Computational Materials Science 101 (2015) 211–215

Contents lists available at ScienceDirect

Computational Materials Science

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

Calculating probability densities associated with grain-size distributions

J.M. Rickman [a,b,*], A. Lawrence [a], A.D. Rollett [c], M.P. Harmer [a]

[a] Department of Materials Science and Engineering, Lehigh University, Bethlehem, PA 18015, United States
[b] Department of Physics, Lehigh University, Bethlehem, PA 18015, United States
[c] Department of Materials Science and Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, United States

ARTICLE INFO

Article history:
Received 24 November 2014
Received in revised form 17 January 2015
Accepted 21 January 2015
Available online 18 February 2015

Keywords:
Microstructure analysis
Statistics

ABSTRACT

We describe a methodology for calculating approximate, yet accurate analytical expressions for the probability density function of grain diameter as obtained from experimental microstructures. This methodology relies on a novel cumulant expansion that is tailored to the lognormal distribution and provides a systematic description of departures from lognormality. We test our methodology by characterizing two data sets obtained from the microstructures associated with polycrystalline, high-purity Al_2O_3 samples. The utillity of this approach is demonstrated by a detailed statistical analysis.

1. Introduction

The microstructures of polycrystals typically have a significant impact on their mechanical and electronic properties, and therefore a quantitative characterization of salient microstructural features is especially important in such systems [1]. One such characterization is the determination of a probability density function (pdf) for an important microstructural descriptor, such as the grain diameter. Indeed, several probability densities have been employed to describe coarsened microstructures. For example, in early work Feltham [2] suggested that grain diameters are lognormally distributed, and others [3] have demonstrated that this distribution can be derived from an assignment of grains to topological classes. By contrast, Louat proposed a model in which grain-boundary motion is essentially random in nature, with the resulting grain diameter following a Rayleigh distribution [2]. It should be noted, however, that some stochastic models, such as the one proposed by Louat, appear to have no theoretical justification since the associated Fokker–Planck equation does not contain the requisite diffusive terms [4]. In recent work, Barmak et al. [5] advanced our understanding of other microstructural descriptors, such as the grain-boundary character distribution, by formulating an entropy-based theory that indicates that the evolution of this distribution obeys a Fokker–Planck equation.

There have been various studies that tried to assess the validity of particular grain-diameter pdfs generated both experimentally and by computer simulation. In early work, Palmer et al. [6] characterized normal and secondary grain growth in Ge thin films. In this study the authors fit the grain-size distribution obtained for a 30-nm-thick thin film to the lognormal, Louat [7] and Hillert [8] distributions, and found that the lognormal density provided the best fit at small grain sizes. In addition, Barmak et al. [9] studied grain growth and ordering kinetics in CoPt alloy films by quantification of the grain-size distribution, finding that the grain-size distribution is lognormal at stagnation. More recently, Donegan et al. [10] investigated deviations from lognormality in the upper tails of grain-size distributions compiled from two- and three-dimensional microstructures. They were able to quantify these deviations by employing extreme-value statistics to the upper tails of distributions. By contrast, Carpenter et al. [11] performed a grain-size analysis on a large grain population ($\sim 10^4$ grains) associated with an Al thin film and were unable to find acceptable agreement with several plausible theoretical densities (i.e., lognormal density, gamma density and the Rayleigh density) [11].

Given the need to quantify the distribution of grain sizes, we describe here a methodology for calculating analytical expressions for the grain-diameter pdf as obtained from experimental microstructures. Our strategy is to use shape features of pdfs that are compiled from microstructural data to find an approximate pdf that characterizes the data. In particular, by assuming that the distribution of grain diameters is "close" to a known pdf (e.g., lognormal density), the compiled pdf will be expressed in terms of an expansion that is tailored to the known pdf. The formalism for such an expansion will be considered in general. In the case of a lognormal density in particular, the requisite expansion will be written in terms of log-cumulants that arise in so-called "second-kind" statistics [12]. As will be seen below, in this formulation a Mellin

* Corresponding author at: Department of Materials Science and Engineering, Lehigh University, Bethlehem, PA 18015, United States.
E-mail address: jmr6@lehigh.edu (J.M. Rickman).

http://dx.doi.org/10.1016/j.commatsci.2015.01.030
0927-0256/

transform (as opposed to the usual Fourier or Laplace transform) is used to generate the moments (or cumulants) of a distribution based on the logarithm of a random variable. It has been employed, to a limited degree, in the analysis of radar images and mixtures of gamma distributions [12]. The novel, associated cumulant expansion is tailored to the lognormal distribution and provides a systematic description of departures from lognormality. Before describing our analysis of microstructural data, we first outline the more conventional analysis of densities that are approximately normal using the Gram–Charlier series [13]. We then validate our approach by characterizing the grain populations for two experimentally obtained microstructures.

2. Background

2.1. Formalism and normal density

Consider a grain-size distribution, $p(G)$, reflecting the probability density of finding a grain of diameter, G, in a population associated with a microstructure. The nth moment of this probability density function, $\langle G^n \rangle$, can be conveniently calculated from the characteristic function [16]

$$\tilde{p}(t) = \int_{-\infty}^{\infty} dG\, e^{itG}\, p(G), \tag{1}$$

by differentiation, noting that

$$\langle G^n \rangle = (-i)^n \frac{\partial^n \tilde{p}(t)}{\partial t^n}\Big|_{t=0}. \tag{2}$$

It is also useful to regard the characteristic function as a cumulant generating function via the relation

$$\tilde{p}(t) = \exp\left[\sum_{n=1}^{\infty} \frac{(it)^n}{n!} C_n\right], \tag{3}$$

where C_n is the nth-order cumulant [19]. Cumulants are combinations of moments and provide another means to characterize a distribution [17,18]. The cumulants can therefore be calculated directly from Eq. (3) from

$$C_n = (-i)^n \frac{\partial^n \ln \tilde{p}(t)}{\partial t^n}\Big|_{t=0}. \tag{4}$$

To see the utility of the cumulant-generating function, suppose that the probability density is given by the Gaussian (i.e., normal density)

$$p^{Gauss}(G) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right) \exp\left[-\frac{(G-\mu)^2}{2\sigma^2}\right], \tag{5}$$

with mean, μ, and standard deviation, σ. Since the corresponding characteristic function is given by

$$\tilde{p}^{Gauss}(t) = \exp\left[it\mu - \frac{t^2\sigma^2}{2}\right], \tag{6}$$

one can see that this probability density is fully described by its first two cumulants, namely $C_1^{Gauss} = \mu$ and $C_2^{Gauss} = \sigma^2$. In other words, for a Gaussian, all cumulants beyond the first two vanish.

It is also possible to find an expression for a probability density that is "close" to a Gaussian in terms of a cumulant expansion, the so-called Gram–Charlier series [13,14]. More specifically, for the characteristic function $\tilde{p}(t)$, one can write

$$\tilde{p}(t) = \exp\left[\sum_{n=1}^{\infty} \frac{(it)^n}{n!}\left(C_n - C_n^{Gauss}\right)\right] \tilde{p}^{Gauss}(t). \tag{7}$$

Now, if $C_n^{Gauss} = C_n$ for $n = 1, 2$, then

$$\tilde{p}(t) = \exp\left[\sum_{n=3}^{\infty} \frac{(it)^n}{n!} C_n\right] \tilde{p}^{Gauss}(t). \tag{8}$$

By taking the inverse Fourier transform of Eq. (8), one can obtain the desired distribution, namely $p(G)$. In practice, given the complexity of Eq. (8), one can only obtain an approximation to this distribution.

To obtain this approximation, consider a term-by-term inversion. This can be accomplished by using an integral representation for the Hermite polynomial of order n, H_n, given by [15]

$$H_n(x) = \frac{(-2i)^n}{\sqrt{\pi}} \int_{-\infty}^{\infty} dt\, e^{-(t-ix)^2} t^n. \tag{9}$$

One then finds that

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} du\, e^{-u^2\sigma^2/2}\, e^{i(\mu-G)u}\, u^n = i^n \left[\frac{p^{Gauss}(G)}{\sigma^n}\right] He_n\left(\frac{\mu-G}{\sigma}\right), \tag{10}$$

where the probabilist's Hermite polynomial, $He(x)$, is related to $H_n(x)$ by the rescaling identity $He(x) = 2^{-n/2} H_n\left(x/\sqrt{2}\right)$. Then, one can invert the lowest-order terms in Eq. (8) to obtain

$$\begin{aligned}\frac{p(G)}{p^{Gauss}(G)} = 1 &+ \frac{1}{3!}\left(\frac{C_3}{\sigma^3}\right) He_3\left(\frac{G-\mu}{\sigma}\right) + \frac{1}{4!}\left(\frac{C_4}{\sigma^4}\right) He_4\left(\frac{G-\mu}{\sigma}\right)\\ &+ \frac{1}{5!}\left(\frac{C_5}{\sigma^5}\right) He_5\left(\frac{G-\mu}{\sigma}\right)\\ &+ \frac{1}{6!}\left(\frac{C_6}{\sigma^6} + 10\frac{C_3^2}{\sigma^6}\right) He_6\left(\frac{G-\mu}{\sigma}\right) + \cdots,\end{aligned} \tag{11}$$

where the symmetry property $He_n(-x) = (-1)^n He_n(x)$ was used to simplify this expression.

2.2. Lognormal density

A similar formalism, called second-kind statistics, can be obtained from a different characteristic function and will be useful for analyzing distributions that are "close" to the lognormal distribution [12,20]. An approach, similar to that outlined below, has already been employed in the astrophysics literature to quantify the pdf associated with the distribution of particles in cold dark matter simulations [21]. In this approach a second-kind characteristic function is given as the Mellin transform of the probability density by

$$\hat{p}(s) = \int_0^{\infty} dG\, G^{s-1}\, p(G), \tag{12}$$

where s is a complex quantity in Mellin space, and the corresponding log-moments are obtained via differentiation by

$$\langle (\ln G)^n \rangle = \frac{\partial^n \hat{p}(s)}{\partial s^n}\Big|_{s=1}. \tag{13}$$

By analogy with the development above, one can regard the second-kind characteristic function as a log-cumulant generating function using the relation

$$\hat{p}(s) = \exp\left[\sum_{n=1}^{\infty} \frac{(s-1)^n}{n!} K_n\right], \tag{14}$$

where K_n is the nth-order log-cumulant.

As an illustration of the application of this formalism, suppose that one has the lognormal density

$$p^{log}(G) = \left(\frac{1}{\sqrt{2\pi}\sigma G}\right) \exp\left[-\frac{(\ln G - \mu)^2}{2\sigma^2}\right], \tag{15}$$

with parameters μ and σ. The corresponding second-kind characteristic function is given by