Plastic Relaxation at Crack Tip: from Brittle to Ductile Behaviour

Gérard Michot\textsuperscript{1} and M. A. Loyola de Oliveira\textsuperscript{2}

\textsuperscript{1}Ecole des Mines de Nancy, Parc de Saurupt, 54042 Nancy, France
\textsuperscript{2}UFES, C.T., Av. Fernando Ferrari, Goiaberas, Vitoria, 29000 E.S., Brazil

Experimental analysis of plastic zone growth in silicon have recently underlined the fact that two different populations of dislocations sources have to be considered. A small number of primary ones, linked to defects of the crack front, emit tens of dislocations in primary planes. A large number of secondary ones, resulting from cross-slip of primary dislocations coupled with the mechanism of stimulated emission, only emit a few dislocations. This paper shows that primary dislocations cannot account for observed brittle to ductile transition.

(Received July 14, 2000; Accepted August 28, 2000)

Keywords: brittle to ductile transition, crack, dislocation source, shielding, silicon

1. Introduction

At the point where higher stresses than those foreseen during the design of the structure arise, failure of load bearing structures will take place. Most of the time, material deficiencies in the form of pre-existing flaws act as stress risers able to induce such cracking. Failures can either be of the yielding dominant or the fracture-dominant types. The former is preferred because of the associated higher energy consumption rate linked to plasticity. On the microscopic level, plastic deformation rate depends i) on dislocation density (which is function of the ability of the material to nucleate and to multiply dislocations) ii) on dislocation mobility (which is function of both stress level and temperature). For any kind of material, it must be possible to determine a low enough loading rate and/or a high enough temperature above which plastic deformation around the defect will prevent the defect from undergoing brittle fracture. If the so defined brittle to ductile transition temperature (BDTT), for a given loading rate, is lower than the service temperature, the material can then be qualified for the intended application. Apart from this engineering point of view, it would be worthwhile to ask several fundamental questions: i) why do not all materials exhibit a brittle to ductile transition, ii) is this transition an intrinsic property of the material? iii) through which physical mechanisms does plastic deformation reinforce the material? iv) what is the relevant scale parameter adapted to the transition description, i.e., is a dislocation micro-mechanical treatment more worthy than a continuum mechanics approach? This article aims to, at least partly, answer to these questions.

2. The Static Approach

2.1 Stress field around a crack tip

The stress field around a sharp crack in a perfectly elastic medium is usually characterised by a single parameter, the applied stress intensity factor $K_A$ (s.i.f.) which measures the strength of the singular term of the elastic solution. If one considers a structural piece (of width $W$) submitted to a constant external load $\sigma$, the introduction of a flaw of size $a$ induces a s.i.f. $K_A$ given by:

$$K_A = \sigma \sqrt{\pi a} f(a/W)$$

where $f(a/W)$ is a calibration function accounting for the finite width of the sample. As it is well-known in fracture mechanics, the load at fracture $\sigma$, as derived from the critical condition for brittle fracture ($K_A = K_C$), is a function of the defect size. $K_C$ is the material’s toughness.

The situation is radically different for materials in which plastic zones (p.z.) develop at crack tips. The stress distribution can be obtained through different methods. One of these is to model the p.z. by a continuous distribution of infinitesimal dislocations.\textsuperscript{1} The equilibrium solution is then characterized by the fact that the singular term in $r^{-1/2}$ vanishes, i.e., by a complete plastic relaxation. Since the stress within the p.z. is uniform and equal to the yield stress, even at crack tip, a fracture criterion based on local stresses cannot be proposed. An alternative one based on critical opening displacement at the crack tip (CTOD) is more relevant to this problem.

It is more difficult to account for the behaviour of materials which show transitions from brittle to ductile (BDT). Since full relaxation results from any model assuming there is physical continuity between the crack and the plastic zone, a continuity solution was introduced:\textsuperscript{2} such an elastic enclave (dislocation free zone) could be justified by both the necessary average distance between dislocations and the action of image forces. However, a more valuable argument can be put forward. The effect of the plastic zone is to reduce the applied s.i.f. $K_A$ to an effective s.i.f. $K_E$ by the shielding $K_D$ of the crack: $K_E = K_A + K_D$ ($K_D < 0$). An equilibrium situation is reached when $K_E$ falls below $K_{MIN}$ under which dislocation sources are inactive. Such threshold values for dislocation emission could account for some hierarchy between the different materials: near zero value for ductile materials, $K_C$ for brittle ones, intermediate values corresponding to semi-brittle materials.

2.2 Static models of brittle to ductile transition

In their first theoretical paper, Thomson and Rice attempted to evaluate a threshold value for dislocation nucleation at a crack tip.\textsuperscript{31} They defined a material as being intrinsically brittle or ductile according to whether dislocations were or were not emitted at a stress intensity factor lower than the tough-
ness of the material. To be emitted, a dislocation has to overcome the energy barrier resulting from the balance between the applied external work and the dislocation self energy. This model has the merit of making it possible to distinguish two categories of materials: very brittle and very ductile ones (the FCC, some HCP metals, a BCC one, Ta). However, the validity of this model is limited in the sense that the calculated energy barrier is prohibitive for semi brittle materials (for covalent materials and most of the BCC metals). This barrier can be lowered (but not enough) through introduction of a Peierls-Nabarro potential along the slip plane.\(^{4,5}\) Even if atomistic calculations\(^{6-9}\) bring a complementary insight into this nucleation problem they do not solve this discrepancy.

The generation of dislocations at crack tip has been successfully investigated experimentally in silicon through \textit{in-situ} X-ray topography. It has been shown that the crack’s ability to generate dislocations dramatically decreases when the crack’s quality is improved.\(^{10-13}\) This means that nucleation takes place heterogeneously on defects along the crack front. Introduction of defects in the emission models is relatively recent\(^{14,15}\) and still partially implemented. One can derive from this fundamental observation that, as far as the crack front is the only source of dislocations, a material’s property, as important as the BDT, should be extrinsic.

3. Dynamic Models of Brittle to Ductile Transition

It is implicitly assumed with the static model that once a dislocation is emitted under a \(K_{\text{MIN}}\) value, the crack is shielded (eventually blunted). Hence, unless the applied load is raised, no more dislocation can be emitted. By gradually increasing the load, a progressive emission of new dislocations is induced and, thus, shielding will steadily increase. When the shielding rate is lower than the loading rate the material will fail by cleavage (when existing) when the effective s.i.f. reaches \(K_C\). Otherwise the crack will plastically open. Thus, the material’s behaviour will depend on the competition between loading rate and shielding rate. The latter is controlled by the nucleation rate and by the rate at which dislocations move away from the crack, i.e., by the dislocation mobility. Since dislocation mobility is strongly temperature related, one can expect a change in the balance between loading rate and shielding rate according to the temperature. For any given loading rate, one can find a critical temperature above which brittle fracture is avoided, the so-called BDT.

The amplitude of two dimensional shielding was quantified long ago.\(^{3,16-18}\) Hirsch \textit{et al.}\(^{19}\) have chosen a non equilibrium 2D simulation route in order to appreciate the respective influences of the loading and shielding rate parameters, introducing explicitly dislocation mobility. Since the effective s.i.f. \(K_E\) does not increase much above \(K_{\text{MIN}}\) between two consecutive emissions, the authors concluded that such 2D representation could not represent the transition. They then turned their model to a pseudo 3D situation introducing the fundamental concept of a critical distance between sources, \(d_c\), below which all vulnerable points on the crack profile can be shielded \((K_E < K_C)\), under given experimental conditions. The authors then consider that as soon as the distance travelled by the first emitted dislocation (within the scope of their 2D model) is comparable to \(d_c\), that the crack is then totally shielded.

The \(d_c\) parameter which schematises the distance between active slip planes makes it possible to intuitively see the roles played by the dislocations and their discrete sources in fracture. The paragraph 4 will emphasise about its influence on the BDT.

4. Experimental Procedure

Tapered double-cantilever beam samples of approximately 0.6 mm thickness were machined in silicon wafers. Pre-cracking was achieved at room temperature using St. John’s wedge technique.\(^{20}\) Two possible cleavage planes have been used: \{111\} (Fig. 1) and \{110\}.

Pre-cracked samples were loaded in mode I in a high temperature stage mounted on the two axis spectrometer of the topography beam port of the synchrotron facility LURE-DCI. Experiments were conducted under a reducing atmosphere (10% H\(_2\), 90% N\(_2\)). Samples were loaded at constant rate \(dK/dt\) and were at the same time observed \textit{in-situ} by synchrotron X-ray topography. Experiments were stopped at fracture (in the case of brittle fracture) or at a conventional \(K\) value \(\sim 1.7K_C\) (in the case of plastic opening).

Samples which did not break during \textit{in-situ} loading were cooled down, with the applied load, in order to freeze-in the dislocation pattern. After characterisation by conventional X-ray topography the samples were loaded to fracture. Broken samples were etched and examined in an optical or scanning electron microscope to determine the actual crack tips’ geometry and the dislocations’ distributions.

5. Dislocation Sources in Silicon

Little is known about the nature of nucleation sites of dislocations in general. In the absence of bulk sources in near perfect materials such as silicon, emission is restricted to the crack tip. Experiments performed on silicon showed that the ability of a crack front to generate dislocations is dramatically reduced by improving the cracks’ quality.\(^{10,11}\) underlining

---

Fig. 1 Tapered double cantilever beam sample (left) machined in a [\(\overline{1}2\)] silicon wafer. The (111) cleavage plane propagates toward [110]. Corresponding orientation of the Thompson tetrahedron with respect to the crack position (right).
the inhomogeneous nature of nucleation on the crack front. Hence, defects are required to authorize dislocation emission. Recent calculations\(^{15}\) confirm our observations.

However, only in some cases, it is possible to link the origin of dislocations with an observable ledge of the crack front. We call this site "primary sources". What role do these sources play on the BDT?

5.1 Density of primary sources

Typical dislocation patterns around sources as observed on the crack plane for (100) and (112) samples are given respectively in Figs. 2(a) and (b). In the former case only one glide system was activated, giving birth to 90 dislocations. The latter case exhibits typical source configurations with activation of the three possible slip planes (a detailed analysis of this configuration can be found in\(^{21}\)). Figure 3(a) shows the distribution of the primary sources along a complete crack front where six primary sources are well identified ([111] cleavage). This sample was loaded at \( T = 885 \text{ K} \) under \( dK/dt = 62 \text{ Pa}\text{\sqrt{m}}\text{s}^{-1} \) and failed at \( K_A = 1.11 K_{IC} \). The maximum dislocation loop extension (as measured on the cleavage plane) was \( \sim 80 \mu\text{m} \).

Two parallel (1[1]) slip planes, emanating from two sources are drawn. If further development were allowed (Fig. 3(b)), the actual loops would grow until they touched the (112) sample boundary. The intersection points must be aligned along a [132] direction. The number of primary sources can then be deduced from the number of such parallel rows.

Observation of the etch pattern on the (112) face of a sample (Fig. 4) loaded 2 hours at \( T = 1023 \text{ K} \), under \( K_A = 0.66 \text{ MPa}\text{\sqrt{m}} \), indicates that three well defined sources have emitted more than five loops. This observation is quite usual in our experiments. With less than ten primary sources, the minimum distance between sources is of about 50 \( \mu\text{m} \) (N.B.: one must also point out that sources are often detected at the intersection of the crack front with the free (112) surfaces. This is probably due to the plane stress state which prevails at these places).

5.2 Contribution of the primary dislocations to crack shielding

Is the contribution to shielding due to primary dislocations important enough to account for the brittle to ductile transition? It is now possible to evaluate the amplitude of the shielding induced by an individual dislocation loop in the 3D geometry\(^{22,23}\). The result depends on the orientation of the slip system with respect to the sample geometry (on the loop size and on its position in its glide plane).

One defines this latter parameter by the emission angle \( \phi_0 \) for which the work performed by the external loading is maximum. The resolved shear stress \( \tau \) acting on a given slip system can be derived from the mode I stress tensor\(^{21}\) in

---

Fig. 2  Etch patterns observed on [110] (a) and [111] (b) cleavage planes after complete fracture of the sample (optical microscopy, markers: 10 \( \mu\text{m} \)). The position of the crack front during plastic deformation is materialized by a dashed line on the micrograph. During room temperature propagation its position has moved downwards.
the following form: \( \tau = K_A H(\Phi) / \sqrt{2\pi R} \), where \( \Phi \) and \( R \) are the polar co-ordinates in the slip plane (with \( \Phi = 0 \) along the intersection of this plane and the cleavage plane). Results are expressed in the form of equal-stress contours \( G(\Phi) = H(\Phi)^2 \) (Ref. 24). Figure 5 represents such a curve for the [011](111) glide system. For loops of small size, one may reasonably assume a circular shape. Since the observed large hexagonal loops are attached to the crack tip, we can assume that this will hold true for emerging loops. Under this condition the external work spent to grow a circular dislocation loop (like the one shown Fig. 5) from zero to a radius \( R \) is equal to:

\[
W = \eta K_A b R^{3/2}
\]

with

\[
\eta = \frac{4}{3\sqrt{\pi}} \int_{-\pi/2}^{\pi/2} H(\Phi) \cos^{3/2}(\Phi) \, d\Phi
\]

Let us consider the emission of [011](111) glide system in the negative \( \Phi \) sector. According to the equal stress contour (Fig. 5), the emission angle \( \Phi_0 \) should be a slightly smaller than \(-\pi/2\) in order to maximize the work. Since the emission angle \( \Phi_0 \) sector is restricted to \(-\pi/2\) and \(+\pi/2\) if a full circular loop is considered, a minimum value of \(-\pi/2\) should be chosen for \( \Phi_0 \). In order to avoid numerical complications arising from long dislocation segments parallel to the crack plane in the case of large loops one should choose an angle \( \Phi_0 = -85^\circ \). It has been verified that this does not significantly change the shielding amplitude.

The evolution of mode I shielding \( K_D \) as a function of the position \( z_0 \) along the crack front is given in Fig. 6. One can notice an anti-shielding effect along the negative side of the \( z \) axis, \( i.e. \) at the left of the emission point \( S_0 \) of Fig. 2(b), and a larger shielding effect along the positive side. The maximum shielding amplitude \( K_{D\text{MAX}} \) is \( K_{D\text{MAX}} \approx -0.12K_0 \), \( i.e. \) \(-0.19K_{IC} \) (with \( K_{IC} \approx 0.93 \text{MPa}\sqrt{\text{m}} \) for (111) cleavages in silicon and \( K_0 = \mu\sqrt{b} = 1.5 \text{MPa}\sqrt{\text{m}}, \) where \( \mu \) is the shear modulus). Hence the dislocation has an huge effect on the crack front.
This maximum shielding amplitude $K_{D_{\text{MAX}}}$ is, at first approximation, practically independent of the loop size. However, the localization of the effect varies. The crack front feels the influence of the dislocation loop on a distance $L$ called the influence range. According to the St Venant principle, its value should be something like three or four times the loop size, depending on the choice of the stress amplitude below which the effect is considered as negligible. The boundary stresses chosen for the integration in our numerical calculations are $0.1\mu$ in the inner dislocation core, and $10^{-5}\mu$ at the outer integration boundary. Under these precise conditions the evolution of $L$ with the loop radius $R$ can be plotted (Fig. 7). The numerical approximation $L = 6.18 \cdot R^{0.627}$ holds up to 50,000b. It means that for a small loop (diameter 100b) the crack front is perturbed on four times the loop size on each side of the emission point, while for a large loop (diameter 10^4b) the crack front is perturbed on each side only on a distance close to a half loop size. For loop sizes larger than 50,000b, the crack feels the influence of a nearly straight infinite dislocation, the influence range thus tends toward a constant. Outside the influence range, at a distance $L$ from the emission point, the $K_D$ value is less than 0.1% of the maximum shielding amplitude, i.e., $\sim 2 \cdot 10^{-4} K_{IC}$.

A third parameter easier to handle is the shielding amplitude averaged along the distance $L$, $(K_D)$. It has been introduced in order to evaluate the resulting shielding induced by a population of dislocations. Its evolution with the loop radius is given in Fig. 8. It first decreases and then tends toward a constant value for loop sizes larger than $\sim 50,000b$.

The sample shown Fig. 3(a) failed under an external loading of 1.1$K_{IC}$. Is this 10% increase representative of shielding? The sources are roughly separated by 100μm. In order to shield the crack front between two sources, one needs dislocations whose influence range is of about 100μm, i.e. 260,000b. It is unlikely that this value can be reached according to the saturation tendency noticed in Fig. 7. Furthermore, if one adds up the contribution of all the dislocations emitted by one source (like the one observed in Fig. 2(b) which includes about 200 dislocations), one can obtain a maximum upper bound for the total shielding amplitude outside the influence range of about 0.04$K_{IC}$. This value is still boldly overestimated. Hence, one can definitely rule out the possibility of some influence of these sources on the fracture stress of this sample. The noticed 10% toughness increase remains within the limits of experimental scattering (on both toughness and applied load evaluations).

Since the number of primary sources in our experiments is always under ten, this conclusion will always hold true. Crack front shielding at a mid-distance from two primary sources induced by dislocations emitted from these sources is insufficient to avoid brittle fracture. Other sources are needed.

5.3 Secondary sources

If well defined rows of dislocations emanating from primary sources are observed on the outer part of the plastic zone (Fig. 4), it is difficult on the contrary to find slip traces along the primary plane in the inner part (left side of the same figure). One can conclude that plastic relaxation starts by the triggering of a few primary sources which emit tens of dislocations (well defined rows), rapidly followed by the activation of many secondary sources which emit only few dislocations on other planes. This is confirmed by AFM observations of the crack surfaces after extensive plastic deformation which show that the distance between slip lines is much smaller than the distance between primary sources. A mechanism of dislocation source multiplication has been proposed. One of the dislocations emitted at a primary source cross-slips to a plane where it can be attracted by the crack. The intersection of this dislocation with the crack front at a non-shielded spot triggers a secondary source through the stimulated emission process. The process then repeats itself giving rise to an "avalanche multiplication" of dislocations which strongly shield the crack. Could we get an idea of the creation rate of
such secondary sources?

5.4 Production rate of dislocations at temperatures close to the BDTT

For a given loading rate, brittle fracture takes place below the transition temperature when the effective s.f. \( K_F \) overcomes the material’s toughness \( K_{IC} \) at the weakest point of the crack tip. Above this temperature \( K_F \) is necessarily smaller than \( K_{IC} \). In other words, \( K_F \) must go through its maximum at the transition temperature. The decrease in crack extension force due to plasticity, \( (K_F^2 - K_F^2)/H \) must be equal to the force applied to the dislocations of the plastic zone (with \( H \) an elastic constant). Considering the plastic zone as a super dislocation Nb submitted to a mean resolved shear stress \( \tau \), this force equals \( N \tau b \). In situ, kinetic measurements of plastic zone growth under a constant loading rate at a temperature only slightly above the BDTT proves that the stress acting on the external dislocation is constant during the test (Fig. 9).\(^{25} \)

Since \( K_F \) is also nearly constant (close to \( K_{IC} \)) one can assume that the shear stress distribution does not evolve with time within the plastic zone. The derivation versus time of the decrease of the crack extension force leads to:

\[
K_A dK_A/dt \sim \tau b dN/dt.
\]

That means that under a constant loading rate the dislocation emission’s rate increases linearly with time (such a conclusion is incompatible with the results of the dislocation dynamics models which are based on a constant number of sources and which will rather show a saturation of sources’ activity). Since each secondary source emits roughly the same number of dislocations one may assume that the rate of creation of secondary sources also increases linearly with time (this point will be verified through simulation experiments). Because they cover the whole crack front, one can expect such secondary dislocations to have a strong shielding effect and hence to be at the origin of the BDTT.

6. Conclusions

In every experiment achieved on nearly intrinsic silicon the brittle to ductile transition in silicon is loading-rate dependent with an activation energy close to 2 eV over a wide temperature range. Since this value is the same as that measured for dislocation mobility, it has long ago been recognised that the transition was controlled by dislocation mobility. However, large shifts (several hundred of degrees) are observed in the Arrhenious plot according to the experimental conditions.\(^{20} \)

They are attributed to differences in the source’s activities. The smaller the number of primary sources, the less efficient will the multiplication mechanism of secondary sources be: the more brittle the material will be. In absence of primary sources due to the crack front perfection, as it is the case for GaAs,\(^{10} \) the material remains totally brittle.

What is the influence of the primary sources in usual materials? In the case where these sources are only needed to trigger secondary sources their role will be negligible in the presence of bulk sources. As a matter of fact, in materials containing grown-in dislocations, activation of Frank-Read sources close to the front stimulates emission at the crack tip, triggering the avalanche mechanism which shields the crack.

In ductile materials where dislocation mobility is not limited the effect is quasi instantaneous, no BDT is expected. In semi-brittle materials kinetic limitations lead to the competition between loading and shielding rates at the origin of a BDT.

Dislocation-free silicon appears to be a special case.

REFERENCES