

Comparison of two methods for the ODF modeling for magnesium alloy with low symmetry

T M Ivanova¹ and V N Serebryany²

¹National Research Nuclear University, 31 Kashirskoe Shosse, Moscow 115409, Russia

²Laboratory of Crystal Structural Research, Institute of Metallurgy and Material Science, RAS, 49 Leninsky Prospect, Moscow 119991, Russia

E-mail: ivatatiana@gmail.com, vns@imet.ac.ru

Abstract. In this work we compare two methods for the reconstruction of orientation distribution function (ODF) from experimental pole figures. The magnesium alloy sample Mg-4.5%Nd is used for the investigation. The component fit method with the central normal distributions on $SO(3)$ and the orientation grid method demonstrate similar reconstructed ODF. The component fit method is more robust to experimental errors, however it relies on non-linear optimization with a set of initial approximations provided by a complicated selection problem. The orientation grid method is more sensitive, but easier to implement.

1. Introduction

In quantitative texture analysis, the orientation distribution function provides researches with the most complete information about the texture of the polycrystalline materials. There is a lot of literature on the problem of the ODF modeling from a set of the experimental pole figures (PFs) (for example, [1]–[9]). The problem is known to be ill-posed [5],[6] due to the non-uniqueness of the solution and its instability. The purpose of the paper is to discuss the advantages and disadvantages of the two methods for the ODF modeling: the component fit method with central normal distribution on $SO(3)$ [6] and the orientation grid method [6]. With these methods, the solution is searched in a certain classes of correctness, so the uniqueness is achieved.

2. Fundamentals

Suppose that an ODF can be represented as a finite sum of some standard functions depending on unknown parameters on the rotation group $SO(3)$. This ODF produces pole figures (PFs) as combinations of corresponding functions on the unit sphere S^2 with the same parameters. Orientation distribution function $f(g)$ and pole $P_{\vec{h}}(\vec{y})$ are related by the following integral equation

$$P_{\vec{h}}(\vec{y}) = \frac{1}{2\pi} \int_0^{2\pi} \left\{ f\left(\left\{\vec{h}, \phi\right\}^{-1} \left\{\vec{y}, 0\right\}\right) + f\left(\left\{-\vec{h}, \phi\right\}^{-1} \left\{\vec{y}, 0\right\}\right) \right\} d\phi. \quad (1)$$



These are the PFs which are usually measured in the experiment, so we can estimate unknown parameters mostly by fitting the set of measured PFs. Let $\{\vec{h}_\lambda, \lambda = 1, \dots, \Lambda\}$ and $\{\vec{y}_{ij}, i = 1, \dots, I, j = 1, \dots, J\}$ be directions in the crystallite and in the sample correspondingly, λ is an index of pole figure and ij are indices of measured points, then

$$P_{\vec{h}_\lambda}(\vec{y}_{ij}) = P^{model}(\vec{c}, \vec{t}, \vec{h}_\lambda, \vec{y}_{ij}) + \text{err}(\lambda, i, j). \quad (2)$$

In equation (2) vector \vec{c} is a set of linear parameters of a model, \vec{t} is a vector of nonlinear parameters, and $\text{err}(\lambda, i, j)$ are the measurement errors which are influenced by many factors.

The maximum likelihood estimators for unknown parameters of the model are the values which minimize the sum of weighted squared residuals

$$r^2 = \sum_{\lambda=1}^{\Lambda} \sum_{i=1}^I \sum_{j=1}^J W(\lambda, i, j) \left\{ N_\lambda P_{\vec{h}_\lambda}(\vec{y}_{ij}) - P^{model}(\vec{c}, \vec{t}, \vec{h}_\lambda, \vec{y}_{ij}) \right\}^2 \rightarrow \min. \quad (3)$$

Here the statistical weight $W(\lambda, i, j)$ reflects the degree of the significance of the measurement at point \vec{y}_{ij} on the pole figure $P_{\vec{h}_\lambda}(\vec{y})$. The parameter N_λ should be chosen to normalize the model (and, of course, experimental pole figures too). Let $S = \Lambda I J$ be the total number of the experimental data. We will refer to a particular set of $\{\lambda, i, j\}$ by single number s , where $s = 1, \dots, S$.

First, consider the linear problem with model function in (3) depending on $\vec{c} = (c_1, \dots, c_K)$, \vec{h}_λ and \vec{y}_{ij} . Then the number of parameters to be fitted is K . Let $A = (A_{sk})$ be the design matrix [10] of this problem, vector \vec{b} be the representative of experimental PF data, that is

$$A_{sk} = W(s) \frac{\partial P^{model}}{\partial c_k}, \quad \vec{b} = W(s) \left(N_\lambda P_{\vec{h}_\lambda}(\vec{y}_{ij}) \right)_s. \quad (4)$$

Then the minimization of r^2 yields normal equations of the least square problem (3):

$$(A^T A) \vec{c}^T = A^T \vec{b}^T. \quad (5)$$

In many cases, normal equations are very close to singularity, so a regularization step for solution of problem (5) is necessary. The regularization can be achieved by applying an orthogonal decomposition to the design matrix A , such as SVD [10],[11].

In the case with nonlinear parameters \vec{t} in the problem (3), the minimization has to be employed iteratively. Starting from $\vec{t}_{cur} = \vec{t}_0$ in each step of the iterations, we update the values $\vec{t}_{cur} \rightarrow \vec{t}_{next}$. The initial value set \vec{t}_0 is usually selected by a researcher (i.e. human factor). Consider the quadratic approximation of the common residual

$$r^2(\vec{t}_{next}) \approx r^2(\vec{t}_{cur}) + (\vec{t}_{next} - \vec{t}_{cur}) \nabla r^2 + 0.5 (\vec{t}_{next} - \vec{t}_{cur}) D (\vec{t}_{next} - \vec{t}_{cur})^T, \quad (6)$$

where $D = (D_{pq})$ is Hessian matrix $D_{pq} = \frac{\partial^2 r^2}{\partial t_p \partial t_q}$. In Levenberg–Marquard method [12] for

solution of the problem (3) Hessian D is replaced by \tilde{D} : $\tilde{D}_{pq} = \frac{\partial r^2}{\partial t_p} \frac{\partial r^2}{\partial t_q} (1 + \mu \cdot \delta_{pq})$ with some

heuristic parameter μ , and updating procedure looks as follows

$$\vec{t}_{next}^T = \vec{t}_{cur}^T - \tilde{D}^{-1} \nabla r^2(\vec{t}_{cur}). \quad (7)$$

If $r^2(\vec{t}_{next}) < r^2(\vec{t}_{cur})$, then \vec{t} gets the new value \vec{t}_{next} and μ decreases by some factor, if it is not the case, then \vec{t} remains the old value and μ increases by some factor. Iterations stop when common residual value becomes stable. It should be noticed that matrix \tilde{D} may degenerate if the dependence on some parameter t_p is rather weak, and in some other cases when μ is very small. That is why the regularization of the same type is desirable in the nonlinear case too.

3. The component fit method with central normal distributions

The component fit method is well known [2],[4],[6]. It's heavily based on the idea of ideal orientations for the texture interpretation and can be treated as smoothing and high compressing of the experimental data. The ODF model with modest number of standard functions describes the main features of the texture. Various analogs of normal distributions for mathematical modeling of ODF and PFs are used in quantitative texture analysis. The canonical normal distributions obeying the central limit theorem on the rotation group SO(3) are suggested in [2],[3]. The modified method of components with the central normal distributions is considered in [6]. Unlike linear methods, the component fit method involves a moderate number of parameters (about 50). The texture is supposed to consist of K normal components (axial with an axis \vec{n}_{0k} or peak with center) with volume fractions c_k , parameters of dispersion ε_k , centers g_{0k} , and a volume fraction of the nonstructural component c_0 . This model ODF generates model pole density as follows

$$P^{model}(\vec{c}, \vec{t}, \vec{h}, \vec{y}) = c_0 + \sum_{k=1}^K c_k P^{cnd}(\vec{h}, \vec{y}, g_{0k}, \vec{n}_k, \varepsilon_k), \quad (8)$$

where P^{cnd} can be either the peak (9) or the axial (10) component of reduced CND represented by the restricted series over Legendre polynomials

$$P^p(\vec{h}_\lambda, \vec{y}, g_0, \varepsilon^2) = \sum_{l=0(2)}^L (2l+1) \exp(-l(l+1)\varepsilon^2) P_l(\vec{h}_\lambda g_0 \vec{y}), \quad (9)$$

$$P^a(\vec{h}_\lambda, \vec{y}, g_0, \varepsilon^2, \vec{n}) = \sum_{l=0(2)}^L (2l+1) \exp(-l(l+1)\varepsilon^2) P_l(\vec{y}\vec{n}) P_l(\vec{h}_\lambda g_0 \vec{n}). \quad (10)$$

For unknown adjustable parameters $\{c = (c_0, \dots, c_K), t = (g_{01}, \varepsilon_1, \dots, g_{0K}, n_K, \varepsilon_K)\}$ and the set of measured PFs $\{P_{\vec{h}_\lambda}(\vec{y}_{ij}), \lambda = 1, \dots, \Lambda\}$, we get nonlinear problem (3). The statistical weights $W(\lambda, i, j)$ are supposed to be proportional to the element of area of S^2 .

The minimization of common residual r^2 in this case refers to the problems of nonlinear optimization (6)-(7) and is solved by Levenberg–Marquard method [12]. The singular value decomposition [10],[11] for ill-conditioned matrix \tilde{D} in (7) is applied. Regularization boils down to zeroing of small singular numbers. This causes condition number of \tilde{D} to decrease and improves statistical properties of the solution: it becomes smoother and robust to the experimental errors [10]-[12]. This method recovers the main texture features even in case when experimental errors are considerable. For estimating the adequacy of the model $RP(\Delta)$ –factor [1]-[4] is used in this work.

4. The ODF modelling by the grid of orientations

Another possible representation of $f(g)$ is a large set of standard distributions with identical small dispersion ε and centers $\{g_{0k}\}$, forming a grid of orientations on SO(3) [7]-[9]. The model ODF is

represented as a superposition of these distributions. The central normal distribution [2],[3] $f^G(g, g_0, \varepsilon)$ as a standard function is offered. For narrow peaks ($\varepsilon < 0.3$) Gaussian distribution has a good approximation [2]:

$$f^s(g, g_0, \varepsilon) \equiv f(\omega, \varepsilon) = \left[\frac{\sqrt{\pi}}{\varepsilon^3} \exp(\varepsilon^2 / 4) \operatorname{erfc}(\varepsilon / 2) + 1/\varepsilon^2 \right] \times \frac{\omega/2}{\sin(\omega/2)} \exp(-\omega^2/4\varepsilon^2) \quad (11)$$

with $\omega_k(g_0, g) = \arccos(0.5(\operatorname{Tr}(g_0^{-1}g) - 1))$. The distribution (11) generates the reduced standard pole density

$$P^s(\vec{h}, \vec{y}, g_0, \varepsilon) \equiv P(\theta, \varepsilon) = \frac{1}{2\varepsilon^2} \left[\exp(-\theta^2/4\varepsilon^2) + \exp(-(\pi - \theta)^2/4\varepsilon^2) \right], \quad (12)$$

where $\cos \theta$ is a scalar product of vectors \vec{h} and $g_0\vec{y}$: $\cos \theta = (\vec{h}g_0\vec{y})$.

Since the model ODF is a superposition of K standard functions (11) with volume fractions $c_k > 0$, the corresponding model PFs is a superposition of (12) with the same parameters

$$P(\vec{h}, \vec{y}) = \sum_{k=1}^K c_k P^s(\vec{h}, \vec{y}, g_{0k}, \varepsilon), \quad (13)$$

Here, the centers g_{0k} are not adjustable parameters. They form some regular grid in $SO(3)$. The choice of the dispersion parameter ε depends on the angular distance ψ between grid nodes. An initial value of ε can be based on the relation $\psi \approx 4\varepsilon\sqrt{\ln 2}$. The true value minimizes common residual (3).

The design matrix of the problem has elements $A_{sk} = P^G(\{\vec{h}, \vec{y}\}_s, g_{0k}, \varepsilon)$, and the solution of the problem (3) is obtained by the iterative method of projections [7]. The condition of the volume fraction positivity and normalization are taken into account during iteration. The result of this method always converges to the estimations of the least square method [10]. The solution with minimum norm ($\sum c_k^2 \rightarrow \min$) is obtained. This method produces complete nonnegative ODF.

5. Results and discussion

The methods discussed were applied for the sample of magnesium alloy Mg-4,5%Nd. Experimental pole figures $[0004]$, $[11\bar{2}0]$, $[10\bar{1}0]$, $[10\bar{1}1]$, $[10\bar{1}2]$, $[10\bar{1}3]$ are produced by x-ray scattering method. It is obtained by the component fit method that the texture can be adequately described (the value $RP(0.5)$ for the recalculated PFs is about 14%) by seven axial components with following parameters:

$$\begin{aligned} c_1 &= 29.8, \varepsilon_1 = 0.010, \vec{n}_1 = \{259.2^\circ, 23.0^\circ\}, \rho_1 = 0.94^\circ, c_2 = 18.7, \varepsilon_2 = 0.012, \vec{n}_2 = \{295.2^\circ, 30.1^\circ\}, \rho_2 = 2.17^\circ, \\ c_3 &= 13.3, \varepsilon_3 = 0.007, \vec{n}_3 = \{223.2^\circ, 32.6^\circ\}, \rho_3 = 6.86^\circ, c_4 = 8.7, \varepsilon_4 = 0.024, \vec{n}_4 = \{259.2^\circ, 43.7^\circ\}, \rho_4 = 6.62^\circ, \\ c_5 &= 6.3, \varepsilon_5 = 0.014, \vec{n}_5 = \{74.7^\circ, 36.5^\circ\}, \rho_5 = 6.84^\circ, c_6 = 4.9, \varepsilon_6 = 0.010, \vec{n}_6 = \{114.3^\circ, 37.3^\circ\}, \rho_6 = 6.81^\circ, \\ c_7 &= 3.4, \varepsilon_7 = 0.008, \vec{n}_7 = \{90.0^\circ, 32.7^\circ\}, \rho_7 = 3.20^\circ, c_0 = 14.9. \end{aligned}$$

In the orientation grid method the ODF model is represented by a superposition of 3000 standard peaks with $\psi = 8^\circ$. The value $RP(0.5)$ for the recalculated PFs is about 11%.

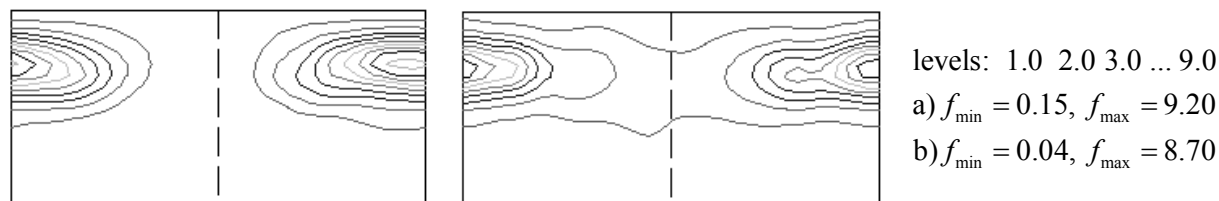


Figure 1. The sections of model ODF a) in the component fit method, b) in the orientation grid method. Maximum and minimum values of the model ODF in both methods and levels are given.

The Figure 1 represents zero section of the ODF function obtained by the aforementioned methods. Both methods lead to similar results, but non-texturized component's value in the orientation grid method is much smaller than the corresponding value in the component fit method, and the max value of ODF in grid method is less than the one in component fit method. This is due to the fact that the non-texturized component is approximated by a large number of standard functions. The following method development requires taking into account the adjustable constant c_0 , which is equal to the volume fraction of the non-texturized component. The component fit method is more stable, but definitely requires human assistance in order to select initial values of parameters and also the number and type of components. The orientation grid method is fully automatic, but unfortunately less stable. Moreover, it does not provide the researcher with information needed to investigate the texture development process.

References

- [1] Bunge H J 1982 *Texture Analysis in Material Science* (Butterworths:London)
- [2] Bucharova T I and Savyolova T I 1993 Application of normal distribution on SO(3) and S2 for orientation distribution function approximation *Text. Microstruct.* **23** 161–176
- [3] Nikolayev D I and Savyolova T I 1997 Normal distribution on the rotation group SO(3), *Text. Microstruct.* **29** 201–233
- [4] Helming K and Eschner T 1990 A new approach to texture analysis of multiphase materials using a texture component *Model Cryst. Res. Technol.* **25** K203–K208
- [5] Matthies S 1979 On the reproducibility of the orientation distribution function from pole figures (ghost phenomena) *phys. stat. sol. (b)* **92** N2 K135–137
- [6] Ivanova T M and Savelova T I 2006 Robust method of approximating the orientation distribution function by canonical normal distributions *Phys. Met. Metallog.* **101** N2 114–118
- [7] Kourtasov S F 2007 *Indust. Lab.* **73** N7 41–44
- [8] Savyolova T I and Kourtasov S F 2005 *Proc. Int. Conf. on Textures of Materials (Louvain)* (Switzerland: Trans Tech Publication) pp 301–306
- [9] Serebryany V N, Kourtasov S F and Savyolova T I 2005 *Proc. Int. Conf. on Textures of Materials (Louvain)* (Switzerland: Trans Tech Publication) pp 1693–1697
- [10] Kahaner D, Moler C, Nash S 1989 *Numerical Methods and Software* (Englewood Cliffs, NJ: Prentice Hall)
- [11] Golub G and Van Loan Ch 1996 *Matrix Computations* (Johns Hopkins University Press, Baltimore)
- [12] Dennis J E and Schnabel R B 1983 *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Englewood Cliffs, NJ: Prentice Hall)