Journal of
Applied Crystallography

ISSN 0021-8898

Received 20 December 2007
Accepted 18 September 2008

A novel pole figure inversion method: specification of the *MTEX* algorithm

R. Hielscher[a,b]* and H. Schaeben[a,b]

[a]Helmholtz Zentrum München, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany, and [b]Technische Universität Bergakademie Freiberg, Institut für Geologie, B. v. Cotta Strasse 2, 09599 Freiberg, Germany. Correspondence e-mail: ralf.hielscher@gsf.de

A novel algorithm for ODF (orientation density function) estimation from diffraction pole figures is presented which is especially well suited for sharp textures and high-resolution pole figures measured with respect to arbitrarily scattered specimen directions, *e.g.* by area detectors. The estimated ODF is computed as the solution of a minimization problem which is based on a model of the diffraction counts as a Poisson process. The algorithm applies discretization by radially symmetric functions and fast Fourier techniques to guarantee smooth approximation and high performance. An implementation of the algorithm is freely available as part of the texture analysis software *MTEX*.

1. Introduction

The analysis of crystallographic preferred orientations in polycrystalline specimens is a common technique in materials science (Bunge, 1982), geology (Wenk, 1985) and biology (Rodriguez-Navarro *et al.*, 2006). Mathematically, the statistical distribution of crystallographic orientations is modelled by the so-called 'orientation density function' (ODF), a function that is defined as the relative frequency of crystal orientations within a specimen by volume. Experimentally, one distinguishes between methods that measure individual crystallographic orientations directly, *e.g. via* optical methods or electron backscatter diffraction (EBSD), and methods based on neutron, synchrotron or X-ray diffraction which provide only integral information about the ODF called 'diffraction pole figures'. In the latter context, the mathematical and computational determination of the ODF from diffraction pole figures used to be a major challenge. This problem is often referred to as the 'pole figure inversion problem' (Bunge, 1969).

Various methods for the resolution of the pole figure inversion problem have been proposed, the earliest dating back to 1965 (*e.g.* Bunge, 1965; Roe, 1965; Pospiech & Jura, 1974; Jura *et al.*, 1976; Ruer, 1976; Matthies & Pospiech, 1980; Pospiech *et al.*, 1981; Van Houtte, 1983; Imhof, 1983; Bunge & Esling, 1985; Pawlik, 1986; Schaeben, 1988; Matthies, 1988; Helming & Eschner, 1990; Vadon & Heizmann, 1991; Van Houtte, 1991; van den Boogaart *et al.*, 2006; Bernier *et al.*, 2006; Hielscher, 2008). However, most of these methods, and their implementation in a particular software code, do not allow for high-resolution diffraction pole figures measured with respect to arbitrarily scattered specimen directions, *e.g.* by area detectors.

In fact, most of them apply to experimental intensities along a regular grid covering a specified part of the pole sphere, *e.g.* $\rho \in [0^\circ, 360^\circ]$, $\theta \in [0^\circ, 70^\circ]$, and require pre-processing by an interpolation or approximation method to transform irregularly spaced data to a regular grid with constant step size in ρ and θ, respectively. They are in no way suitable for analysing data of successively locally refined resolution as introduced by Schaeben *et al.* (2007). Moreover, most of the methods do not apply, computationally, to sharp textures, *e.g.* the classical harmonic method involving F coefficients (Bunge, 1965 1969, 1982). Other methods, *e.g.* eWIMV (Latteroti *et al.*, 1999), which is applied to sharp textures by decreasing the size of the orientational boxes to say $1 \times 1 \times 1^\circ$ or less, are not actually applicable for reasons of numerical complexity, ill-posedness or computational efficiency. A practical application of any of these methods or related software to a texture as pronounced as the hematite texture presented in this paper (see §4) has not yet been reported.

In this paper we present a novel method for pole figure inversion that allows for arbitrarily scattered pole figure data with a spatial resolution of up to 1°. This method is implemented as part of the freely available *MATLAB* toolbox for quantitative texture analysis, *MTEX* (Hielscher, 2008), and satisfies the following key requirements: (i) the estimated ODF is non-negative; (ii) the method allows for arbitrary crystal and diffraction geometries; (iii) the method allows for superposed pole figures; and (iv) the unknown normalization coefficients of the pole figures are determined simultaneously with the ODF.

Additionally, the method considers the measurement error affecting the diffraction data. In this paper, we model the measured diffraction intensities as a random sample of a Poisson process parameterized by the unknown ODF. Then we look for the best estimator of the unknown parameter given the observed diffraction intensities. Eventually we derive an estimator that differs from the regularized least-squares estimator (Bernier *et al.*, 2006) by the characteristic that devia-

tions from small diffraction intensities are more severely penalized than deviations from large diffraction intensities. In this way, our method accounts for the standard deviation of measurement errors which increase with the square root of the diffraction counts (Wenk, 1985, p. 37).

Another remarkable feature of our method is the discretization by radially symmetric functions, *i.e.* the unknown ODF is approximated by a linear combination of up to 100 000 unimodal bell-shaped standard ODFs. Approximation by radially symmetric functions is a well established method (*cf.* Buhmann, 2003; Hubbert & Baxter, 2001; Böhlke *et al.*, 2006; Schaeben, 1996, and references therein). It results in smooth yet possibly sharp ODFs for which the corresponding pole figures can be calculated explicitly. In order to perform computation with ODFs given as the linear combination of radially symmetric functions, numerically we apply fast Fourier techniques on the rotation group SO(3) and the sphere $\mathbb{S}^2$, *cf.* Kunis & Potts (2003) and Keiner *et al.* (2006). Our method allows us to set the harmonic degree up to about 500 which keeps the truncation error small.

The objective of this communication is to give a complete, concise and self-contained specification of the new pole figure inversion algorithm implemented in *MTEX*. For this reason, the paper starts with the definition of the most fundamental notations in quantitative texture analysis, *i.e.* orientation, Euler angles, and spherical and generalized spherical harmonics. Since our ODF estimator is defined as the solution of a minimization problem, much effort in this paper is devoted to the description of an adapted iterative solver. Furthermore, this communication includes specifications of algorithms that compute the Fourier coefficients, the pole figures and the volume portions of the estimated ODF.

A single example with a fairly sharp texture is presented in §4. There are no example computations, applications or numerical tests included in this communication – numerical results are given by Schaeben *et al.* (2007) and Hielscher (2007). An extensive numerical analysis of the algorithm as well as a comparison with existing algorithms will be part of a forthcoming paper.

2. Notations and general setting

2.1. Crystallographic background

We commence with the definition of the orientation of a crystal within a polycrystalline specimen. Let $\mathcal{K}_S = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ be a right-handed orthogonal specimen coordinate system, and let $\mathcal{K}_C = \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ be a right-handed orthogonal crystal coordinate system. Then, we call a rotation $\mathbf{g} \in \mathrm{SO}(3)$ the *orientation* of the crystal if it rotates the specimen coordinate system onto the crystal coordinate system, *i.e.* $\mathbf{gx} = \mathbf{a}$, $\mathbf{gy} = \mathbf{b}$, $\mathbf{gz} = \mathbf{c}$. Let $\mathbf{r} = (u, v, w)^T$ be a coordinate vector with respect to the specimen coordinate systems and let $\mathbf{h} = (h, k, l)^T$ be the corresponding coordinate vector with respect to the crystal coordinate system, *i.e.* both coordinate vectors represent the same direction, and we have

$$u\mathbf{x} + v\mathbf{y} + w\mathbf{z} = h\mathbf{a} + k\mathbf{b} + l\mathbf{c}. \tag{1}$$

Then, the orientation $\mathbf{g} \in \mathrm{SO}(3)$ interpreted as a matrix realizes the basis transformation between the coordinate systems, and we have the equation

$$\mathbf{gh} = \mathbf{r}. \tag{2}$$

Since the crystal coordinate system can be assigned to the crystal only modulo crystal symmetry, every orientation $\mathbf{g} \in \mathrm{SO}(3)$ is associated with a whole class of crystallographically equivalent orientations. Crystallographic symmetries are commonly described by symmetry groups. When analysing diffraction data for preferred crystallographic orientation, it is sufficient to consider the restriction of the Laue group $G_{\mathrm{Laue}} \subset \mathrm{O}(3)$ to its purely rotational part $\overline{G}_{\mathrm{Laue}} = G_{\mathrm{Laue}} \cap \mathrm{SO}(3)$. Two orientations $\mathbf{g}, \mathbf{g}' \in \mathrm{SO}(3)$ are described as 'crystallographically equivalent' if there is a symmetry element $\mathbf{q} \in \overline{G}_{\mathrm{Laue}}$ such that $\mathbf{gq} = \mathbf{g}'$. Analogously, two crystallographic directions $\mathbf{h}, \mathbf{h}' \in \mathbb{S}^2$ are termed 'crystallographically equivalent' if there is a symmetry element $\mathbf{q} \in \overline{G}_{\mathrm{Laue}}$ such that $\mathbf{qh} = \mathbf{h}'$.

According to Bunge (1982), we define the ODF of a specimen as the function

$$f : \mathrm{SO}(3) \to \mathbb{R} \tag{3}$$

which models the relative frequencies of crystal orientations within the specimen by volume. The ODF possesses the symmetry property

$$f(\mathbf{g}) = f(\mathbf{gq}), \ \ \mathbf{g} \in \mathrm{SO}(3), \ \mathbf{q} \in \overline{G}_{\mathrm{Laue}} \tag{4}$$

and is normalized to

$$\int_{\mathrm{SO}(3)} f(\mathbf{g}) \, \mathrm{d}\mathbf{g} = 8\pi^2, \tag{5}$$

where $\mathrm{d}\mathbf{g}$ denotes the rotational-invariant measure on SO(3). The axis distribution function (Bunge, 1982) or pole density function (PDF) (Matthies *et al.*, 1987) of a specimen is defined as the function

$$P : \mathbb{S}^2 \times \mathbb{S}^2 \to \mathbb{R} \tag{6}$$

which models the relative frequencies of lattice plane orientations, *i.e.* the relative frequencies of normal vectors, within the specimen by volume. Mathematically, the PDF P corresponding to an ODF f is characterized by the 'fundamental equation of texture analysis' (Bunge, 1982, §4.2)

$$\begin{aligned} P(\mathbf{h}, \mathbf{r}) &= \frac{1}{2}[\mathcal{R}f(\mathbf{h}, \mathbf{r}) + \mathcal{R}f(-\mathbf{h}, \mathbf{r})], \\ \mathcal{R}f(\mathbf{h}, \mathbf{r}) &= \frac{1}{2\pi} \int_{G(\mathbf{h}, \mathbf{r})} f(\mathbf{g}) \, \mathrm{d}\mathbf{g}, \end{aligned} \tag{7}$$

where the path of integration $G(\mathbf{h}, \mathbf{r}) := \{\mathbf{g} \in \mathrm{SO}(3) \mid \mathbf{gh} = \mathbf{r}\}$ is defined as the set of all rotations that map the crystallographic direction $\mathbf{h} \in \mathbb{S}^2$ onto the specimen direction $\mathbf{r} \in \mathbb{S}^2$. The fundamental equation of texture analysis involves the integral operator $\mathcal{R}$, recently recognized as the totally geodesic Radon transform (Schaeben & van den Boogaart, 2003). The properties [equations (4) and (5)] of an ODF f imply the following properties of the corresponding PDF P,

$$P(\mathbf{h}, \mathbf{r}) = P(\mathbf{qh}, \mathbf{r}), \;\; \mathbf{h}, \mathbf{r} \in \mathbb{S}^2, \; \mathbf{q} \in \overline{G}_{\text{Laue}}, \tag{8}$$

and

$$\int_{\mathbb{S}^2} P(\mathbf{h}, \mathbf{r}) \, \mathrm{d}\mathbf{h} = \int_{\mathbb{S}^2} P(\mathbf{h}, \mathbf{r}) \, \mathrm{d}\mathbf{r} = 4\pi. \tag{9}$$

The relationship between an ODF f and its Radon transform $\mathcal{R}f$ is one to one, whereas the PDF P determines only the even-order Fourier coefficients of the corresponding ODF f. The odd-order Fourier coefficients are only partly accessible by the non-negativity property of the ODF (Matthies, 1979) or other additional modelling assumptions.

2.2. Diffraction at polycrystalline materials

A detailed introduction to diffraction at a single crystal is given by Forsyth (1988) and Randle & Engler (2000, §3). For diffraction at polycrystalline materials, we refer readers to Randle & Engler (2000, §4). Here, we only give the basic facts necessary for a simple mathematical model upon which our inversion algorithm is based.

For any wavelength λ and any Bragg angle θ we denote by $H(\lambda, \theta) \subset \mathbb{S}^2$ the set of normal vectors to lattice planes responsible for diffraction under the specific combination of parameters. The set $H(\lambda, \theta)$ is symmetric with respect to the restricted Laue group $\overline{G}_{\text{Laue}} \subset \mathrm{SO}(3)$, *i.e.* with $\mathbf{h} \in H(\lambda, \theta)$ we have $\mathbf{gh} \in H(\lambda, \theta)$ for any symmetry element $\mathbf{g} \in \overline{G}_{\text{Laue}}$. Additionally, Friedel's law implies $-\mathbf{h} \in H(\lambda, \theta)$ for any $\mathbf{h} \in H(\lambda, \theta)$.

In most cases, the set $H(\lambda, \theta)$ contains at most one class of crystallographically equivalent directions. However, especially for lower-order crystal symmetries, superposition of beams diffracted by different lattice planes occurs. In this case the structure coefficients $\rho(\mathbf{h})$ of the diffracting lattice planes $\mathbf{h} \in H(\lambda, \theta)$ describe the superposition of the reflection intensities with respect to the lattice planes in $H(\lambda, \theta)$ (Randle & Engler, 2000, §3.3).

Multiplying the structure coefficients $\rho(\mathbf{h})$ with the PDF $P(\mathbf{h}, \mathbf{r})$, *i.e.* with the relative volume of all crystal orientations satisfying Bragg's law with respect to a fixed specimen direction $\mathbf{r} \in \mathbb{S}^2$, we obtain the relative intensity of the beam diffracted at the lattice plane $\mathbf{h} \in H(\lambda, \theta)$. Summing over all $\mathbf{h} \in H(\theta, \lambda)$ we achieve a normalized deterministic model for the diffraction intensity $I_{\text{norm}}(\lambda, \theta, \mathbf{r})$ caused by a polycrystalline specimen given a certain wavelength λ, a Bragg angle θ and a specimen direction $\mathbf{r}$,

$$I_{\text{norm}}(\lambda, \theta, \mathbf{r}) = \left[\sum_{\mathbf{h} \in H(\lambda, \theta)} \rho(\mathbf{h}) \right]^{-1} \sum_{\mathbf{h} \in H(\lambda, \theta)} \rho(\mathbf{h}) P(\mathbf{h}, \mathbf{r}). \tag{10}$$

In practice, normalized diffraction intensities are not accessible and only particle counts are measured, which are affected by background radiation and measurement errors and differ from the normalized intensities by an unknown normalization factor. Let us denote the known background radiation by $I^b(\lambda, \theta, \mathbf{r})$ and the unknown normalization factor by $\nu(\lambda, \theta)$. Then we have, for the absolute intensity,

$$I_{\text{abs}}(\lambda, \theta, \mathbf{r}) = \nu(\lambda, \theta) \sum_{\mathbf{h} \in H(\lambda, \theta)} \rho(\mathbf{h}) P(\mathbf{h}, \mathbf{r}) + I^b(\lambda, \theta, \mathbf{r}). \tag{11}$$

Since particle counts can be modelled with the Poisson distribution with its mean value set to the absolute intensity (Randle & Engler, 2000, §4.3.5), we obtain, for an individual diffraction measurement, the statistical model

$$I(\lambda, \theta, \mathbf{r}) \simeq \mathrm{Pois}\left[\nu(\lambda, \theta) \sum_{\mathbf{h} \in H(\lambda, \theta)} \rho(\mathbf{h}) P(\mathbf{h}, \mathbf{r}) + I^b(\lambda, \theta, \mathbf{r}) \right], \tag{12}$$

characterizing it as a one-element random sample of a parameterized Poisson process.

It should be noted that equation (12) is only a very simple model for experimental diffraction counts. First of all, Bragg's law itself is only a rough simplification of much more sophisticated models explaining diffraction (*e.g.* Cowley, 1995). Second, the diffraction counts commonly used for texture determination are obtained by processing a spectrum of diffraction counts for varying Bragg angles θ or wavelengths λ (Hammond, 1997; Randle & Engler, 2000, §4). However, even the simple model implies that the standard deviation of diffraction counts is approximately

$$\sigma_{I(\lambda, \theta, \mathbf{r})} = \left[I_{\text{abs}}(\lambda, \theta, \mathbf{r}) \right]^{1/2} \simeq \left[I(\lambda, \theta, \mathbf{r}) \right]^{1/2}, \tag{13}$$

(*cf.* Wenk, 1985, p. 37).

2.3. The diffraction experiment

In a complete diffraction experiment for the purpose of ODF determination, diffraction counts with respect to several diffraction parameters and several specimen directions are measured. We assume the parameters and the data are organized in the following way. We enumerate the sequence of diffraction parameters (λ_i, θ_i) and the corresponding sequence of lattice planes $H_i = H(\lambda, \theta) \subset \mathbb{S}^2$ by the index $i = 1, \ldots, N$. If no superposed pole figures are measured, the sequence H_i, $i = 1, \ldots, N$, corresponds simply to the sequence of measured lattice planes. For each pair (λ_i, θ_i), $i = 1, \ldots, N$, of a wavelength and a Bragg angle, we enumerate the specimen directions $\mathbf{r}_{ij} \in \mathbb{S}^2$ with respect to which diffraction intensities are measured by the indices $j = 1, \ldots, N_i$. Correspondingly, we abbreviate the diffraction counts by $I_{ij} = I(\lambda_i, \theta_i, \mathbf{r}_{ij})$ and the background intensities by $I^b_{ij} = I^b(\lambda_i, \theta_i, \mathbf{r}_{ij})$. The indexing ij, $i = 1, \ldots, N$, $j = 1, \ldots, N_i$, is in fact short for ij_i. Thus, it in no way implies any specific grid but applies to arbitrarily scattered specimen directions in each pole figure.

In order to derive a concise representation, we abbreviate the unknown normalization coefficients by the vector $\boldsymbol{\nu}_{\text{true}} \in \mathbb{R}^N_+$, $[\boldsymbol{\nu}_{\text{true}}]_i = \nu(\lambda_i, \theta_i)$, $i = 1, \ldots, N$, and collect the diffraction counts I_{ij} in the single vector

$$\mathbf{I} = \Big(\underbrace{I_{11}, \ldots, I_{1N_1}}_{\mathbf{I}_1^T \in \mathbb{R}^{N_1}}, \underbrace{I_{21}, \ldots, I_{2N_2}}_{\mathbf{I}_2^T \in \mathbb{R}^{N_2}}, \ldots, \underbrace{I_{N1}, \ldots, I_{NN_N}}_{\mathbf{I}_N^T \in \mathbb{R}^{N_N}} \Big)^T \in \mathbb{R}^{\overline{N}}. \tag{14}$$

The diffraction counts $\mathbf{I}_i = (I_{i1}, \ldots, I_{iNi})^T \in \mathbb{R}^{N_i}_+$ corresponding to the ith set of lattice planes H_i are called the ith pole figure.

Table 1
List of parameters of a diffraction experiment.

Symbol	Description
$N \in \mathbb{N}$	Number of pole figures
$N_i \in \mathbb{N}, i = 1, \ldots, N$	Number of specimen directions
$G_{\text{Laue}} \subset \text{SO}(3)$	Laue group
$H_i = H(\lambda_i, \theta_i) \subset \mathbb{S}^2, i = 1, \ldots, N$	Superposed lattice planes
$\rho_i : H_i \rightarrow \mathbb{R}+$	Structure coefficients
$\mathbf{r}_{ij} \in \mathbb{S}^2, i = 1 \ldots, N, j = 1, \ldots, N_i$	Specimen directions
$I_{ij} \in \mathbb{R}_+, i = 1 \ldots, N, j = 1, \ldots, N_i$	Diffraction counts
$I^b_{ij} \in \mathbb{R}_+, i = 1 \ldots, N, j = 1, \ldots, N_i$	Background intensities

In our notation, we have N pole figures with a total number of $\overline{N} = \sum_{i=1}^{N} N_i$ measured diffraction data.

Introducing finally for any ODF $f : \text{SO}(3) \rightarrow \mathbb{R}$ the notation

$$\mathcal{R}f(H_i, \mathbf{r}_{ij}) = \sum_{\mathbf{h} \in H_i} \rho_i(\mathbf{H}) \mathcal{R}f(H_i, \mathbf{r}_{ij}), \tag{15}$$

we rewrite our statistical model [equation (12)] for a complete diffraction experiment as

$$\begin{aligned} I_{ij} &\simeq \text{Pois}\left[I^b_{ij} + [\boldsymbol{\nu}_{\text{true}}]_i \, \mathcal{R}f_{\text{true}}(H_i, \mathbf{r}_{ij})\right], \\ & i = 1, \ldots, N, \;\; j = 1, \ldots, N, \end{aligned} \tag{16}$$

using the fundamental equation of texture analysis [equation (7)].

A complete overview of all parameters of a diffraction experiment is given in Table 1. From the point of view of the statistical model [equation (16)], the measured diffraction counts $\mathbf{I} \in \mathbb{R}^{\overline{N}}$ occur as a one-element random sample of a parameterized Poisson process. Hence, the objective of quantitative texture analysis is to estimate the unknown ODF f_{true} and unknown normalization coefficients $\boldsymbol{\nu}_{\text{true}}$ from the random sample $\mathbf{I}$.

2.4. Harmonic functions on $\mathbb{S}^2$ and SO(3)

An important tool for the analysis of ODFs and PDFs is harmonic functions on the two-dimensional sphere $\mathbb{S}^2$ and on the rotation group SO(3). Given the series expansion of an ODF into harmonic functions on SO(3), the corresponding PDF has a well known series expansion in terms of spherical harmonics (Bunge, 1982, §4.2). Furthermore, its Fourier coefficients are instrumental in calculating the mean macroscopic properties of the corresponding specimen, *e.g.* its thermal expansion, optical refraction index, electrical conductivity or elastic properties.

We render an explicit definition of harmonics as there are many slightly different ways to define them, *e.g.* with respect to normalization, the disastrous impact of which is only revealed in the course of writing and checking software code. Therefore, it is our hope that the reader and practitioner of texture analysis and open-source software development will appreciate a comprehensive and consistent view of a method with unique features.

Harmonic analysis on the sphere is based on the Legendre polynomials $\mathcal{P}_l : [-1, 1] \rightarrow \mathbb{R}, l \in \mathbb{N}_0$, where

$$\mathcal{P}_l(t) = (1/2^l l!)\,(\mathrm{d}^l/\mathrm{d}t^l)[(t^2 - 1)^l] \tag{17}$$

and on the associated Legendre functions $\mathcal{P}^k_l : [-1, 1] \rightarrow \mathbb{R}, l \in \mathbb{N}_0, \; k = 0, \ldots, l$,

$$\mathcal{P}^k_l(t) = [(l-k)!/(l+k)!]^{1/2}\,(1 - t^2)^{k/2}\,(\mathrm{d}^k/\mathrm{d}t^k)\mathcal{P}_l(t). \tag{18}$$

In terms of the associated Legendre functions, we define the spherical harmonics $\mathcal{Y}^k_l(\mathbf{r}), l \in \mathbb{N}_0, k = -l, \ldots, l$, by

$$\mathcal{Y}^k_l(\mathbf{r}) = [(2l+1)/4\pi]^{1/2}\,\mathcal{P}^{|k|}_l(\cos\theta)\exp(ik\rho), \tag{19}$$

where $\theta, \rho \in \mathbb{R}$ are the polar coordinates $\rho \in [0, 2\pi], \theta \in [0, \pi]$ of the vector

$$\mathbf{r} = (\cos\rho\sin\theta, \sin\rho\sin\theta, \cos\theta)^T \in \mathbb{S}^2. \tag{20}$$

By this definition, the spherical harmonics are normed to

$$\int_{\mathbb{S}^2} \mathcal{Y}^k_l(\mathbf{r})\overline{\mathcal{Y}^{k'}_{l'}(\mathbf{r})}\,\mathrm{d}\mathbf{r} = \int_{\mathbb{S}^2} \mathcal{Y}^k_l(\theta, \rho)\overline{\mathcal{Y}^{k'}_{l'}(\theta, \rho)}\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\rho = \delta_{ll'}\delta_{kk'} \tag{21}$$

and, hence, provide an orthonormal basis in $L^2(\mathbb{S}^2)$.

In order to define harmonic functions on SO(3) we use the parameterization of a rotation $\mathbf{g} \in \text{SO}(3)$ in terms of Euler angles by Matthies *et al.* (1987):

$$\begin{aligned} \mathbf{g} = \mathbf{g}(\alpha, \beta, \gamma) &= \text{Rot}(\mathbf{z}, \alpha)\,\text{Rot}(\mathbf{y}, \beta)\,\text{Rot}(\mathbf{z}, \gamma), \\ & \alpha, \gamma \in [0, 2\pi], \beta \in [0, \pi], \end{aligned} \tag{22}$$

where $\text{Rot}(\mathbf{r}, \omega)$ denotes the rotation about the axis $\mathbf{r} \in \mathbb{S}^2$ and angle $\omega \in [0, \pi]$. Now we follow Nikiforov & Uvarov (1988) (see also Varshalovich *et al.*, 1988; Kostelec & Rockmore, 2003; Vollrath, 2006) and define for $l \in \mathbb{N}_0, k, k' = -l, \ldots, k$, the generalized spherical harmonics or Wigner D functions as

$$D^{kk'}_l(\alpha, \beta, \gamma) = \exp(-ik\alpha)\,d^{kk'}_l(\cos\beta)\exp(-ik'\gamma), \tag{23}$$

where

$$\begin{aligned} d^{kk'}_l(t) = s_{kk'}\,\frac{(-1)^{l-k'}}{2^l}\left[\frac{(l+k')!}{(l-k')!(l+k)!(l-k)!}\right]^{1/2} \\ \times \left[\frac{(1-t)^{k-k'}}{(1+t)^{k+k'}}\right]^{1/2} \frac{\mathrm{d}^{l-k'}}{\mathrm{d}t^{l-k'}}\,(1-t)^{l-k}(1+t)^{l+k} \end{aligned} \tag{24}$$

and

$$s_{kk'} = \begin{cases} 1 & k, k' \geq 0, \\ (-1)^k & k' \geq 0, k<0, \\ (-1)^{k'} & k \geq 0, k'<0, \\ (-1)^{k+k'} & k, k'<0. \end{cases} \tag{25}$$

The last term $s_{kk'}$ corrects for the normalization we used for the spherical harmonics, and is slightly different from that used by Nikiforov & Uvarov (1988). The Wigner D functions are orthogonal in $L^2[\text{SO}(3)]$, *i.e.*

$$\begin{aligned} \int_0^{2\pi}\int_0^{\pi}\int_0^{2\pi} D^{mn}_l(\alpha, \beta, \gamma)\overline{D^{m'n'}_{l'}(\alpha, \beta, \gamma)}\,\mathrm{d}\alpha\,\sin\beta\,\mathrm{d}\beta\,\mathrm{d}\gamma \\ = \frac{8\pi^2}{2l+1}\,\delta_{ll'}\delta_{mm'}\delta_{nn'}, \end{aligned} \tag{26}$$

and are related to the spherical harmonics by the representation property

$$\sum_{k=-l}^{k} D_l^{kk'}(\mathbf{g})\mathcal{Y}_l^{k'}(\mathbf{h}) = \mathcal{Y}_l^{k}(\mathbf{gh}),\ \ \mathbf{g} \in \mathrm{SO}(3), \mathbf{h} \in \mathbb{S}^2. \tag{27}$$

Moreover, any ODF $f \in L^2$ [SO(3)] has an associated Fourier expansion of the form

$$f \simeq \sum_{l=0}^{\infty} \sum_{kk'=-l}^{l} \frac{(l+\frac{1}{2})^{1/2}}{2\pi} \hat{f}(l,k,k') D_l^{kk'}, \tag{28}$$

with Fourier coefficients $\hat{f}(l,k,k'), l \in \mathbb{N}_0, k, k' = -l, \ldots, l$.

The following well known central theorem characterizes the relationship between the Fourier expansion of an ODF and its corresponding PDF.

Theorem 1. Let $f \in L^2$ [SO(3)] be an ODF with Fourier expansion

$$f \simeq \sum_{l=0}^{\infty} \sum_{kk'=-l}^{l} \frac{(l+\frac{1}{2})^{1/2}}{2\pi} \hat{f}(l,k,k') D_l^{kk'}. \tag{29}$$

Then the corresponding PDF $P = \in L^2(\mathbb{S}^2 \times \mathbb{S})$, $P(\mathbf{h},\mathbf{r}) = (1/2)[\mathcal{R}f(\mathbf{h},\mathbf{r}) + \mathcal{R}f(-\mathbf{h},\mathbf{r})]$ possesses the associated Fourier expansion

$$P(\mathbf{h},\mathbf{r}) \simeq \sum_{l \in 2\mathbb{N}_0} \sum_{k,k'=-l}^{l} \frac{1}{(l+\frac{1}{2})^{1/2}} \hat{f}(l,k,k') \mathcal{Y}_l^{k'}(\mathbf{h})\, \overline{\mathcal{Y}_l^{k}(\mathbf{r})}. \tag{30}$$

2.5. Fourier transforms

The operator that transforms any function into its corresponding Fourier sequence is generally called a Fourier transform. However, on the domains $\mathbb{S}^2$ and SO(3) a discrete Fourier transform in the above sense requires a quadrature rule for the specific set of nodes. That is the reason why, in numerical harmonic analysis, the term Fourier transform is often used for the inverse operator (Kunis & Potts, 2003). We adapt this denomination and define the discrete Fourier transform on the sphere $\mathbb{S}^2$ as the operator

$$\begin{aligned} &\mathbf{F}_{\mathbf{r},L} : \mathbb{C}^{(2L+1)^2} \to \mathbb{C}^N, \\ &[\mathbf{F}_{\mathbf{r},L}\hat{\mathbf{P}}]_j = \sum_{l=0}^{L} \sum_{k=-l}^{l} \hat{P}_{lk} \mathcal{Y}_l^k(\mathbf{r}_j),\ j = 1, \ldots N, \end{aligned} \tag{31}$$

which evaluates the Fourier series given by the Fourier coefficients $\hat{\mathbf{P}} \in \mathbb{C}^{(2L+1)^2}$ with bandwidth, *i.e.* series expansion degree, $L \in \mathbb{N}$ at the vector $\mathbf{r}$ of nodes $\mathbf{r}_j \in \mathbb{S}^2, j = 1, \ldots, N$.

Analogously, we define the discrete Fourier transform on the rotation group SO(3) as the operator

$$\begin{aligned} &\mathbf{F}_{\mathbf{g},L} : \mathbb{C}^{1/3(L+1)(2L+1)(2L+3)} \to \mathbb{C}^M, \\ &[\mathbf{F}_{\mathbf{g},L}\hat{\mathbf{f}}]_m = \sum_{l=0}^{L} \sum_{k,k'=-l}^{l} \frac{(l+\frac{1}{2})^{1/2}}{2\pi} \hat{f}_{lkk'} D_l^{k,k'}(\mathbf{g}_m),\ m = 1, \ldots, N, \end{aligned} \tag{32}$$

where $\mathbf{g} = (\mathbf{g}_1, \ldots, \mathbf{g}_M)$ denotes a vector of $M \in \mathbb{N}$ arbitrary nodes $\mathbf{g}_i \in \mathrm{SO}(3)$ and $\hat{\mathbf{f}} \in \mathbb{C}^{1/3(L+1)(2L+1)(2L+3)}$ denotes a vector of Fourier coefficients with bandwidth $L \in \mathbb{N}_0$.

For our algorithms we will also need the adjoint Fourier transforms $\mathbf{F}_{L,\mathbf{r}}^H$ and $\mathbf{F}_{L,\mathbf{g}}^H$ which are defined as

$$\mathbf{F}_{L,\mathbf{r}}^H : \mathbb{C}^N \to \mathbb{C}^{(2L+1)^2},\ [\mathbf{F}_{L,\mathbf{r}}^H \mathbf{c}]_{l,k} = \sum_{j=1}^{N} c_j \overline{\mathcal{Y}_l^k(\mathbf{r}_j)} \tag{33}$$

and

$$\begin{aligned} &\mathbf{F}_{L,\mathbf{g}}^H : \mathbb{C}^N \to \mathbb{C}^{1/3(L+1)(2L+1)(2L+3)}, \\ &[\mathbf{F}_{L,\mathbf{g}}^H \mathbf{c}]_{lkk'} = \frac{(l+\frac{1}{2})^{1/2}}{2\pi} \sum_{j=1}^{N} c_j \overline{D_l^{kk'}(\mathbf{g}_j)}, \end{aligned} \tag{34}$$

2.6. Radially symmetric functions

In texture analysis, radially symmetric functions appear as unimodal bell-shaped model ODFs. Mathematically, they are defined as functions $\psi : \mathrm{SO}(3) \to \mathbb{R}$ or $\varphi : \mathbb{S}^2 \to \mathbb{R}$ that depend only on the distance to a centre rotation $\mathbf{g}_0 \in \mathrm{SO}(3)$ or a centre direction $\mathbf{r}_0 \in \mathbb{S}^2$, respectively, *i.e.* we have

$$\psi(\mathbf{g}) = \psi(\mathbf{g}') \text{ and } \varphi(\mathbf{r}) = \varphi(\mathbf{r}') \tag{35}$$

for all rotations $\mathbf{g}, \mathbf{g}' \in \mathrm{SO}(3)$ with $\measuredangle(\mathbf{g}\mathbf{g}_0^{-1}) = \measuredangle(\mathbf{g}'\mathbf{g}_0^{-1})$ and all directions $\mathbf{r}, \mathbf{r}' \in \mathbb{S}^2$ with $\measuredangle(\mathbf{r}, \mathbf{r}_0) = \measuredangle(\mathbf{r}', \mathbf{r}_0)$. Here, $\measuredangle(\mathbf{g}\mathbf{g}_0^{-1})$ denotes the rotational angle of the rotation $\mathbf{g}, \mathbf{g}_0^{-1}$.

Both radially symmetric functions on the rotation group as well as on the sphere have characteristic Fourier series expansions. More precisely, there exist Chebyshev coefficients $\hat{\psi}(l)$ and Legendre coefficients $\hat{\psi}(l), l \in \mathbb{N}$, respectively, such that

$$\begin{aligned} \psi(\mathbf{g}) &\simeq \sum_{l=0}^{\infty} \hat{\psi}(l) \sum_{k,k'=-l}^{l} D_l^{kk'}(\mathbf{g}) \overline{D_l^{kk'}(\mathbf{g}_0)} \\ &\simeq \sum_{l=0}^{\infty} \hat{\psi}(l) \mathcal{U}_{2l}\left[\cos\frac{\measuredangle(\mathbf{g}\mathbf{g}_0^{-1})}{2}\right] \end{aligned} \tag{36}$$

and

$$\begin{aligned} \varphi(\mathbf{r}) &\simeq \sum_{l=0}^{\infty} \hat{\varphi}(l) \frac{4\pi}{2l+1} \sum_{k=-l}^{l} \mathcal{Y}_l^k(\mathbf{r}) \overline{\mathcal{Y}_l^k(\mathbf{r}_0)} \\ &\simeq \sum_{l=0}^{\infty} \hat{\varphi}(l) \mathcal{P}_l(\mathbf{r} \cdot \mathbf{r}_0). \end{aligned} \tag{37}$$

Here $\mathcal{U}_l, l \in \mathbb{N}$, denote the Chebyshev polynomials of the second kind.

It is well known that all pole figures of radially symmetric ODFs are radially symmetric. In particular, the Chebyshev coefficients $\hat{\psi}(l)$ of a radially symmetric ODF ψ coincide with the Legendre coefficients of its Radon transform $\mathcal{R}\psi(\mathbf{h},\cdot)$, *i.e.*

$$\mathcal{R}\psi(\mathbf{h},\mathbf{r}) \simeq \sum_{l=0}^{\infty} \hat{\psi}(l) \mathcal{P}_l(\mathbf{g}_0\mathbf{h} \cdot \mathbf{r}),\ \mathbf{h}, \mathbf{r} \in \mathbb{S}^2. \tag{38}$$

Example 2. An example of a well localized, non-negative, radially symmetric function on SO(3) is the de la Vallée Poussin kernel (Schaeben, 1997). It is given for any $k \in \mathbb{N}$ by

$$\psi(\mathbf{g}) = \frac{B(\frac{3}{2}, \frac{1}{2})}{B(\frac{3}{2}, \kappa + \frac{1}{2})} \cos^{2\kappa} \frac{\measuredangle \mathbf{g}}{2}. \tag{39}$$

Its Radon transform is calculated as

$$\mathcal{R}\psi(\mathbf{h}, \mathbf{r}) = \frac{1+\kappa}{2^\kappa} (1 + \mathbf{h} \cdot \mathbf{r})^\kappa = (1+\kappa) \cos^{2\kappa} \measuredangle(\mathbf{h}, \mathbf{r}). \tag{40}$$

The parameter κ controls the halfwidth of the kernel. The de la Vallée Poussin kernel ψ, its Radon transform $\mathcal{R}\psi$ and its Chebyshev coefficients $\hat{\psi}$ are plotted in Fig. 1.

3. Numerical ODF estimation

In this section we present an estimator of the unknown ODF that is based on the statistical model [equation (16)] and discuss its numerical implementation. Throughout all of this section we denote by $\overline{G}_{\text{Laue}} \subset \text{SO}(3)$, $\mathbf{I}_i$, $\mathbf{I}_i^b \in \mathbb{R}_+^{N_i}$, $H_i \subset \mathbb{S}^2$, $\rho_i : H_i \to \mathbb{R}$ and $\mathbf{r}_i = (\mathbf{r}_{i1}, \ldots, \mathbf{r}_{iN_i})$, $\mathbf{r}_{ij} \in \mathbb{S}^2$, $i = 1, \ldots, N$, the known parameters of a diffraction experiment as summarized in Table 1.

3.1. The weighted least-squares estimator

In §2 we have shown that the pole figure inversion problem can be interpreted as a classical parameter estimation problem, with the ODF and the normalization coefficients as the unknown parameters, given the random sample $I_{ij}, i = 1, \ldots, N, j = 1, \ldots, N_i$ of diffraction counts. We consider two popular estimators – the maximum likelihood estimator

$$(f_{\text{ME}}, \boldsymbol{\nu}_{\text{ME}}) = \text{argmin} \sum_{i=1}^{N} \sum_{j=1}^{N_i} I_{ij} \ln[\nu_i \mathcal{R}f(H_i, \mathbf{r}_{ij}) + I_{ij}^b] - \nu_i \mathcal{R}f(H_i, \mathbf{r}_{ij}) \tag{41}$$

and the least-squares (LS) estimator

$$(f_{\text{LS}}, \boldsymbol{\nu}_{\text{LS}}) = \text{argmin} \sum_{i=1}^{N} \sum_{j=1}^{N_i} I_{ij}^{-1} [\nu_i \mathcal{R}f(H_i, \mathbf{r}_{ij}) + I_{ij}^b - I_{ij}]^2, \tag{42}$$

where both minimization problems are subject to the constraints

$$\boldsymbol{\nu} \geq 0, \ f \geq 0 \text{ and } \int_{\text{SO}(3)} f(\mathbf{g}) \, \mathrm{d}\mathbf{g} = 8\pi^2. \tag{43}$$

Both estimators may be regularized by adding a Sobolev norm $\lambda ||f||_{\mathcal{H}[\text{SO}(3)]}$ of the ODF f as a penalty term. This penalty term can be interpreted as the prior information on the unknown ODF to be smooth.

Because of the unknown normalization coefficients, $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_N)$ are neither linear nor convex and hence both minimization problems may have several solutions (Hielscher, 2007). A second problem inherent in the minimization problems [equations (41) and (42), respectively] is that common algorithms result in solutions that tend to be unstable and largely dependent on the initial guess of the unknown parameters f and $\boldsymbol{\nu}$ (Hielscher, 2007). A more robust estimator is obtained when the unknown normalization coefficients $\nu_i, i = 1, \ldots, N$, in the LS estimator are replaced by their quadrature rule estimators

$$\tilde{\nu}_i(f) = \left(\sum_{j=1}^{N_i} I_{ij} - I_{ij}^b \right) \Big/ \left[\sum_{j=1}^{N_i} \mathcal{R}f(H_i, \mathbf{r}_{ij}) \right]. \tag{44}$$

We call the resulting estimator

$$f_{\text{MLS}} = \text{argmin} \sum_{i=1}^{N} \sum_{j=1}^{N_i} \left\{ [\tilde{\nu}_i(f) \, \mathcal{R}f(H_i, \mathbf{r}_{ij}) + I_{ij}^b - I_{ij}]^2 / I_{ij} \right\} + \lambda ||f||_{\mathcal{H}[\text{SO}(3)]}^2 \tag{45}$$

the modified least-squares estimator.

The representation of the ODF estimation problem as a least-squares problem which includes the unknown normalization coefficients as unknown variables dates back to Houtte (1983). Weighing the least-squares functionals [equations (42) and (45)] with the inverse expected variance I_{ij}^{-1} of the measurement error [equation (13)] we ensure the homoscedasticity of the underlying regression problem. The regularization term $\lambda ||f||_{\mathcal{H}[\text{SO}(3)]}$ was first suggested by Bernier *et al.* (2006) and van den Boogaart *et al.* (2006), and can be interpreted as a model assumption on the true ODF that biases the estimator towards smoother ODFs.

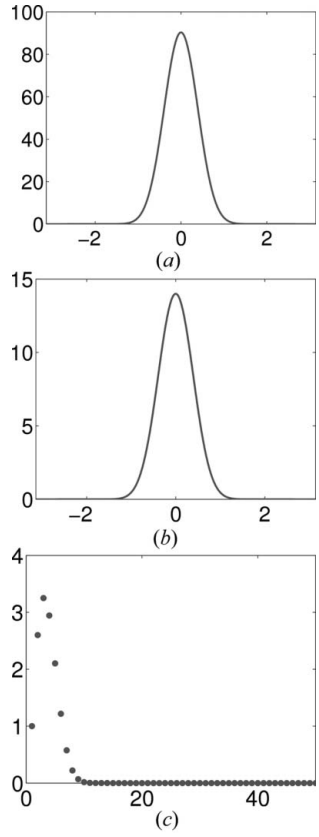

Figure 1
(*a*) The de la Vallée Poussin kernel ψ for $\kappa = 7$, (*b*) its Radon transform $\mathcal{R}\psi$ and (*c*) its Chebyshev coefficients.

3.2. Discretization

Solving the minimization problem [equation (45)] numerically requires a discretization of the parameter space, *i.e.* of the space of all ODFs. In the traditional harmonic method (Bunge, 1969) the ODF is approximated by its truncated Fourier series. The drawback of this discretization is that it is hard to ensure the non-negativity constraint. Another approach is the approximation by piecewise constant or linear ODFs (Bernier *et al.*, 2006). In this case, ensuring non-negativity is straightforward, but the calculation of the corresponding PDF is mathematically more involved and numerically slower in comparison to the Fourier series approach.

A compromise between the two approaches is the approximation by finite linear combinations of radially symmetric functions, *i.e.* by functions of the form

$$f(\mathbf{g}) = \sum_{m=1}^{M} c_m \psi(\mathbf{g}\mathbf{g}_m^{-1}), \tag{46}$$

where

$$\psi(\mathbf{g}) = \tilde{\psi}(\measuredangle \mathbf{g}) = \sum_{l=0}^{\infty} \hat{\psi}(l)\, \mathcal{U}_{2l}[\cos(\measuredangle \mathbf{g}/2)],\ \mathbf{g} \in \mathrm{SO}(3) \tag{47}$$

is a non-negative radially symmetric function and $\mathbf{g}_1, \ldots, \mathbf{g}_M$ is a set of nodes in the domain of rotations. Approximation by radially symmetric functions is a well known technique in approximation theory on the sphere and other manifolds. The resulting functions are smooth given the *Ansatz* function ψ is smooth and the non-negativity of the coefficients c_m, $m = 1, \ldots, M$, immediately implies the non-negativity of f.

In order to consider crystal symmetry, we look for approximations of the ODF by linear combinations

$$f(\mathbf{g}) = \sum_{m=1}^{M} c_m \psi_{\overline{G}_{\mathrm{Laue}}}(\mathbf{g}\mathbf{g}_m^{-1}) \tag{48}$$

of symmetrized radially symmetric functions

$$\psi_{\overline{G}_{\mathrm{Laue}}}(\mathbf{q}) = \frac{1}{|\overline{G}_{\mathrm{Laue}}|} \sum_{\mathbf{q}' \in \overline{G}_{\mathrm{Laue}}} \psi(\mathbf{q}\mathbf{q}'). \tag{49}$$

If the nodes $\mathbf{g} = \mathbf{g}_1, \ldots, \mathbf{g}_M$ are almost uniformly distributed in the domain of orientations with resolution δ, and the *Ansatz* function ψ is fairly well localized in spatial and in frequency domain with halfwidth $(3/2)\delta$, then any sufficiently smooth ODF can be approximated by a function of the form of equation (48) at resolution δ. A family of such *Ansatz* functions is formed by the de la Vallée Poussin kernel as introduced in Example 2. If the ODF is known to be concentrated in a certain region of the domain of orientations, then restricting the nodes $\mathbf{g} = \mathbf{g}_1, \ldots, \mathbf{g}_M$ to this region may largely improve the computational performance of the discretization. Such a region might be automatically computed by the zero range method (Bunge & Esling, 1979).

The first notable property of the function f is that its Radon transform is the linear combination of radially symmetric functions on the sphere and can be computed using the fast spherical Fourier transform. Assume that the *Ansatz* function ψ has a finite Fourier expansion with bandwidth $L \in \mathbb{N}_0$. Then we obtain from equations (38) and (37) the Fourier expansion

$$\begin{aligned}
\mathcal{R}f(\mathbf{h}, \mathbf{r}) &= \left[\mathcal{R} \sum_{m=1}^{M} c_m \psi_{\overline{G}_{\mathrm{Laue}}}(\circ \mathbf{g}_m^{-1}) \right](\mathbf{h}, \mathbf{r}) \\
&= \sum_{m=1}^{M} \frac{c_m}{|\overline{G}_{\mathrm{Laue}}|} \sum_{\mathbf{q} \in \overline{G}_{\mathrm{Laue}}} \mathcal{R}\psi(\mathbf{g}_m \mathbf{q}\mathbf{h}, \mathbf{r}) \\
&= \sum_{m=1}^{M} \frac{c_m}{|\overline{G}_{\mathrm{Laue}}|} \sum_{\mathbf{q} \in \overline{G}_{\mathrm{Laue}}} \sum_{l=0}^{L} \hat{\psi}(l) \frac{4\pi}{2l+1} \sum_{k=-l}^{l} \mathcal{Y}_l^k(\mathbf{g}_m \mathbf{q}\mathbf{h}) \overline{Y_l^k(\mathbf{r})} \\
&= \sum_{l=0}^{L} \sum_{k=-l}^{l} \hat{\psi}(l) \frac{4\pi}{2l+1} \mathcal{Y}_l^k(\mathbf{r}) \sum_{\mathbf{q} \in \overline{G}_{\mathrm{Laue}}} \sum_{m=1}^{M} \frac{c_m}{|\overline{G}_{\mathrm{Laue}}|} \overline{\mathcal{Y}_l^k(\mathbf{g}_m \mathbf{q}\mathbf{h})}.
\end{aligned} \tag{50}$$

Identifying the left-hand sum as a spherical Fourier transform and the right-hand sum as an adjoint spherical Fourier transform, we derive the following result.

Proposition 3. The Radon transform of the function f can be represented as the composition of a direct and an adjoint spherical Fourier transform,

$$\mathcal{R}f(\mathbf{h}, \mathbf{r}) = \mathbf{F}_{\mathbf{r},L}\left(\mathbf{v} \odot \sum_{\mathbf{q} \in \overline{G}_{\mathrm{Laue}}} \mathbf{F}^H_{\mathbf{q}\mathbf{h},L}\, \mathbf{c} \right),\ v_{lk} = \frac{4\pi}{2l+1} \hat{\psi}(l), \tag{51}$$

where $\mathbf{v}$ is a $(2L+1)$-dimensional vector and $\odot$ denotes the componentwise multiplication.

A second notable property of the function f is that its Fourier series can be calculated by an adjoint discrete SO(3) Fourier transform $\mathbf{F}^H_{\mathbf{g},L}$ of the coefficients c_m, $m = 1, \ldots, M$. More precisely, we obtain from equations (32) and (36) the following equality.

Proposition 4. The Fourier coefficients of the function f can be computed by

$$\hat{f}_{lkk'} = \mathbf{w} \odot \mathbf{F}^H_{\mathbf{g},L}\, \mathbf{c},\ \ w_{lkk'} = \frac{4\pi^2}{l + \frac{1}{2}} \hat{\psi}\,(2l), \tag{52}$$

where $\mathbf{w}$ is a $(1/3)(L+1)(2L+1)(2L+3)$-dimensional vector.

Since the Sobolev norm $\|f\|_{\mathcal{H}[\mathrm{SO}(3)]}$ in equation (45) is just the ℓ^2 norm $\|\omega \odot \hat{\mathbf{f}}\|_2$ of the Fourier coefficients of f with weights $\omega_{lkk'}$, $l = 0, \ldots, \infty$, $k, k' = -l, \ldots, l$, given by the Sobolev space $\mathcal{H}[\mathrm{SO}(3)]$ (Freeden *et al.*, 1998; Hielscher, 2007), we obtain the following representation in terms of the coefficients c_m, $m = 1, \ldots, M$,

$$\|f\|_{\mathcal{H}[\mathrm{SO}(3)]} = \|\mathbf{w} \odot \mathbf{F}^H_{L,\mathbf{g}}\, \mathbf{c}\|_2,\ \ w_{lkk'} = \frac{4\pi^2}{l + \frac{1}{2}} \omega_{lkk'}\, \hat{\psi}(2l). \tag{53}$$

Next we are going to restrict the estimator [equation (45)] to the finite-dimensional space of functions of the form of equation (48).

Proposition 5. The restriction of the estimator [equation (45)] to functions of the form of equation (48) is equivalent to the minimization problem

$$\mathbf{c}_{\mathrm{MLS}} = \mathrm{argmin}_{\mathbf{c}\in\mathbb{R}^M} J(\mathbf{c}),\ \text{subject to}\ \mathbf{c}\geq 0,\ \mathbf{c}\neq 0 \tag{54a}$$

with

$$J(\mathbf{c}) = \sum_{i=1}^{M} \left\| \mathbf{I}^{-1/2} \odot \left(\frac{\Psi_i \mathbf{c}}{\mathbf{c}^T \mathbf{a}_i} + \mathbf{I}_i^b - \mathbf{I}_i \right) \right\|^2 + \left\| \mathbf{w} \odot \frac{\mathbf{F}_{\mathbf{g},L}^H \mathbf{c}}{\mathbf{c}^T \mathbf{a}_0} \right\|^2, \tag{54b}$$

where the matrices $\boldsymbol{\Psi} \in \mathbb{R}^{M,N_i}$ and the vectors $\mathbf{a}_0, \mathbf{a}_i \in \mathbb{R}^M$, $i = 1, \ldots, N$, are defined as

$$\boldsymbol{\Psi}_{ij,m} = [\mathcal{R}\psi_{\overline{G}_{\mathrm{Laue}}}(\circ\mathbf{g}_m^{-1})](H_i, \mathbf{r}_{ij}),\ \ \mathbf{a}_0 = \mathbf{1}_M,\ \ \mathbf{a}_i = \frac{\boldsymbol{\Psi}_i^T \mathbf{1}_{N_i}}{\| \mathbf{I}_i - \mathbf{I}_i^b \|_1}, \tag{55}$$

and the weights $\mathbf{w} \in \mathbb{R}^{1/3(L+1)(2L+1)(2L+3)}$

$$[\mathbf{w}_{\lambda,s}]_{lkk'} = \frac{4\pi^2}{l+\frac{1}{2}} \omega_{lkk'}\, \hat{\psi}(2l) \tag{56}$$

are chosen according to the Sobolev space $\mathcal{H}[\mathrm{SO}(3)]$ and the *Ansatz* function ψ.

Proof. By equations (50) and (53) the right-hand side of equation (45) equals for all $\mathbf{c} \in \mathbb{R}_+^M$ and

$$f(\mathbf{g}) = \sum_{m=1}^{M} c_m \psi_{\overline{G}_{\mathrm{Laue}}}(\mathbf{g}\mathbf{g}_m^{-1}) \tag{57}$$

the functional $J(\mathbf{c})$. Hence, the restriction of the estimator [equation (45)] to functions of the form of equation (48) is equivalent to the restriction of the minimization problem [equation (54)] to the domain $\{\mathbf{c} \in \mathbb{R}_+^M \mid \|\mathbf{c}\|_1 = 1\}$. Since the functional J is scaling invariant, *i.e.* $J(\mathbf{c}) = J(\mu\mathbf{c})$ for all $\mu > 0$, the constraint $\|\mathbf{c}\|_1 = 1$ can be replaced by the constraint $\mathbf{c} \neq \mathbf{0}$.

It is not difficult to show that the functional J is differentiable and that the minimization problem [equation (54)] has a solution. However, the solution is in general not unique. An example of this behavior as well as a proof of the preceding assertion are given by Hielscher (2007).

Remark 6. The matrix vector multiplication $\boldsymbol{\Psi}_i \mathbf{c}$ essentially requires the computation of the sum [equation (50)] for all specimen directions $\mathbf{r}_{ij}$, $j = 1, \ldots, N_i$. According to Proposition 3, this sum can be computed with numerical complexity $\mathcal{O}(M + N_i + L^2 \log^2 L)$ using the fast spherical Fourier transform (Keiner *et al.*, 2007).

In Proposition 4 and equation (53), we have shown that the evaluation of the second term in equation (54) essentially requires a discrete Fourier transform on SO(3). A fast algorithm performing this transformation with numerical complexity $\mathcal{O}(L^3 \log^2 L)$ is described by Vollrath (2006).

Consequently, using these algorithms, the evaluation of the functional $J(\mathbf{c})$ has the total numerical complexity $\mathcal{O}(\overline{N} + NL^2 \log^2 L + L^3 \log^2 L)$.

3.3. The modified steepest descent algorithm

There are several algorithms for nonlinear, non-negatively constrained minimization, *e.g.* projected steepest descent, modified steepest descent, gradient projection residual norm conjugated gradients *etc.* (*cf*. Vogel, 2002). Here, we confine ourselves to the modified steepest descent algorithm (Kim, 2002, §4.2.1; Bardsley & Nagy, 2005) since it is appropriate to large-scale problems and combines simplicity and fast convergence. Moreover, it is especially well suited for problems where the unknown vector is sparse (Bardsley & Nagy, 2005).

Let J be some arbitrary differentiable function on $\mathbb{R}^M$. We are looking for solutions of the non-negatively constrained minimization problem

$$\mathbf{c}_{\mathrm{est}} = \mathrm{argmin}_{\mathbf{c}\in\mathbb{R}_+^M} J(\mathbf{c}). \tag{58}$$

The modified steepest descent algorithm is an iterative method based on the fix point iteration

$$\mathbf{c}^{(n+1)} = \mathbf{c}^{(n)} + \tau^{(n)} \tilde{\mathbf{c}}^{(n)} \tag{59}$$

where $\mathbf{c}^{(n)}$ is the actual estimate in the nth iteration, $\tilde{\mathbf{c}}^{(n)} \in \mathbb{R}^M$ is some descent direction, and $\tau > 0$ is the step size. In contrast to the ordinary steepest descent algorithm, the descent direction $\tilde{\mathbf{c}}^{(n)}$ is fixed as the negative gradient of $J[\mathbf{c}^{(n)}]$ componentwise multiplied with the actual estimate $\mathbf{c}^{(n)}$, *i.e.*

$$\tilde{\mathbf{c}}^{(n)} = -\mathbf{c}^{(n)} \odot \mathrm{grad}\, J[\mathbf{c}^{(n)}]. \tag{60}$$

The step length $\tau^{(n)}$ is calculated by a line search. In order to ensure the non-negativity of $\mathbf{c}^{(n+1)}$ at each iteration $n \in \mathbb{N}$ the step size has to be restricted to $[0, \tau_{\max}^{(n)}]$ with

$$\begin{aligned} \tau_{\max}^{(n)} &= \max\{\tau > 0 \mid \mathbf{c}^{(n)} + \tau \tilde{\mathbf{c}}^{(n)} \geq 0\} \\ &= \min\left\{ -\frac{[\mathbf{c}^{(n)}]_m}{[\tilde{\mathbf{c}}^{(n)}]_m} \mid m = 1, \ldots, M, [\tilde{\mathbf{c}}^{(n)}]_m < 0 \right\}. \end{aligned} \tag{61}$$

Algorithm 1 (shown in Fig. 2) outlines the modified steepest descent algorithm.

3.4. Adaptation to the functional *J*

Next, we are going to apply the modified steepest descent algorithm to solve the minimization problem [equation (54)]. We will use the following abbreviations.

Definition 7. Let $n \in \mathbb{N}$ and $\mathbf{c}^{(n)}, \tilde{\mathbf{c}}^{(n)} \in \mathbb{R}^M$. Then we define for $i = 0, \ldots, N$ the coefficients $\nu_i^{(n)}, \tilde{\nu}_i^{(n)} \in \mathbb{R}$ as

$$\nu_i^{(n)} = 1/[\mathbf{a}_i^T \mathbf{c}^{(n)}] \text{ and } \tilde{\nu}_i^{(n)} = 1/[\mathbf{a}_i^T \tilde{\mathbf{c}}^{(n)}]. \tag{62}$$

Moreover, we define the residuals $\mathbf{u}_i^{(n)}, \tilde{\mathbf{u}}_i^{(n)} \in \mathbb{R}^{N_i}$, $i = 1, \ldots, N$ as

$$\begin{aligned} \mathbf{u}_i^{(n)} &= \mathbf{I}^{-1/2} \odot [\nu_i^{(n)} \boldsymbol{\Psi}_i \mathbf{c}^{(n)} + \mathbf{I}_i^b - \mathbf{I}_i] \text{ and} \\ \tilde{\mathbf{u}}_i^{(n)} &= \mathbf{I}^{-1/2} \odot [\tilde{\nu}_i^{(n)} \boldsymbol{\Psi}_i \tilde{\mathbf{c}}_i^{(n)} + \mathbf{I}_i^b - \mathbf{I}_i] \end{aligned} \tag{63}$$

and set for completeness the vectors $\mathbf{u}_0^{(n)}, \tilde{\mathbf{u}}_0^{(n)} \in \mathbb{C}^{1/3(L+1)(2L+1)(2L+3)}$ to

input: $\mathbf{c}_0 \in \mathbb{R}_+^M$ /* initial guess */

output: $\mathbf{c} \in \mathbb{R}_+^M$ /* minimizer */

$k \leftarrow 0$

while *no convergence* **do**

$\mathbf{u}^{(n)} \leftarrow \operatorname{grad} J[\mathbf{c}^{(n)}]$ /* calculate gradient */

$\tilde{\mathbf{c}}^{(n)} \leftarrow -\mathbf{c}^{(n)} \odot \mathbf{u}^{(n)}$ /* calculate descent direction */

$\tau_{\text{opt}} \leftarrow \operatorname{argmin} J[\mathbf{c}^{(n)} + \tau\tilde{\mathbf{c}}^{(n)}]$ /* line search */

$\tau_{\text{bndry}} \leftarrow \min\left\{-\dfrac{[\mathbf{c}^{(n)}]_m}{[\tilde{\mathbf{c}}^{(n)}]_m} \mid m = 1, \ldots, M,\ [\tilde{\mathbf{c}}^{(n)}]_m < 0\right\}$

$\tau_{(n)} \leftarrow \min\{\tau_{\text{opt}}, \tau_{\text{bndry}}\}$

$\mathbf{c}^{(n+1)} \leftarrow \mathbf{c}^{(n)} + \tau_{(n)}\tilde{\mathbf{c}}^{(n)}$ /* update $\mathbf{c}$ */

$k \leftarrow k + 1$

end

$\mathbf{c} \leftarrow \mathbf{c}^{(n)}$

Figure 2
Algorithm 1: modified steepest descent.

$$\mathbf{u}_0^{(n)} = \nu_0^{(n)}\mathbf{w} \odot \mathbf{F}_{\mathbf{g},L}^H \mathbf{c}^{(n)} \text{ and } \tilde{\mathbf{u}}_0^{(n)} = \tilde{\nu}_0^{(n)}\mathbf{w} \odot \mathbf{F}_{\mathbf{g},L}^H \tilde{\mathbf{c}}^{(n)}. \quad (64)$$

Here we make use of the Fourier weights $\mathbf{w} \in \mathbb{R}^{1/3(L+1)(2L+1)(2L+3)}$ as defined in Proposition 5.

With these abbreviations we have:

Proposition 8. Let $n \in \mathbb{N}$ and $\mathbf{c}^{(n)} \in \mathbb{R}^M$. Then the functional J as defined in Proposition 5 simplifies to

$$J[\mathbf{c}^{(n)}] = \sum_{i=0}^{N} \| \mathbf{u}_i^{(n)} \|_2^2 . \quad (65)$$

Let, furthermore, $\tilde{\mathbf{c}}^{(n)} \in \mathbb{R}^M$ and $\mathbf{c}^{(n+1)} = \mathbf{c}^{(n)} + \tau^{(n)}\tilde{\mathbf{c}}^{(n)}$ for some $\tau^{(n)} \in \mathbb{R}$. Then we have, for all $i = 0, \ldots, N$, the recurrence formulae

$$\nu_i^{(n+1)} = \frac{\nu_i^{(n)}\tilde{\nu}_i^{(n)}}{\tau^{(n)}\nu_i^{(n)} + \tilde{\nu}_i^{(n)}} \text{ and } \mathbf{u}_i^{(n+1)} = \frac{\tilde{\nu}_i^{(n)}\mathbf{u}_i^{(n)} + \tau^{(n)}\nu_i^{(n)}\tilde{\mathbf{u}}_i^{(n)}}{\tau^{(n)}\nu_i^{(n)} + \tilde{\nu}_i^{(n)}}. \quad (66)$$

Proof. By Definition 7 we obtain for any $i = 0, \ldots, N$ the relationship

$$\begin{aligned}\nu_i^{(n+1)} &= \frac{1}{\mathbf{a}_i^T[\mathbf{c}^{(n)} + \tau^{(n)}\tilde{\mathbf{c}}^{(n)}]} = \frac{1}{[1/\nu_i^{(n)} + \tau^{(n)}/\tilde{\nu}_i^{(n)}]} \\ &= \frac{\nu_i^{(n)}\tilde{\nu}_i^{(n)}}{\tau^{(n)}\nu_i^{(n)} + \tilde{\nu}_i^{(n)}} \end{aligned} \quad (67)$$

and for any $i = 1, \ldots, N$ the equality

$$\begin{aligned}\mathbf{u}_i^{(n+1)} \odot \mathbf{I}_i^{1/2} &= \nu_i^{(n+1)}\boldsymbol{\Psi}_i\left[\mathbf{c}^{(n)} + \tau^{(n)}\tilde{\mathbf{c}}^{(n)}\right] + \mathbf{I}_i^b - \mathbf{I}_i \\ &= \frac{\nu_i^{(n)}\tilde{\nu}_i^{(n)}}{\tau^{(n)}\nu_i^{(n)} + \tilde{\nu}_i^{(n)}}\boldsymbol{\Psi}_i\left[\mathbf{c}^{(n)} + \tau^{(n)}\tilde{\mathbf{c}}^{(n)}\right] + \mathbf{I}_i^b - \mathbf{I}_i \\ &= \frac{\tilde{\nu}_i^{(n)}[\nu_i^{(n)}\boldsymbol{\Psi}_i\mathbf{c}^{(n)} + \mathbf{I}_i^b - \mathbf{I}_i] + \tau^{(n)}\nu_i^{(n)}[\tilde{\nu}_i^{(n)}\boldsymbol{\Psi}_i\tilde{\mathbf{c}}^{(n)} + \mathbf{I}_i^b - \mathbf{I}_i]}{\tau^{(n)}\nu_i^{(n)} + \tilde{\nu}_i^{(n)}} \\ &= \frac{\tilde{\nu}_i^{(n)}\mathbf{u}_i^{(n)} + \tau^{(n)}\nu_i^{(n)}\tilde{\mathbf{u}}_i^{(n)}}{\tau^{(n)}\nu_i^{(n)} + \tilde{\nu}_i^{(n)}} \odot \mathbf{I}_i^{1/2}. \end{aligned} \quad (68)$$

For $i = 0$ the proof of equation (66) is analogous.

With the abbreviations of Definition 7, we find the following expression for the gradient of the functional J.

Lemma 9. Let $n \in \mathbb{N}$. Then the gradient of the functional $J[\mathbf{c}^{(n)}] = \sum_{i=0}^N \| \mathbf{u}_i^{(n)} \|_2^2$ is given by

$$\begin{aligned}(1/2)\operatorname{grad} J[\mathbf{c}^{(n)}] = \nu_0^{(n)}\left\{\mathbf{F}_{\mathbf{g},L}[\mathbf{u}_0^{(n)} \odot \mathbf{w}] - \| \mathbf{u}_0^{(n)} \|_2^2\, \mathbf{a}_0\right\} \\ + \sum_{i=1}^{N} \nu_i^{(n)}\left[\mathbf{v}_i - \nu_i^{(n)}\mathbf{v}_i^T\mathbf{c}^{(n)}\mathbf{a}_i\right], \end{aligned} \quad (69)$$

where $\mathbf{v}_i = \boldsymbol{\Psi}_i^T[\mathbf{u}_i^{(n)} \odot \mathbf{I}_i^{-1/2}]$.

Proof. By the chain rule we obtain

$$\begin{aligned}&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\mathbf{c}^{(n)}}\left\| \frac{\boldsymbol{\Psi}_i\mathbf{c}^{(n)}}{\mathbf{a}_i^T\mathbf{c}^{(n)}} + \mathbf{I}_i^b - \mathbf{I}_i \right\|_{\mathbf{I}_i^{-1}}^2 \\ &\quad = \frac{\mathbf{a}_i^T\mathbf{c}^{(n)}\boldsymbol{\Psi}_i^T - \mathbf{a}_i[\boldsymbol{\Psi}_i\mathbf{c}^{(n)}]^T}{[\mathbf{a}_i^T\mathbf{c}^{(n)}]^2}\left\{\left[\frac{\boldsymbol{\Psi}_i\mathbf{c}^{(n)}}{\mathbf{a}_i^T\mathbf{c}^{(n)}} + \mathbf{I}_i^b - \mathbf{I}_i\right] \odot \mathbf{I}_i^{-1}\right\} \\ &\quad = \frac{\boldsymbol{\Psi}_i^T[\mathbf{u}_i^{(n)} \odot \mathbf{I}_i^{-1}]}{\mathbf{a}_i^T\mathbf{c}^{(n)}} - \frac{[\mathbf{u}_i^{(n)} \odot \mathbf{I}_i^{1/2}]^T\boldsymbol{\Psi}_i\mathbf{c}^{(n)}}{[\mathbf{a}_i^T\mathbf{c}^{(n)}]^2}\mathbf{a}_i. \end{aligned} \quad (70)$$

On the other hand, the gradient of the regularization term is

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\mathbf{c}^{(n)}}\left\| \frac{\mathbf{w} \odot \mathbf{F}_{\mathbf{g},L}^H\mathbf{c}^{(n)}}{\mathbf{a}_0^T\mathbf{c}^{(n)}} \right\|_2^2 &= \frac{\mathbf{a}_0^T\mathbf{c}^{(n)}\mathbf{F}_{\mathbf{g},L} - \mathbf{a}_0(\mathbf{F}_{\mathbf{g},L}^H\mathbf{c}^{(n)})^T}{(\mathbf{a}_0^T\mathbf{c})^2} \\ &\quad \times \frac{\mathbf{w}_{\lambda,s}^2 \odot \mathbf{F}_{\mathbf{g},L}^H\mathbf{c}\mathbf{F}_{\mathbf{g},L}}{\mathbf{a}_0^T\mathbf{c}^{(n)}} \\ &= \nu_0^{(n)}\left\{\mathbf{F}_{\mathbf{g},L}[\mathbf{u}_0^{(n)} \odot \mathbf{w}] - \| \mathbf{u}_0^{(n)} \|_2^2\, \mathbf{a}_0\right\}. \end{aligned} \quad (71)$$

We will also need the following representation of the function $\tau \mapsto J[\mathbf{c}^{(n)} + \tau\tilde{\mathbf{c}}^{(n)}]$ as a simple rational function based on the quantities $\mathbf{u}_0^{(n)}$ and $\tilde{\mathbf{u}}_0^{(n)}$ as defined in Definition 7.

Lemma 10. Let $\mathbf{c}^{(n)}, \tilde{\mathbf{c}}^{(n)} \in \mathbb{R}^M$. Then $\tau \mapsto J[\mathbf{c}^{(n)} + \tilde{\mathbf{c}}^{(n)}]$ is a rational function in τ. More precisely, we have

$$J[\mathbf{c}^{(n)} + \tau\tilde{\mathbf{c}}^{(n)}] = \sum_{i=0}^{N} \frac{A_i + 2\tau B_i + \tau^2 C_i}{[\tau\nu_i^{(n)} + \tilde{\nu}_i^{(n)}]^2}, \quad (72)$$

where we have set for any $i = 0, \ldots, N$,

$$A_i = \| \tilde{\nu}_i^{(n)} \mathbf{u}_i^{(n)} \|_2^2,\ B_i = \langle \tilde{\nu}_i^{(n)} \mathbf{u}_i^{(n)}, \nu_i^{(n)} \tilde{\mathbf{u}}_i^{(n)} \rangle,\ C_i = \| \nu_i^{(n)} \tilde{\mathbf{u}}_i^{(n)} \|_2^2 . \tag{73}$$

Proof. By Proposition 8 the function $\tau \mapsto J(\mathbf{c} + \tau\tilde{\mathbf{c}})$ can be rewritten as

$$J[\mathbf{c}^{(n)} + \tau\tilde{\mathbf{c}}^{(n)}] = \sum_{i=0}^{N} \left\| \frac{\tilde{\nu}_i^{(n)} \mathbf{u}_i^{(n)} + \tau \nu_i^{(n)} \tilde{\mathbf{u}}_i^{(n)}}{\tau \nu_i^{(n)} + \tilde{\nu}_i^{(n)}} \right\|_2^2$$
$$= \sum_{i=1}^{N} \frac{\| \tilde{\nu}_i^{(n)} \mathbf{u}_i^{(n)} \|_2^2 + 2\tau \langle \tilde{\nu}_i^{(n)} \mathbf{u}_i^{(n)}, \nu_i^{(n)} \tilde{\mathbf{u}}_i^{(n)} \rangle + \tau^2 \| \nu_i^{(n)} \tilde{\mathbf{u}}_i^{(n)} \|_2^2}{[\tau \nu_i^{(n)} + \tilde{\nu}_i^{(n)}]^2}. \tag{74}$$

The next lemma shows that one can choose the upper bound $\tau_{\max}^{(n)}$ as defined in equation (61) as the maximum stepsize.

Lemma 11. Let $\mathbf{c}^{(n)} \in \mathbb{R}^M$ and let $\tilde{\mathbf{c}}^{(n)} = \mathbf{c}^{(n)} \odot \operatorname{grad} J[\mathbf{c}^{(n)}]$ be the modified gradient of the functional J. Then, the maximum step length as defined in equation (61),

$$\tau_{\max}^{(n)} = \min \left\{ -\frac{[\mathbf{c}^{(n)}]_i}{[\tilde{\mathbf{c}}^{(n)}]_i} \mid m = 1, \ldots, M, [\tilde{\mathbf{c}}^{(n)}]_i < 0 \right\}, \tag{75}$$

is finite.

Proof. Since $J[\mathbf{c}^{(n)}]$ does not depend on the scale of $\mathbf{c}^{(n)}$, *i.e.* $J[\mathbf{c}^{(n)}] = J[\mu\mathbf{c}^{(n)}]$ for all $\mu > 0$, the gradient of $J[\mathbf{c}^{(n)}]$ is orthogonal to $\mathbf{c}^{(n)}$. Taking into account that $\mathbf{c}^{(n)} \geq 0$ we conclude that the modified descent search direction $\tilde{\mathbf{c}}^{(n)} = -\operatorname{grad} J[\mathbf{c}^{(n)}] \odot \mathbf{c}^{(n)}$ is either zero or has at least one negative component. Hence, $\tau_{\max}^{(n)}$ is finite.

Merging Remark 6, Proposition 8, Lemmas 9, 10 and 11, we obtain the following theorem.

Theorem 12. Algorithm 3 (shown in Figs. 3 and 4) implements the modified steepest descent (MSD) algorithm for the minimization problem [equation (54)]. The numerical complexity of each iteration corresponds to the numerical complexity of the matrix vector multiplications $\mathbf{\Psi}_i \mathbf{c}$, $i = 1, \ldots, N$, and $\mathbf{F}_{L,\mathbf{g}} \mathbf{c}$.

Proof. Algorithm 3 implements the modified steepest descent Algorithm 1.

In lines 1– 5 the vectors $\mathbf{a}_i \in \mathbb{R}^M$, $i = 1, \ldots, N$, required for the calculation of the normalization coefficients (*cf.* Proposition 5), the initial values of the residuals $\mathbf{u}_i^{(0)}$ and the normalization coefficients $\nu_i^{(0)}$, $i = 1, \ldots, N$ (*cf.* Definition 7), are computed. These calculations require the matrix vector multiplications $\mathbf{\Psi}_i^T \mathbf{1}_{N_i}$ and $\mathbf{\Psi}_i \mathbf{c}$, $i = 1, \ldots, N$.

In lines 8–10 the gradient of the functional J in $\mathbf{c}^{(n)}$ is computed according to Lemma 9. This essentially requires the matrix vector multiplications $\mathbf{\Psi}_i^T \mathbf{u}_i^{(n)}$, $i = 1, \ldots N$.

In line 11 the modified descent direction of the MSD algorithm is computed according to equation (60).

In lines 12–14 the updates $\tilde{\mathbf{u}}_i^{(n)}$ and $\tilde{\boldsymbol{\nu}}_i^{(n)}$, $i = 1, \ldots, N$, of the residuals and the normalization coefficients are computed (*cf.* Definition 7). Again, this requires the matrix vector multiplications $\mathbf{\Psi}_i \mathbf{c}$, $i = 1, \ldots, N$.

In lines 15 and 16 the step size is computed using Algorithm 2 and the initial step length as approved in Lemma 11.

The updating of the coefficient vector is performed in line 17, whereas the residuals and the normalization coefficients are updated in lines 18 and 19 (*cf.* Proposition 8).

Corollary 13. Combining Remark 6 and Theorem 12 we conclude that the numerical complexity of each iteration of algorithm 3 is $\mathcal{O}(\overline{N} + M + L^3)$.

Since the estimated ODF is a superposition of radially symmetric functions, it is straightforward to calculate its Fourier coefficients (*cf.* Proposition 4) or arbitrary pole figures

input: $\tau_{\max}^{(n)} \in \mathbb{R}$ /* maximum step length */

$\mathbf{u}_0^{(n)} \in \mathbb{R}^{1/3(L+1)(2L+1)(2L+3)}$

$\tilde{\mathbf{u}}_0^{(n)} \in \mathbb{R}^{1/3(L+1)(2L+1)(2L+3)}$

$\mathbf{u}_i^{(n)} \in \mathbb{R}^{N_i}, i = 1, \ldots, N$

$\tilde{\mathbf{u}}_i^{(n)} \in \mathbb{R}^{N_i}, i = 1, \ldots, N$

$\boldsymbol{\nu}^{(n)} \in \mathbb{R}^{N+1}$

$\tilde{\boldsymbol{\nu}}^{(n)} \in \mathbb{R}^{N+1}$

output: $\tau^{(n)} \in \mathbb{R}_+$ /* optimum step length */

for $i \leftarrow 0, \ldots, N$ **do** /* precomputation */

$A_i \leftarrow ||\tilde{\nu}_i^{(n)} \mathbf{u}_i^{(n)}||_2^2$

$B_i \leftarrow \langle \tilde{\nu}_i^{(n)} \mathbf{u}_i^{(n)}, \nu_i^{(n)} \tilde{\mathbf{u}}_i^{(n)} \rangle$

$C_i \leftarrow ||\nu_i^{(n)} \tilde{\mathbf{u}}_i^{(n)}||_2^2$

end

$J_0 \leftarrow \sum_{i=0}^{N} A_i [\tilde{\nu}_i^{(n)}]^{-2}$ /* current value of $J[\mathbf{c}^{(n)}]$ */

$\tau \leftarrow \tau_{\max}^{(n)}$

$J \leftarrow \sum_{i=0}^{N} \frac{A_i + 2\tau B_i + \tau^2 C_i}{[\tau \nu_i^{(n)} + \tilde{\nu}_i^{(n)}]^2}$ /* value of $J[\mathbf{c}^{(n)} + \tau\tilde{\mathbf{c}}^{(n)}]$ */

while $J > J_0$ **do**

$\tau \leftarrow \frac{1}{2}\tau$ /* reduce step length */

$J \leftarrow \sum_{i=0}^{N} \frac{A_i + 2\tau B_i + \tau^2 C_i}{[\tau \nu_i^{(n)} + \tilde{\nu}_i^{(n)}]^2}$ /* update value of $J(\mathbf{c} + \tau\tilde{\mathbf{c}})$*/

end

$\tau^{(n)} \leftarrow \tau$

Figure 3
Algorithm 2: line search.

input: $\mathbf{c}_0 \in \mathbb{R}^M$ /* initial vector */

$\mathbf{I}_i \in \mathbb{R}^{N_i}, i = 1, \ldots, N$ /* diffraction counts */

$\mathbf{I}_i^b \in \mathbb{R}^{N_i}, i = 1, \ldots, N$ /* background intensities */

$\mathbf{w} \in \mathbb{R}^{1/3(L+1)(2L+1)(2L+3)}$ /* regularization weights */

output: $\mathbf{c} \in \mathbb{R}^M$

$\mathbf{a}_0 \leftarrow 1_M$

for $i \leftarrow 1, \ldots, N$ **do** $\mathbf{a}_i \leftarrow \dfrac{\boldsymbol{\Psi}_i^T \mathbf{1}_{N_i}}{(\mathbf{I}_i^T - \mathbf{I}_i^b)^T \mathbf{1}_{N_i}}$

for $i \leftarrow 1, \ldots, N$ **do** $\nu_i^{(0)} \leftarrow \dfrac{1}{\mathbf{a}_i^T \mathbf{c}^0}$ /* normalization coefficients*/

$\mathbf{u}_0^{(0)} \leftarrow \nu_0^{(0)} \mathbf{w} \odot \mathbf{F}_{\mathbf{g},L} \mathbf{c}^0$

for $i \leftarrow 1, \ldots, N$ **do** $\mathbf{u}_i^{(0)} \leftarrow [\nu_i^{(0)} \boldsymbol{\Psi}_i \mathbf{c}^{(0)} + \mathbf{I}_i^b - \mathbf{I}_i] \odot \mathbf{I}_i^{1/2}$

$n \leftarrow 1$

while *no convergence* **do**

$\tilde{\mathbf{v}}_0^{(n)} \leftarrow -\nu_0^{(n)} \left\{ \mathbf{F}_{\mathbf{g},L} [\mathbf{w} \odot \mathbf{u}_0^{(n)}] - \| \mathbf{u}_0^{(n)} \|_2^2 \, \mathbf{a}_0 \right\}$

for $i \leftarrow 1, \ldots, N$ **do** $\tilde{\mathbf{v}}_i^{(n)} \leftarrow -\nu_i^{(n)} \boldsymbol{\Psi}_i^T [\mathbf{u}_i^{(n)} \odot \mathbf{I}_i^{-1/2}]$

$\mathbf{v}^{(n)} \leftarrow \tilde{\mathbf{v}}_0^{(n)} + \sum_{i=1}^{N} \tilde{\mathbf{v}}_i^{(n)} - \nu_{(i)}^{(n)} \langle \tilde{\mathbf{v}}_i^{(n)}, \mathbf{c}^{(n)} \rangle \mathbf{a}_i$ /* gradient */

$\tilde{\mathbf{c}}^{(n)} \leftarrow \mathbf{v}^{(n)} \odot \mathbf{c}^{(n)}$ /* descent direction */

for $i \leftarrow 0, \ldots, N$ **do** $\tilde{\nu}_i^{(n)} \leftarrow \dfrac{1}{\mathbf{a}_i^T \tilde{\mathbf{c}}^{(n)}}$

$\tilde{\mathbf{u}}_0^{(n)} \leftarrow \tilde{\nu}_0^{(n)} \mathbf{w} \odot \mathbf{F}_{\mathbf{g},L} \tilde{\mathbf{c}}^{(n)}$

for $i \leftarrow 1, \ldots, N$ **do** $\tilde{\mathbf{u}}_i^{(n)} \leftarrow [\tilde{\nu}_i^{(n)} \boldsymbol{\Psi}_i \tilde{\mathbf{c}}^{(n)} + \mathbf{I}_i^b - \mathbf{I}_i] \odot \mathbf{I}_i^{-1/2}$

$\tau_{\max}^{(n)} \leftarrow \min_{m=1, \ldots, M} \left\{ -\dfrac{\mathbf{c}_m^{(n)}}{\tilde{\mathbf{c}}_m^{(n)}} \mid \tilde{\mathbf{c}}_m^{(n)} < 0, m = 1, \ldots, M \right\}$

$\tau^{(n)} \leftarrow \text{LineSearch} \left[\tau_{\max}^{(n)}, \nu_i^{(n)}, \tilde{\nu}_i^{(n)}, \mathbf{u}_i^{(n)}, \tilde{\mathbf{u}}_i^{(n)}, i = 0, \ldots, N \right]$

$\mathbf{c}^{(n)} \leftarrow \mathbf{c}^{(n)} + \tau^{(n)} \tilde{\mathbf{c}}^{(n)}$ /* update solution */

for $i \leftarrow 0, \ldots, N$ **do** $\mathbf{u}_i^{(n)} \leftarrow \dfrac{\tilde{\nu}_i^{(n)}}{\tau^{(n)} \nu_i^{(n)} + \tilde{\nu}_i^{(n)}} \mathbf{u}_i^{(n)} + \dfrac{\tau^{(n)} \nu_i^{(n)}}{\tau^{(n)} \nu_i^{(n)} + \tilde{\nu}_i^{(n)}} \tilde{\mathbf{u}}_i^{(n)}$

for $i \leftarrow 0, \ldots, N$ **do** $\nu_i^{(n)} \leftarrow \dfrac{\nu_i^{(n)} \tilde{\nu}_i^{(n)}}{\tau^{(n)} \nu_i^{(n)} + \tilde{\nu}_i^{(n)}}$

$n \leftarrow n + 1$

end

$\mathbf{c} \leftarrow \mathbf{c}^{(n)}$

Figure 4
Algorithm 3: modified least-squares ODF estimator.

[see equations (38) and (50)]. In order to calculate the volume portion of the ODF in a certain region of the domain of orientations, one can use the fast Fourier transform on SO(3) to evaluate the estimated ODF at a number of nodes within the region and apply a quadrature rule to calculate the volume.

4. Practical example

To demonstrate the features of *MTEX*, we elaborate on the analysis of preferred crystallographic orientation of a hematite specimen H43C1 which has been interpreted in terms of experimental deformation by Siemes *et al.* (2008). The subject of their communication is the critical resolved shear stresses of hematite crystals and glide modes of hematite. Hematite, Fe_2O_3, is a trigonal mineral with corundum structure and a hexagonal cell with $a_0 = 0.5038$, $c_0 = 1.3772$ nm, to which the symbols for planes (*hkil*), sets of symmetry-related planes {*hkil*}, directions [*uvtw*] and sets of symmetry-related directions ⟨*uvtw*⟩ refer. Tetragonal prismatic specimens, sized 7 × 7 × 14 mm, were prepared in different crystallographic orientations with their top plane either parallel to *c*(0001), *r*(01$\bar{1}$2), *f*(10$\bar{1}$1), *a*(11$\bar{2}$0) or *m*(10$\bar{1}$0). Complete pole figures representing the crystallographic preferred orientation of the

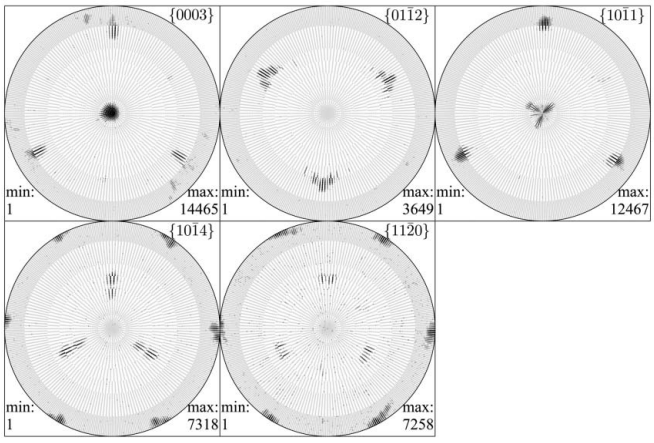

Figure 5
Experimental pole figures (equal area projection, upper hemisphere) of the reflections $c(0001)$, $f(10\bar{1}1)$, $r(01\bar{1}2)$, $e(10\bar{1}4)$ and $a(11\bar{2}0)$, measured by neutron diffraction with specimen H43C1 deformed by 3.4% in compression perpendicular to $c(0001)$ at 873 K and 300 MPa confining pressure, with compression direction perpendicular to plane of projection.

bulk volume were measured with a neutron texture diffractometer at the Research Centre Jülich (Jansen *et al.*, 2000). Since this diffractometer is equipped with a position-sensitive detector, the reflections of $c(0001)$, $f(10\bar{1}1)$, $r(01\bar{1}2)$, $e(10\bar{1}4)$ and $a(11\bar{2}0)$ were simultaneously measured for a 2θ range of 50° using a wavelength of 0.2332 nm, with $c(0001)$ obtained from the third-order reflection (0003). To detect the peaks in the pole figures of these reflections, the standard scanning grid comprising about 500 sample positions was refined to 14 616 positions with a mean distance of 1.5°. The total measuring time was about 3 d. The intensity data were corrected for background intensities according to the Jülich data-processing procedure.

The five experimental pole figures are displayed in Fig. 5. To compute an ODF explaining the data, the de la Vallée Poussin kernel with a halfwidth of 1.25° corresponding to the finite bandwidth of $L = 416$ of the series expansion into generalized spherical harmonics was used. Applying the 'zero range' option of *MTEX* the kernel was centred at 1294 orientations with a mean distance of 1.2°. Owing to the use of the 'zero range' option of *MTEX*, the time elapsed to compute the ODF was 427 s with a notebook equipped with a Core 2 Duo CPU with 1.86 GH c.p.u. frequency and 2 GB RAM.

The computed orientation density is displayed in 12σ sections of Fig. 6. Its texture index is approximately 5000; correspondingly, the entropy is approximately −7.8.

There is a major mode at (150°, 0°, 60°) computed by *MTEX* (in terms of Matthies' Euler angles). The corresponding volume portion in the 5° neighbourhood of the major mode is approximately 55%, the value of the ODF at the major mode is 9100. Three minor modes are recognized by visual inspection at about (90°, 65°, 59°), (30°, 115°, 1°) and (150°, 115°, 1°), respectively. The corresponding volume portions in the 5° neighbourhood of the three minor modes are approximately 8, 13.5 and 8.5%; they sum to approximately 30%. The values of the ODF at the three minor modes are 660, 545 and 580. It is once more confirmed that an interpretation of an orientation density function in terms of its values may be deceiving. A proper interpretation is accomplished in terms of volume portions only.

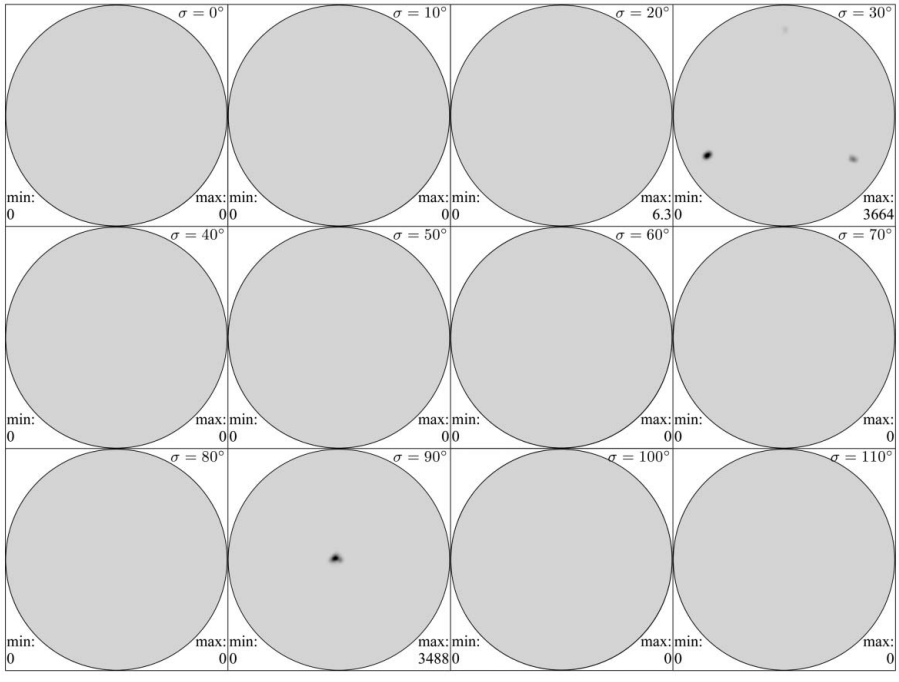

Figure 6
MTEX's recovered ODF based on experimental pole figures of the reflections $c(0001)$, $f(10\bar{1}1)$, $r(01\bar{1}2)$, $e(10\bar{1}4)$ and $a(11\bar{2}0)$.

The recalculated pole figures are diplayed in Fig. 7. Then the recalculated pole figures are augmented with the pole points corresponding to the major and minor modes and shown in Fig. 8, where the major mode depicted in black represents the parent crystal orientation, and the three minor modes depicted in blue, red and green, respectively, indicate three *r*-twin orientations.

Eventually, the pole figure RP errors have values of RP(0001) = 0.5, RP($10\bar{1}1$) = 0.8, RP($01\bar{1}2$) = 0.8, RP($10\bar{1}4$) = 0.6, and RP($11\bar{2}0$) = 0.8.

5. Conclusions

A novel algorithm for ODF estimation from diffraction pole figures has been presented, which is especially well suited for sharp textures and high-resolution pole figures measured with respect to arbitrarily scattered specimen directions, *e.g.* by area detectors. The estimator has been

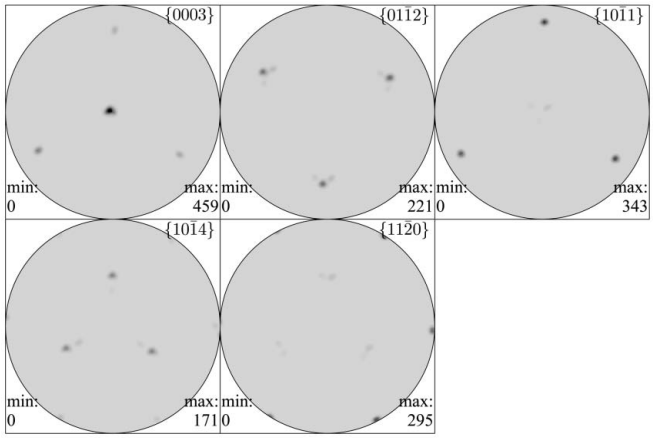

Figure 7
Recalculated pole figures (equal area projection, upper hemisphere) of the reflections $c(0001)$, $f(10\bar{1}1)$, $r(01\bar{1}2)$, $e(10\bar{1}4)$ and $a(11\bar{2}0)$ computed and displayed with *MTEX*.

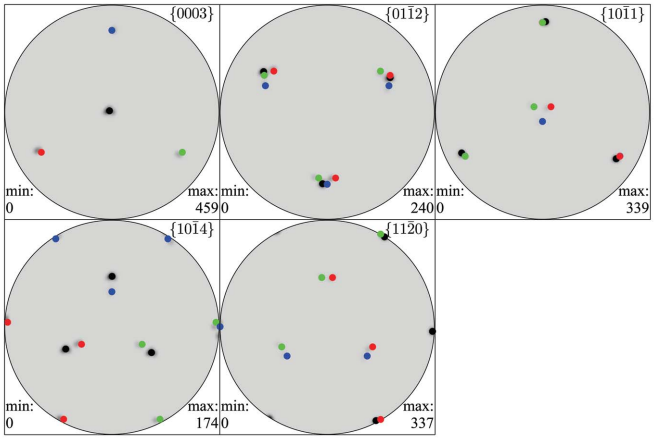

Figure 8
Recalculated pole figures (equal area projection, upper hemisphere) of the reflections $c(0001)$, $f(10\bar{1}1)$, $r(01\bar{1}2)$, $e(10\bar{1}4)$ and $a(11\bar{2}0)$ augmented with major mode (150°, 0°, 60°) (black), and minor modes (90°, 65°, 59°) (blue), (30°, 115°, 1°) (red), and (150°, 115°, 1°) (green), respectively, computed and displayed with *MTEX*.

characterized as the solution of a nonlinear minimization problem considering the unknown orientation density values and the unknown normalization coefficients simultaneously. Neglecting the regularizing Sobolev norm term, it is emphasized that the objective function aims at minimization of the relative error and thus takes statistical distribution of the measurement errors into proper account. To solve the minimization problem numerically, an iterative algorithm has been presented that applies a discretization provided by radially symmetric functions and fast Fourier techniques.

The authors would like to thank the editor Daniel Chateigner, Caen, for his encouragement and patience, and in particular Heinrich Siemes, RWTH Aachen, for his sympathetic help in improving the manuscript by including a real-world example considering his hematite data.

References

Bardsley, J. M. & Nagy, J. G. (2005). *SIAM J. Matrix Anal. Appl.* **27**, 118–119.
Bernier, J. V., Miller, M. P. & Boyce, D. E. (2006). *J. Appl. Cryst.* **39**, 697–713.
Böhlke, T., Haus, U. & Schulze, V. (2006). *Acta Mater.* **54**, 1359–1368.
Boogaart, K. G. van den, Hielscher, R., Prestin, J. & Schaeben, H. (2006). *J. Comput. Appl. Math.* **199**, 122–140.
Buhmann, M. D. (2003). *Radial Basis Functions: Theory and Implementations*. Cambridge University Press.
Bunge, H. J. (1965). *Z. Metallkd.* **56**, 872–874.
Bunge, H. J. (1969). *Mathematische Methoden der Texturanalyse*. Berlin: Akademie Verlag.
Bunge, H. J. (1982). *Texture Analysis in Materials Science, Mathematical Methods*. London: Butterworths.
Bunge, H. J. & Esling, C. (1979). *J. Phys. Lett.* **40**, L627.
Bunge, H. J. & Esling, C. (1985). *Preferred Orientations in Deformed Metals and Rocks – An Introduction to Modern Texture Analysis*, edited by H. R. Wenk, pp. 109–122. Orlando: Academic Press.
Cowley, J. M. (1995). *Diffraction Physics*, 3rd ed. Amsterdam, New York: North-Holland Personal Library.
Forsyth, J. B. (1988). *Int. Ser. Numer. Math.* **137**, 177–188.
Freeden, W., Gervens, T. & Schreiner, M. (1998). *Constructive Approximation on the Sphere*. Oxford: Clarendon Press.
Hammond, C. (1997). *The Basics of Crystallography and Diffraction*. Oxford University Press.
Helming, K. & Eschner, T. (1990). *Cryst. Res. Technol.* **25**, K203–K208.
Hielscher, R. (2007). PhD thesis, TU Bergakademie Freiberg, Germany.
Hielscher, R. (2008). *MTEX*. Version 1.0. MATLAB and C subroutine library. http: //code.google.com/p/mtex/.
Hubbert, S. & Baxter, B. (2001). *Recent Progress in Multivariate Approximation*. 4th International Conference, September 2000, Witten–Bommerholz., edited by W. H. K. Jetter & M. Reimer, International Series of Numerical Mathematics, Vol. 137, pp. 33–47. Basel: Birkhäuser.
Imhof, J. (1983). *Phys. Status Solidi A*, **75**, K187–K189.
Jansen, E., Schäfer, W. & Kirfel, A. (2000). *J. Struct. Geol.* **22**, 1559–1564.
Jura, J., Pospiech, J. & Bunge, H. J. (1976). *Prace Komisji Metalurgiczno-Odlewniczej PAN, Metalurgia*, **24**, 111.
Keiner, J., Kunis, S. & Potts, D. (2006). *Computing*, **78**, 1–15.
Keiner, J., Kunis, S. & Potts, D. (2007). *J. Fourier Anal. Appl.* **13**, 435–458.
Kim, B. (2002). PhD thesis, Stanford University, CA, USA.
Latteroti, L., Matthies, S. &Wenk, H.-R. (1999). *IUCr Commission on Powder Diffraction Newsletter*, No. 21, pp. 14–15.
Kostelec, P. J. & Rockmore, D. N. (2003). Santa Fe Institute Working Papers Series, Paper 03-11-060.
Kunis, S. & Potts, D. (2003). *J. Comput. Appl. Math.* **161**, 75–98.
Matthies, S. (1979). *Phys. Status Solidi B*, **92**, K135–K138.
Matthies, S. (1988). *ICOTOM8*, edited by J. S. Kallend & G. Gottstein, pp. 37–48. Santa Fe: The Metallurgical Society.
Matthies, S. & Pospiech, J. (1980). *Phys. Status Solidi B*, **97**, 547.
Matthies, S., Vinel, G. W. & Helmig, K. (1987). *Standard Distributions in Texture Analysis*, Vol. 1. Berlin: Akademie-Verlag.
Nikiforov, A. F. & Uvarov, V. B. (1988). *Special Functions in Mathematical Physics*. Basal: Birkhäuser.
Pawlik, K. (1986). *Phys. Status Solidi*, **134**, 477.
Pospiech, J. & Jura, J. (1974). *Z. Metallkd.* **65**, 324–330.
Pospiech, J., Ruer, D. & Baro, R. (1981). *J. Appl. Cryst.* **14**, 230–233.
Randle, V. & Engler, O. (2000). *Texture Analysis.* Amsterdam: Gordon and Breach Science Publishers.
Rodriguez-Navarro, A. B., Cabralde Melo, C., Batista, N., Morimoto, N., Alvarez-Lloret, P., Ortega-Huertas, M., Fuenzalida, V. M., Arias, J. I., Wiff, J. P. & Arias, J. L. (2006). *J. Struct. Biol.* **156**, 355–362.

Roe, R. J. (1965). *J. Appl. Phys.* **36**, 2024–2031.
Ruer, D. (1976). *Methode Vectorielle d'Analyse de la Texture*. Universite de Metz, France.
Schaeben, H. (1988). *J. Appl. Phys.* **64**, 2236–2237.
Schaeben, H. (1996). *J. Appl. Cryst.* **29**, 516–525.
Schaeben, H. (1997). *Phys. Status Solidi B*, **200**, 367–376.
Schaeben, H. & van den Boogaart, K. G. (2003). *Tectonophysics*, **370**, 253–268.
Schaeben, H., Hielscher, R., Fundenberger, J.-J., Potts, D. & Prestin, J. (2007). *J. Appl. Cryst.* **40**, 570–579.
Siemes, H., Klingenberg, B., Rybacki, E., Naumann, M., Schäfer, W., Jansen, E. & Kunze, K. (2008). *Ore Geol. Rev.* **33**, 255–279.
Vadon, A. & Heizmann, J. J. (1991). *Textures Microstruct.* **14**, 37–44.
Van Houtte, P. (1983). *Textures Microstruct.* **6**, 1–19.
Van Houtte, P. (1991). *Textures Microstruct.* **13**, 199–212.
Varshalovich, D., Moskalev, A. & Khersonski, V. (1988). *Quantum Theory of Angular Momentum*. Singapore: World Scientifc Publishing.
Vogel, C. R. (2002). *Computational Methods for Inverse Problems*, Frontiers in Applied Mathematics. Philadelphia: Society for Industrial and Applied Mathematics (SIAM).
Vollrath, A. (2006). Diploma thesis, Universiẗat zu Lübeck, Germany.
Wenk, H. R. (1985). *Preferred Orientation in Deformed Metals and Rocks: An Introduction to Modern Texture Analysis*. New York: Academic Press.