really complete, I think it might be the best we can do, and possibly sufficient for our purposes.

The idea of this approach, based on ideas above, is that we look first at an "ideal structure" and compute the set of types we expect to find in it. For example, in the close-packed lattice, in a perfect crystal, or small perturbation of a perfect crystal, the Voronoi cell of every particle should be a hexagon, and it's neighbors should all be hexagons, described by the code (6,6,6,6,6,6). Any other Voronoi cells or codes indicate some defect kind of structure.

The type of structure seen in Figure 6 has a substantially larger class of types associated with it, but a set which we can still enumerate. It will be important to precompute this set of codes.

When analyzing a particular structure, we should compute the Voronoi cell topology of every particle in the system, and also the codes for every particle, each of which captures the "local structure" nearby that particle. Regions in which only the (6,6,6,6,6,6) code appears should be understood to be of close-packed phase. Those with a large number of these other codes might be considered a second kind of phase. Particles with codes that we have not expected likely indicate some sort of defect in the structure, or perhaps boundaries between two different phases.

References

[1] Lazar, E. A., Han, J. & Srolovitz, D. J. Topological framework for local structure analysis in condensed matter. *Proceedings of the National Academy of Sciences* **112**, E5769–E5776 (2015).

[2] Leipold, H., Lazar, E. A., Brakke, K. A. & Srolovitz, D. J. Statistical topology of perturbed two-dimensional lattices. *arXiv preprint arXiv:1509.03280* (2015).

Department of Mathematics, University of Pennsylvania, Philadelphia, PA 19104
E-mail address: mlazar@seas.upenn.edu
URL: www.seas.upenn.edu/~mlazar/