Investigation of the indentation size effect through the measurement of the geometrically necessary dislocations beneath small indents of different depths using EBSD tomography

Eralp Demir *, Dierk Raabe, Nader Zaafarani, Stefan Zaefferer

Max-Planck-Institut für Eisenforschung, 40237 Düsseldorf, Germany

Received 11 June 2008; received in revised form 27 September 2008; accepted 27 September 2008
Available online 24 October 2008

Abstract

We study the link between the indentation size effect and the density of geometrically necessary dislocations (GNDs) through the following approach: four indents of different depth and hardness were placed in a Cu single crystal using a conical indenter with a spherical tip. The deformation-induced lattice rotations below the indents were monitored via a three-dimensional electron backscattering diffraction method with a step size of 50 nm. From these data we calculated the first-order gradients of strain and the GND densities below the indents. This approach allowed us to quantify both the mechanical parameters (depth, hardness) and the lattice defects (GNDs) that are believed to be responsible for the indentation size effect. We find that the GND density does not increase with decreasing indentation depth but rather drops instead. More precisely, while the hardness increases from 2.08 GPa for the largest indent (1230 nm depth) to 2.45 GPa for the smallest one (460 nm depth) the GND density decreases from $\approx 2.34 \times 10^{15}$ m$^{-2}$ (largest indent) to $\approx 1.85 \times 10^{15}$ m$^{-2}$ (smallest indent).

Crown Copyright © 2008 Published by Elsevier Ltd on behalf of Acta Materialia. All rights reserved.

Keywords: GND; Size effect; 3-D EBSD; Orientation gradient

1. Introduction

Advances in microtechnology require better understanding of materials at small scales. This includes mechanical characterization at small dimensions using methods such as nanoindentation. In this context it was observed that a material’s resistance to plastic flow depends on the size of the sample [1,2] and on the deformed volume [3,4]. These phenomena are referred to as mechanical size effects.

Physically based models attribute strengthening at small dimensions to the existence of immobile dislocations of two types: geometrically necessary dislocations (GNDs) and statistically stored dislocations (SSDs). GNDs provide the lattice continuity in the presence of curvature [5–8], whereas SSDs evolve through random interactions among dislocations with plastic strain [9,10].

The average spacing between immobile dislocations, which is inversely proportional to the square root of the dislocation density, determines the yield strength of the material [9]. Introduction of additional dislocations such as GNDs thus increases the flow strength of the material by reducing the mean spacing between dislocations:

$$\tau^a = cGh\sqrt{\rho_{\text{gnd}}^a + \rho_{\text{ssd}}^a}$$

(1)

where $c$ is a geometrical constant ranging between 0.5 to 1, $G$ is the shear modulus, $b$ is the magnitude of the Burgers vector, and $\rho^a$ and $\tau^a$ are the dislocation density and yield strength on slip system “a”, respectively.

GNDs that create rotational gradients of strain are used to explain size-dependent plasticity at small dimensions [11–15]. For instance, the increase in strength with a decrease in diameter of copper torsion wires was attributed to strain gradients. Ashby and Fleck introduced a material length scale parameter that is coupled to the rotational gradients of strain [15]. Later, Fleck and Hutchinson applied...
additional length scale parameters to stretch gradients in order to capture size effects during nanoindentation [11]. Nix and Gao introduced a physically based intrinsic length scale assuming a uniform distribution of GNDs below indents [16,17]. The relative contribution of this effect to the overall resistance of a material to plastic flow becomes particularly visible in the nanometer regime [18].

The crystallographic misorientation between two neighboring points can be used as an approximate measure for the GNDs as shown below in Eq. (2), where $\rho_{\text{gnd}}$ is the GND density that is required to accommodate a curvature $\omega$ [19,20,28]:

$$
\rho_{\text{gnd}} = \frac{\omega}{b} 
$$

Parks calculated the GNDs by minimizing their total length for different possible geometrical arrangements [21,22] in face-centered cubic (fcc) crystals. Sun and Rollett applied this method to orientation maps obtained via electron backscattering diffraction (EBSD) from cross-sections of deformed crystals [23–25]. They used the misorientations to identify the GND density.

An appropriate selection of the EBSD step size is a critical issue for the calculation of GND densities. EBSD studies on deformed crystals have demonstrated the effect of the step size on the accuracy of measured misorientations [26,27]. The noise in an EBSD analysis strongly depends on the step size, the type of deformation and the amount of strain. On the other hand, the step size determines the resolution for a Frank loop analysis [28]. Hence, the step size used for the GND analysis has to be low enough to be of microscopic relevance and high enough to filter out noise.

The size dependence of indentation hardness has been associated with strain gradients which are realized in the lattice through GNDs. This motivated us to directly measure lattice rotations below indents with the aim of quantifying the density of these defects. Monitoring the evolution of lattice rotations, GND densities and hardness in the same set of experiments may help to better understand indentation size effects. For this purpose we map the orientation distribution below four nanoindents of different depths. The experiment makes use of a tomographic high-resolution electron backscatter diffraction orientation microscope in conjunction with a focused ion beam instrument for automated serial sectioning (3-D EBSD) [33]. The GND densities are estimated below the four indents using EBSD data considering both the 2-D and the 3-D rotational gradients.

2. Methods

2.1. Theory

The translation symmetry of the crystal lattice remains intact during plastic deformation. The line integral of the displacement, $du_i$, therefore, has to vanish around any area normal in the crystal, Eq. (3). The same equality holds when displacements are calculated in terms of gradients in a reference frame.

$$
\int du_i = \int \beta_{ij} dx_j = 0 
$$

The displacement gradient or distortion tensor, $\beta_{ij}$, consists of an elastic and a plastic part, Eq. (4). Additive decomposition of distortion allows the use of either the elastic distortion, $\beta_{ij}^{\text{el}}$, or the plastic distortion, $\beta_{ij}^{\text{pl}}$, as a measure of incompatibility.

$$
\beta_{ij} = \frac{\partial u_i}{\partial x_j} = \beta_{ij}^{\text{el}} + \beta_{ij}^{\text{pl}} 
$$

Using the elastic distortion as a measure of incompatibility in conjunction with Stoke’s theorem, the dislocation tensor can be obtained from the curl ($\nabla \times$) of the elastic distortion, Eq. (5). Elastic distortion consists of both an elastic stretch, $\epsilon^e_{ij}$, and a rotation, $g$. This is shown in Eq. (6) where “$e$” indicates the permutation symbol. Ignoring the elastic stretch, the rotational gradients are used to identify the components of the dislocation tensor, $a$, Eq. (7).

$$
a = \nabla \times \epsilon^e_{ij} 
$$

$$
a_{pi} = \epsilon_{pij} (\epsilon^e_{ij} + g_{ijk}) 
$$

$$
a_{pi} = \epsilon_{pij} g_{ijk} 
$$

The misorientation, $\Delta \phi$, between two points is used to calculate orientation gradients, $g_{ij,k}$. In the following we explain the procedure for two arbitrary points with given orientations $\phi(1)$ and $\phi(2)$. The 24 crystal symmetry operators ($O'^{\text{si}}$) are applied to both orientations to identify the minimum misorientation.

$$
|\Delta \phi| = \min \left\{ \cos^{-1}\left\{ tr[(O'^{\text{si}})\phi(1)O'^{\text{si}}]\phi(2) \right\} \right\} 
$$

$$
i = 1 \ldots 24, \quad j = 1 \ldots 24 
$$

$$
\Delta \phi = \phi(2)^{-1} \phi(1) 
$$

The orientation difference between the two points is estimated in terms of Eq. (10) where $I$ is the second-order identity matrix.

$$
\phi(2) - \phi(1) = (\Delta \phi - I) \phi(1) 
$$

The orientation gradient, $g_{ij,k}$, is the orientation difference over the distance $d$ between the two orientations, Eq. (11). The minimum distance is in this study prescribed by the EBSD step size.

$$
g_{ij,k} = \frac{\phi(2)_{ij} - \phi(1)_{ij}}{dk} 
$$

GNDs are characterized by two vectors, the Burgers vector, $b$, representing the slip direction, and the normalized tangent vector, $t$, indicating the dislocation line direction.

1 Index notation is used: i.e. $A_{ij}$ denotes the components of a second-order tensor; bold letters indicate vectors, $b = b_k \hat{e}_k$ where $\hat{e}_k$ is a Cartesian unit vector; tensor products are represented as $b \otimes t$; the dot product of two vectors is $t \cdot r = r^t t$. 

---

There are 18 possible geometrical configurations of GNDs in fcc crystals (see Table 1).

The orientation gradients can be related to GNDs with the aid of a Frank’s loop construction around an area normal, \( r \). This is also the vector around which the incompatibility is estimated, Eq. (5). The resulting net Burgers vector, \( B \), is the non-zero amount of dislocation lines piercing the area with the normal vector \( r \), Eq. (12).

\[
B = b (\mathbf{t} \cdot \mathbf{r}) = (b \otimes \mathbf{t}) \mathbf{r} \quad (12)
\]

The Frank’s loop compensates the amount of incompatibility that results from the measured orientation gradients, Eq. (13). However, an ambiguity exists when selecting 18 scalar unknown densities \( \rho_{\text{gnd}} \) from the nine given components of the dislocation tensor. In Eqs. (13) and (14), the superscript “\( a \)” represents the slip system.

\[
z_{ij} = \sum_{a=1}^{18} \rho_{gnd}^{a} b_{i}^{a} b_{j}^{a} \quad (13)
\]

A minimization method very similar to Taylor’s minimum work approach [29] is used to identify nine unknown dislocation densities that accommodate a given lattice curvature (see details in Appendix A).

\[
z_{ij} = \sum_{a=1}^{9} \rho_{gnd}^{a} b_{i}^{a} b_{j}^{a} \quad (14)
\]

The method is applied to both 2-D and 3-D tomographic EBSD maps. In the 2-D analysis, the gradients \( g_{ij} \) along the sectioning direction (normal to each 2-D slice), are assumed to vanish, i.e. only the gradients in the two in-plane directions are considered. In the 3-D analysis, the full orientation gradients are estimated by the use of additional orientation information from the neighboring EBSD slices.

2.2. Experimental procedure

The copper single crystals were produced by directional solidification in a Bridgeman furnace. The (111)-oriented surface was wet ground, diamond polished and electropolished. The normal of the cross-section has a [112] crystallographic direction (Z-direction in Fig. 1) [30,31].

The indentation experiments were performed using a Hysitron TriboScope indenter in load-controlled mode. A 60° conical indenter was used with a spherical tip of radius \( R = 1 \mu m \). A loading rate of 1.82 mN s\(^{-1}\) was used with loads of 4000, 6000, 8000 and 10,000 \( \mu \)N producing four indents of different depths. The corresponding depths are 0.46, 0.69, 0.85 and 1.23 \( \mu m \), respectively. The conical surface (60° cone angle) of the indenter is in contact with the material at depths greater than half of the tool radius (0.5 \( \mu m \)) during indentation. Therefore, the tool remains self-similar during indentation only in case of the smallest depth (0.46 \( \mu m \)).

Fig. 2 shows the hardness values obtained for each indent. The area function was determined using PMMA (polymethylmethacrylate) since the indentation depths were too large to obtain a satisfactory calibration from fused quartz.

3-D EBSD measurements of the deformed sample were conducted using a dual-beam high-resolution field emission scanning electron microscopy EBSD set-up together with a focused ion beam (FIB) system [33]. The surface was

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Slip and line directions of dislocations for GNDs in a fcc crystal.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sqrt{2} b ):</td>
<td>110 101 01 110 101 01 110 101 01 110 101 01</td>
</tr>
<tr>
<td>( \sqrt{6} t ):</td>
<td>112 121 211 112 121 211 112 121 211 110 101 01 110 101 01</td>
</tr>
</tbody>
</table>
coated with Pt to avoid Ga+ damage during sectioning. Layers 100 nm thick were successively milled and EBSD measurements were carried out before each milling step [34,35]. The FIB milling procedure required for the tomo-graphic EBSD method leads to minor edge effects at the surface. Therefore, we avoided probing very shallow indents, i.e. the depths were chosen to assemble reliable EBSD data (more details are given in Refs. [30,31]).

2.3. EBSD data

Fig. 3 shows the orientation patterns in three different slices that are equally spaced around the center position (section 39). The patterns before and after the center section are not symmetric with respect to the center of the indent. Section 33 (−600 nm from center layer) reveals sharper orientation changes compared to section 45 (+600 nm from center layer). Also the magnitude of the orientation changes are slightly different among the two outer sections. The difference in the texture evolution with respect to the sectioning direction is attributed to the crystallographic asymmetry of the orientation relative to loading.

3. Experimental results and discussion

3.1. Selection of step size for the GND analysis

The selection of the gradient step size plays an important role in establishing a reliable analysis for the GND density (see $d_k$ in Eq. (11)). As the orientations are integrated and averaged over a larger region, the noise in the data typically drops with increasing step size. On the other hand, a large step size shifts the analysis to a more mesoscopic level so that one might miss the regime where strain gradients become effective. In addition, the small-scale nature of indentation prohibits the use of a large step size approaching the depths of the indents.

The 2-D analysis explained in Section 2 is applied to the EBSD map obtained from the center cross-section using three different step sizes for comparison, i.e. 50, 100 and 200 nm (Fig. 4). A step size above 200 nm is not considered since it approaches the depth of the indents (∼400 nm for the smallest indent). The step size used during EBSD data acquisition was 50 nm. The coarse graining of the data was performed using TSL software.\(^3\)

The step size used for the gradient analysis influences the distribution and the total magnitude of the GNDs. First, due to the averaging of neighboring orientations during coarse graining, the GND density is locally averaged into larger volumes. Second, higher orientation gradients occur in the case of a larger step size (200 nm) when compared to those obtained in case of smaller step sizes. Higher orientation gradients mean higher magnitudes of GNDs. Third, some local peak information may be lost if a coarse step size is selected due to averaging (200 nm in this case). For step sizes of 50 and 100 nm, the peak values remain the same ($3.5 \times 10^{16} \text{ m}^{-2}$), indicating no loss of essential information. The gradient peak value drops to $2.4 \times 10^{16} \text{ m}^{-2}$ for the 200 nm step size which indicates a peak smoothing effect. However, the average overall GND density becomes larger for this step size (Fig. 4).

\(^3\) TSL is a trademark of EDAX/TSL, Draper, UT.
3.2. 2-D GND analysis of five EBSD sections

First, the 2-D analysis is individually applied to five equally spaced cross-sections, i.e. the center slice (section 39) and the two nearest sections on either side (±100 and ±200 nm). The comparison of the four indents in the center section is reasonable since the cone shape effects are expected to be minimum close to the center (Fig. 5).

The sum of the GND densities beneath each indent was calculated by dividing their number by the affected reference volumes (see details in Appendix A). This normalization step renders the GND dislocation number value a density measure. The reference volumes were determined by using a threshold value for the GNDs. Fig. 6 shows a corresponding example calculation containing the total length of the GNDs beneath each indent.

Fig. 7 shows the sum of the GND densities below each indent that are extracted individually for each of the five 2-D EBSD maps. The main observation is that the total GND density decreases with decreasing depth and increasing hardness (Figs. 2 and 7). The GND density variation before and after the center section is attributed to the asymmetry of deformation. The GND densities are larger in the front sections 37 and 38 compared to the sections behind the center.

3.3. 2-D GND analysis of all EBSD sections and comparison to the hardness

Fig. 8 shows the total 2-D GND density below each indent together with the measured hardness values. The total 2-D GND density is obtained by summing the GNDs over all 50 individual 2-D EBSD sections and dividing this number by the total reference volume (Fig. 9). For this purpose the GND density was first calculated separately for each 2-D slice and subsequently the overall density was determined by summing over the 50 sections (see Appendix A). The data show that the GND density decreases with decreasing indentation depth and increasing hardness. The density of the GNDs decreases faster with decreasing indentation depth than the affected reference volume. This leads to an overall decrease in the GND density for smaller indents.

3.4. 3-D GND analysis

Five sections closest to the center are used for the calculation of the GND densities considering also the out-of-plane orientation gradients, $g_{ij} \neq 0$ (Fig. 10). It has to be emphasized, however, that the full tomographic analysis requires the neighboring EBSD slices to be perfectly...
aligned with respect to each other to avoid artificial gradients due to misalignment. The results of the full 3-D analysis (see Fig. 10) are similar to those obtained from the accumulative 2-D analysis (Fig. 8). A slightly higher GND distribution is found underneath the indents in the 3-D analysis compared to the 2-D results. This is due to the small misalignment between the neighboring sections as discussed above.

3.5. Relationship between GND density, indentation depth and hardness

The results reveal three main points. Firstly, the GNDs are not arranged homogeneously but in the form of patterns below and around the indents. Secondly the distribution of the GND density is anisotropic with respect to the center slice. Thirdly, and most importantly, the density of the GNDs does not increase with decreasing indentation depth but rather drops. The GND density decreases from $\approx 2.34 \times 10^{15}$ m$^{-2}$ (largest indent) to $\approx 1.85 \times 10^{15}$ m$^{-2}$ (smallest indent) while the hardness increases from $\approx 2.08$ GPa for the largest indent (1230 nm) to $\approx 2.43$ GPa for the smallest one (460 nm) (Fig. 8).

The first two observations, namely GND patterning and anisotropy of the lattice curvature, were reported before [30–33]. The third observation, however, contradicts the commonly expected inverse relationship between the indentation depth and the density of the GNDs. The GND-based strain gradient theories that are usually applied to analyze indentation size effects associate larger GND densities with smaller indents but not with larger indents [15–17]. Our analysis shows the opposite trend. Of course these results are only valid for the selected indentation depths and type of indenter used. It must be noted in that context that depending on the details of the indenter and tip geometry, the theoretically expected GND density may in special cases also be independent rather than inversely related to the indentation depth. However, a steep decay of the GND density for smaller indents as found in this study was neither experimentally observed nor theoretically predicted up to now.

Another aspect that deserves attention is the indentation size regime. The indents under investigation are between 1230 and 460 nm deep. The indentation size effect is, however, more pronounced for shallower indents. For scales below 460 nm, however, the EBSD method would be of limited help since it is has a lateral resolution of 40–50 nm in copper [33,30,31]. Irrespective of the rather large size of the indents and the comparably weak size effect in this depth regime, the current observations document the occurrence of an indentation size effect (Figs. 2 and 8).
Fig. 5. Five equally spaced cross-sections (center slice, ±100 nm, ±200 nm) through the four indents. Color code: GND density in decadic logarithmic scale ($m^2/C_0^2$). The noise level for the determination of the GND density is $10^{14} m^2/C_0^2$ for a step size of 100 nm.
This means that the GND density cannot in the current case explain the indentation size effect. An alternative explanation could be that the SSDs are responsible for the increase in hardness at small scales. This explanation is, however, also not plausible since the density of the SSDs is linked to the strain and in the current case the total strain decreases with decreasing indentation depth. Another explanation might be a dislocation source limitation effect. Conventional dislocation multiplication and hardening may be expected to be highly localized close to the contact zone between indenter and specimen. The increase in the overall dislocation density (including GNDs) leads to a rapid drop in the free dislocation segment length, which in turn is the main parameter for the activation of dislocation sources in terms of the Frank–Read stress (the activation stress for a dislocation source is inversely proportional to the dislocation segment length). For small indents the interface-to-volume ratio is higher compared to large ones. Assuming a rapid drop in the free dislocation segment length in these narrow contact zones, therefore, might qualitatively explain why smaller indents yield a higher hardness than large ones. The reason, however, would not be a larger density of GNDs alone but their contribution to the reduction in the free dislocation segment length. According to this approach, the mechanical size effect observed in this work would not be due to the hardening.
contribution of the GNDs but due to a dislocation multiplication limitation effect. This idea of source strengthening in conjunction with the indentation size effects has also been discussed by Pippan et al. [32].

4. Conclusions

We studied the effect of the indentation size effect and the GND density by using the following approach. Four indents of different depth and hardness were placed in a Cu single crystal using a conical indenter with spherical tip of radius 1 μm. The deformation-induced lattice rotations were monitored using a tomographic EBSD method. The GND densities were calculated from the measured rotations using the dislocation density tensor which connects lattice curvature to GNDs. This approach allowed us to directly quantify in one experiment both the mechanical parameters (depth, hardness) and the lattice defects (GNDs) that are held responsible for the indentation size effect. The main results are:

Fig. 8. The total GND densities below each indent obtained by summation over all 50 individual 2-D EBSD sections together with the measured hardness.

Fig. 9. 3-D view of the volume around the four indents. Upper row: EBSD pattern quality. The bright gray scale indicates large local lattice distortions. Bottom row: local orientation deviation. The color code is chosen to indicate the largest local orientation changes (15°) in red.
The GNDs have an inhomogeneous distribution underneath the indents with very high local density values. This fact questions the homogeneous distribution assumptions frequently used in the literature. The heterogeneous distribution in GND density might play the key role in strengthening since a long-range back stress field develops due to imbalance in signed dislocation density and resists dislocation motion. The total GND density below the indents decreases with decreasing indentation depth. This observation contradicts the strain gradient theories attributing size-dependent material properties to GNDs. The amount of deformation imposed reduces with decreasing indentation depth. Therefore, SSDs that evolve through strain do not account for the increasing hardness values with decreasing indentation depth. The decreasing dislocation segment lengths associated with decreasing indentation depth are believed to account for the increase in hardness. Explaining size-dependent material strengthening effects by using average density measures for both GNDs and SSDs is not sufficient to understand the indentation size effect.

Acknowledgements

We thank Dr. F. Roters for his valuable contribution to the calculation of gradients. Comments of Dr. Y. Cao on data interpretation are greatly appreciated. One of the authors (E.D.) also acknowledges the lectures of Prof. A.D. Rollett at Carnegie Mellon University.

Appendix A. Details of the analysis of GNDs from EBSD data

The procedure to extract GNDs from EBSD data was implemented in MatLab. The EBSD output file (TSL format) is processed to identify a GND density for every data point. The solution procedure uses five steps: (i) Nine slip systems (corresponding to the nine components of the dislocation density tensor) are selected arbitrarily from the 18 possible GND systems (48,620 possible combinations). (ii) The slip system combinations are checked for geometrical independence and 21,464 linearly independent solutions are found. (iii) A solution is calculated for each possible combination of nine slip systems. (iv) The total GND density, which is the sum of the absolute values of the nine individual GND densities, is calculated. (v) The solution yielding the minimum total GND density among all possible combinations is selected.

There are three main input quantities to the analysis, i.e. the Burgers vector, the gradient step sizes in the \( x \), \( y \), and \( z \) directions, and the lower threshold limit for the misorientations. The Burgers vector is \( 2.5 \times 10^{-4} \) \( \mu \)m in the case of copper. The influence of the step size used for the gradient calculation has been discussed in Section 3.1. A lower limit for misorientations is used to eliminate noise and to save analysis time. On the contrary, some information may be lost depending on the value of the lower limit. Therefore, the lower limit is selected slightly above the noise level (0.5°) so that a minimum amount of information is lost.

The GNDs have to be summed over the deformed region in order to identify the total GND density under an indent. The total GND density is estimated by summing the number of all GNDs and dividing this value by the total reference volume (Fig. 6). Firstly, the region affected by an indent is selected. Secondly, the densities of each point that exceeds the lower bound are summed and the total number of all data points are counted. Finally, the total density of the region is found by dividing the sum of the GND densities by the number of points, Eq. (15). (In Eq. (15), \( N \) represents the number of points and \( A_0 \) stands for the unit area affected.)
\[ \rho_{\text{gnd}}^{\text{total}} = \frac{\sum_{\ell} \rho_{\text{gnd}}^{\ell} A_{\ell}}{\sum_{\ell} N_{\ell} A_{\ell}} = \frac{\sum_{\ell} \rho_{\text{gnd}}^{\ell}}{\sum_{\ell} N_{\ell}} \]  

(15)

References