Journal: Acta Materialia, Paper: A-21-0401
Authors: Sterling G. Baird, Eric R. Homer, David T. Fullwood and Oliver K. Johnson
Title: Five degree-of-freedom property interpolation of arbitrary grain boundaries via Voronoi fundamental zone octonion framework.

The manuscript describes methods for time-efficient commutation of distances between boundaries in the 'space' of macroscopic boundary parameters and for interpolation of grain boundary energy as a function of these parameters. Some statements made in the manuscript are not correct. In my opinion, the manuscript does not meet requirements of Acta Materialia.

A. Key for the computation of distances between grain boundaries is accounting for symmetrically equivalent boundary representations. There are a number of indications that the authors do not fully understand this subject.

1. According to the manuscript, the cardinality of the group O_h is $N_p = 24$ (p.4, l.5), (p.4, l.27, R), (p.18, l.22, R).[1] It is well known that the number of elements in the octahedral point group O_h is 48; see, International Tables for Crystallography. Twenty four is the number of proper rotations in the O_h group. 1

2. According to the manuscript, the number of symmetrically equivalent representations of a boundary (number of SEOs) is N_p^2 (p.3, l.40, R). This number is not correct. Assuming crystals are invariant with respect to inversion, N_p is the number of proper symmetry operations of the point group, and grain exchange symmetry is allowed, then the number of symmetrically equivalent representations of a boundary is $4N_p^2$. 2

3. According to the manuscript, the proposed methods "can be applied to any of the 32 crystallographic point groups" (p.1, l.38), (p.21, l.25, R). Such a general approach would have to encompass noncentrosymmetric crystals. No details are given in the manuscript, and I doubt in correctness of that approach. 3

4. According to the manuscript, "to perform a traditional distance calculation it is necessary to compare all SEOs of o_1 to all of the SEOs of o_2 and take the smallest distance." (p.3, ll.22–25, R). Again, this is not correct. With the considered metric, a transformation between equivalent representations is an isometry, and therefore, to calculate the distance between two boundaries, it is enough to compare a *single* representation of one boundary to all representations of the other boundary. See, e.g., section 7 in [50]. 4 5

The first three points on the list above can be viewed as mere mistakes. The last point, however, affects large parts of the manuscript devoted to speeding up computations of distances between boundaries. 6

B. Using Voronoi based fundamental zone in texture related analyzes is not really new. Such fundamental zones are well defined in the case of orientations. In the case of grain boundaries (or misorientations), the construction is not that elegant. Contrary to what is written in the manuscript (p.3, l.25, L), the grain boundary representation taken as the reference point cannot be arbitrary. If the reference point is invariant with respect to a symmetry operation other than identity, then it is at the same smallest distance to *multiple* representatives of another boundary, and the procedure does not work; the resulting Voronoi cell is a multiple of the fundamental zone. 7 8 9

[1] L and R denote left and right columns, respectively.

Summary of comments: A-21-0401-Review 2 Attachment.pdf

Page:1

Number: 1 Author: sterg Subject: Highlight Date: 2021-03-27 21:07:02

Text now uses the phrase proper rotations instead of cardinality.

Number: 2 Author: sterg Subject: Highlight Date: 2021-03-27 21:13:06

Switching between values and Big O notation (https://en.wikipedia.org/wiki/Big_O_notation) likely led to this mistake in the text. The text has been updated accordingly.

Number: 3 Author: sterg Subject: Highlight Date: 2021-03-27 21:33:01

From a practical standpoint, the symmetry operators for the 32 point groups are available via https://github.com/ichesser/GB_octonion_code/tree/master/TutorialCode/crystal_symmetry_ops. In the VFZO repository, point group is an input parameter (i.e. easily changed to accommodate the user's needs). Lack of an inversion center should be reflected in the symmetry operators. From a theoretical standpoint, we believe this would translate to differently shaped/larger VFZs rather than present a fundamental challenge. Given the already lengthy manuscript, this is a topic of future work. We have updated the text with a footnote clarifying this.

Number: 4 Author: sterg Subject: Highlight Date: 2021-03-27 22:11:28

The authors intended this to mean a traditional, symmetrized GBO distance calculation (see last paragraph of Section 2.1 of [49] DOI: doi: 10.1016/j.actamat.2020.05.024.) This was ambiguous and we have updated the text accordingly.

Number: 5 Author: sterg Subject: Highlight Date: 2021-03-30 00:32:00

Prior to submission of the manuscript, one author noted similarity between what was discussed in section 7 of [50] (DOI: 10.1016/j.actamat.2019.09.032) and the approach developed within the VFZO framework (esp. the significant computational gain). We agree with reviewer #2 that "it is enough to compare a single representation of one boundary to all representations of the other boundary" and have updated the manuscript accordingly, including several references to [50] w.r.t. isometry equivalence. By numerical testing, we find that the estimate of 6.6 CPU years is reduced to 153 CPU days when the isometry equivalence assumption is applied and that the computed distances are nearly identical to those reported in [49] (DOI: 10.1016/j.actamat.2020.05.024), with an RMSE of 1.6566e-07 degrees for the 388x388 pairwise distance matrix. This still represents a speed-up by several orders of magnitude relative to a 10-component ensemble and has major implications for reducing computation time of e.g. grain growth simulations.

Number: 6 Author: sterg Subject: Highlight Date: 2021-03-30 01:54:39

The manuscript has been updated to reflect comparison not only with the computational runtime reported in [49], but also with an in-house runtime using the isometry equivalence assumption of [50] and the same underlying distance function (GBdist4.m) that's used in the VFZO framework (see get_pd_fix.m).

Number: 7 Author: sterg Subject: Highlight Date: 2021-03-30 00:29:12

We agree. In fact, the description of this work in terms of Voronoi Fundamental Zones originated via an author's suggestion after the majority of the framework had been developed. We intended this description and corresponding visualizations to serve as an effective teaching tool especially given the expected background of readers.

Number: 8 Author: sterg Subject: Highlight Date: 2021-03-30 01:33:32

Regrettably, a clarification was removed from the text before submission which likely resulted in confusion. This has been re-added to the manuscript as the last paragraph of Section 2.1.1.

Number: 9 Author: sterg Subject: Highlight Date: 2021-03-30 01:40:39

The re-added content clarifies that we have an internal check within the codebase that verifies a unique minimum-distance SEO has been found. On average, the likelihood of a second equivalent distance SEO being found is ~0.002%. Thus, we reject the notion that the Voronoi cell is a multiple of the fundamental zone in the case of an arbitrary, *low-symmetry* GB (see first sentence of section 2.1.1).

Moreover, with boundary representatives limited to the fundamental zone, close boundaries may have distant representatives. The authors addressed the problem by performing computations for a number of different reference points, i.e., for a number of different fundamental zones.[2] This approach is novel, but not fully reliable; if the number of different fundamental zones is small, the data may still contain some incorrect distances. 1

C. The unquestionable new contribution of the manuscript is the use of metric-based all-purpose multivariate interpolation methods to predicting symmetric energy function from known energies of a given set of boundaries. The method is different from and simpler than that described in [66], but the resulting energy functions will be smooth[3] (no cusps). 2

In conclusion, the manuscript A-21-0401 does report some new science (**C** and **B**), but extensive fragments of it are devoted to solving a non-existent problem (**A**). Therefore, I recommend against publication of A-21-0401 in Acta Materialia. 3

In case my recommendation is overruled or the manuscript is submitted elsewhere, I would like add the following two comments.

First, the text is long and difficult to read. Here are a couple of measures which (in my opinion) would make the paper more readable. 4

• I would remove most abbreviations. Some are used only once (3DOF, GBED, kFCV, LOOCV) or twice (BPFZ, TJ), and they do not help. Moreover, I would check for repetitions. E.g., fragments (p.18, ll.30–37, L) and (p.18, ll.26–33, R) carry practically the same message. 5

• The text is overloaded with overcomplicated expressions. For instance, there is no need to talk about Riemannian manifolds (p.1, l.20), (p.3, l.11, L) as the considered metric is not intrinsic. Or take the word 'octonion' repeated so many times explicitly and in acronyms. Since the octonion product is never used, referring to a pair of unit quaternions as an octonion is a misuse of a well defined mathematical term.[4] The authors followed terminology introduced in other papers, but repeating mistakes of others adds to the confusion. When seeing the word 'octonion' repeated 200 times, a reader reaches the false conclusion that the formalism relies on true octonions. 6

Second, most parts of the work, in particular considerations concerning interpolation, are independent of boundary parameterization and are applicable to other properly defined metrics. There is no reason for limiting these considerations to a particular grain boundary parameterization or a particular distance function. 7

[2] The method known to me was to use a larger region = a single fundamental zone plus a symmetry-dependent margin.

[3] This is my guess. No relevant figures are shown.

[4] A simple analogue would be to refer to unit real numbers as complex numbers. Elements of the pair (r_1, r_2) of unit reals satisfy $|r_i| = 1$, i.e., (r_1, r_2) is one of $(\pm 1, \pm 1)$. The multiplication rule for the pairs (r_1, r_2) and (r'_1, r'_2) is $(r_1, r_2) \star (r'_1, r'_2) = (r_1 r'_1, r_2 r'_2)$, whereas for complex numbers it is $(r_1, r_2) \star (r'_1, r'_2) = (r_1 r'_1 - r_2 r'_2, r_1 r'_2 + r_2 r'_1)$. Clearly, using the term 'complex numbers' for the pairs $(\pm 1, \pm 1)$ is out of place. There is no need for complex numbers to multiply, say, 1 by -1.

Page:2

Number: 1 Author: sterg Subject: Highlight Date: 2021-03-30 01:46:34

For the purpose of interpolation, such infrequent overestimations (again, considering the use of logarithmically scaled color in Figure 3d) will have an insignificant effect on interpolation accuracy (see discussion near end of Section 2.1.3).

Number: 2 Author: sterg Subject: Highlight Date: 2021-03-30 01:05:44

This is only the case for GPR when a finite smoothness length is used. A close-to-zero smoothness length will resolve cusps at the expense of not handling noise well.

Number: 3 Author: sterg Subject: Highlight Date: 2021-03-30 01:17:40

Even when using the isometry equivalence assumption, fitting a model for 50,000 GBs could take as long 150 CPU days. Not only that, but sampling the model for another 50,000 GBs would take an additional 150 CPU days. Considering esp. grain growth simulations which are iterative, this is a non-negligible cost. Use of the VFZO framework results in several orders of magnitude reduction in computational requirements. Thus, we assert that the VFZO framework solves a problem that does exist.

Number: 4 Author: sterg Subject: Highlight Date: 2021-03-30 01:56:27

We will remove infrequently used abbreviations, remove redundant content where appropriate, and remove unnecessarily complicated expressions.

Number: 5 Author: sterg Subject: Highlight Date: 2021-04-01 10:47:36

Abbreviations that appear less than 3 times have been changed to non-abbreviated form.

Number: 6 Author: sterg Subject: Highlight Date: 2021-03-30 02:17:46

Does reviewer #2 have any suggestion(s) for alternatives to the use of the word 'octonion'?

Number: 7 Author: sterg Subject: Highlight Date: 2021-03-30 00:58:37

We are aware of other metrics discussed in [50] that have closed-form expressions for GB distance. GBO is the only one to our knowledge with publicly available and tested machinery (i.e. a codebase) and was the natural choice for our purposes. We are eager to compare VFZO with metrics other than GBO in terms of accuracy and computational runtime for computing distance matrices and performing interpolation. Without existing machinery that has been validated/optimized, this is a large endeavor that is outside the scope of this work. We are interested in pursuing this topic for future study.

However, we think it is likely that interpolation strategies such as GPR and barycentric interpolation (the two with highest accuracy) will be difficult to implement for the other distance metrics which are not Euclidean nor approximately Euclidean to our knowledge. For example, implementation of custom covariance kernels in fitrgp() relies on numerical differentiation for which computation time is dramatically increased. In other coding languages, one can specify analytical derivatives for the covariance matrix using established machinery; however, these analytical expressions would need to be derived. Barycentric interpolation represents an even more difficult challenge as each simplex would likely need to be constructed individually and via specialized distance-only simplex construction algorithms (which would also need to be implemented, validated, and optimized). NN and IDW can be performed using pairwise distance matrices; however, we expect that computational requirements will be exorbitant compared with ensemble VFZO distances when these are used as surrogate models for iterative, large-scale simulations where 1000's of new GBs are probed at each timestep. While we expect these limitations with many distance metrics other than VFZO, we would be happy to see results that show otherwise.