NANOINDENTATION-INDUCED DEFECT–INTERFACE INTERACTIONS: PHENOMENA, METHODS AND LIMITATIONS

W. W. GERBERICH†, D. E. KRAMER, N. I. T YMI AK, A. A. VOLINSKY, D. F. BAHR‡, and M. D. KRIESE§

Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN 55455, U.S.A.

Abstract—Nanoindentation for measuring thin film mechanical properties is probably the most popular yet ill-understood method due to its inherent complexities. As opposed to burst pressure or microtensile tests of lithographed structures, where relatively uniform stress fields may be generated, the indentation-induced stress gradients can produce unique challenges. Because of the test’s simplicity and ability to mechanically probe the smallest of scales, it is becoming increasingly applied. Five possible stages of deformation are suggested from Hertzian elastic to film delamination and double buckling. In particular metal films on harder substrates are emphasized where it is shown that dislocation nucleation and arrest are only partially understood. Later stages of film delamination are illustrated with Cu/SiO2/Si where it is shown that the true work of adhesion is 0.6 J/m². Current limitations of indentation-induced delamination measures of toughness involve large scatter associated with sensitivity of the fracture radius to the contact radius ratio.

© 1999 Acta Metallurgica Inc. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Copper; Thin films; Yield phenomena; Interface fracture

1. INTRODUCTION

Understanding how various defect types nucleate, grow and produce material instabilities is essential to the accurate measure of flow and fracture by nanoindentation. The present study is a review mostly of our own work drawing upon others where possible to illustrate a few of the unsolved problems. For example, how do dislocations and cracks nucleate under an indenter at a thin film interface? These are not necessarily uncoupled events since a diamond tip may nucleate dislocations at an oxide film/metal interface or a metal/substrate interface or in the substrate itself influencing either film cracking and/or delamination. Five stages of flow and fracture are suggested in Fig. 1. The first stage of elastic Hertzian contact is well understood [1–4] as suggested by the time-line chart of Table 1. Starting with Hertz’s original solutions [1], Hill [2], Tabor [3] and later Johnson [4] further developed the elastic theory into an elastic–plastic theory for describing yielding and work hardening. In particular the plasticity analyses have been shown to work well in the nanoindentation regime for many materials [5–9]. Not so clear is how scale dependent yield nucleation and work hardening might be, i.e. stages II and III of Fig. 1. This paper will particularly address these in terms of dislocation nucleation and yield instability arrest. We will then only address stages IV and V in terms of film decohesion, as measured by a recently developed indentation-induced superlayer technique [10, 11].

2. DISLOCATION NUCLEATION

Initial observations [6, 7] and later extensive papers [12–16] have now dealt exclusively with the fact that nanoindentation can detect the onset of dislocation nucleation. While not nanoindentation, this was first suggested by Gane and Bowden’s [17] and Pethica and Tabor’s [18] seminal experiments exhibiting yield instabilities of sharp gold point contacts via TEM [17] and resistivity [18] techniques. However, as late as 1994, there were persistent reports that the “pop-in” or displacement excursion in metals was most often attributed to the breaking of a surface oxide or passivating layer [19]. In that paper the alternative suggestion was that dislocation nucleation might be responsible. At about the same time [20], experiments in Fe–3 wt% Si single crystals demonstrated a significant time delay for the “pop-in” event when holding at a point well below the yield load for a displacement excursion under constantly increasing load. This hold time as a function of contact pressure and hence local shear stress was suggestive of a thermally activated process for dislocation nucleation. In addition,
Tangyunyong et al. [21] had shown a yield point in gold. More in-depth studies of that phenomenon [15, 16] have clearly shown that up to the yield point in gold Hertzian behavior has been followed. A monolayer of thiol, in the case of the former experiments [15], would not represent the same kind of barrier to dislocation nucleation as an oxide film and from this it appeared unlikely that film breaking was responsible.

Further convincing evidence as to the source of most “pop-ins” or yield excursions in metals come from our own recent experiments on mechanically and electropolished surfaces of single crystals. In (111) NiAl, indentation of the electropolished surface followed Hertzian behavior to a yield point of 3.3 mN and a displacement of 65 nm as seen in Fig. 2. At that point a displacement excursion of approximately 35 nm occurred as the load dropped.

Fig. 1. Various stages of nanoindentation-induced plasticity and interface fracture at a film interface.
No such excursion occurred in the mechanically polished sample which had a high density of pre-existing dislocations due to abrasive wear. As the tip encountered these on contact, no Hertzian regime was experienced with a plastic response at the lowest detected loads. A similar set of coupled experiments was conducted on W single crystals with similar results [14].

One could argue that in both the NiAl and the W that dislocations were already present near the surface of the mechanically polished material, requiring no film breakthrough event. The further argument would be that in the electropolished material the few dislocations emitted under an oxide film or a passivating layer would not cause detectable deviation from Hertzian behavior. However, that could provide the inverse pile-up stress concentration to break an oxide film or a passivating layer as we had once speculated [21]. Nevertheless, we think the film-breaking mechanism not to be the controlling case based upon two observations. The first comes from a staircase yielding phenomenon as seen in Fig. 3 and discussed in more detail elsewhere for Au [13, 15, 16, 20] and W [14]. For gold, Corcoran et al. [16] found the force for the first displacement excursion to depend heavily on crystal orientation. Houston and co-workers [13, 15] have found similar effects. The gold surfaces in these two studies were very different in terms of passivation layers and yet the maximum shear stress for the \(h_{111}\) orientation was very similar as indicated in Table 2, about 3 GPa. If passivation layer breakdown were controlling this would not be the case. In a more detailed study, Kiely and Houston [15] have shown that the resolved shear stress on all three orientations indicated in Fig. 3 was determined to be 1.8 GPa. If passivation layer breakdown were controlling this would not be the case. In a more detailed study, Kiely and Houston [15] have shown that the resolved shear stress on all three orientations indicated in Fig. 3 was determined to be 1.8 GPa. If one compares this with the two model calculations in Table 2, the homogeneous dislocation loop model [13] is about a factor of two low and the unstable stacking model using straight dislocations [12], about a factor of three high.

The second observation negating a film breakdown model is based upon an oxide film assessment alluded to in Table 2. While not necessary in gold studies, inclusion of an oxide film effect in Fe–3 WT% Si AND W with thermally grown films was necessary [14]. The physical principles and simplifying assumptions are that:

1. nearly all metallic surfaces have a protective oxide film;
2. a dislocation loop is nucleated at the film/metal interface due to the preponderance of point defect and/or step sources and the fact that shear stresses increase into the oxide film;
3. the energy barrier is a modified Frenkel sinusoid in terms of Rice’s unstable stacking energy concept [12];
4. a balance of forces for loop nucleation is proposed in terms of the indenter tip forcing the loop into the substrate with appropriate back forces from line tension and image forces from the oxide film and vacuum;
5. to model the stress fields, a homogeneous half-space is considered with isotropic elasticity [20] although crystal plasticity and the critical resolve shear stress was later included [12].

The model for dislocation nucleation at a metallic surface, with further details elsewhere (see Fig. 7 of Ref. [12]) is given by

\[
P = \frac{4}{3} E^{*} R^{2} \left\{ \frac{(h_{0}/R)^{2}}{(h_{0}/R) - \frac{\pi \gamma_{s}}{2 E^{*}}} \right\}^{2} - \frac{h_{0}}{R} \right\}^{3/2}
\]
where $E^*$ is the reduced modulus of the metallic film, $R$ is the tip radius, $h_{fo}$ is the oxide film thickness, and $\tau_c$ is the critical shear stress for nucleating dislocations at the film/substrate interface. This analysis is basically continuum theory for the maximum shear stress under a short cylindrical contact. In these experiments the plan view shape of the tip was slightly elliptical making it as much a short cylindrical contact as a spherical one. For the unstable stacking concept at nucleation, a substitution for the shear stress term becomes

$$\pi \tau_c = \frac{\pi \mu_s b}{2E^*} \sin\beta \cos\alpha$$

where $\mu_s$ is the shear modulus on the active slip plane, $\eta$ is a stress concentration associated with steps at the interface, $h$ is the spacing between slip planes and $\beta^{\text{vac}}$ are image force constants. For Fe–3 wt% Si (100) crystals, the appropriate values are (see Ref. [12] for details):

$$E^* = 195 \text{ GPa}, \quad \beta^{\text{vac}} = 0.1137$$

$$\mu_s = 69.3 \text{ GPa}, \quad \beta^{\text{mic}} = -0.0242$$

Substituting these into equations (1) and (2), the load, $P$, was calculated for dislocation nucleation into iron with thermally grown oxide films ranging from 40 to 175 Å. These are compared to the load at the displacement excursion in Fig. 4. The numbers in parentheses indicate the numbers of indentations and the error bars the full range of values obtained. This relatively tight scatter and the similar slope for the data compared to the theory was encouraging but not definitive regarding dislocation nucleation vs oxide film cracking. Applying the same model to (100) tungsten was disappointing but instructive in two ways. First, the data and theory do not agree even qualitatively as seen in Fig. 5. Second and more importantly the load at the first displacement excursion was independent of the film thickness for the (110) and (111) orientations. This is highly unlikely unless the critical fracture stress for different thickness films was identical. However, since the tungsten oxide modulus is less than tungsten, the interface image force term is of opposite sign compared to iron, and increasing film thickness does not make it necessary to increase forces for dislocation nucleation. In addition, this possibly relaxes the requirement for nucleation at the oxide/film interface suggesting a different model for the data of Fig. 5. Whatever the model, these
experimental observations are consistent with image force concepts rather than film breakdown.

Summarizing this section, the preponderance of evidence is that yield excursions in metal surfaces are largely due to dislocation nucleation. For gold surfaces a reasonable understanding is emerging but for most metals with oxide films, an inadequate treatment exists.

3. DISLOCATION ARREST

After dislocations nucleate at a critical resolved shear stress they continue to emit until a sufficient back force establishes a local equilibrium. In addition to the physical principles and assumptions used for nucleation, it was further assumed that:

1. dislocation velocities are sufficiently rapid so that a quasi-static solution of equilibrium is appropriate at arrest;
2. no interaction occurs between active slip bands;
3. for tip–dislocation and dislocation–dislocation interactions only straight edge segments are considered for summation of forces at arrest;
4. no tip-emission condition is considered.

This has been discussed elsewhere in some detail [12, 20] and if the unstable stacking model for short cylindrical contacts is invoked, the relationship between the displacement excursion, δexc, and the load, P, at which it occurs may be determined. Using a geometric contact radius with equation (19) of Ref. [12] this becomes

\[ P = \frac{3\pi(2\delta_{exc}R)^{1/2}\mu b}{4(\pi - 2^2(1 + e^2)^{-1/2})} \left[ 1 + \frac{\delta_{exc}}{\beta\pi(z - 1)} \right] \]  

where \( \mu b \) is as usual and \( z = r_t \) with \( r_t \) the superdislocation position of the emitted pile-up and \( a \) the contact radius. Note that qualitatively equation (3) gives the displacement excursion increasing with increasing load for a constant tip radius or with decreasing tip radii for a constant load. For a large number of observations in different orientations of tungsten, this is seen to be the case in Fig. 6(a). Since the only unknown is \( z \) this can be tested by making the simplifying assumption that the superdislocation position is at half the plastic zone radius, \( R_p \). From other studies [14, 30] Johnson’s analysis has been simplified to give

\[ R_p = \sqrt{\frac{3P}{2\pi\sigma_{ys}}} \]  

which with the above assumption gives

\[ \frac{r_s}{a} \simeq \frac{R_p}{2} = \sqrt{\frac{3P}{8\pi\sigma_{ys}}} \]  

From this, the only unknown in equation (3), \( \alpha \), is given by

\[ \alpha = \frac{r_s}{a} = \sqrt{\frac{3P}{16\pi\sigma_{ys}\delta_{exc}R_{tip}}} \]  

Originally, we used a constant value of \( \alpha = 3 \) and determined an excursion length vs applied load at yield from equation (3). While not exact, the curves in Fig. 6(b) reasonably well represent the data of Fig. 6(a). Note that the measured parameters \( P, R \) and \( \delta_{exc} \) are quite accurate and the material parameters \( \mu, b \) are well known. The only adjustable parameter is \( \alpha \). Later, when we evaluated the data through equation (6), we found that the average values of the normalized superdislocation position from equation (6) were 2.53, 2.77, 2.05 and 2.23 for the four sets of data representing tip radii of 85, 360, 1800 and 5000 nm, respectively. While the phenomena appear to be well ordered, the summary in Table 3 suggests several limitations to the present model descriptions. In particular, the ill-defined assumptions on either dislocation fraction going into producing the excursion [14] or definition of dislocation positions [13] could be better defined by more sophisticated discretized models.
4. SUPERLAYER INDUCED FILM DECOHESION

With the descriptions of how localized slip may nucleate and arrest, it is appropriate to discuss stages IV and V as illustrated in Fig. 1. Film decohesion has been extensively studied [22–24] but few techniques have allowed very thin films to be decohered. Three exceptions are a microscratch technique [25] and two fine-line techniques involving either a superlayer residual stress [26] or micro-wedge indentation [27] for triggering interface fracture. The latter two are preferred due to more firmly established theoretical foundations. However, these lithographic-based techniques require both masks for specimen fabrication and several steps in preparation and/or testing for quantitative results. This is often time consuming and costly. To circumvent this, we have combined the best features of these techniques [25–27] to produce a superlayer-induced film decohesion as triggered by nanoindentation. As discussed in more detail elsewhere [10, 11], this involves a superlayer, typically 1 μm of tungsten, to drive the delamination once the crack has been nucleated by nanoindentation. The effect of extra stored elastic energy in the superlayer is seen in Fig. 7(a). Here, a comparison of nanoindentation-induced delamination with and without a superlayer is shown. The material system was Cu deposited on (001) silicon wafers which have a thermally oxidized SiO₂ surface, one wafer additionally having a sputter-deposited superlayer of W. See Ref. [10] for more details. For a 250-mN indentation it is seen in Fig. 7(a) that the delamination

<table>
<thead>
<tr>
<th>Investigators</th>
<th>Phenomena, δexc</th>
<th>Method</th>
<th>Theory</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bahr et al. [14]</td>
<td>Load–displacement excursion</td>
<td>Superdislocation equilibrium</td>
<td>( F_e(\beta, \theta, \mu_r, E^*, R, P, \tau_t) )</td>
<td>(i) One-dimensional (ii) One ill-defined assumption required (iii) Three measured parameters required ( \mu_r, E^*, \tau_t )</td>
</tr>
<tr>
<td>This workshop</td>
<td>( P-\delta ) excursion</td>
<td>Superdislocation equilibrium</td>
<td>( F_e(\beta, \theta, \mu_r, R, P, b) )</td>
<td>(i) One-dimensional (ii) One ill-defined assumption required (iii) No friction stress included</td>
</tr>
</tbody>
</table>

\( a = r/a, \perp \) position/contact radius; \( \beta = \perp \) fraction; \( \theta = \) slip plane function; \( \mu_r, E^* = \) moduli; \( P = \) load; \( R = \) tip radius; \( b = \) burgers vector.

Fig. 7. Optical images of delamination radii produced after indenting (a) to 250 mN of load, without and with a tungsten superlayer on Cu: note the factor of four difference in scale; (b) for delamination of Al on sapphire with and without a carbon interlayer (courtesy J. Schneider, Sandia, Livermore).
radius, \( D_r \), is about four times as large utilizing the superlayer. As large values of \( D_r/a \) are preferred with respect to obtaining precise and reproducible interfacial energies, the advantage is obvious. Clear differences are seen in the results of Schneider et al. [28], where fairly narrow error bars were observed in the measure of aluminum/sapphire interfaces. With an interlayer of carbon to promote poor adhesion, \( G_i \) was 0.8–1.3 J/m\(^2\) whereas the adhesion energy was 5.1–6.1 J/m\(^2\) for the Al/Al\(_2\)O\(_3\) interface. In both cases, a tantalum superlayer was used to assist the delamination process during nanoindentation.

To examine the resistance side of the energy balance, we have determined both the mechanical properties and interfacial energies over a larger range of Cu film thicknesses on SiO\(_2\)/Si substrates. An example of grain size determination by AFM with both height mode (left) and deflection mode (right) imaging is shown in Fig. 8 for a 505-nm-thick film. Coupled with standard techniques for determining modulus [7] and yield strength [29, 30], a Hall–Petch type plot is shown in Fig. 9. It is seen that neither a \( 1/d \) nor a \( 1/d^{1/2} \) functional dependence results, but nevertheless, the measured yield strength values determined can be used for resistance calculations. As discussed in more detail elsewhere [31], an upper bound determination of plastic energy dissipation assuming a fully plastic metal layer is given by

\[
G_i \approx h_i \sigma_{ys}^2 \left\{ \ln \left( \frac{h_i}{b} \right) - 1 \right\}
\]

where \( h_i \) is the film thickness, \( \sigma_{ys} \) is yield strength, \( E_i \) the film modulus and \( b \) is the burgers vector.

The simplifying assumptions here are that the burgers vector is an appropriate cut-off radius for integrating energy density within the plastic zone which extends in an elastic–perfectly plastic material from the cut-off to the copper/superlayer interface. Also, no contribution from \( W \) is added at the crack arrest point which is a reasonable distance from the indentation site. Finally, equation (7) applies to \( h_i \gg b \) which is the case here. Since the thickness has been accurately measured by either ellipsometry, AFM or step-profilometry, and \( E_i, b \) are well known, the only parameter with much variance is the yield stress as seen in Fig. 9. Compared to previous [10] data and data from this investigation, experimental values are seen to fall nearly an order of magnitude below this upper bound in Fig. 10. In fact this is expected if one converts the \( G_i \) values shown to a plastic zone size and compares these to the film thickness. For example, at a 2000-nm thickness, the 790 nm plastic zone size calculated is only about 40% of the film thickness while at a 200-nm thickness, the calculated zone of 33 nm is only 17% of the film thickness. Using only a portion of the metal film for plastic energy absorption clearly would place the data below the upper bound of equation (7).

A further point of interest from Fig. 10 is that the strain energy release appears to plateau near a thickness of about 80 nm, being

\[
G \approx \frac{h \sigma_{ys}^2}{E_i} \left[ \ln \left( \frac{h_i}{b} \right) - 1 \right]
\]

Fig. 8. Grain size assessment of four different Cu film thicknesses. Left column are height mode and right column are deflection mode AFM images.

Fig. 9. Yield stress as a function of copper film grain sizes.

Fig. 10. Measured interfacial strain energy release rates of Cu films as a function of copper thickness.
In Table 4, a summary of the phenomena, methods and limitations using superlayers to measure adhesion energies is given. This suggests that considerable progress has been made but that appropriate theoretical refinements of plasticity contributions and near-tip stress field interactions are yet to be accomplished. One can point to the large variation in strain energy release rate for relatively small changes in delamination [36, 37] radius to indentation contact radius as being the major contributor to the scatter indicated in Fig. 10. This is discussed in more detail elsewhere [31].

5. CONCLUSIONS AND SUMMARY

In the course of this we conclude that:

1. Discontinuous yield excursions in metallic systems are most often associated with dislocation nucleation and not passivation or oxide film breakdown.
2. In metallic substrates and films yield excursions appear to be well ordered with dislocation field back forces exerting equilibrium with tip forces.
3. For deeper indentations into films precipitating delamination superlayer techniques appear best suited both for reproducibility and the ability to measure the true work of adhesion, 0.6 J/m² in the case of Cu/SiO₂/Si.
4. All of the above models addressing these phenomena are in their initial stages of formulation and currently have severe limitations.

Acknowledgements—The authors gratefully appreciate support from the Office of Naval Research under Grants E-25-T46-51/ONR and N00014-91-J-1998 for three of us (W.W.G., D.E.K. and D.F.B.) from the Department of Energy under grants DOE/DE-FG02-96ER45574 and DOE/DEFG02-85ER13433A011 for four of us (W.W.G., A.F., N.T. and M.D.K.) and under grant NSF/CDR-8721551 for instrumentation usage. We would also like to thank Judy Schneider and Neville Moody of Sandia National Laboratories, Livermore, for the use of Fig. 7(b).
REFERENCES

34. Li, J. C. M., Scripta metall., 1986, 209, 1477.
35. Rice, J. and Thomson, R., Phil. Mag., 1974, 29, 73.