Physica D 334 (2016) 60–81

Contents lists available at ScienceDirect

Physica D

journal homepage: www.elsevier.com/locate/physd

Topological microstructure analysis using persistence landscapes

Paweł Dłotko [a,*], Thomas Wanner [b]

[a] Inria Saclay – Ile-de-France, 1 rue Honoré d'Estienne d'Orves, 91120 Palaiseau, France
[b] Department of Mathematical Sciences, George Mason University, Fairfax, VA 22030, USA

HIGHLIGHTS

- Computing persistent homology of phase separation patterns.
- Persistence keeps information about total mass and the stage of decomposition.
- Persistence landscapes and topological processes used for statistics.

ARTICLE INFO

Article history:
Received 25 August 2015
Received in revised form
19 April 2016
Accepted 20 April 2016
Available online 28 April 2016

Keywords:
Phase separation mechanisms
Computational topology
Inverse analysis

ABSTRACT

Phase separation mechanisms can produce a variety of complicated and intricate microstructures, which often can be difficult to characterize in a quantitative way. In recent years, a number of novel topological metrics for microstructures have been proposed, which measure essential connectivity information and are based on techniques from algebraic topology. Such metrics are inherently computable using computational homology, provided the microstructures are discretized using a thresholding process. However, while in many cases the thresholding is straightforward, noise and measurement errors can lead to misleading metric values. In such situations, persistence landscapes have been proposed as a natural topology metric. Common to all of these approaches is the enormous data reduction, which passes from complicated patterns to discrete information. It is therefore natural to wonder what type of information is actually retained by the topology. In the present paper, we demonstrate that averaged persistence landscapes can be used to recover central system information in the Cahn–Hilliard theory of phase separation. More precisely, we show that topological information of evolving microstructures alone suffices to accurately detect both concentration information and the actual decomposition stage of a data snapshot. Considering that persistent homology only measures discrete connectivity information, regardless of the size of the topological features, these results indicate that the system parameters in a phase separation process affect the topology considerably more than anticipated. We believe that the methods discussed in this paper could provide a valuable tool for relating experimental data to model simulations.

1. Introduction

Complicated patterns which evolve with time occur in a variety of applied contexts, and quantifying or even just distinguishing such patterns can pose serious challenges. Over the last decade, computational topology has emerged as a tool which on the one hand allows for significant data reduction, while at the same time focusing on and retaining essential connectivity information of the studied patterns. One particularly interesting application area is materials science, where complex evolving patterns are frequently created through complicated phase separation processes. Computational topology encompasses a wide variety of possible tools [1], and most of them in one way or another are based on homology theory. The deeper reason for this can be found in the inherent computability of homology, and in recent years powerful algorithms have been devised which enable fast homology computations for very large data sets, see for example [2,3], as well as the references therein.

Homology has been used in a number of materials science contexts. Earlier studies have made use of the easily computable Euler characteristic, see for example the references in the recent survey [4]. However, it has been pointed out that in some situations the information retained by the Euler characteristic is not enough to distinguish certain important pattern features.

* Corresponding author.
E-mail address: pawel.dlotko@inria.fr (P. Dłotko).

http://dx.doi.org/10.1016/j.physd.2016.04.015
0167-2789/

In contrast, the Betti numbers, which are associated with the homology groups and will be described below, provide a finer metric. They were used for example in [5] to relate the pattern complexity evolution as described by averaged Betti number evolution curves to the amount of stochasticity inherent in a phase field model for binary phase separation in metal alloys. This study was the first to use homology information in the context of model validation. Based on available experimental data it was shown that if the noise in the system is too low, the observed Betti number evolution curves are qualitatively different from the experimental ones. In addition, it was demonstrated in [5] that while Betti numbers can be used to separate bulk from boundary behavior, the averaged Euler characteristic can only describe the boundary effects. Similar materials science studies in the context of polycrystals can be found in [6,7]. While the first of these papers uses homology to study the connectivity properties of grain-boundary networks in planar sections of polycrystals, the second paper employs Betti numbers as a means to describe the thermal-elastic response of calcite-based polycrystalline materials such as marble. More precisely, in [7] homological techniques are used to characterize not only the elastic energy density and maximum principal stress response fields in a polycrystal, but also the respective grain-boundary misorientation distributions which generated these response fields. It was shown that this topological analysis can quantitatively distinguish between different types of grain-boundary misorientations, and relate them to differences in the resulting response fields.

In all of the applications described so far, the numerical or experimental data is given in the form of a field, or in other words, a real-valued function $u : \Omega \to \mathbb{R}$ defined on some domain $\Omega \subset \mathbb{R}^d$. The associated patterns are subsets of the domain Ω, and usually created through a thresholding process. For example, after selecting a suitable threshold θ, one can consider the sub- and super-level sets

$$\mathcal{M}^{\pm} = \{x \in \Omega : \pm (u(x) - \theta) \geq 0\},$$

which in some contexts are called nodal domains of u. As subsets of $\Omega \subset \mathbb{R}^d$, the sets $\mathcal{M}^{\pm}$ have well-defined singular homology groups, and if the function u is sufficiently regular, these groups can be computed precisely; see for example [8,9] for the two-dimensional case $d = 2$. However, if the function u is not smooth, or if the thresholding process involves a field created from experimental or noisy data, then the above thresholding process may not capture the correct topology of the actual underlying pattern. Moreover, in certain applications the thresholding approach itself might not be appropriate, for example if there is no obvious or physically relevant choice of threshold.

An extension of the concept of homology to situations involving noise or the lack of a clear thresholding process has been proposed some fifteen years ago and is called persistent homology [1]. While this extension is described in more detail in the next section, persistent homology is a dimension reduction technique which provides a topological description of the evolution of sublevel sets of a mapping u as a function of the thresholding level θ. It gives rise to intervals $[\theta_1, \theta_2]$ over which certain topological features persist, and the length $\theta_2 - \theta_1$ can in some sense be viewed as a measure of importance of the specific feature. Thus, topological features with small interval lengths are usually considered as "noise" or "unimportant", while features with long intervals are deemed "significant". Needless to say, the precise meaning of these notions will change from application to application. The concept of persistent homology has already been used in a number of materials science contexts, such as for example in the analysis of granular media [10,11], in the study of protein compressibility [12], as well as in the classification of amorphous structures [13] and glass [14]. Efficient algorithms for computing persistent homology are described in [1,15].

Despite the success of the above uses of computational topology in applications, one immediate question is the extent of the resulting dimension reduction. Through homology, patterns or microstructures are basically reduced to a finite set of integers, and it is therefore natural to wonder what information is still encoded in this reduced measurement, beyond the obvious connectivity information. It was pointed out in [16] that even small Betti number counts, which taken in isolation do not provide much in terms of pattern differentiation, can provide significant information when viewed in a stochastic, i.e., averaged setting. More precisely, it was shown in [16] that during the phase separation process called nucleation, averaged droplet counts on small domains give extremely precise projections for the observed averaged droplet counts on large domains, where the droplet count is just the zero-dimensional Betti number.

In the present paper, we demonstrate that when viewed in a stochastic and time evolving framework, topological information encodes considerably more than anticipated. This will be accomplished in the setting of instantaneous phase separation in binary metal alloys, as modeled by the Cahn–Hilliard theory of spinodal decomposition [17,18]. This phase separation phenomenon is initiated immediately after a high-temperature melt of two uniformly mixed metal components is quenched, i.e., rapidly cooled. Depending on the concentrations of the involved components, they will quickly separate to form complicated microstructures which contain some apparent element of randomness. Some of the resulting patterns in two space dimensions are shown in Fig. 1. These patterns evolve with time, and after the initial phase separation process a coarsening stage sets in, during which the characteristic length scale of the patterns increases.

The first mathematical model for spinodal decomposition was introduced by Cahn and Hilliard [19,20], who proposed a nonlinear evolution equation for the relative concentration difference $u = \rho_A - \rho_B$, where ρ_A and ρ_B denote the relative concentrations of the two components, i.e., $\rho_A + \rho_B = 1$. Their model is based on the Ginzburg–Landau free energy given by

$$E_\epsilon(u) = \int_\Omega \left(\frac{\epsilon^2}{2} |\nabla u|^2 + F(u) \right) dx, \tag{1}$$

where $\Omega \subset \mathbb{R}^d$ is a bounded domain, and the positive parameter ϵ models interaction distance. The bulk free energy F is a double well potential, which for the purposes of this paper is taken as

$$F(u) = \frac{1}{4} \left(u^2 - 1\right)^2. \tag{2}$$

Taking the variational derivative $\delta E_\epsilon / \delta u$ of the Ginzburg–Landau free energy (1) with respect to the concentration variable u, one then obtains first the chemical potential $w = -\epsilon^2 \Delta u + F'(u)$, and then the associated Cahn–Hilliard equation $\partial u / \partial t = \Delta w$, i.e., the fourth-order partial differential equation

$$\frac{\partial u}{\partial t} = -\Delta \left(\epsilon^2 \Delta u - F'(u) \right), \tag{3}$$

subject to homogeneous Neumann boundary conditions for both w and u. Due to these boundary conditions, any mass flux through the boundary is prohibited, and therefore mass is conserved. We generally consider initial conditions for (3) which are small-amplitude random perturbations of a spatially homogeneous state, i.e., we assume that $u(0, x) \approx \mu$ for all $x \in \Omega$, as well as $\int_\Omega u(0, x)\, dx / |\Omega| = \mu$, where we use the standard abbreviation $|\Omega| = \int_\Omega 1\, dx$. One can easily see that such initial conditions lead to instantaneous phase separation in the Cahn–Hilliard equation