Crack Nucleation Related Dislocation Dynamics
A Numerical Study on the FCC Metal

by

Harold Wing Hei Kwok

Thesis submitted to
The Faculty of Graduate and Postdoctoral Studies
In partial fulfilment of the requirements for the Degree of
Master of Science

uOttawa

Summary

The present thesis consists of a number of studies on the fatigue related to dislocation dynamics phenomena. These studies were intended to provide an in-depth understanding of the dynamics at the atomic scale. Until now, our understanding on such atomistic dynamics remains sparse and incomplete. Special attention was paid to two particular dislocation processes, which were the dislocation dipoles disintegration and the triple junction deformation. Both scenarios are closely related to a fatigue crack initiation at the cycle fatigue loading condition. Studies were carried out by means of numerical molecular dynamics simulation. The software used in these studies was developed by the author, and was tailored for superior performance and efficiency.

One of the most important questions addressed in this thesis was the intermediate disintegration pathway of a group of dislocation dipoles. The disintegration of dipoles was realised in experiments for decades. It was known to produce massive amount of point defects as a result. However, the process cannot be observed directly. The steps between dislocation dipole accumulation to point defects production are in the missing puzzle. Molecular dynamics simulation with a precise force description was deployed to investigate the scenario in a pure aluminum sample. The analysis suggests that cross-slip is the major mechanism for the disintegration. The cross-slip leads to the formation of stacking fault tetrahedrons, which then collapse to form vacancies clusters.

Another question addressed in the thesis was the deformation mechanism at the vicinity of a triple junction. It is known that such process leads to the embrittlement of a material and is related to the intergranular crack initiation. Models in the past did not consider the dynamical nature at the molecular scale. Investigation was therefore carried out to address this issue. The study illustrates that triple junction can deform via dynamic recrystallization and amorphization. Such illustration can be useful for the development of a more comprehensive model of crack initiation at a triple junction.

Statement of Contribution

The research studies and writings presented in this thesis are my original works. Studies discussed in chapters 4 and 5 are, to the best of my knowledge, new and of scientific values. Research was executed under the supervision of Dr. Kuiying Chen at the National Research Council of Canada (NRC). Dr. Chen has provided me with valuable advices on scientific and social conformity. The simulation software deployed in this research is developed solely by me. It was written from scratch with FORTRAN. Research ideas are original. They were inspired by a collection of previous works discussed in the literature.

Table of Contents

List of Acronyms

2NN	Second Nearest Neighbour
AEAM	Analytic Embedded Atom Method
BCC	Body-Centered Cubic
CPU	Central Processing Unit
Cr	Chromium
CSL	Coincident Site Lattice
CV	Coefficient of Variation
DDRX	Discontinuous Dynamics Recrystallization
DFT	Density Functional Theory
DRX	Dynamics Recrystallization
EAM	Embedded Atom Method/Embedded Atom Model
FCC	Face-Centered Cubic
FD	Force Decomposition
FM	Force-Matching
GB	Grain Boundary
GDB	Grain Boundary Dislocation
GPU	Graphic Processing Unit
HCF	High Cycle Fatigue
HCP	Hexagonal Close Packed
LAMMPS	Large-Scale Atomic/Molecular Massively Parallel Simulator
LCF	Low Cycle Fatigue
LDA	Local Density Approximation
MD	Molecular Dynamics
MPI	Message Passing Interface
MPMDS	Multi-Parallel Molecular Dynamics Simulator
Nd	Neodymium
NDE	Non-Destructive Evaluation
PSB	Persistent Slip Band
RAM	Random Access Memory
SEM	Scanning Electron Microscopy
SFT	Stacking Fault Tetrahedron
SIMD	Simple Instruction Multiple Data
TJ	Triple Junction
ZSK	Zener-Stroh-Koehler

Chapter 1

Micromechanics of Fatigue Crack Nucleation and Growth: a Literature Review

[1]*Harold Wing Hei Kwok,* [2*]*Kuiying Chen*

(published in NRC/IAR report. LTR-SMPL-2009-0236)

[1] *Department of Physics, University of Ottawa*
[2] *Structures & Materials Performance Laboratory, Institute for Aerospace Research*
National Research Council Canada, Ottawa

Abstract

Fatigue crack nucleation in gas turbine engine components is a complicated process that plays an important role in determining the life of these components. It involves dynamic interactions among various types of microstructures such as dislocations, grain boundaries, surfaces, vacancies, as well as interstitial and external obstacles. This process is also associated with creep, diffusivity, and microstructure evolution of alloys at elevated temperatures. Fatigue crack nucleation essentially involves multiple-length scale physics that ranges from the atomistic scale to the microstructure level. It is also related to bond breaking and reformation during the processing. This report presents an overview on current status of theories and models on fatigue crack nucleation. It identifies and describes different approaches on evaluating fatigue crack nucleation processes in ductile alloys. The scope of this report includes (1) dislocation-based analytical models of crack nucleation, (2) probabilistic-based empirical models, and (3) finite-element simulations based numerical models. Specific theories and models on fatigue crack nucleation are analyzed for both low cycle fatigue (LCF) and high cycle fatigue (HCF). The models emphasize subsurface crack nucleation at a grain boundary or at the interface between different phases. The objective of this report is aimed at relating the scientific description of materials behaviour with engineering challenges of gas turbine engine design and operation.

1.1 Introduction

Advancements in gas turbine engines for military aircraft rely heavily on leading edge material technology. The pursuance of performance, reliability, cost efficiency and better engineering design protocol is the driving forces behind the progress. In an aircraft gas turbine engine design, the efficiency and reliability are limited by the material's capability to tolerate the operating temperatures and fatigue loading conditions.

Cyclic fatigue is the dominant cause of component failure. Figure 1.1 shows statistics from the US Air Force on the damage mode distribution of jet engine components[1]. It illustrates the significance of both high and low cyclic fatigue (HCF, LCF) when compared to other failure modes. The complexity of the problem and the high occurrence of incidences signified the need for in-depth investigation and thorough fundamental research. Particularly reliable fatigue life prediction model incorporating crack nucleation is needed.

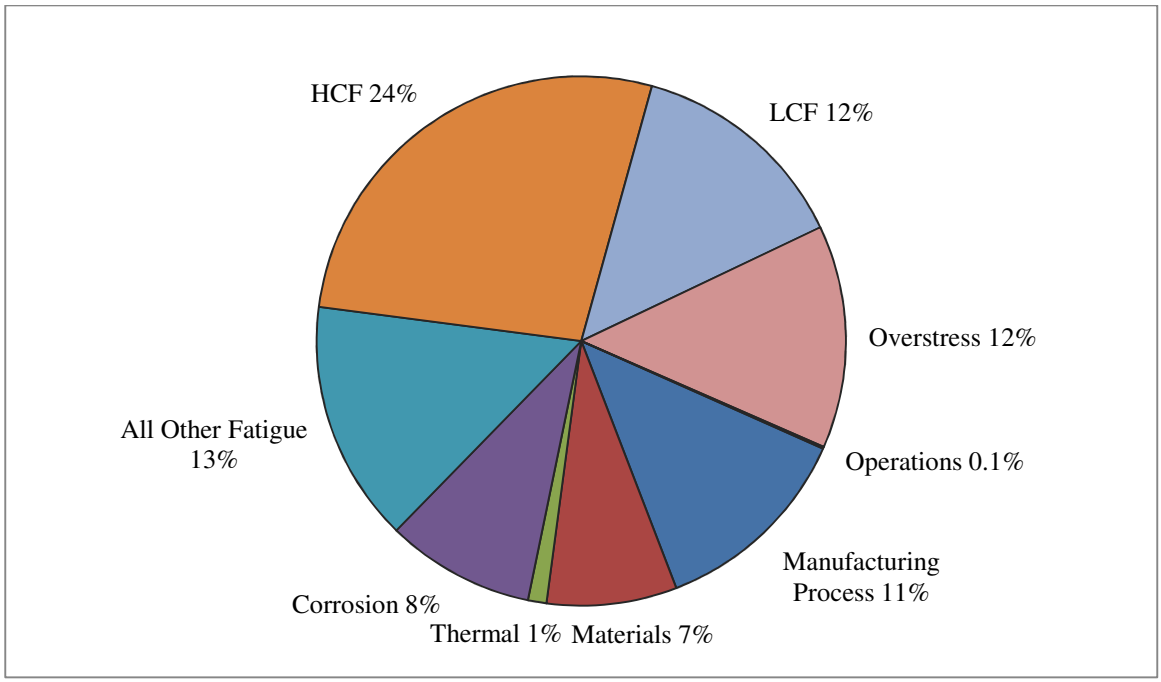


Figure 1.1: US air force jet engine component distress mode statistics [1].

Unlike LCF, cracks caused by HCF spend the majority of their life time on crack nucleation and short crack growth, followed by a rapid component failure at the end of its life, as illustrated in Figure 1.2. This poses challenges in experimental failure prediction and detection. The conventional implementation of the fracture mechanics based damage tolerance approach is not applicable for HCF for the reason that neither flaw-size nor crack propagation rate is observable at short crack scale using current non-destructive evaluation (NDE)

technology. This laid the burden on engineering design to minimize HCF damage. Applicable engineering practices include identifying and minimizing excitatory vibrational stresses, maximizing material tolerance to HCF and incorporating fatigue consideration in structure design. These practices are made possible only by fundamental fracture mechanics research addressing key questions and advancing the understanding of fracture mechanics, in particular, fatigue crack nucleation and other related phenomena.

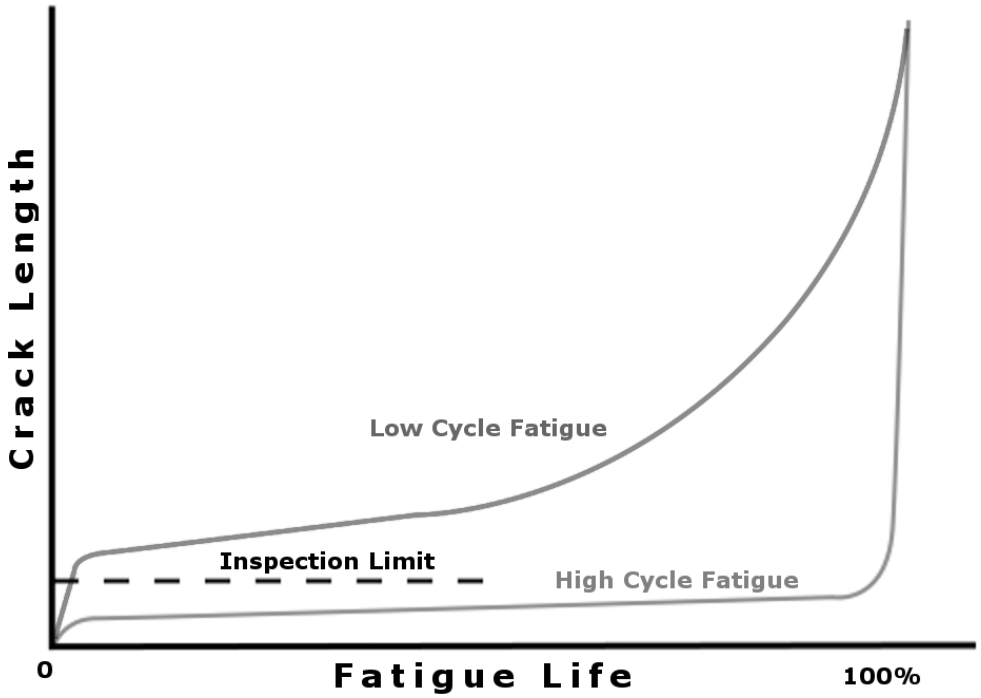


Figure 1.2: Schematic of LCF and HCF life. Majority of HCF life is below inspection limit [1].

The fatigue failure phenomenon is generally considered as a four stage process: 1) crack nucleation, 2) short crack growth, 3) crack propagation and 4) final failure or unstable crack propagation. Amongst these four stages, crack nucleation is the least understood. Modeling such phenomenon is proven extremely challenging. This process is thought to be dependent on intrinsic material parameters and microscopic structures (such as grain boundaries and inclusions). It is also dependent on loading conditions and dynamical dislocation interaction. For materials microstructure, the current understanding of fatigue crack nucleation is still lacking essential details. Most models used are heavily based on assumptions that are not fully rationalized. This report reviews essential models and examines their rationale on the understanding of fatigue crack nucleation.

1.2 Fatigue Nucleation Concept

1.2.1 Overview of Fatigue Crack Nucleation

The term crack *nucleation* and crack *initiation* are often used interchangeably[17]. In this report, we define crack nucleation as a process that describes the localized thermal dynamics aspect of the process, including but not limited to states of system, dislocation dynamics, diffusion and coalescence, energy consideration, statistical phenomenon, etc, while crack *initiation* is defined in a more general term as an event covering all features including the mechanical and the phenomenal aspect, as well as the early crack growth stage. Of course, the scope of thermal dynamics is itself dependent on the approach and perspective.

Fatigue crack nucleation begins when a ductile material is subjected to a repetitive loading. The loading stresses are distributed in a complicated fashion in each individual grain. These stresses are resolved as a shear stress in each grain, activating its slip systems. These slip systems, upon being activated, nucleate dislocations to accommodate slip motions. Dislocations then glide along their slip plane, following the stress conditions. Upon being set into motions, they interact dynamically with surrounding microstructures. Microstructures they interact with include interfaces, vacancies, surfaces, grain boundaries, inclusions and other dislocations or discontinuities. The interaction can transform, halt, translate, multiply or annihilate a dislocation. Eventually, some of these dislocations will pile up at specific sites, most preferably at an obstacle like a grain boundary. The pile-up can manifest in numerous observable forms, such as dislocation dipole, slip bands (of persistent type or secondary type), dislocation vein, extrusion or intrusion. The pile-up intensifies after each loading cycle, when the critical state of stress is reached then a crack will form to release the locally accumulated energy.

1.2.2 Crack Nucleation Site

The location of crack nucleation is important in modeling the fatigue mechanism, as efficient nucleation sites can result in different fracture processes. A fracture at the surface, for example, can be exploited by a very different fracture mechanism when compared to the fracture inside the body. The nucleation site is a material parameter and is a stress/strain condition dependent. The subsurface cracks, as of this review's focus, are found to nucleate along persistent slip bands (*PSB*) (transgranular) or grain boundary (*GB*) (intergranular) for pure ductile metals like copper and aluminum. The choice between these two options is dependent on the strain amplitude. For copper bicrystal at strain amplitudes lower than 0.1% HCF, Huang[3] found that cracks nucleate mainly at PSB. As the strain amplitude increases, cracks nucleating at GB become more significant. Beyond a strain amplitude of 0.2% (LCF), GB becomes the dominate site for crack nucleation. It is also noted that, even at very low strain amplitudes, some cracks do nucleate at GB

where PSB impinges, presumably for high angle GB. This finding is consistent with reports from several other research groups[3-7].

In the case of highly doped alloys, second phase particles also provide a common crack nucleation site[8,9]. It was found that these particles are usually harder than the matrix. While being more brittle, cracks preferentially nucleate inside the particle first, and then propagate into the matrix, as illustrated in Figure 1.3 and 1.4. In the case where the bonding between the particles and the matrix is weak, cracks can also nucleate along the interface, through debonding mechanism.

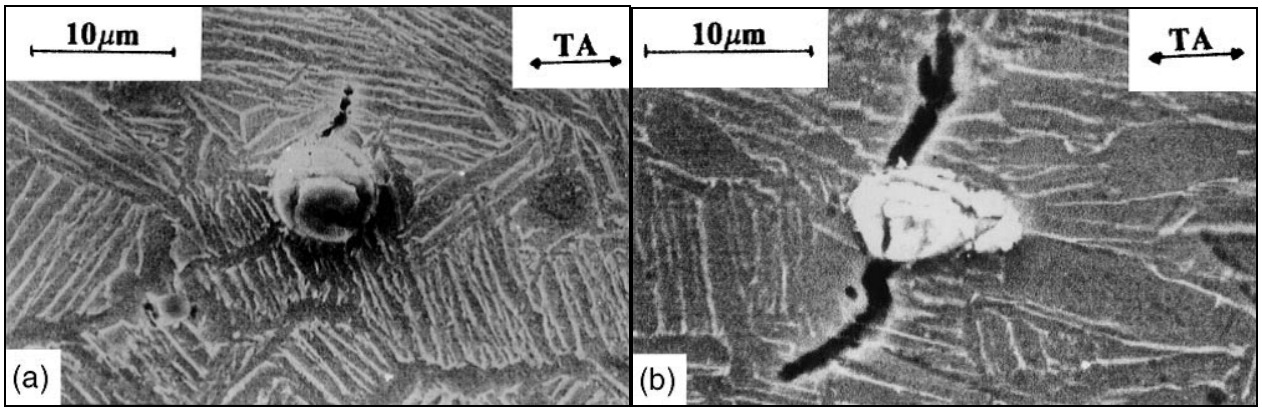


Figure 1.3: A series of SEM photographs indicating the fatigue crack initiation at an Nd-rich particle. 'TA' indicates the tension axis. Figure reproduced with permission from [8], © 1997 Cambridge University Press.

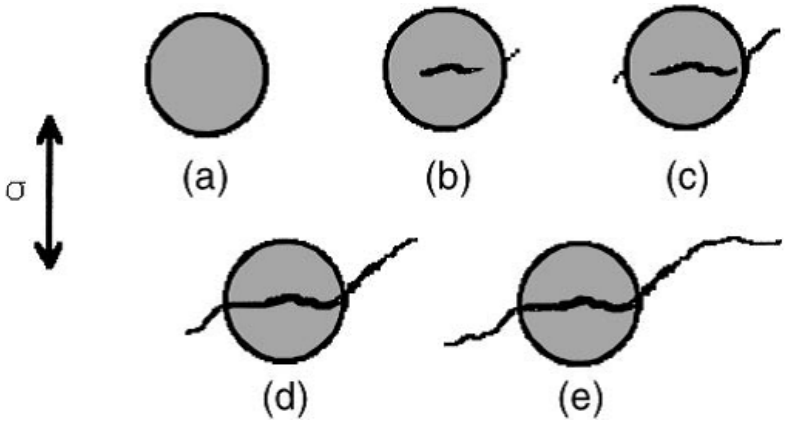


Figure 1.4: Suggested model of fatigue crack initiation at a soft Nd-rich
(a) Intact particle. (b) Slight particle cracking and surface defects.
(c) Further particle cracking and microcrack formation.
(d) Microcrack expanding at about 45° to the tensile axis, and nucleation of another microcrack at the other side of the particle.
(e) The crack growth direction changes to be normal to the tensile. Figure reproduced with permission from [8], © 1997 Cambridge University Press

1.2.3 Dislocation Pile-up Criteria

It is believed that most dislocation based crack nucleation models rely on the idea of dislocation pile-up at grain boundaries. It is therefore important to model accurately the interaction dynamics of such phenomenon. The interaction between dislocations and grain boundaries depends heavily on the slip system configuration. A dislocation gliding toward a grain boundary can be halted, absorbed, transmitted to the adjacent grain or reflected back to the incident grain. For the transmission event, dislocations can undergo transformation while entering another slip system. A residual Burgers vector is often formed during this transformation process to satisfy the conservation of Burgers vectors, namely the Frank's rule. These phenomena are the fundamentals for the bulk physical properties; i.e., they explain the Hall-Petch relation of yield strength and grain size. In the case of an absorption or the left behind of residue Burgers vectors, internal stress accumulates along the grain boundary providing a weak spot for crack nucleation. It is therefore increasingly important to be able to predict slip transfer and layout the condition at which it will occur. Shen and Wagoner [10,11] formulated a set of criteria for slip transfer prediction by examining the interaction between dislocations and grain boundaries in metals under *in situ* high voltage electron microscopy and static transmission electron microscopy. Later, Lee, Robertson and Birnbaum[12,13] further developed Shen's work and published their set of modified conditions, which are subsequently referred to as LRB criteria. They stated that, for dislocations to be transferred across the grain boundary[12,13]:

1) the angle between the incident and the emitting slip plan should be small;
2) residual Burgers vectors remaining at the grain boundary after emission should be small;
3) the resolved shear stress acting on the emitted dislocation should be very high.

The LRB criteria impose some geometric constraints on the system. Criterion 1), for example, limits the relative orientation of the two grains and their active slip systems. In criterion 2), the residue Burgers vector in a transmission event can be determined by the following interaction equation:

$$\vec{b}_r = \vec{b}_1 - R\vec{b}_2, \tag{1.2.1}$$

where $\vec{b}_1$ and $\vec{b}_2$ express the Burgers vectors of the incident and transmission dislocations respectively, while R is the rotation matrix that describes the misorientation between the two grains. Criterion 2) implies that edge dislocations are usually more difficult to transfer across the GB, as they usually have larger residue Burgers vector compared to screw or mixed dislocations. This phenomenon was also observed experimentally, as reported by Kashihara *et al*[14] in their SEM study on aluminum. It is interesting to note that the resolved shear stress described in criterion 3) is not static. After each transmission, the stress state changes due to the build-up of the residue stress. This indicates that the prior interactions between

dislocations and grain boundaries influence the kinetics of the reactions. The build up of the residue Burgers vector can lead to the activation of new slip systems, emission or desorption of dislocations and possibly to other dynamical behaviors.

1.2.4 Comparison between High Cycle Fatigue and Low Cycle Fatigue

The major difference between HCF and LCF is the loading condition at which these two fatigue mechanisms are activated. HCF is commonly being thought of as fatigue driven by low repetitive strain at which the bulk materials respond elastically. It usually has a total life time exceeding 10^4 cycles. In gas turbine engines, HCF can be driven by aerodynamic fluctuations, mechanical or acoustic vibrations and other excitatory cyclic loads. In contrary, LCF usually happens due to a high cyclic strain, close to or beyond the material's yield point. In addition to mechanical cyclic loading, cyclic stress induced by operational thermal cycling is another example of such excitatory source for LCF.

An important aspect of HCF is the manifestation of a macroscopic applied stress on the microscopic heterogeneity. At a low stress amplitude, a very limited number of slip systems are activated. Therefore, the plastic deformation is highly localized. In a contrary, the plastic deformation in LCF is a bulk response to the external load, rendering a relatively uniform deformation pattern. This by no means implies a microscopic homogeneity. The stochastic nature of the crystallography in polycrystalline materials introduces a regional statistical variation.

In engineering practice, empirical formulations are employed for fatigue life prediction in both scenarios. For HCF, fatigue life can be predicted by Basquin's rule. Basquin observed that the stress-life data could be approximated using a power relationship, which resulted in a straight line on a log-log plot. This observation corresponds to the elastic material behavior in the strain-life approach. The Basquin equation can be expressed in terms of the true elastic strain amplitude as

$$\varepsilon_e = \frac{\sigma'_f}{E}(2N)^b, \tag{1.2.2}$$

where ε_e is the elastic component of the strain amplitude, E is the elastic modulus, σ'_f and b are material parameters referred to as the *fatigue strength coefficient* and *fatigue strength exponent*, respectively. In the case of LCF, the Manson-Coffin relationship is the most applicable. Similar to Basquin's model, the

Manson-Coffin model addresses the plastic component of the strain response, in a form of a power relationship in strain-life plot. This relationship can be written as:

$$\varepsilon_p = \varepsilon_f' (2N)^c, \tag{1.2.3}$$

where p denotes the plastic component of strain, while ε_f' and c are the *fatigue ductility coefficient and fatigue ductility exponent.* These two models are widely applied in most engineering applications.

1.3 Dislocation-Based Fatigue Nucleation Models

1.3.1 Persistent Slip Band – Grain Boundary Interaction

As discussed in previous sections, PSB-GB interaction plays a significant role in low strain crack nucleation of transgranular and intergranular situations. A PSB-GB interaction model was first proposed by Mughrabi[15,16] and then developed by Christ *et al* [4]. Their models of interaction are like this: when a Burgers vector ***b*** of the edge dislocations of a PSB has a major component parallel to the surface, the forces caused by the pile-up of the interface dislocations act against the grain boundaries on either side of the PSB[4]. Figure 1.5 illustrates this PSB-GB interaction scenario. The parameter α, is defined as the angle between the stress axis and the trace of the PSB, while β is the angle between the trace of the PSB and GB. The third angle, γ, is supplementary to the other two, and therefore is a dependent variable. When these angles vary, the stress acting on the PSB, directed from the external stress, is adjusted. Following this idea, one would expect the existence at least one set of angles at which the required external stress for crack nucleation is minimized. Smith and Barnby[36] derived an equation relating the angles with such stress:.

$$\sigma = \left[\frac{\pi \gamma G}{2(1-\nu)d} \right]^{\frac{1}{2}} f(\alpha, \beta), \tag{1.3.1}$$

where γ is the surface energy of the crack, υ is Poisson's ratio, G is the shear modulus and d the grain size, with,

$$f(\alpha, \beta) = \left(5 + 2\cos\beta - 3\cos^2\beta\right)^{-\frac{1}{2}} \left(\cos\alpha \sin\alpha\right)^{-1}, \tag{1.3.2}$$

This equation plotted in Figure 1.6 has a local minimum at α = 45° and β=70.5°. For these angles, σ is

minimized, thus cracks are most easily formed. One can see that the divergence of σ is much larger along the α domain, signifying a strong orientation dependence. In order words, cracks are most easily nucleated when the slip system is inclined approximately 45° with respect to the axis of the applied stress.

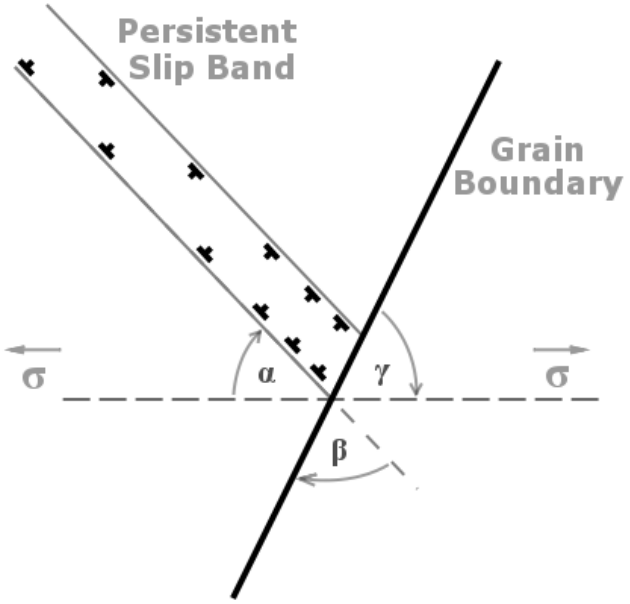


Figure 1.5: Schematic representation of the formation of a PSB-GB crack [4].

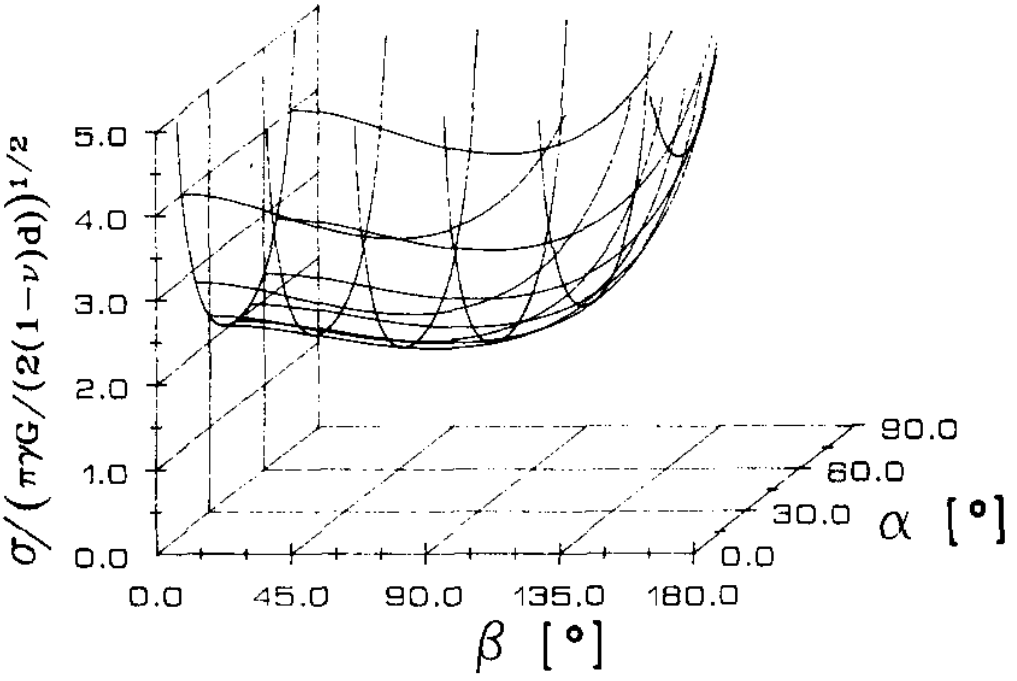


Figure 1.6: Dependency of the stress necessary to from a crack, see definition for angles α and β
Figure reproduced with permission from [4], © 1989 Elsevier.

1.3.2 Dislocation Based Crack Nucleation Model

Zener (1948) [41] was the first to suggest a crack nucleation mechanism based on dislocation pile-up. Later, Stroh and Koehler [42] advanced on his idea, establishing a detailed and consistent mathematical model on this topic. They illustrated that edge dislocations pile-up at an obstacle, likely a grain boundary, can coalesce into a crack nucleus. Weertman[17] in his book named this type of crack after these three pioneers in acknowledgement of their contribution on this topic. Therefore, we will refer to such a crack as the *Zener-Stroh-Koehler* (ZSK) crack. During dislocation pile-up, the set of dislocations on the same slip plane can be viewed as one superdislocation. While all dislocations have the same orientation, the resulting plastic stress field is antisymmetric across the slip plane, leaving one crack tip under tensile stress while the other is under compressive stress. Therefore, a ZSK crack, unlike a Griffith crack, does not close completely at the absence of an external load. In 1958, Cottrell[18] suggested a variant version of the ZSK crack nucleation mechanism specific to the body centred cubic (BCC) structure. His idea involved the interaction of two intersecting <111> glide systems, where <111> dislocations at the two planes combine at the intersection and form <100> dislocations. While such resulting dislocations are sessile, a pile-up occurs, that leads to crack nucleation. The two versions of this mechanism are shown in Figure 1.7.

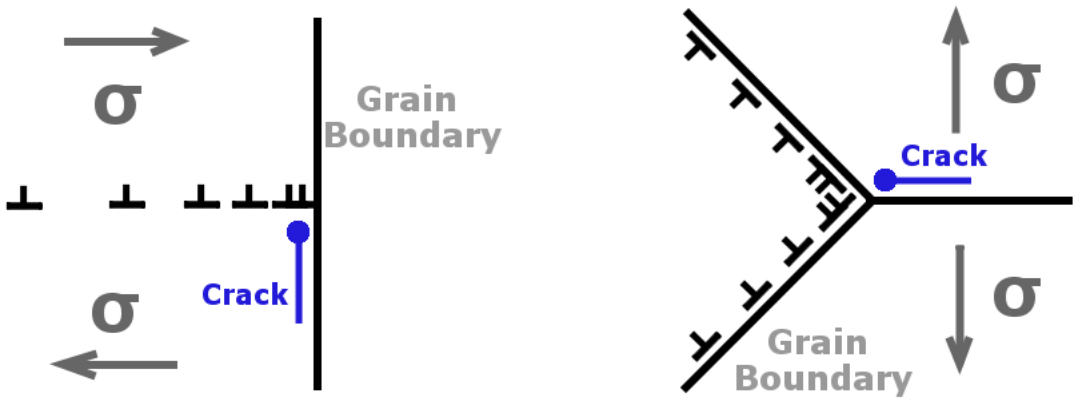


Figure 1.7: Schematic of two possible Mode I Zener-Stroh-Koehler crack initiation at Left) grain boundary, Right) triple junction [17].

When a crack is formed, the two crack faces are separated by a finite distance. It is possible to define a continuous distribution of infinitesimal dislocations on the PSB plane such that the displacement $D(x)$ is given by:

$$D(x) = \int_{x}^{\infty} B(x)dx. \tag{1.3.3}$$

For a ZSK crack of length 2a, its dislocation (density) distribution is given by:

$$B(x) = \frac{b_T}{\pi\sqrt{a^2 - x^2}} \quad \text{for } -a < x < a. \tag{1.3.4}$$

While this function is symmetrical, the resulting displacement function is antisymmetrical in shape, namely:

$$D(x) = (b_T / \pi)\left[(\pi/2) - \sin^{-1}(x/a)\right]. \tag{1.3.5}$$

The term b_T is an integration constant corresponding to the total Burgers vector along the crack plane. Since *D(x)* is antisymmetric, the right positive side of the crack tip is sharp while the other crack tip is blunt. To evaluate the stress components at the point (x, y) away from the crack, one can consider adding the stress component of each infinitesimal dislocation, scaled by its distribution function, which gives:

$$\sigma_{ij}(x, y) = \int_{-a}^{a} B(x')\sigma_{ij}^{'}(x - x', y)dx'. \tag{1.3.6}$$

Substituting *eq 1.3.4* into *eq 1.3.6* and evaluating the integral in the complex plane result in the stress field of a ZSK crack. For mode I crack, it is given by[17]:

$$\sigma_{xx} = \left(Gb_T / 4\alpha\pi r^*\right)\left[2\cos\theta^* - \sin\theta_R \sin(\theta^* + \theta_R) - \sin\theta_L \sin(\theta^* + \theta_L)\right], \tag{1.3.7}$$

$$\sigma_{yy} = \left(Gb_T / 4\alpha\pi r^*\right)\left[2\cos\theta^* + \sin\theta_R \sin(\theta^* + \theta_R) + \sin\theta_L \sin(\theta^* + \theta_L)\right], \tag{1.3.8}$$

$$\sigma_{zz} = \nu(\sigma_{xx} + \sigma_{yy}), \tag{1.3.9}$$

where $r^* = \sqrt{r_L r_R}$ and $\theta^* = \frac{1}{2}(\theta_L + \theta_R)$, other parameters are defined in Figure 1.8. It is important to note that the stress field mentioned above is only valid during crack nucleation. Once the crack begins to grow through an applied external load, and the crack tip advances away from the nucleation site, the stress field approaches the limit of Griffith's crack.

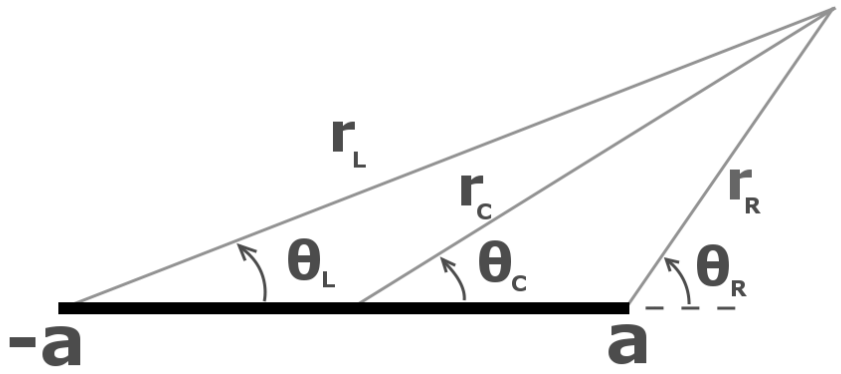


Figure 1.8: Schematic of a Zener-Stroh-Koehler crack configuration [17].

Because the ZSK crack originates from the stress imposed solely by dislocations, it is still opened without the presence of external force. From that, we can expect an equilibrium state where a crack ceases to advance. It is convenient to define *a* as the half crack length for that state. To evaluate this quantity, consider the plastic zone, which extends a distance *c* away from the crack center. It is related to physical parameters of the material by:

$$c^2 - a^2 = \left(b_T G / 4\alpha\sigma_T\right)^2. \tag{1.3.10}$$

At the ZSK crack tip where it is blunt, a finite displacement D(a) that correlates *c* and *a* is given as

$$D(a) = \frac{2\alpha\sigma_t}{\pi G}\left[\sqrt{(c^2 - a^2)}\left\{\pi - 2\sin^{-1}(a/c)\right\} - 2a\ln(c/a)\right] \tag{1.3.11}$$

Lets define d_t as the critical crack tip displacement which marks the onset of the further crack growth, then, at the limit $(c-a)/a \ll 1$, *eq 1.3.11* is reduced to

$$D(a)/d_t = \left(K_{ZSK}/K_{IC}\right)^2, \tag{1.3.12}$$

where K_{IC} and K_{ZSK} are the fracture toughness of Griffith crack and ZSK crack. According to the linear

elastic fracture mechanics, the fracture toughness of a mode I Griffith crack is $K_{IC} \propto \sqrt{a}$ and increases as the crack length increases. On the contrary, K_{ZSK} decreases as the crack tip advances away from the crack center at early growth. K_{ZSK} has the form given by:

$$K_{ZSK} = \frac{b_T G}{2\alpha\sqrt{\pi a}}. \tag{1.3.13}$$

At equilibrium, $D(a_{eq}) = d_t$ and therefore $K_{ZSK} = K_{IC}$, yielding the solution:

$$a_{eq} = \frac{(b_T G)^2}{4\pi\alpha E}. \tag{1.3.14}$$

This is the equilibrium half length of a ZSK crack under the internal stress. When an external load is present, the crack becomes a composite of a ZSK and Griffith crack. If the plastic zone is relatively small, such that $(c-a)/a << 1$, the equilibrium half-length can be determined by:

$$K = K_{IC} - K_{ZSK}, \tag{1.3.15}$$

where

$$a^{1/2} = \frac{K_{IC}}{2\sigma\sqrt{\pi}}\left[1 \pm \sqrt{\left(1 - \frac{2Gb_T\sigma}{\alpha K_{IC}^2}\right)}\right]. \tag{1.3.16}$$

These two solutions correspond to the stable (negative sign) and the unstable (positive sign) equilibrium half length. A physically meaningful solution of eq 1.3.16 is only available when

$$\sigma_C = \frac{\alpha K_{IC}^2}{2Gb_T} = \frac{2\gamma}{b_T}, \tag{1.3.17}$$

beyond which equilibrium is not defined. In this case, a crack would undergo unconstrained growth, resulting in a catastrophic failure.

1.3.3 Fatigue Crack Nucleation Life

Considering a dislocation pile-up at GB as the sole crack nucleation mechanism at low strain, *Tanaka* and *Mura*[19] formulated a dislocation based model for fatigue crack nucleation life by considering a systematic increase in the amount of dislocation pileups during cyclic loading.

Their model considers slip bands as a composite of dislocation monopoles in pairs or triplets formation referred to as a vacancy/interstitial dipole or a tripole, as shown in Figure 1.9. At each half of the cyclic loading, dislocations are nucleated on the preferred monopole layer, constituting a pile-up at the grain boundary. At the same time, these newly created dislocations exert a back stress to reinforce the pile-up of the adjacent dislocation layers of the opposite sign. Cyclic loading leads to a constant increment of dislocation pile-up for each load. A fatigue nucleation life can then be formulated based on this constant increment. This model is applicable to both crack nucleation at grain boundaries and at surfaces. In the case where a crack nucleates at a surface, the plastic deformation manifests itself as strain, forming intrusion and extrusion before a crack appears. Tanaka's model relies on the assumption of a perfectly irreversible plastic flow, which is a reasonable first estimation.

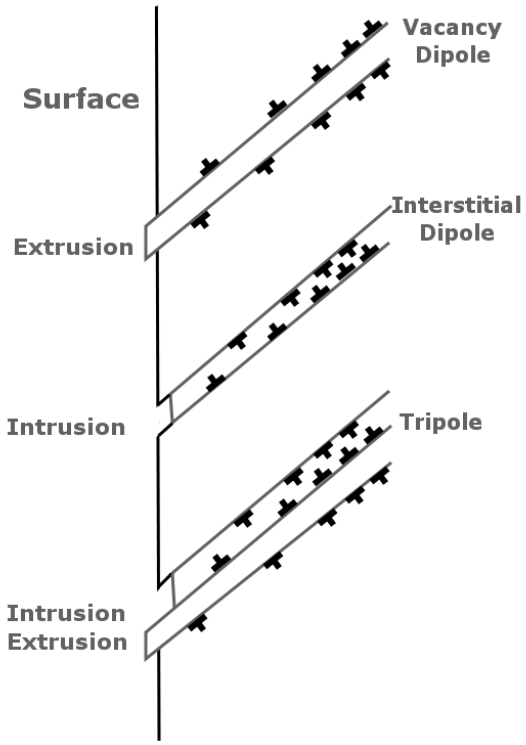


Figure 1.9: Dislocation Pileup in the form of dipole or tripole, forming extrusions or intrusions on the surface [19].

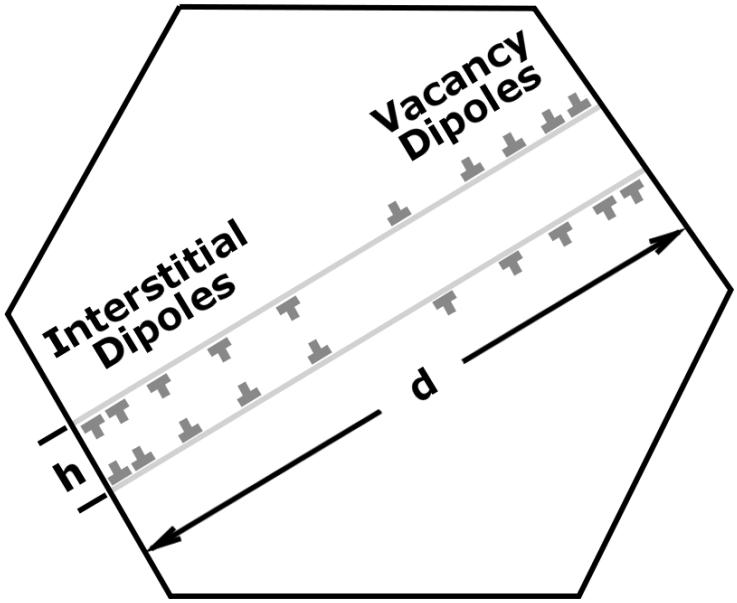


Figure 1.10 Dislocations Pileup in a subsurface grain, impinging a grain boundary [19].

Starting with the equilibrium shear condition for the first generated dislocation monopole layer inside a grain of size *2a*,

$$\tau_1^D + \tau_1 - \tau_f = 0, \tag{1.3.18}$$

where *D* denotes the dislocation stress (back stress) and *f* denotes the frictional stress, then the corresponding equilibrium condition for the subsequently initiated layer I is:

$$\tau_2^D + \tau_1^D + \tau_2 + \tau_f = 0. \tag{1.3.19}$$

Tanaka and Mura[19] then employed the inversion of the formula of Muskhelishvili[42], under the condition of unbounded ends at $x=\pm a$, to solve the dislocation density profile, $D_1(x)$, with *eq 1.3.18* obtaining:

$$D_1(x) = \frac{1}{\pi^2 A(a^2 - x^2)^{1/2}} \int_{-a}^{a} (a^2 - x'^2)^{1/2} \frac{\tau_1 - \tau_f}{x - x'} dx' = (\tau_1 - \tau_f)x / \pi A(a^2 - x^2)^{1/2}, \tag{1.3.20}$$

where $A = Gb/2\pi(1-\upsilon)$. Subsequently, the plastic displacement caused by the dislocations generated at $x = 0$ is:

$$\phi(x) = \int_{x}^{a} bD(x')dx', \tag{1.3.21}$$

and thus, the total plastic displacement, or better recognized as the plastic strain, gives:

$$\gamma_1 = \int_{-a}^{a} \phi(x)dx = \int_{-a}^{a} bD_1(x)xdx = (\tau_1 - \tau_f)ba^2/2A. \tag{1.3.22}$$

The stored energy of dislocations per unit thickness can then be expressed as:

$$U_1 = -\frac{1}{2}\int_{-a}^{a} \tau_1^D \phi(x)dx = -\gamma_1\tau_1^D/2. \tag{1.3.23}$$

Equating *eq 1.3.23* with *eq 1.3.18* yields:

$$U_1 = \gamma_1(\tau_1 - \tau_f)/2. \tag{1.3.24}$$

Once the essential parameters for the first forward loading are solved, the same parameters for the reverse loading and for the subsequent cyclic loadings can be determined in the same manner. For example, the reverse loading yields:

$$D_2(x) = -(\Delta\tau - 2\tau_f)x/\pi A(a^2 - x^2)^{1/2}, \tag{1.3.25}$$

$$\gamma_2 = -(\Delta\tau - 2\tau_f)ba^2/2A, \tag{1.3.26}$$

$$U_2 = -\gamma_2(\Delta\tau - 2\tau_f)/2, \tag{1.3.27}$$

on layer II, where $\Delta\tau = \tau_1 - \tau_2$. The generalized incremental change for each subsequent loading is then:

$$\Delta D(x) = (\Delta\tau - 2\tau_f)x/\pi A(a^2 - x^2)^{1/2}, \tag{1.3.28}$$

$$\Delta\gamma = (\Delta\tau - 2\tau_f)ba^2 / 2A, \tag{1.3.29}$$

$$\Delta U = \Delta\gamma(\Delta\tau - 2\tau_f)/2. \tag{1.3.30}$$

Note that each loading cycle consists of a forward and a reverse load, and so the change in energy for the N^{th} cycle is $2N\Delta U$. At the critical value, denoted by *c*, where a fracture occurs, an energy balance equation can be written as:

$$2N_c\Delta U = 4aw_s, \tag{1.3.31}$$

where w_s is the fracture energy per unit area. Rearranging it and equating it with eq 1.3.29 and eq 1.3.30 gives a crack nucleation life to the crack of size *a* as:

$$N_c = 8Aw_s / ba(\Delta\tau - 2\tau_f)^2, \tag{1.3.32}$$

One of the significant features of this model, as discussed by Tanaka, is that it is conformal to the empirically determined Coffin-Manson relationship, namely:

$$N\Delta\varepsilon^{-2} = Const., \tag{1.3.33}$$

which very well represents metallic materials behaviour in a wide range of fatigue strain amplitudes. What is more significant though, is the formulism of the problem, which was adopted by many authors to develop a wide variation of more powerful models. Two of those variations will be discussed in the following paragraphs.

Mura [20] has been one of the active authors in modifying and extending the mentioned model. In 1990, he presented an advanced modification, with thermal dynamic considerations, relating the fatigue nucleation life with material parameters such as width of dipoles, grain size and surface energy.

Inspired by Griffith's theory of linear elastic fracture mechanics, Mura attempted to develop fracture criteria with the consideration of Gibbs free energy. He evaluated the scenario at which the system contains a vacancy dipole of width *h* spanning across a grain of size *a.* It nucleates a crack of length *c* with the crack surface energy of w_s . Then, the change in Gibbs free energy, measured with respect to a reference energy

level, G_0 can be written as:

$$\Delta G = -W_e - W_m + 2cw_s, \tag{1.3.34}$$

where W_e is the elastic strain energy after *N* cycles of loading, and W_m is the mechanical energy released during the formation of crack surfaces. The crack length *c* is equal to the plastic strain at N^{th} cycle, $c = N\Delta\gamma$ given by eq 1.3.29. The elastic strain energy W_e is evaluated by considering the dislocation distribution that causes a strain field to the system. After some derivation, Mure obtained:

$$W_e = (1-\upsilon)(\Delta\tau - 2\tau_f)^2 N^2 a^2)(1-k)[(3+k)K(k) - 3E(k)]/k\mu, \tag{1.3.35}$$

where

$$k = \left(\sqrt{4a^2 + h^2} + h^2\right)^2 / 4a^2, \tag{1.3.36}$$

$$K(k) = \int_0^1 d\xi / \sqrt{[1-\xi^2][1-(k\xi)^2]}, \tag{1.3.37}$$

$$E(k) = \int_0^1 d\xi \sqrt{[1-(k\xi)^2]/[1-\xi^2]}. \tag{1.3.38}$$

When the width of the dipole is small compared to its length, i.e., $\varepsilon = h/a \approx 0$, then eq 1.3.35 becomes:

$$W_e \approx (1-\upsilon)(\Delta\tau - 2\tau_f)^2 (Na\varepsilon)^2 [\log(8/\varepsilon) - (3/2)]/\mu. \tag{1.3.39}$$

Similarly, the mechanical energy is:

$$W_m \approx (1-v)(\Delta\tau - 2\tau_f)^2 (Nab\Delta\tau)^2 / 8\pi A^2 \mu. \tag{1.3.40}$$

Gibbs free energy can be utilized in crack nucleation. In this scenario, the energy measures the total work the system can release to the surroundings by forming crack surfaces. For the crack nucleation event to be thermally activated (fracture spontaneously), it must lose Gibbs free energy during the crack formation, that is, $dG < 0$. In regards to crack nucleation life, this criteria is:

$$\frac{\partial G}{\partial N} = \frac{\partial}{\partial N}\left(G_0 + \Delta G\right) = \frac{\partial \Delta G}{\partial N} \le 0. \tag{1.3.41}$$

When left side of this equation is equal to zero, *N* marks the onset of crack nucleation. Equating eqs *1.3.34, 1.3.39, 1.3.40* and 1.3.41 results in:

$$-(1-\upsilon)(\Delta\tau - 2\tau_f)\frac{N_c}{\mu}\left[h^2 \log(8a/h) + \frac{(ab\Delta\tau)^2}{8\pi A^2}\right] + \frac{baw_s}{\pi A} = 0, \tag{1.3.42}$$

or

$$N_c = \frac{ba\mu w_s}{\pi A(1-\upsilon)(\Delta\tau - 2\tau_f)}\left[h^2 \log(8a/h) + \frac{(ab\Delta\tau)^2}{8\pi A^2}\right]^{-1}. \tag{1.3.43}$$

This is the fatigue nucleation life calculated by Mura's model. When *h* is sufficiently small, his model approaches the limit of Tanaka's model (*eq 1.3.32*). This formulation interestingly relates the fatigue nucleation life to the surface energy, dipole width, and grain size. It is worth to note that this equation predicts the existence of optimum values for *h* and for *a*, at which the nucleation life is the shortest.

1.3.4 Microstructure Induced Fatigue Crack Nucleation

Second phase particles often exist in heavily doped alloy. These particles include inclusions (unwanted oxides), dispersoids (thermally stable intermetallics) and intermetallics (metal complexes). They are found to nucleate cracks much more readily when compared to the matrix. One of the reasons is that these particles are usually harder than the matrix, which suggests brittleness. Also, these microstructures concentrate stress effectively inside an alloy, providing preferential crack sites. It is therefore essential to include these particles in crack nucleation modeling.

Chan[21] in 2003 published a model which takes into account microstructure parameters. His work is an extension of Tanaka's model, and follows Mura's Gibbs free energy approach. He considers the situation where slip bands formed inside a subsurface inclusion are impinging the interface. The nucleation life calculated in such manner is identical to that of a subsurface grain scenario, which is given by Tanaka:

$$(\Delta\tau - 2k)^2 N_c = \frac{4\mu w_s}{\pi(1-\upsilon)a}. \tag{1.3.44}$$

The Gibbs energy considered is derived from eq 1.3.34 with the elastic strain energy equal to:

$$W_e = \frac{\xi b a^2 (\Delta\tau - 2k)^2 N^2}{4\pi^2 A} = z_1 N^2, \tag{1.3.45}$$

where ξ is the inclusion size. The mechanical energy term is equivalent to Irwin's strain-energy release rate due to crack opening. It is given by[37-39]:

$$W_m = \int \frac{1-\upsilon}{\mu}(K_I^2 + K_{II}^2), \tag{1.3.46}$$

yielding

$$W_m = \frac{(1-\upsilon)N^2}{8\pi\mu}\left[\frac{(\Delta\tau - 2k)\Delta\tau ab}{A}\right]^2 = z_2 N^2. \tag{1.3.47}$$

Finally, the constant *c* in eq 1.3.34 is given by:

$$c = \frac{ab(\Delta\tau - 2k)N}{\pi A} = z_3 N. \tag{1.3.48}$$

Thus,

$$\Delta G = -z_1 N^2 - z_2 N^2 + 2 z_3 \gamma_3 N. \tag{1.3.49}$$

Following the same thermal dynamic criterion in eq 1.3.43 yield:

$$(\Delta\sigma - 2M\tau_f) N_c^{\alpha} = \frac{M\mu h^2}{a(h+a)}\sqrt{\frac{8c(\mu+\mu')}{\lambda\mu'\xi}}, \tag{1.3.50}$$

for the fatigue-crack nucleation life at the presence of an inclusion. In this equation, *M* is the Taylor factor, relating the shear stress range $\Delta\tau$ to the applied stress range $\Delta\sigma$. It is evaluated by adding the reciprocal Schmid factor over all slip systems. The ξ and μ' are the size and the shear modulus of the inclusion. In this equation α is a constant with a value between zero and unity. It is an experimentally determined value governing the dependency on *N*, taking into account all dynamical effects not explicitly included in this

model. Chan[21] refers to the term $2M\tau_f$ as the fatigue limit (or endurance limit), considering that a crack can only be formed when the applied stress range is larger than this term.

A unique feature of Chan's model is its capability of extending the crack nucleation model to short crack growth regime by taking differentiation of crack length *c* with respect to *N*, which results in:

$$\frac{dc}{dN} = 2\alpha\left[\frac{\lambda\pi(1-v)d^3}{8M^2h^2}\left(\frac{\Delta\sigma - 2M\tau_f}{\mu}\right)^2 c^{2\alpha-1}\right]^{\frac{1}{2\alpha}}. \tag{1.3.51}$$

The parameter α effectively couples the growth rate to the crack length. The value of 1/2 in the power position of the expression marks the onset of accelerated growth, beyond which the growth rate increases with the crack length. This interpretation is useful for investigating the criteria at which a crack growth is halted.

1.4 Probabilistic and Numerical Fatigue Nucleation Models

1.4.1 Reliability-based Crack Nucleation Life

Models discussed in the previous sections consider fracture as a deterministic process. In practice, it is important to account for the stochastic nature of crack nucleation. For this purpose, Tryon and Cruse[26,27] developed a probabilistic model for crack nucleation life. Their model addresses the scatter of nucleation life observed in experimental data. This scatter can be quantified by the coefficient of variation (CV), which is a statistical parameter defined as the ratio of standard deviation to the mean of a stochastic variable. This quantity is dependent on materials and can vary from 10 to 500 percent for various steel alloys[28,29]. This probabilistic variation arises from the microstructure inhomogeneity of the materials. Amongst all stochastic microstructure parameters, variations in grain orientation and grain size are experimentally determined to be the most influential[30-33]. This can be rationalized by realizing that the grain orientation determines the resolved shear stress, τ_{rss}, on a slip system, while the grain size determines the dislocation pile-up condition. The resolved shear stress is defined as the projection of the axial applied stress σ_{xx} on a slip direction. The relationship can be written as

$$\tau_{rss} = \frac{\sigma_{xx}}{M}, \tag{1.4.1}$$

where $M = (\cos\phi\cos\lambda)^{-1}$ is the reciprocal Schmid factor (Figure 1.11) and the equation is known as the Schmid law. In an untextured polycrystalline material, the distribution of grain orientations can be assumed to be uniform on the stereographic triangle (Figure 1.12). With the help of the method developed by Tanaka[23] for determination of the grain orientation distribution, the probability mass function of the reciprocal Schmid factor, *P(M)*, can be defined. Therefore, the probabilistic distribution of the resolved shear stress (microscopic) can be related to the axial applied stress (macroscopic) by eq 1.4.1. The probabilistic nature of the grain size distribution is governed by the thermal dynamic of the material production process. In material fabrication by solidification, i.e. casting, the thermal dynamic is thought to be universal and insensitive to the material. Therefore, the scatter of the grain size is similar in a wide variety of materials, including pure metals, complex alloys, and inorganic ceramics[25].

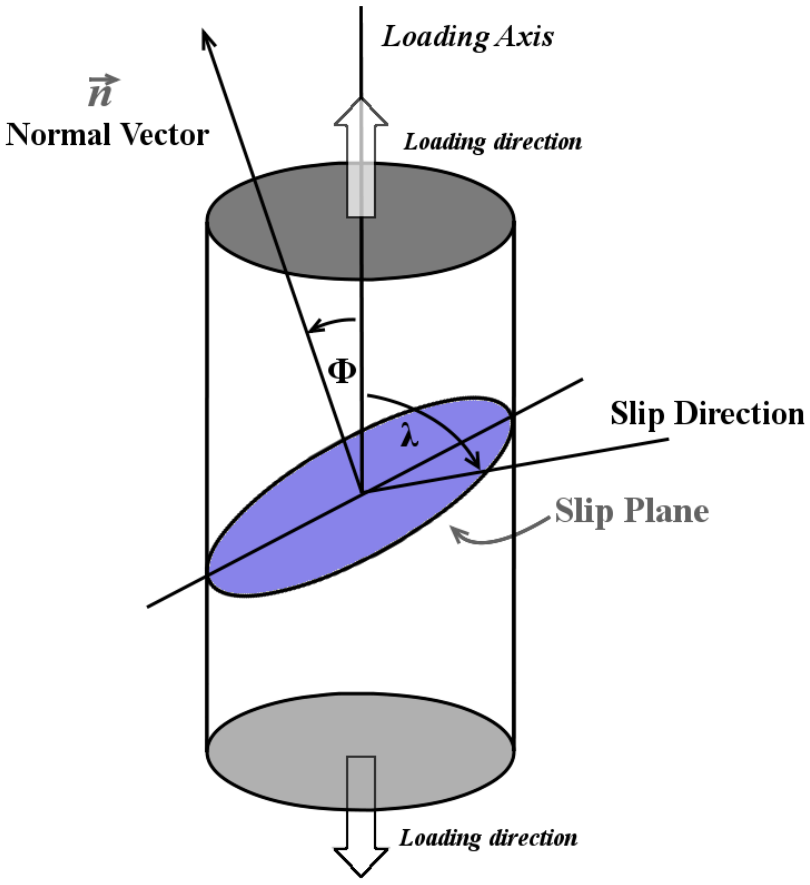


Figure 1.11: Geometric configuration of Schmid Law.

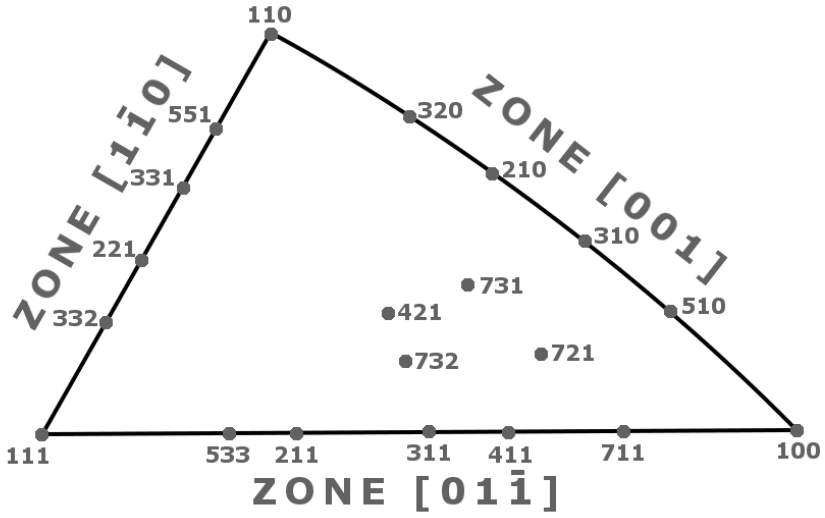


Figure 1.12: Unit stereographic triangle of an FCC single-crystal [24].

The reliability *R* of crack nucleation life is defined as the probability of the observed crack nucleation life N_n that exceeds the specific number of cycles N_s. The reliability is evaluated using the performance function, $g(\vec{X})$, given by

$$g(\vec{X}) = N_n(\vec{X}) - N_s, \tag{1.4.2}$$

where $\vec{X}$ is the vector of all the random variables presented in the system. The performance function is used to depict the occurrences of a failure that meets the condition $g(\vec{X}) \le 0$. The probability of a failure can be evaluated by integrating the joint probability density function of $\vec{X}$ over the phase space at which a failure can occur:

$$P_f = 1 - R = \int_{g(\vec{X}) \le 0} f_X(\vec{X}) d\vec{X}. \tag{1.4.3}$$

The random variables in $\vec{X}$ are not necessarily independent, and this posts difficulties in evaluating eq 1.4.3 in practice. Assuming that there exists a vector $\vec{\mu}^*$ in $\vec{X}$ phase space which satisfies the condition

$g(\vec{\mu}^*) = 0$ and $\beta^2 = \sum_i \mu_i^{*2}$ is minimized, then $\vec{\mu}^*$ is the most probable point. A transformation of variables can then be performed by expanding the performance function to the Taylor series of the first order around $\vec{\mu}$, giving

$$g(\vec{\mu}) = g(\vec{X}) \approx A_0 + \sum_i \frac{\partial g}{\partial X_i}(X_i - \mu_i), \tag{1.4.4}$$

and

$$P_f = \int_{g(\vec{\mu})\le 0} f_X(\vec{\mu}) d\vec{\mu}. \tag{1.4.5}$$

Since each variable μ_i is independent, the central limit theorem is applicable and eq 1.4.5 conforms to the normal distribution Φ at the minimum distance to origin β namely:

$$P_f = \Phi(\beta), \tag{1.4.6}$$

and the reliability becomes

$$R = 1 - \Phi(\beta). \tag{1.4.7}$$

With this first order reliability model, the contribution of each variable toward the scattering of fatigue initiation life can be evaluated. This quantity is defined as the sensitivity factor

$$\alpha_i = \frac{\mu_i^*}{\beta}. \tag{1.4.8}$$

Employing Tanaka's model of fatigue nucleation life, eq 1.2.33, and relating it to the external stress by eq 1.4.1, gives

$$N_i = \frac{4Gw_s}{\left(\frac{\Delta\sigma_{xx}}{M} - 2k\right)^2 \pi(1-\upsilon)d}. \tag{1.4.9}$$

With this equation, the performance function is defined and therefore the probability of failure can be

evaluated numerically. The set of random variables in this system is $\{\Delta\sigma_{xx}, k, M, d\}$. Tryon[27] compared this model with experimental data of austenitic steel, assuming G, w_s, and υ to be deterministic variables with values of 76×10^3 MPa, 0.3, and 440 kN/m respectively. The comparisons are plotted in Figure 1.13 and in [26]. These plots illustrated a good agreement between experiment and the prediction.

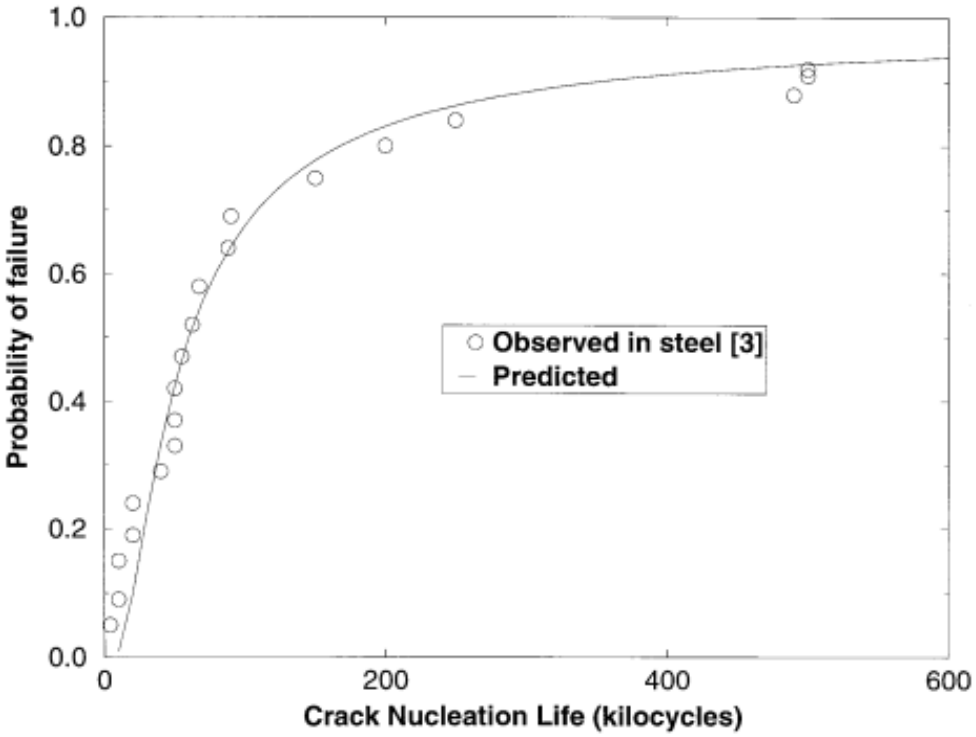


Figure 1.13: Crack Nucleation Life Distribution of individual grain. Figure reproduced with permission from [27], © 1998 John Wiley and Sons.

1.4.2 Fatigue Crack Nucleation Model using the Finite Element Method

At the microscopic level that considers a cluster of grains, both the orientation and the size of each grain are deterministic, and yet the complexity of the system renders analytical model hard to evaluate. Therefore, a numerical solution such as finite-element method is most suitable. Manonukul and Dunne[35] developed a numerical model for nickel-based alloys based on polycrystalline plasticity. This model is applicable to both HCF and LCF conditions, and is a useful tool for further investigating materials behavior under different loading conditions. The model considered the rate-independent plasticity developed by Asaro[34]. Given an applied load, the deformation and rotation of a crystal can be evaluated by the total deformation gradient **F** which is a product of elastic and plastic components:

$$\mathbf{F} = \mathbf{F}^e \mathbf{F}^p. \tag{1.4.10}$$

Here the plastic deformation gradient of a grain α can be considered as a plastic slip γ along a direction vector ***s***, with its normal ***n***, acting on a non-deformed gradient field **I**, giving

$$\mathbf{F}^p = \mathbf{I} + \gamma^\alpha \left(\mathbf{s}^\alpha \mathbf{n}^{\alpha \mathrm{T}} \right). \tag{1.4.11}$$

Consider the velocity gradient **L**, by introducing a pseudo-time step, gives

$$\mathbf{L} = \dot{\mathbf{F}} \mathbf{F}^{-1} = \mathbf{D} + \mathbf{W}, \tag{1.4.12}$$

where **D** is the deformation rate and **W** is the spin tensor. From eq 1.4.11 and 1.4.12, the plastic portion of the velocity gradient is:

$$\mathbf{L}^p = \sum_\alpha \dot{\gamma}^\alpha \mathbf{s}^\alpha \mathbf{n}^{\alpha \mathrm{T}}, \tag{1.4.13}$$

and outward unit normal to the yield facet of a grain is

$$\mathbf{P}^\alpha = \tfrac{1}{2} (\mathbf{s}^\alpha \mathbf{n}^{\alpha \mathrm{T}} + \mathbf{n}^\alpha \mathbf{s}^{\alpha \mathrm{T}}). \tag{1.4.14}$$

The material parameters are introduced to this calculation by the elastic modulus tensor **C** and the hardening modulus matrix **h**. Manonukul and Dunne[35] considered a simple form of **h**:

$$\mathbf{h} = \begin{pmatrix} h & qh & \cdots & \cdots \\ qh & h & qh & \cdots \\ \cdots & qh & h & qh \\ \cdots & \cdots & qh & h \end{pmatrix}, \tag{1.4.15}$$

where h is the self-hardening rate and q is the latent-hardening ratio.

At each time-step during the simulation, a plastic response of each grain is determined by a trial stress rate, $\mathbf{P}^{\alpha} : \mathbf{C} : \mathbf{D}$. The condition for an active slip system is that $\mathbf{P}^{\alpha} : \mathbf{C} : \mathbf{D} > 0$ (i.e.: energy is released though yielding) and the resolved shear stress reaches the critical resolved stress limit, $\tau^{\alpha} = \tau_c^{\alpha}$. This condition can be written as:

$$\left.\begin{array}{ll} \dot{\gamma}^{\alpha} = 0 & \text{if } \mathbf{P}^{\alpha} : \mathbf{C} : \mathbf{D} \le 0 \quad \text{or} \quad \tau^{\alpha} < \tau_c^{\alpha} \\ \dot{\gamma}^{\alpha} \ge 0 & \text{otherwise} \end{array}\right\}. \tag{1.4.16}$$

Manonukul and Dunne[35] applied this model in 2D and 3D finite element simulations. They considered a system of the nickel-based alloy C263 with 62 surface grains that are extrapolated from a heat-treated experimental sample. The grain configuration and the representative mesh system are shown in Figure 1.14. In their calculations, they varied the loading between LCF and HCF conditions. The plastic deformation state was then compared. After a few cycles, a crack was found to nucleate when the accumulated stress exceeded a maximum value, as shown in Figure 1.15. They reported that these results were in good agreement with experimental data. It should be noted that the trend conforms to Basquin's rule for HCF and Coffin-Manson relationship for LCF. According to the authors, the model may also be applicable for estimating the fatigue limit by observing the stress accumulation trend.

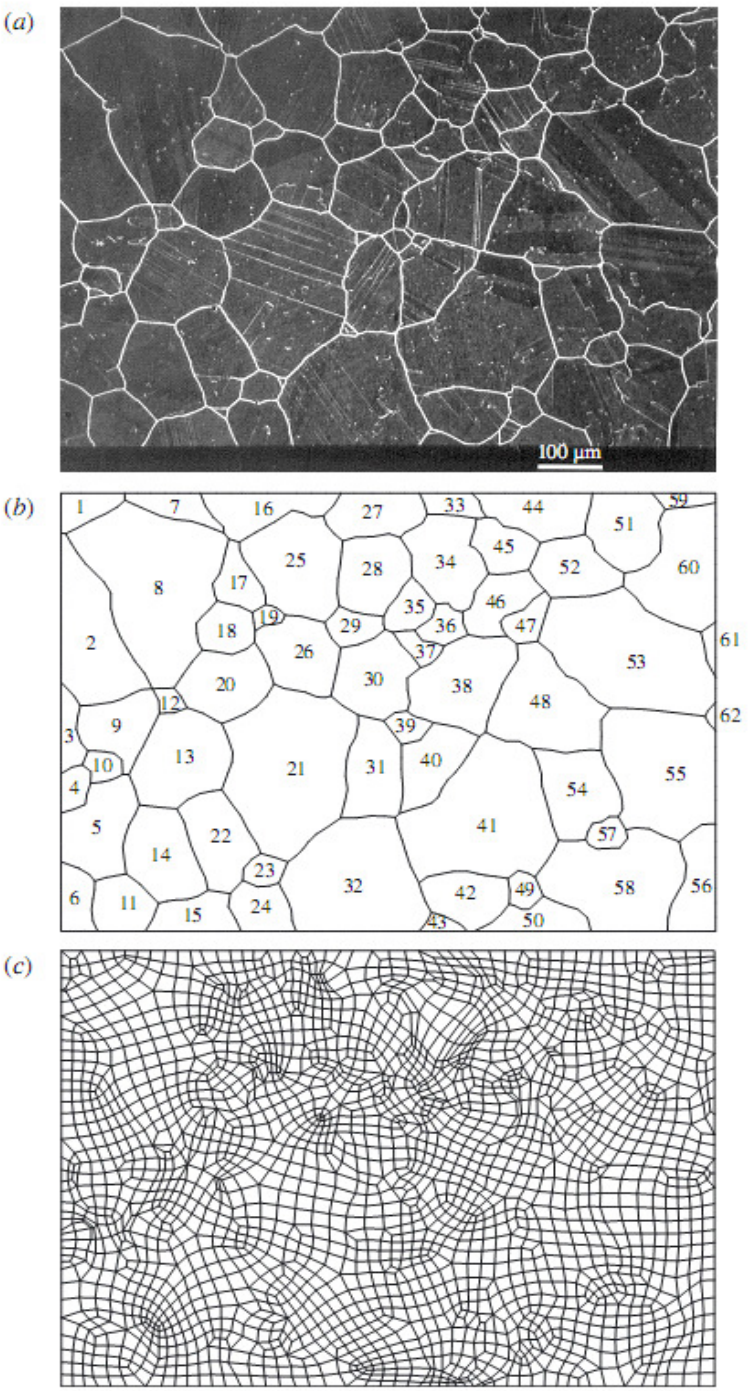


Figure 1.14: a) An surface image of a heat-treated nickel-based alloy C263 sample,
b) A schematic of grain system being simulated,
c) The mesh system for the finite-element simulation.
Figure reproduced with permission from [35], © 2004 The Royal Society.

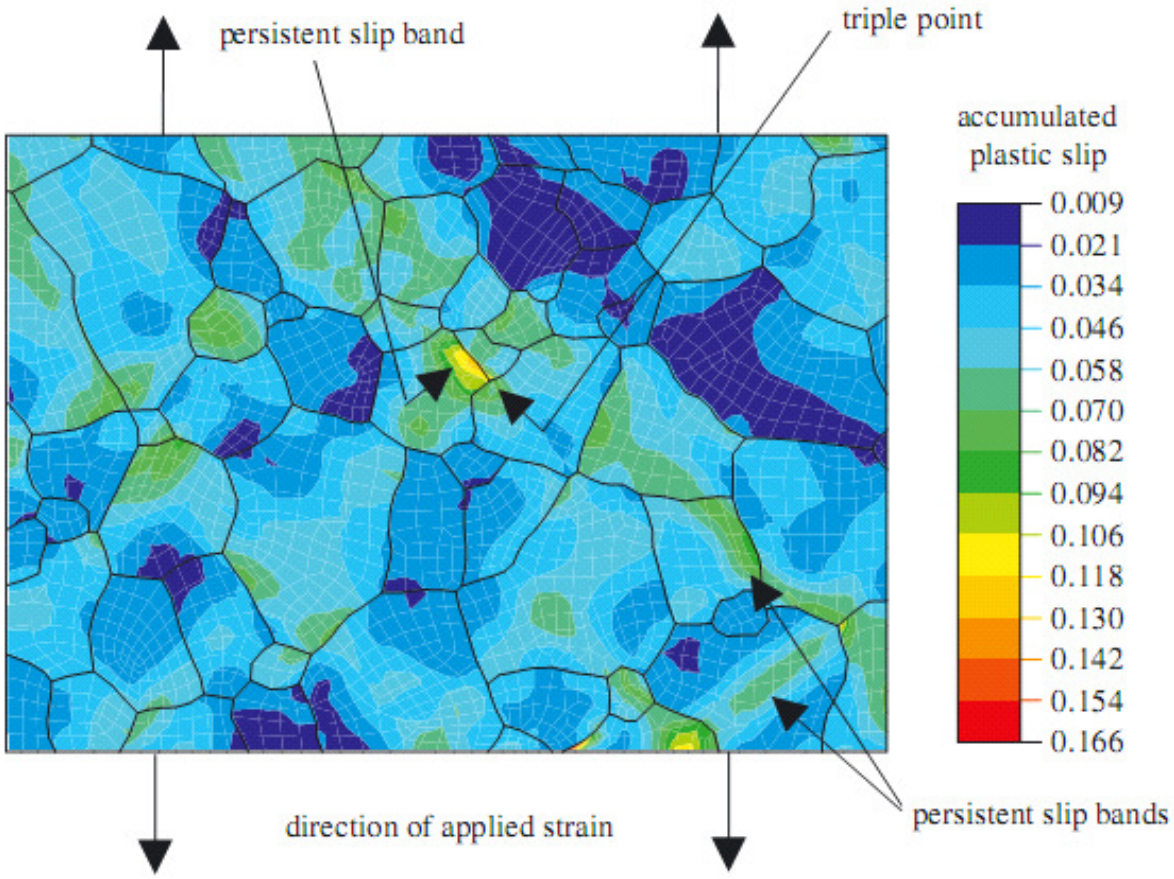


Figure 1.15: Field plot of accumulated plastic slip after two cycles of fully reversed LCF with strain amplitude of 0.5%. Figure reproduced with permission from [35], © 2004 The Royal Society.

1.5. Summary

Analytical and numerical models presented in this review are based on continuous description of microstructures. Although these models are successful in estimating approximate relationships between certain parameters, key details for predicting crack nucleation still remain unclear. For example, the statistical crack nucleation, as a function of cyclic loading or the influence of grain boundary distribution on the nucleation process, cannot be addressed in this fashion. Furthermore, assumptions like irreversibility of plastic flow, and the frictional force in the models described are not fully rationalized. Over the decades, atomistic computations were carried out in many attempts to address these detailed mechanics during a nucleation process. Due to the complexity and the large variation of scenarios possible at atomic level, extensive computational based research at atomic level and at multi-length scales are essential, and will be a growing field of study in the future.

Acknowledgement

This research is supported by DRDC-AVRS project "Life Prediction of Gas Turbine Engine" and DND-TIF project "A Coupled Atomic-Meso-Macroscopic Modeling of Cracks in Aircraft Components"

References

[1] B.A. Cowles, *High cycle fatigue in aircraft gas turbines—an industry perspective*, Int J Fracture. **80**, pp.147 (1996).
[2] T. Nicholas, *Critical issues in high cycle fatigue*, Int J Fatigue. **21**, S221 (1999).
[3] H.L. Huang, N.J. Ho, *The study of fatigue in polycrystalline copper under various strain amplitude at stage I: crack initiation and propagation*, Mat Sci Eng. **A293**, pp.7 (2000).
[4] H.J. Christ, *On the orientation of cyclic-slip-induced intergranular fatigue cracks in face-centered cubic metals*, Mat Sci Eng. **A117**, L25 (1989).
[5] Z.F. Zhang, Z.G. Wang, *Comparison of fatigue cracking possibility along large- and low-angle grain boundaries*, Mat Sci Eng. **A284**, pp.285 (2000).
[6] Y.M Hu, Z.G. Wang, *Fatigue crack initiation and early growth in a copper bicrystal with a grain boundary perpendicular to stress axis*, Scripta Mater. **35**, 8, pp.1019 (1996).
[7] G. Saada, *Stress concentration and crack nucleation at the end of a low-angle tilt boundary*, Phil Mag Lett. **85**, 9, pp.481 (2005).
[8] J.F. Leo, Q.J. Wang, *Fatigue crack initiation at Nd-rich particles in an Nd-containing high-temperature titanium alloy*, J Mater Res. **12**, 10, pp.2571 (1997).
[9] R. Chang, W.L. Morris, *Fatigue crack nucleation at intermetallic particles in alloys — A dislocation pile-up model*, Scripta Metall. **13**, pp.191 (1979).
[10] Z. Shen, R.H. Wagoner, *Dislocation pile-up and grain boundary interactions in 304 stainless steel*, Scripta Metall. **20**, pp.921 (1986).
[11] Z. Shen, R.H. Wagoner, *Dislocation and grain boundary interactions in metals*, Acta Metall. **36**, 12, pp.3231 (1988).
[12] T.C. Lee, I.M Robertson, H.K. Birnbaum, *TEM in situ deformation study of the interaction of lattice dislocations with grain boundaries in metals*, Phil Mag. A **62**, pp.133 (1990).
[13] T.C. Lee, I.M Robertson, H.K. Birnbaum, *An In Situ transmission electron microscope deformation study of the slip transfer mechanisms in metals*, Metall. Trans. A **21A**, 9, pp.2437 (1990).
[14] K Kashihara, F. Inoko, *Effect of piled-up dislocations on strain induced boundary migration (SIBM) in deformed aluminum bicrystals with originally ∑3 twin boundary*, Acta Mater. **49**, pp.3051 (2003).
[15] H. Mughrabi, G.C. Sih, *Defects, Fracture, and Fatigue*, Proc. 2nd Int. Symp. On Defects, Fracture and Fatigue. pp.139 (1982).
[16] H. Mughrabi, R. Wang, *Fatigue Mechanisms: Advances in Quantitative Measurement of Physical Damage,* ASTM Spec. Tech. Publ., **811**, pp.5 (1983).
[17] J. Weertman, *Dislocation Based Fracture Mechanics*, World Scientific. TA409.W44 (1996).
[18] A.H. Cottrell. *Theory of brittle fracture in steel and similar metals*, Trans Metall Soc AIME, **212**, pp.192 (1958).
[19] K. Tanaka, T. Mura, *A dislocation model for fatigue crack initiation*, J Appl Mech. **48**, pp.97 (1981).
[20] T. Mura, Y. Nakasone, *A theory of fatigue crack initiation in solids*, J Appl Mech. **57**, pp.1 (1990).
[21] K.S. Chan, *A microstructure-based fatigue-crack-initiation model*, Metall Mater Trans A, **34**, 1, pp.43 (2003).
[22] W. Elber, *Fatigue crack closure under cyclic tension*, Eng Fract Mech, **2**, pp.37 (1970).
[23] T. Tanaka, M. Kosugi, *Basic Questions in Fatigue Vol 1*, ASTM STP **924**, pp. 98 (1988).
[24] Z.Y. Zhou, N. Tian, *Nanoparticle catalysts with high energy surfaces and enhanced activity synthesized*

by electrochemical method, Faraday Discuss, **140**, pp. 81 (2009).
[25] C.S. Smith, *A Search for Structure*, MIT Press, Cambridge, (1981).
[26] R.G. Tryon, TA Cruse, *Probabilistic mesomechanical fatigue crack nucleation model.* Journal of Engineering Materials and Technology **119**, 1, pp. 65(1997).
[27] R.G. Tryon, TA Cruse, *A reliability-based model to predict scatter in fatigue crack nucleation life*, Fatigue Fact Eng M **21**, pp. 257 (1998).
[28] S. Sasaki, Y. Ochi, *Effects of material structures on statistical scatter in Initiation and growth lives of surface cracks and failure life in fatigue*, JSME Inter J. I, **32**, 1, pp. 155 (1989).
[29] F.A. Bastenaire, *New method for the statistical evaluation of constant stress amplitude fatigue-test results*, Probabilistic Aspects of Fatigue, ASTM STP **511**, pp. 3 (1972).
[30] D. Francois, *Advances in Fatigue Science and Technology*, Academic Publ., pp. 23 (1989).
[31] J. Provan, Z. Zhai, *Fatigue crack initiation and stage-I propagation in polycrystalline materials. I: Micromechanisms*, Int. J. Fat., **13**, 2, pp. 99 (1991).
[32] A. Bruckner-Foit, H. Jackels, *Reliability of components containing microstructural flaws*, Proceedings of ICOSSAR '89, pp. 1499 (1989).
[33] D.L. Davidson, K.S. Chan, *Crystallography of fatigue crack initiation in Astroloy at ambient temperature*, A. Metall, **37**, 4, pp. 1089 (1989).
[34] R.J. Asaro, *Micromechanics of crystals and polycrystals*, Adv Appl Mech. **23**,pp. 1 (1983).
[35] A. Manonukul, F.P.E. Dunne, *High- and low-cycle fatigue crack initiation using polycrystal plasticity*, Proc.R.Soc.Lond. A **460**, pp.1881 (2004).
[36] E. Smith, J. T. Barnby, *Crack nucleation in crystalline solids*, Met Sci J, **1**, pp. 56 (1967).
[37] G.R. Irwin, *Fracturing of Metals*, ASM, Cleveland,OH, pp. 147 (1948).
[38] G. Venkataraman, Y.W. Chung, *Free Energy Formulation of Fatigue Crack Initiation along Persistent Slip Bands: Calculation of S-N Curves and Crack Depths*, Acta Metall Mater, **38**, pp. 31 (1990).
[39] T. Mura, Y. Nakasone, *A theory of fatigue crack initiation in solids*, J Appl Mech, **57**, pp. 1 (1990).
[40] J.S. Koehler, *The production of large stresses by dislocations.* Phys Rev, **85**, pp. 480 (1952).
[41] C. Zener, *The micro-mechanism of fracture*, Fracturing of Metals, American Society for Metals, pp. 3 (1948).
[42] N.I. Muskhelishvili, *Singular Integral Equations 1st ed*, Springer (1977).

.

Chapter 2

Theoretical Rationale and Data Analysis Methodology

2.1 Theory

2.1.1 Embedded Atom Method (EAM)

In Newtonian molecular dynamics simulation, the physical behaviour of atoms is introduced to the system via an interatomic potential. Numerous functional forms for the potential had been suggested throughout the history of atomistic simulation development. They can be classified as *two-body(pair)* or *many-body* potential functions. Further, they can be subcategorised into short-ranged and long-ranged interactions. Amongst those being suggested, the embedded atom method (EAM) and its deviants are most widely adopted in the study of metallic systems. With some appropriate potential, this method is demonstrated to have a good representation on both the bulk properties and atomic trajectory in many metallic and alloy systems.

EAM is a many-body atomic interaction model first proposed by Daw and Baskes in 1983 [1, 2]. Prior to that time, pair potentiasl were used in molecular dynamics (MD) simulations. However, it was found that it is insufficient in many scenarios. In particular, the volume-dependency in pair potentials renders it ambiguous at the presence of surfaces or impurities. This deficiency greatly hindered the applicability of MD simulations in a wide range of topics. To rectify the problem, EAM was constructed on the bases of quasiatom (or effective medium) theory [3,4]. According to the theory, an atom in a host is considered as being embedded in a density of electrons from all surrounding atoms. Since an electron density is definable at any distances from an atom and that a change in an atomic volume is reflected intrinsically in the changes of the density, an arbitrary definition of an atom's volume is not necessary.

The initially proposed form of an EAM potential consists of two independent energy terms, namely:

$$E_i = \frac{1}{2}\sum_{j \neq i} V_{ij}(r_{ij}) + F_i(n_i). \tag{2.1.1}$$

The first term, V_{ij} is a core-core interaction potential. It is dependent on the separation distance, r_{ij}, between atom i and j. The second term, F_i, is the *embedding energy function*. It describes the energy required to embed a charged core into a net electron densities, n_i. For simplicity, the net charge is commonly assumed to be a linear superposition of electron densities from each neighbouring atoms, while each atom's electron density, ρ, is a function of r_{ij} only:

$$n_i = \sum_{j \neq i} \rho(r_{ij}). \tag{2.1.2}$$

This approximation is appropriate in most covalently bonded systems. However, the sole dependency of r_{ij} in the potential energy enforces a spherically symmetric interaction. As such, the effect of polarization or hybridization is completely ignored. Additional treatment to *eq 2.1.1* is required to correctly simulate at those specific conditions. In molecular dynamics simulation, an EAM potential is used to derive the force experienced by an atom. It is the force that is being implemented in an MD simulation program. Taking appropriate differentiation from *eq 2.1.1* and considering reaction force from the neighbours' motion, yields a force equation of:

$$\vec{f}_i = -\sum_{j \neq i} \hat{r}_{ij} \left(V'(r_{ij}) + [F'(n_i) + F'(n_j)]\rho'(r_{ij}) \right). \tag{2.1.3}$$

According this equation, the force of atom i is dependent on n_j, which is contributed from atom j's neighbours. As such, an atom in EAM scheme is effectively interacting within twice the cut-off radius.

2.1.2 Analytical Embedded Atom Method

The functional form of the original EAM potential is purely empirical. Five years after it was first suggested, Johnson *et al*. derived an analytic form of the EAM potential for the FCC system [5]. Subsequently follows were the analytic form for BCC and HCP structures. The analytical approach considers building an EAM potential from some pre-chosen functional form. This method provides a consistent and universal procedure for developing a potential. Johnson's version of analytic EAM (AEAM) functions considered the second nearest neighbour interaction (2NN) and had a smooth termination at the cutoff radius. His model was demonstrated to be applicable in a wide set of metals with few exceptions. The most notable exception was the application on chromium. AEAM in its unmodified form failed to reproduce the negative Cauchy pressure Cr has. An modification to *eq 2.1.1* was proposed to resolve the issue[6]. According to the suggestion, the equation should be rewritten as:

$$E_i = \frac{1}{2}\sum_{j \neq i} V_{ij}(r_{ij}) + F_i(n_i) + M(n_i^*), \tag{2.1.4}$$

and

$$n_i^* = \sum_{j \neq i} \rho^2(r_{ij}). \tag{2.1.5}$$

Here, by introducing non-linearity to the electron density, the term $M(n_i^*)$ can account for both the hybridization and the polarity [6].

2.1.3 Gauge Invariance and its applications in EAM

The EAM model can be recognized in terms of the Hamiltonian formalism. A particular feature of the formulism is the invariance under gauge transformation. Although gauge invariance does not have a physical interpretation, it can be understood as the consequence of redundancy in the information describing a system. In the context of EAM, gauge transformations can be expressed as [7, 8]:

$$\rho(r) \to A\rho(r); F(n) \to F(n/A), \tag{2.1.6}$$

$$V(r) \to V(r) + 2B\rho(r);\ \ V(r) \to V(r) - Bn. \tag{2.1.7}$$

An EAM potential is invariant under these transformations, where A and B are arbitrary constants. It can be understood in terms of linear exchangeability: any linear dependency in the many-body interaction can be reformulated into a purely two-body interaction. It is the non-linearity in the many-body interaction that gives rise to a local volume dependency [5]. Base on this argument, one can establish relationships between different models suggested in the past. For example, Johnson [5] had pointed out that the function used by Finnis and Sinclair[9], which is of square-root dependency, and the functions used by Foiles[10], are all differ with an AMEAM function by a term linear in the superposition variable. As such, he concluded that these models are essentially equivalent. Furthermore, the glue model [8], which is a many-body potential model closely related to the EAM model, can be shown under this transformation as equivalent too. The potential functions developed under these two models can therefore be used interchangeably.

Gauge invariance is often seen as a way of standardization. By inflicting gauge transformation, a standardized electron density n_0 can be defined such that:

$$F'(n_0) = 0. \tag{2.1.8}$$

The value of n_0 is commonly taken as the bulk coordination number. The choice of n_0 does not reflect any physical relevancy. It solely serves the purpose of standardization, such that different potentials can be compared without ambiguity. In the case of alloy systems where multiple single-element potentials are mixed together, this standardization scheme becomes very useful. It allows electron densities of different elements be added in the correct ratio without complication.

2.1.4 Force Matching Potential

The representability of an atomistic simulation employing EAM is dependent heavily on the correctness of the potential energy function. In the early days, potentials were developed by fitting a large set of

experimentally determined bulk parameters. This procedure however raises concerns: potential functions developed under this scheme are not unique. An identical set of bulk properties can be fitted by some distinctively different potential functions. Furthermore, obtusely fitting experimental data disconnect the resulting functions with their physical interpretation. In addition, the precision of a potential is limited by the quality of the experimental data available. Also, a potential reflecting accurate energetic parameters does not necessarily implies a correct description of atomic interaction. These inconsistency and ambiguity renders the rational basis of any simulation result dismissible. A resolution to these deficiencies was introduced by Ercolessi and Adams [7] in 1994. Their proposed procedure is known as the force-matching method, sometimes denote as a prefix '*fm*-' in front of *EAM*. Force-matching method employs first principle calculation as the basis of the construction. First principle or *ab initio* atomistic calculation evaluates the atomistic forces by directly solving the electronic-structure, taking into account the quantum nature of electrons. Most commonly, the atomistic force is evaluated by local density approximation (LDA) pseudo-potential under the density functional theory (DFT). The force-matching method allows a precise evaluation of both the electron densities and the forces at the subatomic scale. Unlike fitting a large set of bulk properties, potential developed via this force-matching scheme provide a good level of confidence on the description of atomic interaction and trajectory. It is important to note that a force-matching procedure often involves assumptions and some small degree of empirical fitting. Therefore, even though it employs *ab initio* calculation as its bases, an *fm-EAM* is not a first-principle model.

2.2 Data Analysis and Visualization

2.2.1 Coordination Number

Coordinate number is an integer assigned to an atom, defined as the number of nearest neighbour an atom has. In practice, when thermal motion and stresses are present, the nearest neighbour is not always well defined. As such, coordination number is usually evaluated as the number of neighbouring atoms present within a cut-off radius. The cut-off radius is adjustable base on its effectiveness in picking out atoms of irregular periodicity. It has a value between the first- and second-nearest neighbours distance. The arbitrary determination of the cut-off radius introduces some minor ambiguity and should be handled with caution.

Coordination number coloring is a powerful way to visualize atomistic configuration. It allows clear numerical and visual identification of majority of the crystalline defects. Although vacancy cannot be directly observed in this coloring scheme, its present can be inferred by the reduced coordination number of its neighbouring atoms. These neighbours appear visually as a sphere of atoms. When multiple vacancies

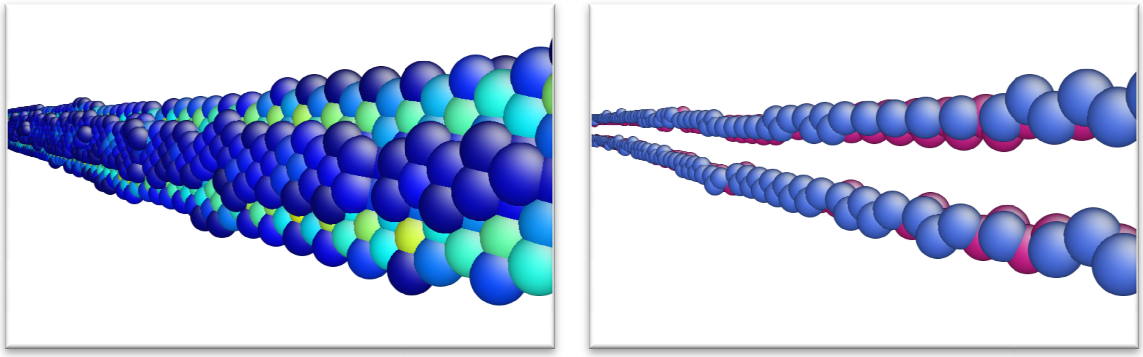
cluster together, the size can be determined by the number of the enveloping atoms. Other crystalline defects, such as dislocations and most of the grain boundaries, are visually obvious in this coloring scheme.

2.2.2 Central Symmetry Parameter

Coordination number coloring is a very powerful visualization tool. However, there are limitations associated with it. The major restriction lies in its incapability in distinguishing stacking fault from regular crystal periodicity. In most crystal structures, stacking fault has a coordination number equivalent to the bulk. An atom on a stacking fault is different from the bulk by its angular relationship with its neighbour. A variable known as the central symmetry parameter can be applied in this scenario. Central symmetry parameter characterizes the degree of inversion-symmetry-breaking of a local set of lattice points [11]. Consider an atom with *N* nearest neighbours, where a neighbour's position with respect to the atom is $\vec{r}_j$, one can define $\lfloor N \rfloor/2$ non-overlapping pairs of neighbours which the pair's positions are closest to the each other's symmetrical image. For example, neighbour j is paired with neighbour i when it minimizes $|R_i|$ amongst all choice of j excluding i, where $R_i = r_i + r_j$. The central symmetry parameter is defined as [11]:

$$c \equiv \frac{\sum_{k}^{\frac{\lfloor N \rfloor}{2}} R_k^2}{2\sum_{j}^{N} r_j^2}. \tag{2.2.1}$$

This quantity is dimensionless and normalized. It has a value in the interval [0, 1]. A value of zero represents a perfect symmetry while a value of one means the opposite. A perfect cubic lattice structure always has a value of zero at every of its points. At the present of moderate thermal fluctuation, its value remains close to zero, with an upper bound at around 0.01. On the other hand, atoms on a stacking fault in a FCC system have a value of around 0.0416, which is clearly distinguishable from that of a perfect crystal. *Figure 2.1* illustrated a stacking fault in the FCC system created by a dissociation of an edge dislocation. The atoms on the two layers of a stacking fault are clearly isolated from the surrounding atoms. It worth to note that there exist other similar parameters used amongst researchers for the same purpose. Of those, the centrosymmetry parameter [12] is a common choice. Several popular MD simulation software have implemented this parameter as a built-in function. It is equivalent in principle to the parameter discussed above, and is equally useful in distinguishing stacking faults. However, its value is not normalized and has a dimension of distance. Elements with different lattice constants or the same element at different compression state will have different filtering criteria for stacking faults identification.



ımeter coloring.

2.2.3 Stress Evaluation and Stress Invariants

The concept of stress outlined the foundation of fracture mechanics. It is a continuum description of the tensorial force (normalized to unit area) subjected on a virtual point within a material. At small deformation, a stress is completely described by the nine components of a Cauchy stress tensor according to Cauchy's stress theorem. The tensor has a form:

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} = \begin{bmatrix} \sigma_x & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_y & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_z \end{bmatrix} = \begin{bmatrix} \mathbf{T}^{\mathbf{e}_1} \\ \mathbf{T}^{\mathbf{e}_2} \\ \mathbf{T}^{\mathbf{e}_3} \end{bmatrix}, \tag{2.2.2}$$

where σ_i and τ_{ij} are the tensile and shear components respectively, while $\mathbf{T}^{\mathbf{e}_i}$ are stress vectors on three imaginary orthogonal faces. This tensor is system orientation dependent and its rotational transformation can be represented by a Mohr's circle. Since stress is a physical quantity, it is possible to extract invariant parameters that are system orientation independent. The first set of invariants, commonly refers to as I_1, I_2 & I_3, can be derived from the characteristic equation of $\boldsymbol{\sigma}$:

$$\det(\boldsymbol{\sigma} - \lambda \mathbf{I}) = \lambda^3 + I_1\lambda^2 + I_2\lambda + I_3 = 0, \tag{2.2.3}$$

where $\mathbf{I}$ is the identity matrix. Thus, yield:

$$\begin{aligned} I_1 &= \sum \sigma_{ii}\,, \\ I_2 &= \frac{1}{2}\sum \left(\sigma_{ii}\sigma_{jj} - \sigma_{ij}\sigma_{ji}\right), \\ I_3 &= \det\left(\sigma_{ij}\right). \end{aligned} \tag{2.2.4}$$

The hydrostatic pressure is equivalent to one third of the first stress invariant, namely: $P = \frac{1}{3}I_1$. From this relationship, the second set of stress invariants $\{J_1\ J_2\ J_3\}$ can be deduced by defining a second order tensor,

$\boldsymbol{s} = \boldsymbol{\sigma} - P\mathbf{I}$. This tensor is known as the stress deviator tensor. Identical to previous derivation, the invariants are obtained from its characteristic equation; resulting with:

$$\begin{aligned} J_1 &= \sum s_{ii} = 0, \\ J_2 &= \frac{1}{2}\sum \left(s_{ii}s_{jj} - s_{ij}s_{ji}\right) = \tfrac{1}{3}I_1^2 - I_2, \\ J_3 &= \det\left(s_{ij}\right) = \tfrac{2}{27}I_1^3 - \tfrac{1}{3}I_1 I_2 + I_3. \end{aligned} \tag{2.2.5}$$

The invariants for the second deviatoric stress are directly related to the yield condition for multi-axial loading condition. According to von Mises yield criterion[13], there exists a critical value of J_2 for an isotropic material at which it begins to yield. This criterion can be expressed in terms of the yield strength, σ_y and the von Mises stress, σ_v, as:

$$\sigma_v = \sqrt{3J_2} \geq \sigma_y, \tag{2.2.6}$$

An equivalent criterion in energy space is suggested by Hencky[13]. He relates the maximum distortion strain energy, W_D with the second deviatoric stress invariant as follows:

$$\frac{J_2}{2G} \geq W_D, \tag{2.2.7}$$

These two criteria can be use to estimate the conditions for dislocation nucleation in molecular dynamics simulations.

2.2.4 Atomistic Stress – Virial Theorem

The original concept of stress is only applicable in an assume-continuous medium. However, it is not well defined at the atomic level where the medium is constituted of distinguishable discrete entities. Atomistic equivalent of the continuum stress description are imminently essential in molecular dynamics simulation. It allows connection between atomic scale phenomena and the continuum mechanics be drawn directly. Attempts have been made to construct such stress description. Of those numerous suggestions, the virial stress derived from the virial theorem is most widely adopted.

Virial theorem stated that the asymptotic average of the derivative of a bounded function is zero. Considering a solid material at finite temperature, both the position and the velocity of atoms in an enclosed volume are confined. Thus, the product of these two quantities is also bounded. Expanding the time average of its time derivative yield [14]:

$$\langle \sum_i \frac{m_i}{V} \frac{d(\boldsymbol{r}_i \cdot \dot{\boldsymbol{r}}_i)}{dt} \rangle = \langle \frac{1}{V} \sum_i \boldsymbol{r}_i \cdot \boldsymbol{F}_i \rangle + \langle \frac{1}{V} \sum_i m_i \dot{\boldsymbol{r}}_i^2 \rangle = 0, \tag{2.2.8}$$

where the force $\boldsymbol{F}_i$ experienced by an atom inside the volume is contributed by both internal and external bodies, ie: $\boldsymbol{F}_i = \boldsymbol{F}_i^{int} + \boldsymbol{F}_i^{ext}$. By considering that $= V^{-1} \sum \boldsymbol{r}_i \cdot \boldsymbol{F}_i^{ext}$, *eq 2.2.8* can be rewritten as:

$$\langle \sigma \rangle = -\langle \frac{1}{V} \sum_i \boldsymbol{r}_i \cdot \boldsymbol{F}_i^{int} \rangle - \langle \frac{1}{V} \sum_i m_i \dot{\boldsymbol{r}}_i^2 \rangle. \tag{2.2.9}$$

During the early last decade, the virial stress described in *eq 2.2.9* was subject to controversy. The main debate centered on its conformity to the macroscopic Cauchy stress. One school of thought [14, 15, 16] asserted that the kinetic term in *eq 2.2.9* should be ignored based on some momenta consideration. It is only the simplified equation that is consistent with the Cauchy stress description. The idea flourished a series of scientific work, including a set of MIT course note [17]. It was only the last few years that the end of the controversy came to light. It was clearly demonstrated that the argument mentioned above is flawed, and that *eq 2.2.9* is indeed correct and complete [17]. The controversy had brought to attention the necessary condition for the virial theorem. It is now realized the emphasis that *eq 2.2.9* only conforms to the Cauchy description when the quantities are averaged sufficiently in the spatial and temporal domain. It is worth to note that scientific works containing flawed stress are still easily accessible in the scientific community. One should proceed with caution.

References

[1] Daw, M. & Baskes, M. Semiempirical, Quantum Mechanical Calculation of Hydrogen Embrittlement in Metals. *Physical Review Letters* **50**, 1285-1288(1983).

[2] Daw, M. & Baskes, M. Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals. *Physical Review B* **29**, 6443-6453(1984).

[3] Stott, M. & Zaremba, E. Quasiatoms: An approach to atoms in nonuniform electronic systems. *Physical Review B* **22**, 1564-1583(1980).

[4] Nørskov, J. Covalent effects in the effective-medium theory of chemical binding: Hydrogen heats of solution in the 3d metals. *Physical Review B* **26**, 2875-2885(1982).

[5] Johnson, R. Analytic nearest-neighbor model for fcc metals. *Physical Review B* **37**, 3924-3931(1988).

[6] Yifang, O. A simple analytical EAM model for bcc metals including Cr and its application. *Zeitschrift fur Physik B Condensed Matter* **101**, 161–168(1996).

[7] Ercolessi, F. & Adams, J.B. Interatomic Potentials from First-Principles Calculations: The Force-Matching Method. *Europhysics Letters (EPL)* **26**, 583-588(1994).

[8] Ercolessi, F., Parrinello, M. & Tosatti, E. Simulation of gold in the glue model. *Philosophical Magazine A* **58**, 213–226(1988).

[9] Finnis, M.W. & Sinclair, J.E. A simple empirical N-body potential for transition metals. *Philosophical Magazine A* **50**, 45-55(1984).

[10] Foiles, S., Baskes, M. & Daw, M. Embedded-atom-method functions for the fcc metals Cu, Ag, Au, Ni, Pd, Pt, and their alloys. *Physical Review B* **33**, 7983-7991(1986).

[11] Yip, S. Handbook of materials modeling. *Kluwer Academic Pub* (2005).

[12] Kelchner, C., Plimpton, S. & Hamilton, J. Dislocation nucleation and defect structure during surface indentation. *Physical Review B* **58**, 11085-11088(1998).

[13] Hill, R. The mathematical theory of plasticity. *Oxford University Press,* USA (1998).

[14] Zhou, M. A new look at the atomic level virial stress: on continuum-molecular system equivalence. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **459**, 2347-2392(2003).

[15] Zhou, M. & McDowell, D. Equivalent continuum for dynamically deforming atomistic particle systems. *Philosophical Magazine A* **82**, 2547-2574(2002).

[16] Zimmerman, J. a. Calculation of stress in atomistic simulation. *Modelling and Simulation in Materials Science and Engineering* **12**, S319-S332(2004).

[17] Subramaniyan, A.K. & Sun, C.T. Continuum interpretation of virial stress in molecular simulations. *International Journal of Solids and Structures* **45**, 4340-4346(2008).

[18] Buehler, M.J. & Gao, H. Atomistic and continuum studies of stress and strain fields near a rapidly propagating crack in a harmonic lattice. *Theoretical and Applied Fracture Mechanics* **41**, 21-42(2004).

[19] Tsai, D. The virial theorem and stress calculation in molecular dynamics. *The Journal of Chemical Physics* **70**, 1375(1979).

Chapter 3

Computation Implementation and Efficiency Optimization

3.1 Computation Implementation

3.1.1 Consideration

A massively parallelized molecular dynamics simulation program was developed for the purpose of the researches discussed in this thesis. The source code will be publicly available and is intended to serve both educational and research purposes for the academia. It is designated to a long-term continuous development. Given the intended purposes, the code was developed from scratch with an emphasis on preeminent clarity and performance. Some simple benchmark results indicated that its performance exceeds another publicly available MD simulation program, LAMMPS, by 74% under comparable conditions, *see section 3.2.5*. Currently the project is assigned a temporary code name: MPMDS, which stands for multi-parallel molecular dynamic simulator. In the following sections, the program will be referred to by this name.

MPMDS is written in *FORTRAN 95*. The language is chosen for several specific reasons: 1) scientific notations and vectors calculations can be expressed naturally with its syntax. 2) It provides a standardized language support for parallelism. 3) Values are intrinsically passed by pointers. 4) The availability of high performance compilers. These features are essential for scientific programming, and are the major reasons why FORTRAN is still a common choice in the scientific community. However, programming in FORTRAN requires special attention on the portability. The support of intrinsic functions and the use of *pragma* are not standardized amongst different compilers. Careful considerations have been taken to address the issue. The code is verified fully portable to both the Microsoft® Window and Linux platform and it can be compile in most of the common FORTRAN compilers without modifications.

3.1.2 Force Evaluation

MPMDS implemented the embedded atom method (EAM) as one of its force evaluation module. The many-body force is described in *eq 2.1.3*. According to the equation, calculating the force for an atom require evaluating both the net charge density of atom *i* and atom *j*, at which the calculation involves the position of a neighbour's neighbour. If the equation is implemented in a straight-forward fashion, it will yield an algorithm with an order of approximately $O(nm^2)$, where n is the number of atoms and m is the average number of neighbours. A simple rearrangement can removes some redundancy and reduce the order to $O(2nm)$. The implementation involves a two-pass algorithm. 'Two-pass' here refers to two loops that pass through every atom. Each loop contains an inner loop going through all neighbours of an atom. The first-pass evaluates the net electron density, n_i, while the second-pass evaluates the force of each individual atom.

This algorithm is compatible with a parallelism technique known as force decomposition (FD)[1], however the technique is only beneficial up to mid-size problem. In the case of large-size problem, a better implementation is achieved by a one-pass algorithm that contains two inner loops. This method requires rewriting eq 2.1.3, as follows:

$$\vec{f}_i = -\sum_{j \neq i} \{\hat{r}_{ij}(V'(r_{ij})) + \vec{\varphi}_{ii} + \vec{\varphi}_{ij}\}, \tag{3.1.1}$$

where

$$\vec{\varphi}_{ii} = F'(n_i)\rho'(r_{ij})\hat{r}_{ij} \quad \text{and} \quad \vec{\varphi}_{ij} = F'(n_j)\rho'(r_{ij})\hat{r}_{ij}. \tag{3.1.2}$$

By considering the skew symmetry, $\hat{r}_{ij} = -\hat{r}_{ji}$, it can be easily deduced that $\vec{\varphi}_{ji} = -\vec{\varphi}_{ii}$. As such, the one-pass algorithm can be illustrated clearly by the following pseudo-code:

```
For i = every atoms  // Outer Loop
   For j = every i's neighbour  //First Inner Loop
      n_i += ρ(r_ij)
      ρ'_net += ρ'(r_ij) * r̂_ij
      V'_net += V'(r_ij) * r̂_ij
   End Loop

   φ_ii = F'(n_i) * ρ'_net
   f_i += -(V'_net + φ_ii)

   For j = every i's neighbour //Second Inner Loop
      f_j += φ_ii    // Inflicting Symmetry φ_ji = -φ_ii
   End Loop
End Loop
```

This implementation exceeds the two-pass algorithm by removing redundancy on the evaluation of $\vec{\varphi}_{ij}$ and by condensing the algorithm into a one-pass loop. In addition, all inner loops are vectorizable under the SIMD (simple instruction, multiple data) architecture. This allows a much higher CPU utilization per cycle and results in a better performance.

3.1.3 Predictor-Corrector Integration

Obtaining the position of an atom from a force require performing numerical time-integration twice over a value. Conventional numerical integration techniques will accumulate significant amount of error as time proceeds. The dynamics and the trajectories in a simulation will therefore quickly become inappreciable. For this reason, a predictor-corrector integration scheme is used in the molecular dynamics simulation program. The idea of this integration scheme is to have an integration error that is bounded and that it would oscillates around zero as time proceeds. This minimizes the accumulative error and improves stability of the

integration. MPMDS incorporated the Gear predictor-corrector algorithm. A fifth order Gear algorithm describes an atom by a set of six generalized coordinates, namely: $[\vec{r}_0, \vec{r}_1, \dots, \vec{r}_5]$, where $\vec{r}_i = \frac{d^i \vec{r}_0}{dt^i}$ and $\vec{r}_0$ is the position. The algorithm evaluates the motion by first predicting the coordinates on the next time increment:

$$\left[\{\vec{r}_i^{\,p}(t+\Delta t)\}\right] = \begin{bmatrix} 1 & a_1 & a_2 & a_3 & a_4 & a_5 \\ 0 & 1 & a_1 & a_2 & a_3 & a_4 \\ 0 & 0 & 1 & a_1 & a_2 & a_3 \\ 0 & 0 & 0 & 1 & a_1 & a_2 \\ 0 & 0 & 0 & 0 & 1 & a_1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} [\{\vec{r}_i(t)\}], \tag{3.1.3}$$

where $a_n = \Delta t^{\mathrm{n}}/\mathrm{n}!$ and the superscript *p* denote a predicted value. Subsequently, the value of $\vec{r}_2^{\,p}$ is compared with that obtained from a force calculation. Let *c* denote a calculated value, then, the difference can be expressed as:

$$\Delta\vec{r}_2 = \vec{r}_2^{\,c} - \vec{r}_2^{\,p}. \tag{3.1.4}$$

Finally, the difference is used to correct the predicted values:

$$\vec{r}_i(t+\Delta t) = \vec{r}_i^{\,p} + c_i \Delta\vec{r}_2. \tag{3.1.5}$$

In this equation, c_i are order dependent constants. For the fifth order Gear algorithm, $c = \left\{\frac{3}{20}, \frac{251}{360}, 1, \frac{11}{18}, \frac{1}{6}, \frac{1}{60}\right\}$. The duration of the numerical stability is determined by the order of the gear algorithm. A fifth order Gear predictor-corrector integration scheme provides a numerical stability longer than most MD simulation required. Higher order terms are rarely necessary. This algorithm scales linearly with the number of atoms, and is fully vectorizable under SIMD framework. As a result, the entire numerical calculation occupies less than 2% of the total MPMDS run time.

3.2 Acceleration Techniques and Efficiency

3.2.1 Overview

In numerical simulation, the available computation power is a limiting factor. It constrained both the simulation size and the level of detail presented in a model. Subsequently, these limitations externalize in the availability of accuracy, precision and complexity of a simulation. A precisely optimized computation program can execute up to two orders of magnitude faster. Techniques such as assembly code injection, vectorization, numerical function tabulation, inline expansion and parallelization are some examples widely implemented in scientific programming. As technologies involved in processors and compilers become extraordinarily sophisticated, optimization for scientific calculations is an intricate topic and is out of the

scope of this discussion. However, there are some optimization techniques specific to molecular dynamics simulation. They will be discussed briefly in the following sections.

3.2.2 Parallelism in Molecular Dynamics Computation – Shared Memory Model

In recent years, there is a sturdy trend toward parallelism in personal computing technology. As advancement in processor clock speed is getting ever challenging, more attention is drawn toward parallelism as a mean to increase computation performance. Parallel computation has been widely utilized in scientific programming for decades. It allows massive physical simulation or complicated calculation be performed in a reasonable time. The capability is particularly necessary in molecular dynamics simulation on material physics applications, where huge number of atoms and neighbour interactions are involved. Parallelism in programming can be realized in several different levels. The parallelism models at different levels are independent and can be implemented cooperatively in a program.

Parallelization in shared-memory model can be realized in hardware with multi-thread or multi-core architecture. It is compatible with all current CPU and most GPU (graphic processing unit) and some specially designed computer cluster. In the context of MD simulation, parallelization at this scale requires decomposing large loops into independent threads of loops. The threads are distributed to parallel processing units for evaluation. The loops must not be recursive or external reference dependent. In terms of parallelisation in multi-core CPU, implementation can be fully automatic at the compile time via auto vectorization. As discussed in previous sections, the core calculation of MPMDS, which compose of EAM force evaluation and Gear integration, are all intrinsically vectorizable. In recent years, highly parallelized processing units like GPU and Cell processor are extending its capability to general purpose calculations. Scientific programming has been shifting rapidly toward such parallelism implementation. The idea of parallelism in GPU is similar to that of CPU in principle. However, the scale of parallelism is massive compare to a CPU. For instance, as of current architecture, a GPU can have several hundreds of processing units running in parallel compare to a few units for a CPU. Further, processing units in GPU are specialized in numeric vector manipulations and transcendental function calculations. For this reason, there have been reports of performance gain of up to two orders of magnitude in some scientific applications utilizing GPU acceleration. Porting scientific calculations onto GPU have contributed the largest performance leap ever recorded in the history. However, parallelism implementation on GPU is not entirely trivial. Currently, auto parallelization technology for GPU is at its infancy and is not yet generally deployable. Besides, under the current GPU interface architecture, memory handling must be controlled manually. A programmer must

provide specific instructions in the code as to what data to pass to a GPU during runtime. In addition, data transfer is tricky to handle because the memory bandwidth between CPU and GPU is very limited. Newest in-development technology might provide a solution to these disadvantages. AMD® is constructing a new class of processor known as an APU (accelerated processing unit) which incorporate a conventional CPU, a general purpose GPU and a built-in memory controller into a unified package. MPMDS will continue to adopt the latest parallelism implementation as the technology becomes available and mature.

3.2.3 Parallelism in Molecular Dynamics Computation - Distributed Computing

Another class of parallelism, refers to as distributed computing, is widely implemented in scientific programming. It involves distributing calculations amongst a cluster of processors; at which each processor maintain a private set of data and memory. Communication between processors is made through a special interconnection network. MPMDS implemented such parallelism via the MPI 2.2 (message passing interface) protocol. MPI allows specific sets of data be distributed to nodes within a cluster. In the context of short-range MD simulation, there are three major choices[1] for such parallel implementation. The three choices can be refers to as: atom-decomposition, force-decomposition and spatial decomposition. In atom-decomposition, atoms in a system are distributed to specific processors for evaluation. The assignment remains unchanged throughout the simulation. In force-decomposition, a subset of force calculations is distributed amongst processors. Each processor must maintain an updated position of all atoms. These two implementations are only beneficial up to a mid-size problem, ie moderate ratio between number of atoms, N and number of processors, P. For large scale problem, spatial-decomposition remains the superior choice. It involves dividing a super-cell into a set of sub-cells. Each processor is responsible for the atoms within the designated sub-cell. Communication is made to exchange borders atoms' information. Super-cell division rules are determined by the shape of the super-cell and the N/P ratio. In principle, one can choose between dividing a super-cell in one-, two- or three-dimensions. However, in practice, having a higher dimension division scheme introduces overhead in calculations and communications. They are only beneficial when division in a dimension saturates, ie: when the thickness of a sub-cell is comparable to four times the force cut-off radius. For this reason, MPMDS employs the linear division scheme. It divides a super-cell along the longest dimension to minimize cross-section area. In this division scheme, each sub-cell has two neighbouring sub-cell, connected by an interface plane. During execution, atoms moving across an interface plane will be passed to the corresponding sub-cell via MPI protocols. Essential information on atoms resided near the interface is also passed to the neighbouring sub-cell for force evaluation purposes. Atoms information is packed into a single package before it is sent to the neighbour. Initiating communication between nodes in a per-atom fashion will introduce unaffordable overhead time. Besides communications

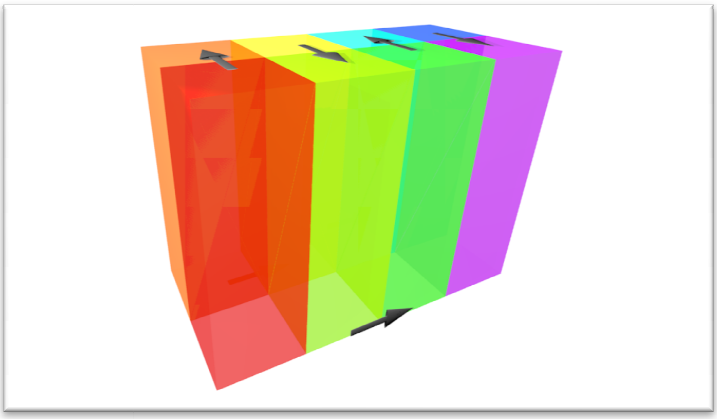
made to satisfy sub-cell boundary conditions, essential information like center of mass motion and temperature are also distributed amongst all computer nodes every cycle. For 3.59 million atoms running on 8 CPU nodes, MPI communications and the related calculations collectively constitute less than 10% of the total run time.

3.2.4 Array Element Reordering

Contrary to common perception, solely optimizing a calculation algorithm does not produce efficient MD codes. Consider the capability of current computer hardware; a CPU core might handle up to millions of atoms each simulation time step. Including intermediate values, there are up to gigabytes of data being referenced each simulation cycle. Memory bandwidth is therefore the most dominant bottleneck in a MD simulation. As a consequence, optimization of memory handling is imminently crucial. In order to accomplish this, the architecture of the current processor must be taken into consideration. As of the latest design, in each processor, there are two or three levels of local and shared memory cache in a pipeline, with the largest size in the order of megabytes shared by multiple cores. The cache implementation is different amongst the two major microprocessor makers and within their series of models. In any case, these caches are designed to provide fast memory service to the core calculation units. When a piece of data is referenced, a specific length of memory, containing the requested data, is transferred from the random access memory (RAM) to the processor's cache. The package of memory being transfer is known as a cache line. It contains the requested data and its neighbouring memory elements. The transfer speed is highly constrained by the available memory bandwidth between RAM and CPU. It is also dependent on the architecture of the CPU. For instances, CPU with built-in memory controller can access RAM relatively faster. However, even with a built-in controller, access speed is still much slower than cache service in a CPU. The event when a piece of data being referenced is already in a cache is known as a 'cache hit'. In that scenario, no RAM access is necessary. Maximizing a cache-hit frequency can therefore optimize a program's execution time by minimizing data transfer. It usually involves reordering elements in a large array by their relevancy with the neighbouring elements. The full picture is slightly more complicated, as programs running under most operating system can only access virtual memory addresses. The ordering of a dataset in a virtual address space does not necessarily be mapped to the physical memory addresses in a direct fashion. The condition can only be assumed to hold true to a good approximation.

MPMDS implemented an optimization scheme that attempts maximize cache-hit frequency. Arrays that contain atoms generalized coordinates are reordered according to the atoms position. The super-cell of the

simulation space is dissembled into a continuous square column, as shown in *figure 3.1*. The square column has a width of three times the force cut-off radius. Priority is assigned to each atom base on its position in the column. For example, atom positioned close to the start of the column will be reordered to the front of an array. This scheme allows related atoms be positioned closely in memory address. A reordering procedure base on priorities is equivalent to a sorting routine. Therefore, an efficient sorting algorithm known as quick-sort is employed for the job. Because quick-sort algorithm only involves comparing the relative priority of two data elements at a time, determination of an absolute priority is not needed prior to the sort. In MPMDS, arrays are reordered in this manner every two hundred simulation cycles. Because quick-sort is a highly efficient algorithm, the sorting time is negligible compare to the total runtime. An optimum reordering frequency is dependent on the rate at which atoms are disordered. When a system is presented a high degree of deformation or thermal motion, higher reordering frequency might be necessary. The performance of this optimization algorithm is tested against an x-y-z reordering scheme. In a particular system, the runtime for 500 simulation steps for each configuration is recorded as 305.8s and 362.2s respectively, which translate to a 18% gain in performance. A detail discussion on its benchmark is discussed in the next section.



3.2.5 Benchmark

The performance of MPMDS is examined by some simple benchmarking procedures. The array element reordering algorithm discussed in the previous section has also been tested during the benchmark. For validation purpose, benchmark results for LAMMPS running under comparable conditions has been

compared. LAMMPS is a popular and powerful molecular dynamics simulator that is freely available to the public. Its benchmark results is listed in its official webpage[4].

For benchmarking purpose, a 3.59 million atoms configuration was simulated with MPMDS under the EAM interaction potential. Both the force and the potential of every atom were evaluated at each time step while stresses were evaluated at every 500 steps. Simulation data were recorded to the hard drive at the same rate as the stresses were evaluated. Debugging modules and error handling protocols were turned off during the benchmark simulation. The program was executed on Intel® Xeon® 2.8 Ghz dual quad-core processors running in parallel. Each CPU was designated a 24GB of RAM. The CPU nodes are connected via QDR InfiniBand interconnect. The cluster runs on the CentOS 5.4 Linux platform. Intel® ifort compiler version 11.1.069 with the 'O2' optimization flag was used to compile MPMDS.

For validation purpose, LAMMPS benchmark results were used to rank the performance of MPMDS. Benchmark on LAMMPS with EAM interaction potential was test on Intel® Xeon® dual quad-core 2.66 Ghz processor clusters under Redhat Linux operating system. LAMMPS is programmed in C++ and was compiled with the Intel® icc compiler. Simulation was performed on some 32000×N atoms configurations where N is the number of processors running in parallel.

Benchmark results shows that MPMDS running with 3.59 millions atoms and 8 processors in parallel requires 305.8 seconds for 500 steps. For LAMMPS running with 32000 atoms on a single processor and 2.56 million atoms on 8 processors in parallel requires 11.58 second and 16.59 second respectively for 100 steps. These values can be compared by translating into a normalized parameter: $\frac{(runtime*CPU)}{(atom*step)}$ as illustrated follows:

	MPMDS (N=8)	LAMMPS (N=1)	LAMMPS (N=8)
Runtime*CPU/(atom*step) (Lower is better)	1.363×10^{-6} s	3.617×10^{-6} s	5.185×10^{-6} s
Ratio to the lowest	1.00	2.65	3.80

(3.2.1)

It is worth to note that MPMDS cannot run in single CPU mode. The benchmark result of MPMDS compares to that of LAMMPS running at 8 processors in parallel shows a 74% reduction in execution runtime.

Another set of benchmark was run to test the endurance of MPMDS performance. Two trials of simulation were executed for 50000 time step. First trial had the in-simulation array reordering function activated, while the second trial had this function turned off. A small section of the atoms were subjected to a constant strain rate. This strain rate induces a continuous deformation and subsequently introduces disorderliness to the dataset. Initially, the configuration of atoms was sorted in the x-y-z order. For example, atoms at the beginning of an array will have the lowest values in the *x* coordinate. This initial order is irrelevant when the in-simulation sorting function is on, as the configuration will be sorted before the first simulation cycle. The arrays will be rearranged every 200 cycles according to the continuous-column-decomposition scheme discussed in previous section.

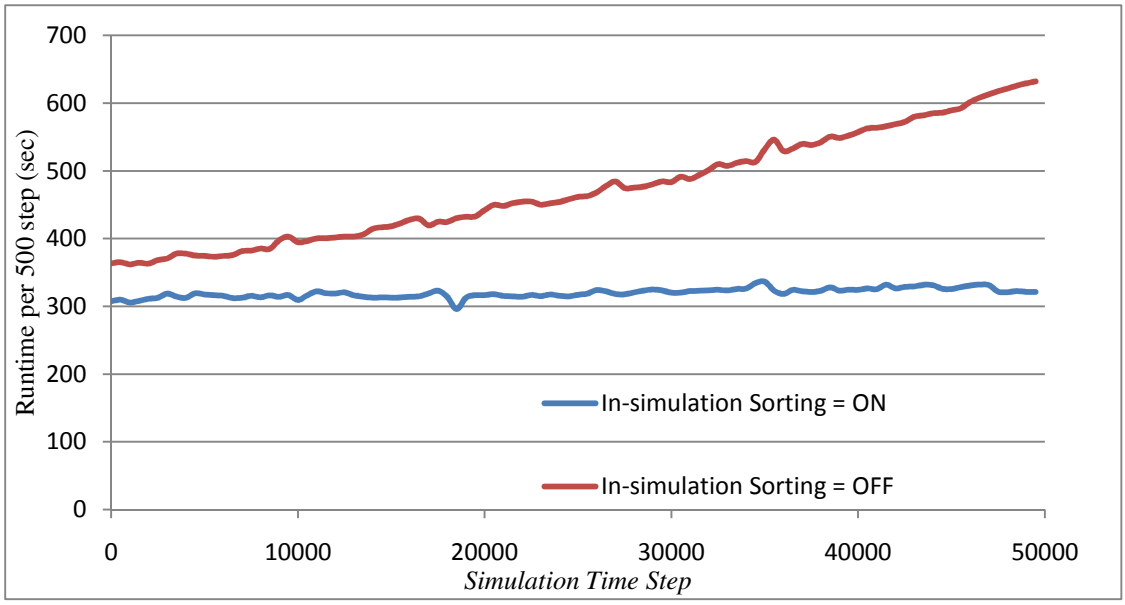


Figure 3.2 Benchmark on endurance of performance. Regularly maintaining the dataset by reordering its elements preserves performance.

Figure 3.2 illustrated the endurance of performance when in-simulation array sorting is on and when it is off. It is clear that performance steadily degenerate when the array elements are not properly maintained during simulation. Initial performance gain by sorting is around 18%. As simulation process, a total runtime was recorded as 31966 *s* and 47382 *s* respectively, equivalent to a 32.5% net gain in performance. Degeneration in performance is negligible when arrays are sorted constantly.

References

[1] Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *Journal of Computational Physics* **117**, 1-19(1995).

[2] Allen, M.P. & Tildesley, D.J. *Computer simulation of liquids*. (Clarendon Press: 1989).

[3] Kaplow, W.K. et al. Run-time reference clustering for cache performance optimization. *Proceedings of IEEE International Symposium on Parallel Algorithms Architecture Synthesis* 42-49(1997).

[4] Plimpton, S. LAMMPS Molecular Dynamics Simulator. *Sandia National Lab* (2007).at <http://lammps.sandia.gov/bench.html>

Chapter 4

Disintegration Dynamics of Dislocation Dipole in Aluminum

[1]*Harold Wing Hei Kwok,* [2*]*Kuiying Chen*

[1] *Department of Physics, University of Ottawa*
[2] *Structures & Materials Performance Laboratory, Institute for Aerospace Research National Research Council Canada, Ottawa*

Abstract

The disintegration process of dislocation dipoles in pure aluminum is studied via large scale molecular dynamics simulations. The investigation concerns a scenario where a series of dislocation dipoles pile up on a grain boundary (GB). It is found that the disintegration process involves an intricate series of dislocation-dynamics events. A variety of crystalline defects are produced during the disintegration, these include stacking fault disks, Lomer-Cottrell junctions, a system of stacking fault tetrahedrons (SFT) and rolls of vacancies. In particular, a batch SFT formation mechanism is observed, where a system of SFT is nucleated simultaneously, and that the group of SFT exhibit uniformity in both position and orientation. A thorough discussion on dislocation dynamics is presented. In addition, a high density of sessile defects is observed to be confined within the dipoles. It is therefore illustrated that the disintegration process constitutes a highly localized hardening mechanism.

Keywords: Dislocation dipole, annihilation, stacking fault tetrahedron, Lomer-Cottrell, Molecular dynamics simulation.

4.1 Introduction

The presence of dislocation dipoles is an essential feature in fatigue mechanisms. In particular, a high density of dipoles is found in ductile materials that are subjected to cyclic loading. These dipoles are bundled in the form of persistent slip band (PSB). The dipoles during a fatigue process are known to undergo annihilation, leaving a large amount of crystalline defects. Such phenomenon is identified as a key process in a fatigue crack initiation. Currently, several theoretical models for fatigue life prediction are built solely on the picture of dislocation dipoles piling up on a grain boundary (GB) and also on free surface. Therefore, the motivation of this study is to investigate the fine details of the disintegration process of dislocation dipoles at the atomic scale.

Gilman and Johnston[1] were the first to infer the presence of dislocation dipoles in fatigued materials. Their experimental study suggested that these dipoles are formed via a double cross-slip mechanism. Later, numerous studies echoed this finding and established an essential role of dislocation dipoles in fatigue crack initiation. On this topic, Christ[2] considered a scenario where dislocations, in the form of dipoles, are piled up on a grain boundary. Based on geometric arguments and the saturation vacancy concentration in a PSB, he established a theoretical model for the stress required for a PSB-GB crack formation. In regards to the nature of dipoles in fatigue samples, Essmann and Mughrabi[3] provided strong experimental evidences that these dipoles undergo annihilation. Their study on fatigued copper using transmission electron microscopy has demonstrated that edge-dislocation dipoles annihilate spontaneously by decomposition into '*invisible debris*'. In their later works[4], a comprehensive discussion was given on the role of edge dislocation annihilation in a PSB. In regards to the kinetics, they proposed an annihilation process caused by dislocation gliding. The net effect is to tilt the effective slip plane away from the crystallographic slip plane. However, Repetto and Ortiz suggested otherwise[5], they proposed that the same annihilation process can be explained by climb instead of glide. Other authors made effort to settle this controversy by suggesting their own annihilation mechanisms. Strunk[6], Quesnel[7] *et al* proposed glide mechanisms for edge dislocation annihilation in different crystal systems. In general, it is now accepted that dipoles annihilate via climb occurring at high temperature in distinction to the glide mechanism.

In regards to fatigue life, Tanaka and Mura[8] evaluated several scenarios where dislocation dipoles pile up on a GB or a surface. They formulated a fatigue life model based on the strain energy accumulation as dislocation dipoles pile-up. Their model and various other extended versions are widely adopted in explaining fatigue problems. These pioneering works have demonstrated the important role of dislocation dipoles in fatigue problems. However, the atomic process of dislocation dipoles accumulation and disintegration has not been investigated in detail. In particular, little is known about the collective dislocation dynamics when a series of dipoles are present.

Internal dislocation dynamics within a narrow dipole are within a few nanometres, it is therefore best to utilize atomistic simulation for the investigation. Most previous numerical studies on this topic focused on specific dislocation dynamics and considered a single dipole system[9,10]. Different from that approach, here we present a study that utilizes large scale molecular dynamics in a wider perspective by incorporating a group of dislocation dipoles in a pile up scenario. It is intended to illustrate the disintegration dynamics of dipoles in a collaborative environment, where the process is initiated by stress concentration.

4.2 Methodology

A massively parallelized molecular dynamics simulation code is developed for the purpose of this study. The program employs three-dimensional Newtonian molecular dynamics algorithm as the simulation platform. Temperature is controlled via kinetic energy rescaling at every 500 steps. The atomic trajectory is evaluated by the 5th order Gear predictor-corrector algorithm[29]. Sample sizes of up to 3.4 millions of aluminum atoms and a dimension of 374Å × 360Å × 434Å are tested. The thicknesses in the z direction are chosen to satisfy the periodic boundary condition. Atoms are arranged into a bi-grain system with relative grain orientation extending a periodicity of a *coincident site lattice* (CSL). The grain boundary type is selected to minimize dislocation transmission. A $\Sigma 11(\bar{7}4\bar{1})$ asymmetric grain boundary is generated with a rotation axis oriented away from the grain boundary plane and its normal. The CSL system can be fully described by the rotational matrix[22], **R**:

$$\mathbf{R} = \frac{1}{11}\begin{bmatrix} 6 & -6 & 7 \\ -9 & -2 & 6 \\ -2 & -9 & -6 \end{bmatrix} = N^{-1}[a_{ij}]. \quad (4.1)$$

From this matrix, the rotation axis, $\boldsymbol{c}$, and the rotation angle, θ can be evaluated by[22]:

$$\boldsymbol{c} = \{a_{32} - a_{23}, a_{13} - a_{31}, a_{21} - a_{12}\}, \quad (4.2)$$

$$2\cos(\theta) + 1 = N^{-1}(a_{11} + a_{22} + a_{33}), \quad (4.3)$$

giving $\boldsymbol{c} = \{-5, 3, -1\}$ and $\theta = 82.2°$. This configuration provides both a tilt and a twist component to the grain boundary and therefore prevents slip systems to adjoin continuously. The arrangement introduces slip system incompatibility and discourages dislocation transmission.

Interactions between atoms in the simulation are governed by a many-body interatomic potential. The form of the potential is of type embedded atom method (EAM) and is described by[21]:

$$E_i = F_i(n_i) + \frac{1}{2}\sum_{j \neq i} V_{ij}(r_{ij}), \quad (4.4)$$

where F_i is the energy required to embed an atom into a local net charge density, n, which is the sum of charge density from each neighbouring atoms, namely: $n_i = \sum \rho_j(r_{ij})$ and V_{ij} is a pair potential describing the direct interaction between two atoms. Before each simulation, the initial configuration of atoms is relaxed according to this energy functional. The conjugate gradient method is employed at 0K for 1000 steps. The given relaxation time is found to be sufficient to relax and stabilize the total energy of the system.

An EAM potential constructed from matching the bulk properties of a material does not guarantee its correctness in describing atomic trajectories. To ensure representability of the dislocation dynamics, the simulation implemented an EAM potential[21] that was developed under the 'force-matching' scheme. This scheme utilizes first-principle numerical calculations, solving electron density in a matrix to evaluate the atomic force experienced by an atom. The force is then translated into an energy potential in the EAM model. This approach is supplementary to the EAM method. It provides a reasonably good description for both the atomic behaviour and the bulk properties. Such force-matched EAM potentials have been applied to the study of dislocation dynamic problems by molecular dynamics simulations and demonstrated a strong validity[21,23].

In the current atomistic simulation techniques, the number of atoms considered in a system is rather limited which the number of atoms has no comparison to the real-life systems. It is therefore critical to evaluate boundary conditions. Steps have been taken to verify the validity of the conditions applied and to eliminate possible artifacts due to size limitation. For the $\pm x$ and $-y$ faces, i.e. the planes with outward normal in the $\pm x$ *and* $-y$ direction (figure 4.1), both fixed and free conditions were tested. These two conditions were selected because they represent the two extreme scenarios. Under real-life condition, the simulated region should be surrounded by other atoms. Assuming the simulated region interacts with the surrounding elastically, one would expect its behaviour lies between the two extremum. The two conditions were examined and no phenomenological differences were found. Some small deviations in quantitative observables like energy and stresses were measured. The differences were found to be less than 2% close to the boundary and 0.1% in the region of interests. These differences would be considered small and negligible. The small difference in stress implies a weak long-range interaction for the dipoles away from the slip planes. Size and boundary condition in z direction is more critical as the z faces always coincide with the region of interest and across most of dislocation lines. Multiple thicknesses were tested under both periodic and free boundary condition. The thickest configuration has a thickness of 434 Å and is four times that of the thinnest. Adjusting the thickness would induce possible size dependent artifact, which if observed, would invalidate the simulation. But for the configuration considered in the present study, no phenomenological differences were observed.

Our simulations were performed under the canonical ensemble with a controlled temperature of 5K. The low temperature is chosen to isolate any thermally-activated processes. Processes such as vacancy diffusion and dislocation climb are not of the focus in this study. Nevertheless, we believe these are important processes in high temperature fatigue mechanisms, and should be dealt with in a separate investigation.

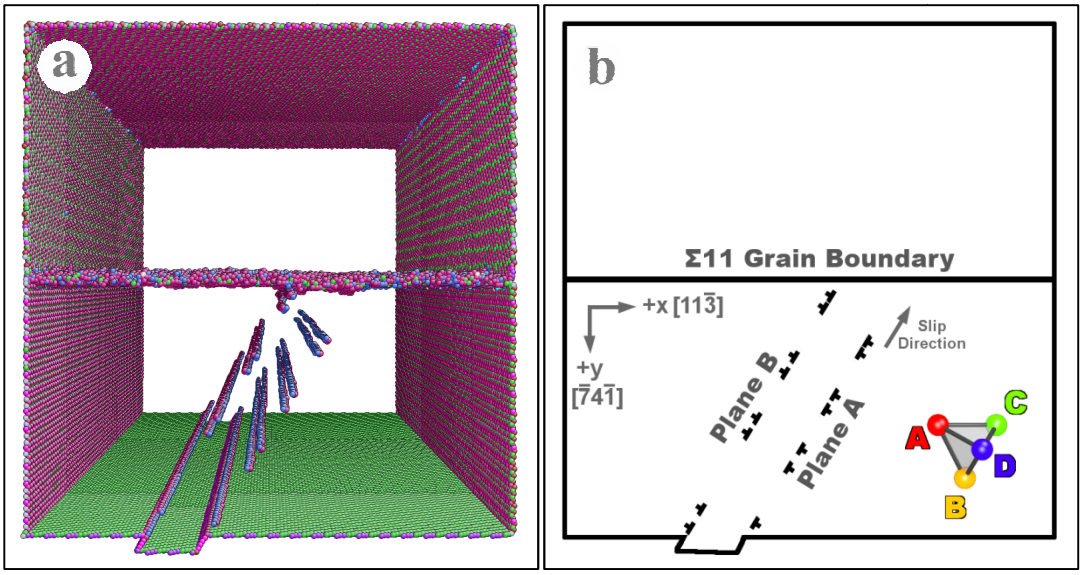


Figure 4.1: Dislocations in dipole formation marching toward the grain boundary. (a) coordination number coloring (b) schematic illustration

4.3 Results and Discussions

4.3.1 Interaction between Dislocation Dipoles and Grain Boundary

To simulate a dislocation dipole, edge dislocations are generated on two slip planes *in-situ* during the simulation. A small section of the atoms along the edge of the system is subjected to a constant strain rate of 0.5mÅ/fs in a direction antiparallel to the slip vector [10$\bar{1}$]. This creates pairs of edge dislocations on the slip planes forming two dislocation walls (figure 4.1). The dipole width is chosen to be 4.55nm. This width, also refer to as the height by some authors[11,25], is consistent with experimental observation on aluminum[11]. The configuration resembles edge dislocation dipoles of vacancy type, defined by having extra half planes on the outward normal direction (figure 4.1b). These edge dislocations are unstable in FCC system and are immediately dissociated upon creation. The dissociation follows the reaction equation, $BC \rightarrow B\alpha + \alpha C$, in terms of Thompson tetrahedron notation, and yields two Shockley partials from each perfect edge dislocation. The pair of partials is connected by a stacking fault. This configuration is sometimes referred to as a stacking fault ribbon. The separation distance between a pair of partials is governed by the stacking fault energy between them. As the stacking fault energy in aluminum is high, the dissociated partials are bounded together closely. An average separation of 12±3Å is measured in

our numerical simulation, which is consistent with first principle calculations and experimental observations by other researchers[12]. These Shockley partials are free to glide on their slip planes and are driven toward the GB by a shear stress. The shear stress is induced by the constant strain rate introduced to the system. These dissociated edge dislocations are primitively assembled in a zigzag pattern along the dipole while gliding toward the grain boundary, as illustrated in *figure 4.1*. Such formation is a result of the attraction from the adjacent plane and the mutual repulsion within the same plane[24]. Consequently, dissociated edge dislocations arrive at the GB in an alternating manner, with dislocations on plane A arriving first. Dislocations that first reached the GB are absorbed. After absorption of several dislocations, back stress start to build up, which impedes further dislocation motion. As a result, ribbons of dislocations are then piled up along the two dislocation walls.

In our simulation, it is found that a few consecutive pairs of partial dislocations were absorbed by the GB before reemission begins. During the reemission process, an array of partials with Burgers vector, γB, nucleate at the GB. *Figure 4.2* illustrates the array of partial dislocations being emitted. Each Shockley partial forms as stacking fault disk and emerging from the grain boundary. Subjected to a strong bowing force, the disks expand radially. The sideway expansion ceases when its perimeter reaches the two dislocation walls. However, these disks continue to expand in the *BD* direction and sweep along the trailing dipoles. As disks continue to expand, a second partial dislocation arc is bow out at the location where the first disk nucleated. The second dislocation arc, having a Burgers vector of $A\gamma$ is complementary to the first disk; removing the stacking faults within its perimeter. *Figure 4.3* demonstrated the unfaulting mechanism. The second dislocation arc, behaving similarly to the first arc, expands radially. While both arcs are expanding anisotropically, as being bounded by the dislocations walls, the second loop cuts the stacking fault disk into two uneven parts. The larger parts continues to sweep along the dipole, crossing and interacting with stacking fault ribbons on the wall while the smaller part is partially being re-absorbed by the GB and becomes stationary for the rest of the simulation time.

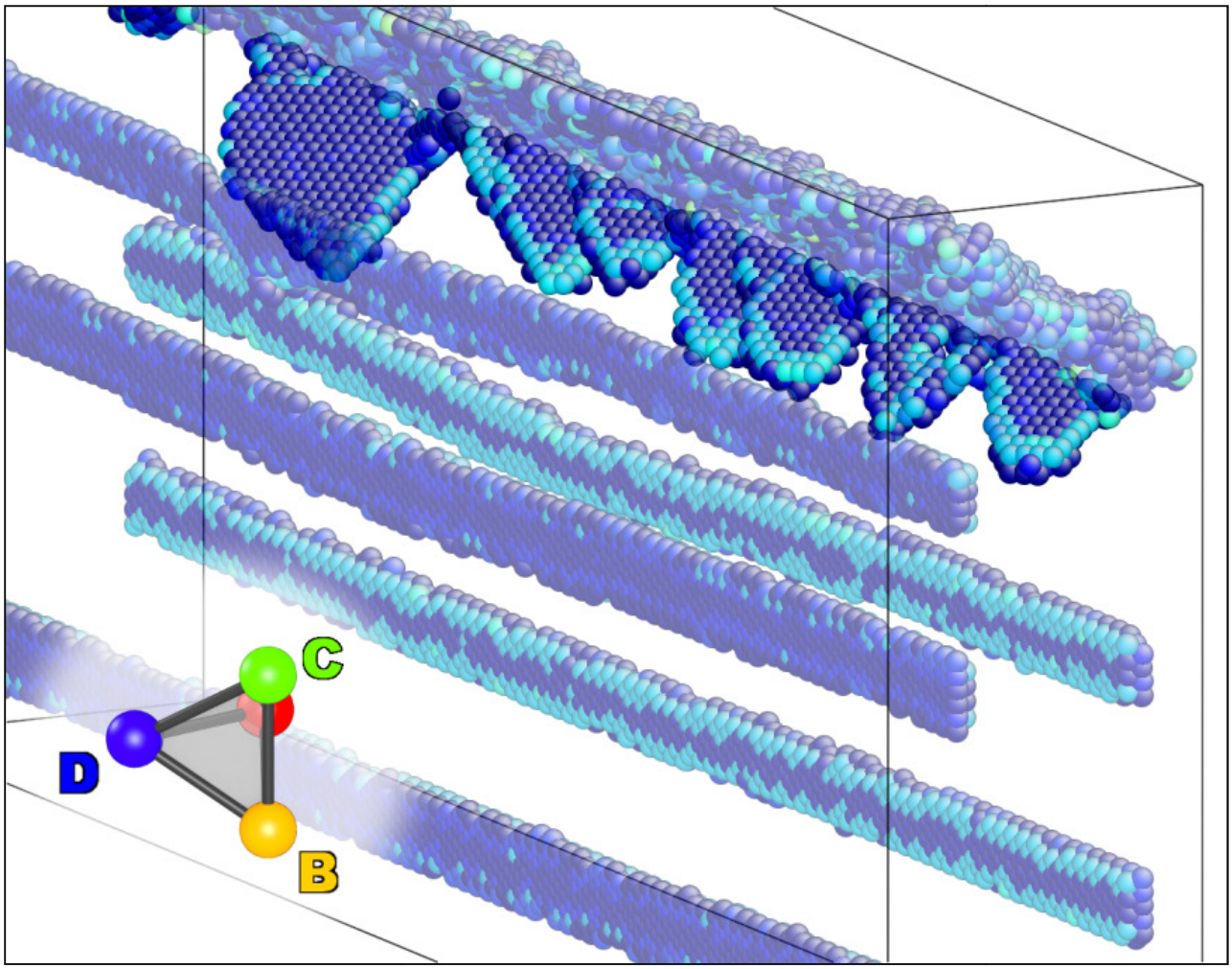


Figure 4.2: Stacking Fault Disks being emitted at the grain boundary. Color in central symmetry parameter.

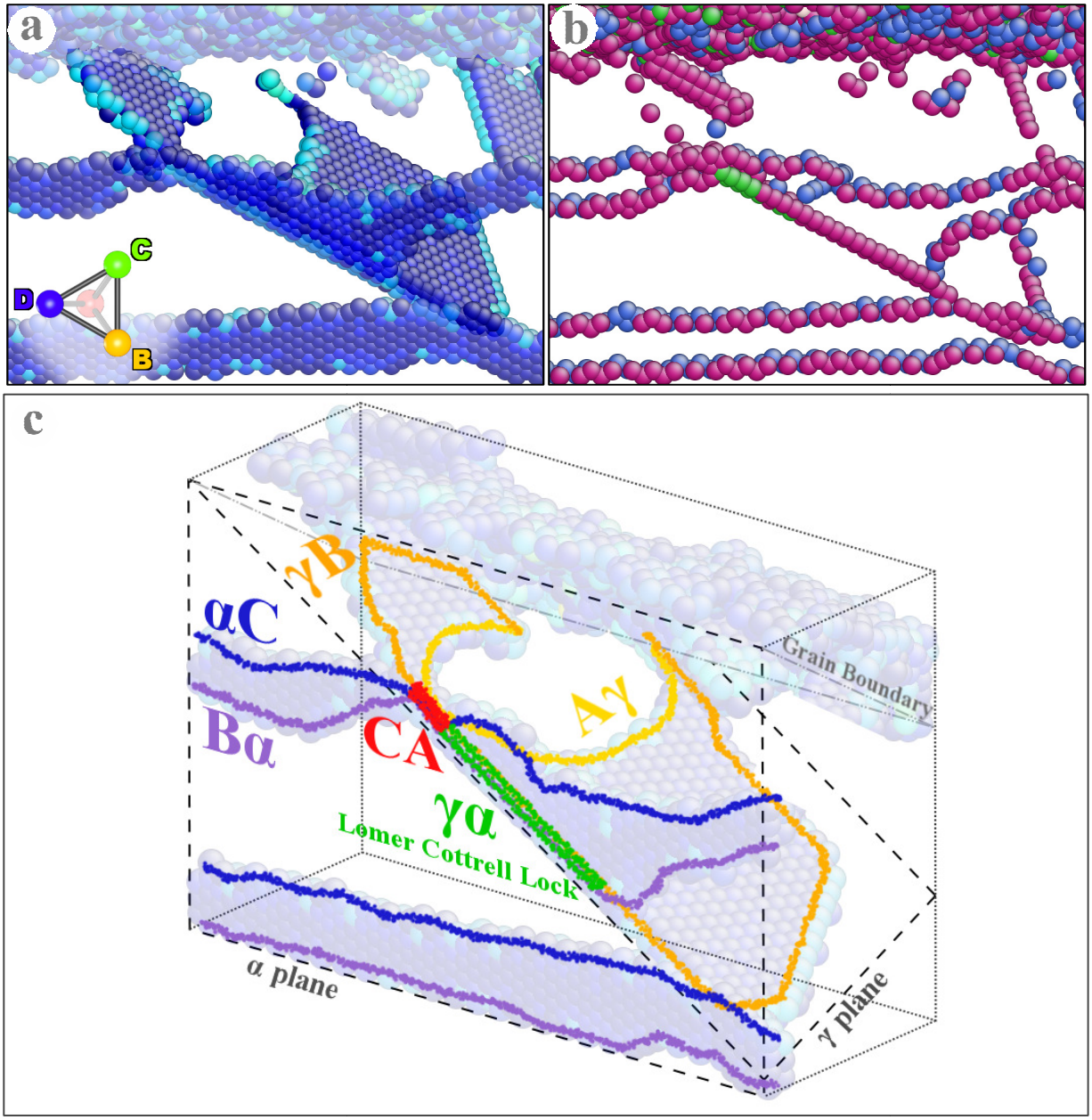


Figure 4.3: A stacking fault disk emerges from the grain boundary and is intersecting a stacking fault ribbon (disassociated edge dislocation). A row of partial point defects nucleate along dislocation CA. Their presence is illustrated by green atoms next to the vacancies under coordinate number coloring. (a) Central symmetry coloring, (b) coordinate number coloring, and (c) Schematic illustration [*tilted view*]

4.3.2 Nucleation of Vacancy Clusters and Lomer-Cottrell Locks

The stacking fault disks that sweep along the dipole are found to cross several dislocations ribbon. They react to the trailing partials on the ribbon according to the equation: $\gamma B + B\alpha \rightarrow \gamma\alpha$ (figure 4.3). The resultant Burgers vector is not on any FCC slip planes, the dislocation is therefore sessile and cannot glide in any direction. This structure is identified as a Lomer-Cottrell junction or Lomer-Cottrell locks as some authors prefer. It is stable and immobile at low temperature and it can only transform or translate via climb or dissociation (unzipping motion)[16]. At the early stage of annihilation process, these junctions are found to be aligned in parallel with the CA direction and they are distributed throughout the dislocation walls. The formation of these Lomer-Cottrell junctions has two eventual consequences: firstly, it provides an anchor where other dislocations can meet, this results in the nucleation of point defects. Secondly, it distorts the dislocation line of piled-up partials, which encourages the formation of stacking fault tetrahedrons.

In regards to vacancy nucleation, it is found that multiple vacancies nucleate into a roll formation along the Lomer-Cottrell junction. Each reactant partial of the sessile dislocation has an unfaulting counterpart which is trailing behind. Under the influence of local stress, short sections of these trailing partials meet at the junction. These trailing partials are unable to cross the sessile junction. As a result, they conjoin together. The conjoined configuration constitutes a large junction structure. The net dislocation reaction is $A\gamma + \gamma B + B\alpha + \alpha C \rightarrow CA$. This reaction resembles a partial annihilation process with a net reduction in Burgers vector. The reduction in magnitude of Burgers vectors manifest in some vacancy spaces. Each vacancy space is slightly smaller than a normal vacancy in volume and can be considered as a partial point defect[20]. It is observed that these vacancy spaces align in a roll formation of up to 8 atomic sites in length, as illustrated in Figure 4.4. It is found that most of these rolls of vacancy persist throughout the simulation and retain their form. As simulation proceeds, some of these vacancy rolls are found to enrol in further dislocation reaction and annihilated into some rolls of whole vacancy.

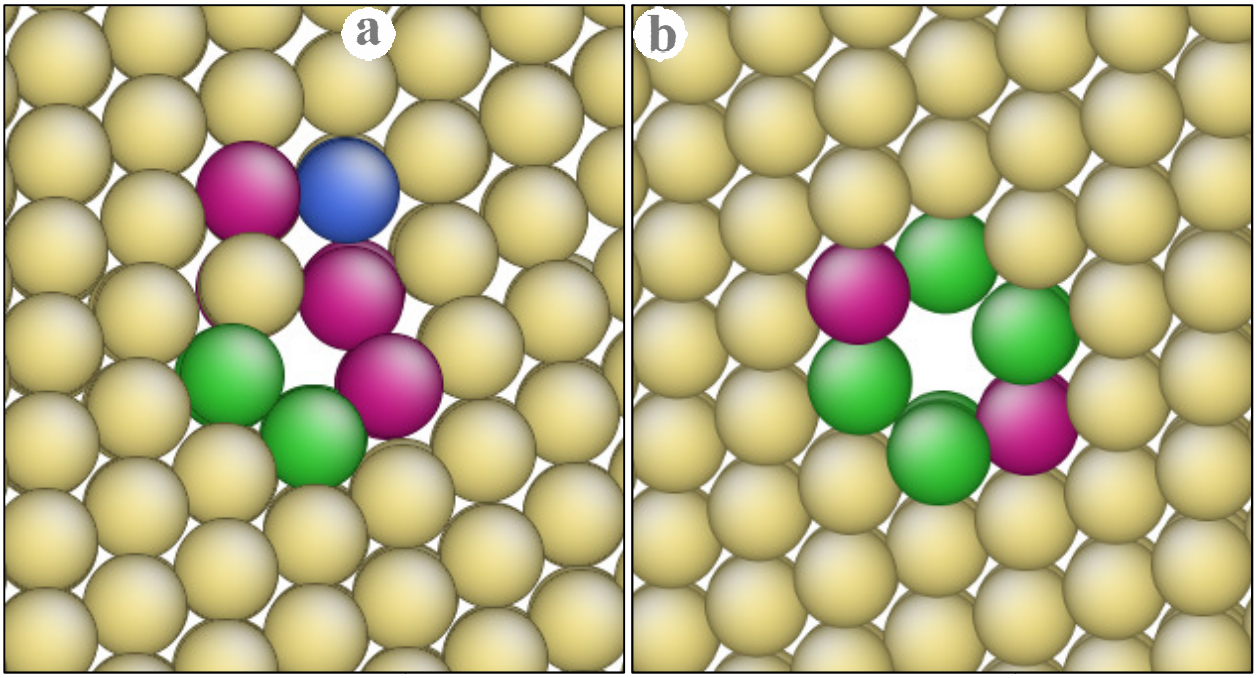


Figure 4.4: (a) A roll of vacancy type partial point defects. (b) A roll of whole vacancies.

4.3.3 Batch Formation of Stacking Fault Tetrahedra

In addition to vacancy nucleation, our simulation also reveals a batch formation process for the stacking fault tetrahedron (SFT). Unlike the mechanisms suggested previously[13,14,15], our proposed model does not involve vacancy clusters aggregation, nor the formation of Frank partial loops. Our simulation is specific to dipole annihilation scenarios where multiple parallel stacking fault disks form as a result of dislocation dipole-GB interaction. It involves the interaction between the arrays of reemitted stacking fault disks and the stacking fault ribbons piled-up on the dislocation walls. It is found that a large number of SFT can be nucleated simultaneously through this process. As illustrated in figure 4.5, the resulting SFTs are positioned systematically and share a common orientation.

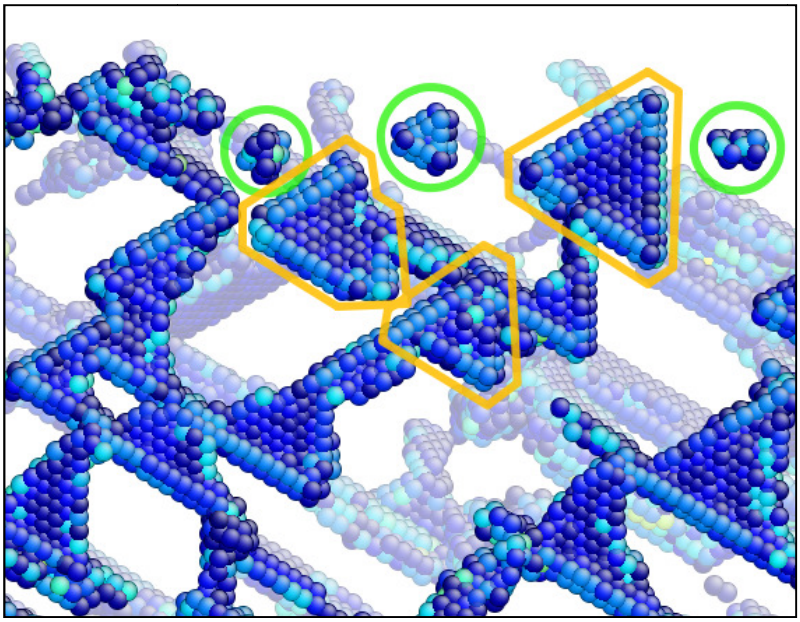


Figure 4.5: Front view of a dislocation wall (coloured in central symmetry parameter), illustrating a system of triangular stacking fault structures. Majority of these triangular structures are connected with each other at their apex. Some of these structures are closed (circled in orange), forming some completed stacking fault tetrahedrons. Circle in green are some collapsed structures manifested as vacancies.

The Lomer-Cottrell junction discussed previously has a Burgers vector of $\gamma\alpha$, which is not perpendicular to any of the partials on the ribbon. When the junction bisects the ribbon, it interacts elastically with partials on the two sides in an opposite manner. As illustrated in *figure 4.3*, on the right side of the ribbon, the sessile dislocation drives the two partials away, enlarging the stacking fault region. On the other side, partials are being driven together. The distorted dislocation lines are under the influence of Peierls potential, favouring the alignment on the edges of the Thompson tetrahedron. As a result, $B\alpha$ and αC partials are forced to align with the *CD* and *BC* direction respectively. In the process, local competitions take place between junctions on the same ribbon. Through maintaining the *CD* alignment, some junctions are pushed downward via an zip-unzip mechanism discussed previously by Rodney[16], Hussein[17], *et al*. Consequently to all these dynamics, a network of aligned triangular stacking fault plates is created. These orderly situated triangular plates and their connected stacking fault disk on the γ plane will become two planes on stacking fault tetrahedrons.

Later in this formation process, upon triangular plates acquired their equilateral shapes and stabilized, the third SFT plate nucleates via a spontaneous emission of a Shockley partial on the *CD*

edge, initiated either from the *C* or *D* apex, figure 4.6e. The emission is equivalent in principle to the cross slip mechanism proposed by Fleischer[18]. It is suggested that a Shockley partial can cross slip at a slip plane intersection via degenerating into a sessile dislocation and another Shockley partial. In our case, the new partial glides from the *CD* edge to the *AD* edge, completing the third plane of an SFT. As Kadoyoshi *et al* pointed out[19] that there exists a critical size at which SFT is stable. In our simulation, it is found that those SFTs with a triangular base less than 32 atomic sites were spontaneously collapsed into vacancies. Those clusters are found to collapse within the dislocation walls. Those that are capable of forming the third plates on the β plane are stable, some of them remain open or partially open throughout the simulation. Some SFTs are able to complete forming the last face at the later period of the simulation, again via spontaneous partial emission at the *C* apex. The last plate of the SFTs remains difficult to close due to the fact that this plate is perpendicular to the strain direction. This orientation receives minimal resolved shear stress, and therefore the process relies on the assistance of local stress state.

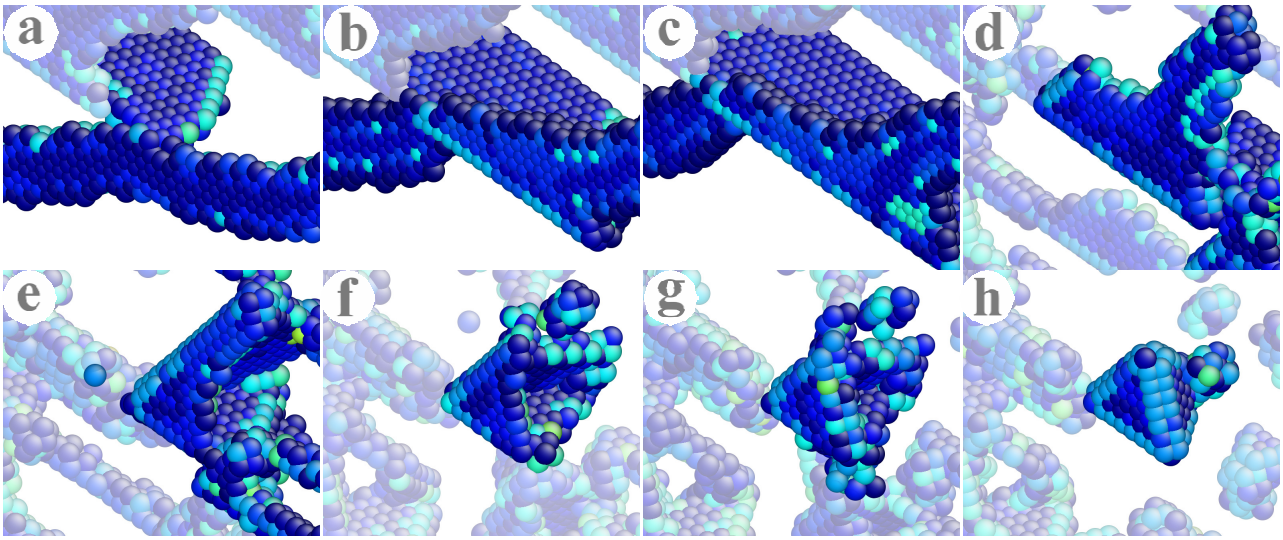


Figure 4.6: Stacking Fault Tetrahedron formation process

a)Stacking fault disk emerging to an stacking fault ribbon (dissociated edge dislocation). b) Stacking fault disk bisected the ribbon. c) Elastic interaction distorded the shape of the ribbon, expanding the stacking fault on the right while abridging the stacking fault on the left. d) Residue stress and perierl force aligned sections of the partial dislocations with the <110> direction. e) Third stacking fault plate is nucleated from the top. The net structure resembles an opened stacking fault tetrahedron. f) Stacking fault plates recede and have their partials aligned with <110> directions. g) A partial cross-sliped to the δ plane, leaving an Lomer-Cottrell junction on the BC edge. h) The cross-sliped partial closed the structure, forming a closed SFT with some vacancies attached onto its A-apex .

The observed SFT formation mechanism has several implications. Firstly, the α plane of a SFT must resides on a dislocation wall. This creates a bias in sessile dislocation density toward the two planes as expected, as a dipole has its dislocations entirely on those two dislocation walls before the disintegration. However, unlike the original Shockley partials, these dislocations are sessile, stable and irreversible. As such, their presence effectively diminishes the shear-ability of these two planes. Secondly, this mechanism illustrates the role of SFT formation as an intermediate step for vacancy nucleation. It is known that vacancies can form during annihilation of a dipole. However, the mechanism is poorly understood, in particular to the question on how two line defects annihilate to form point defects. This simulation illustrated that a stacking fault ribbon is dissected into discrete triangular pieces in the process of SFT formation. As such, creation of these SFT structures become a necessary intermediate step for the formation of discrete vacancy clusters.

4.3.4 Defects Evolution

The evolution of a vacancy dipole, from dislocation pile-up to disintegration, can be realized in terms of the change in potential energy. Figure 4.7 plotted the average potential energy (per atom) of the foremost section of the dipole as the simulation proceeds. The graph exhibits distinctive stages. The two peaks showed in stage-Ø are the signature of by-passing of two edge dislocations. Following is the stage-I when dislocation pile-up begins. The potential energy increases steadily as the number of trailing dislocations accumulates. At this stage, plastic deformation near the grain boundary is refrained as dislocations on the same glide plane are mutually repulsing, confining the position of each dislocation. Upon reaching some critical value, denoted as stage-I(b), the secondary slip plane is activated and an array of stacking fault disk is emitted at the grain boundary. However, the energy does not relax until the stacking fault disks interact with other dislocations on the dipole. Through visualization, it is found that the onset of stage-II is marked by the formation of Lomer-Cottrell locks. Stage-II is the one at which stacking fault ribbons are being distorted into a network of triangular structures. It also includes the formation of vacancy rolls. This stage depicted a major energy relaxation mechanism as the energy drop in this stage cover the dominated portion of energy being relaxed. It is observed that this stage is rapid and short in reference to the complexity of the dislocation dynamics occurred in this interval. At the end of this period, slip pathways inside the dipole are mostly destroyed with the consequence of dislocation motions being

suppressed. The end of major dislocation motion marks the onset of stage-III. This stage is characterized by local atomic rearrangement. The formation of complex defect structures in the previous stage introduces an intricate and localized stress state. Local competition between defect structures dominates the dynamics of this stage. Only small range atomic motions, like the closing of the last face of SFT and the collapse of SFT into vacancies, are observed in this stage. Changes occur slowly because all motions are driven by localized stress, and that the structures are in local competition. This is manifested in the plot as a steady and slow decrease in potential energy. The entire simulation was performed at a low temperature with minimal thermal motion. It is however reasonable to expect that thermal motion at higher temperature can speed up such atomic rearrangements motion.

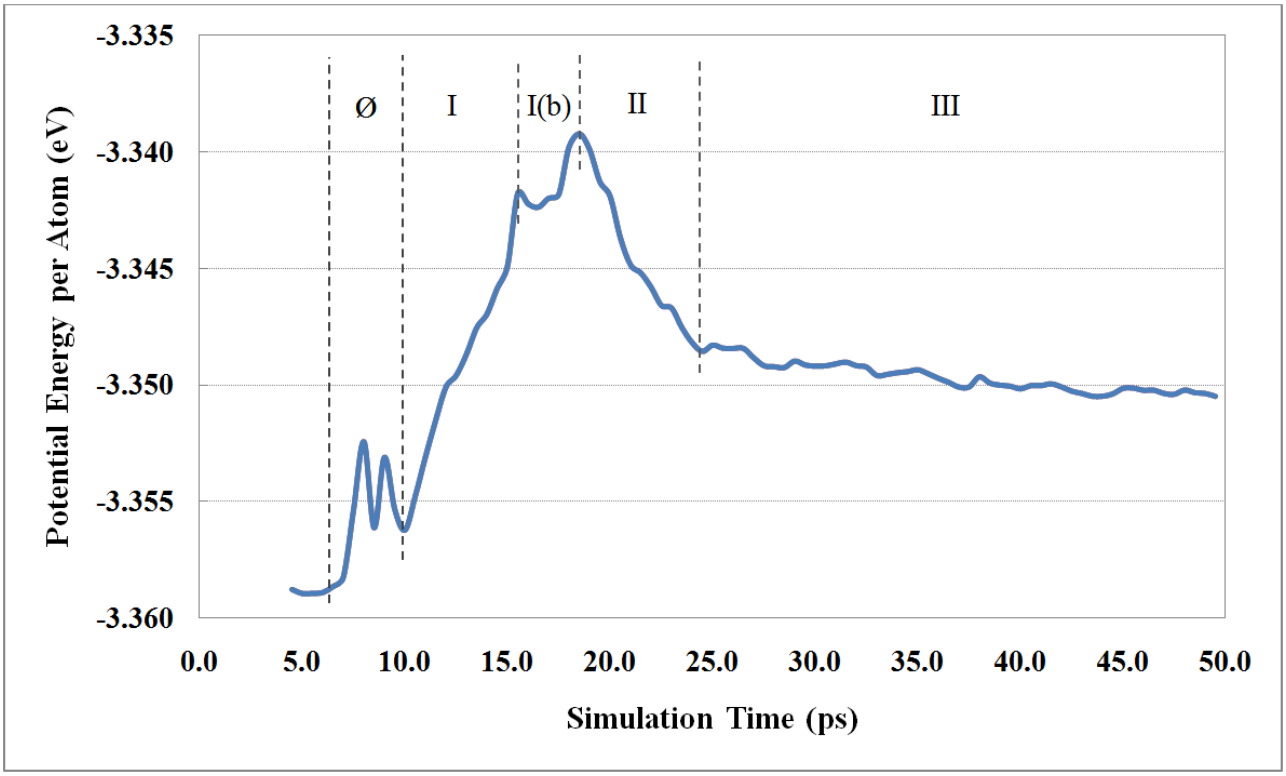


Figure 4.7: Evolution of the disintegration process.

4.3.5 Localized Hardening

It is shown in the simulation that a disintegrated vacancy dipole is composed primarily of vacancy clusters, vacancy rolls, SFT and sessile dislocations. In particular, a high density of sessile Lomer-Cottrell junctions is reside on the two dislocation walls through SFT formation, figure 4.5. These defects are incapable of gliding under shear conditions. Consequentially, their presence impedes further dipole pile-up. As discussed in the references[20,27,28], the existence of sessile defects is inferably a substantiation of work-hardening. As a vacancy dipole disintegrates, these sessile defects are utterly confined within the two dislocation walls, and thus define a work-hardening mechanism that is highly localized. When bundles of these hardened dipoles are present, as in the case of a persistent slip band, it is presumable that the collective stress-strain response will significantly differentiate from the surrounding matrix. Subsequently, the bundle can effectively concentrate stress at the present of external load, providing an essential condition for crack nucleation.

4.4 Conclusion

In summary, when vacancy dipoles pile-up at a GB, annihilation of dipoles can be initiated by a grain boundary via emission of partial dislocations. It is found that the reemitted partials interact with stacking fault ribbon on the dislocation walls to form sessile Lomer-Cottrell junctions. Some sections of these junctions are found to undergo further dislocation reaction, and nucleate rolls of partial point defects and vacancies, while some other sections nucleate into stacking fault tetrahedra. An SFT nucleation process is observed at which it involves the nucleation of a batch of SFT on the dislocation walls simultaneously. These SFT exhibit uniformity in both position and orientation. Some of these SFTs are found to collapse into vacancies inside the dislocation walls. The high concentration of sessile dislocations and vacancies generated will impart a highly localized hardening mechanism.

Acknowledgement

This research is supported by DND-TIF project: A Coupled Atomic-Meso-Macroscopic Modeling of Cracks in Aircraft Components.

References

[1] Gilman J.J. & Johnston W.G, Dislocation multiplication in lithium fluoride crystals. *Journal of Applied Physics*, **31**, 687 (1960).
[2] Christ, H. On the orientation of cyclic-slip-induced intergranular fatigue cracks in face-centered cubic metals. *Materials Science and Engineering: A* **117**, L25–L29(1989).
[3] Essmann, U. & Mughrabi, H. Annihilation of dislocations during tensile and cyclic deformation and limits of dislocation densities. *Philosophical Magazine A* **40**, 731-756(1979).
[4] Essmann, U., Gösele, U. & Mughrabi, H. A model of extrusions and intrusions in fatigued metals I. Point-defect production and the growth of extrusions. *Philosophical Magazine A* **44**, 405-426(1981).
[5] Repetto, E. & Ortiz, M. A micromechanical model of cyclic deformation and fatigue-crack nucleation in fcc single crystals. *Acta materialia* **45**, 2577–2595(1997).
[6] Strunk, H. & Strecker, H. Annihilation of edge dislocation dipoles by a glide mechanism. *Materials Science and Engineering* **62**, 265-269(1984).
[7] Quesnel D.J. & Tsou J.C. Glide Induced Mechanism of Dipole Loop Annihilation. *Scripta Metallurgica* **14** 935 (1980).
[8] Tanaka, K. & Mura, T. A dislocation model for fatigue crack initiation. *Journal of applied mechanics* **48**, 97(1981).
[9] Vegge, T. & Jacobsen, K.W. Atomistic simulations of dislocation processes in copper. *Materials Research* **14**, 2929-2956(2002).
[10] Olmsted, D.L. et al. Atomistic simulations of dislocation mobility in Al, Ni and Al/Mg alloys. *Modelling and Simulation in Materials Science and Engineering* **13**, 371-388(2005).
[11] Kassner, M.E. & Wall, M.A. Microstructure and mechanisms of cyclic deformation in aluminum single crystals at 77 K: Part II. Edge dislocation dipole heights. *Metallurgical and Materials Transactions A* **30**, 777-779(1999).
[12] Woodward, C. et al. Prediction of Dislocation Cores in Aluminum from Density Functional Theory. *Physical Review Letters* **100**, 1-4(2008).
[13] Jøssang, T. & Hirth, J.P. The energies of stacking-fault tetrahedra in f.c.c. metals. *Philosophical Magazine* **13**, 657-670(1966).
[14] Loretto, M.H. The nature of faulted defects in low stacking-fault energy alloys. *Philosophical Magazine* **12**, 125-137(1965).
[15] Silcox, J. & Hirsch, P.B. Direct observations of defects in quenched gold. *Philosophical Magazine* **4**, 72-89(1959).
[16] Rodney, D. & Phillips, R. Structure and Strength of Dislocation Junctions: An Atomic Level Analysis. *Physical Review Letters* **82**, 1704-1707(1999).
[17] Zbib, H. 3D dislocation dynamics: stress–strain behavior and hardening mechanisms in fcc and bcc metals. *Journal of Nuclear Materials* **276**, 154-165(2000).
[18] Fleischer, R. Cross slip of extended dislocations. *Acta Metallurgica* **7**, 134–135(1959).
[19] Kadoyoshi, T. et al. Molecular dynamics study on the formation of stacking fault tetrahedra and unfaulting of Frank loops in fcc metals. *Acta Materialia* **55**, 3073-3080(2007).
[20] Buehler, M.J. et al. The dynamical complexity of work-hardening: a large-scale molecular dynamics simulation. *Acta Mechanica Sinica* **21**, 103-111(2005).
[21] Liu XY, Ohotnicky P, Adams J, Rohrer C, Hyland R. Anisotropic surface segregation in Al-Mg alloys. *Surface science*, **373**(2-3), 357–370(1997).
[22] Grimmer H, Bollmann W, Warrington DH. Coincidence-site lattices and complete pattern-shift in cubic crystals. *Acta Crystallographica Section A*. **30**(2):197-207 (1974).

[23] Dewald, M.P. & Curtin, W. a Multiscale modelling of dislocation/grain-boundary interactions: I. Edge dislocations impinging on Σ11 (1 1 3) tilt boundary in Al. *Modelling and Simulation in Materials Science and Engineering* **15**, S193-S215(2007).
[24] Antonopoulos, J.G., Brown, L.M. & Winter, A.T. Vacancy dipoles in fatigued copper. *Philosophical Magazine* **34**, 549-563(1976).
[25] Kroupa, F. Annihilation of a dislocation dipole. *Czechoslovak Journal of Physics* **17**, 220-226(1967).
[26] Li, J. AtomEye: an efficient atomistic configuration viewer. *Modelling and Simulation in Materials Science and Engineering* **11**, 173-177(2003).
[27] Zbib, H. 3D dislocation dynamics: stress–strain behavior and hardening mechanisms in fcc and bcc metals. *Journal of Nuclear Materials* **276**, 154-165(2000).
[28] Kimura, H., Maddin, R. & Kuhlmann-Wilsdorf, D. Quenched-in vacancies in noble metals–II mechanism of quench hardening. *Acta Metallurgica* **7**, 154–162(1959).
[29] Allen, MP & Tildesley DJ. Computer Simulation of Liquids. *Oxford University Press* (1989).

Chapter 5

Triple Junction Deformation Mechanism and related Dynamic Recrystallization Phenomena

[1]*Harold Wing Hei Kwok,* [2]*Kuiying Chen*

[1] *Department of Physics, University of Ottawa*
[2] *Structures & Materials Performance Laboratory, Institute for Aerospace Research National Research Council Canada, Ottawa*

Abstract

The deformation mechanism of grain boundary triple junction is investigated via numerical atomistic simulation. A symmetrical triple junction system constructed of 4 millions aluminum atoms is studied. A mode II deformation is introduced to the system by an incremental shear strain. Simulation results illustrated that deformation at a triple junction is associated with a rich series of dynamics in a specific order. Phenomena include local amorphization, shear-coupled grain boundary migration and dynamic recrystallization are recognized in the simulation. In particular, the recrystallization process of some nano-sized grains is observed while the junction is being deformed. This article provides a discussion and a detail description on the atomistic picture of a triple junction deformation mechanism.

Keywords: triple junction, amorphization, dynamic recrystallization, grain boundary migration

5.1 Introduction

In a polycrystalline material, the region where a line of atoms is shared by three adjacent grains is refers to as a triple junction (TJ). It is a termination border of a grain boundary plane. Its properties are of interest in a wide field of material research. In particular, its physical property is strongly related to the strength of a material. In this context, the deformation mechanism is an important process to investigate.

Grain boundary sliding and migration motions are widely observed in deforming metallic materials. However, the crystalline characteristic of triple junction is not compatible with such motions. For this reason, triple junction is expected to deform as grain boundary motion proceeds. It is commonly understood that its deformation is closely related to the hardening and crack nucleation mechanism. For instance, it is found that the junctions are preferential crack nucleation sites under creep deformation[1], high cyclic fatigue loading (HCF) or low cycle fatigue loading[2, 3]. Numerous experimental investigations were reported in regards to the relationship between triple junction deformation and crack nucleation. Kobayashi *et at.* studied the hardening of triple junction in molybdenum polycrystals with nanoindentation techniques [4]. He concluded that, amongst all examined triple junctions, those that contain coincident site lattice (CSL) or low-angles grain boundaries have a lower degree of hardening under plastic deformation. And by examining the cracks characteristics of the specimen, he further illustrated that triple junctions compose of random boundaries are the preferential sites for crack nucleation and propagation. In a separate study, he also concluded similar findings for the polycrystalline aluminum [5], and thus illustrated that triple junctions are some preferred crack sites for both ductile and brittle metallic materials. In regards to the deformation process, Walter's [6] experimental observations on high purity aluminum provided a microscopic picture of grain boundary motions and the accompanying triple junction deformation. Besides experimental approaches, theoretical models were also constructed to relate triple junction deformation with crack initiation. Williams [1] derived a model for creep life based on triple junction failure. He considered a scenario where grain boundary dislocations (GDB) accumulate at a triple junction under tensile stress. Similar models [7, 8] for triple junction deformation under shear loading have also been constructed for the nanocrystalline/polycrystalline materials. These models were built on a continuous description of the dislocations interaction and did not take into

consideration the dynamical nature of a deformation at the molecular scale. Theoretical approach to a comprehensive description of triple junction deformation at the molecular scale is proven challenging, in particular to the high deformation or the high shear regime. It is therefore the objective of this study to investigate the atomistic deformation mechanism of a triple junction, with an emphasis on the hardening-related processes.

5.2 Methodology

The study was carried out by newtonian Molecular Dynamics (MD) simulation. The physical behaviour of aluminum atoms was introduced to the system via an embedded atom method (EAM) potential given in ref [9]. The potential has a form:

$$E_i = F_i(n_i) + \frac{1}{2}\sum_{j \neq i} V_{ij}(r_{ij}), \qquad (5.1)$$

where F_i is the many-body potential describing the energy required to embed a charge into a net electron density n_i. The term V_{ij} is a pair potential while r_{ij} is the distance between atom i and j. Atomic motions are evaluated by the 5th order Gear corrector-predictor algorithm[10]. A symmetric triple junction system is constructed as shown in *figure 5.1*. This configuration has a dimension of 720.0Å × 374.1Å × 246.7Å and is composed of 4.08 millions aluminum atoms. A free boundary condition is applied to the x and y faces while periodic boundary conditions is enforced in the z direction. Three layers of atoms on a section of the x faces are pinned and are displaced artificially at a constant rate of ±0.5mÅ/fs in the x direction as shown in *figure 5.1*. This introduces a block shear motion at the symmetry plane. The shear is equivalent in principle to a mode II deformation, which is commonly observed in the short crack growth regime in a fatigue scenario. Initially, the system is relaxed for 1000 steps at 0K. After which it is maintained a temperature of 5K during the simulation. The temperature is controlled via kinetic energy rescaling at every 500 steps. The low temperature is chosen to distinguish dynamical phenomena from thermal processes like: thermal recrystallization and grain boundary diffusion.

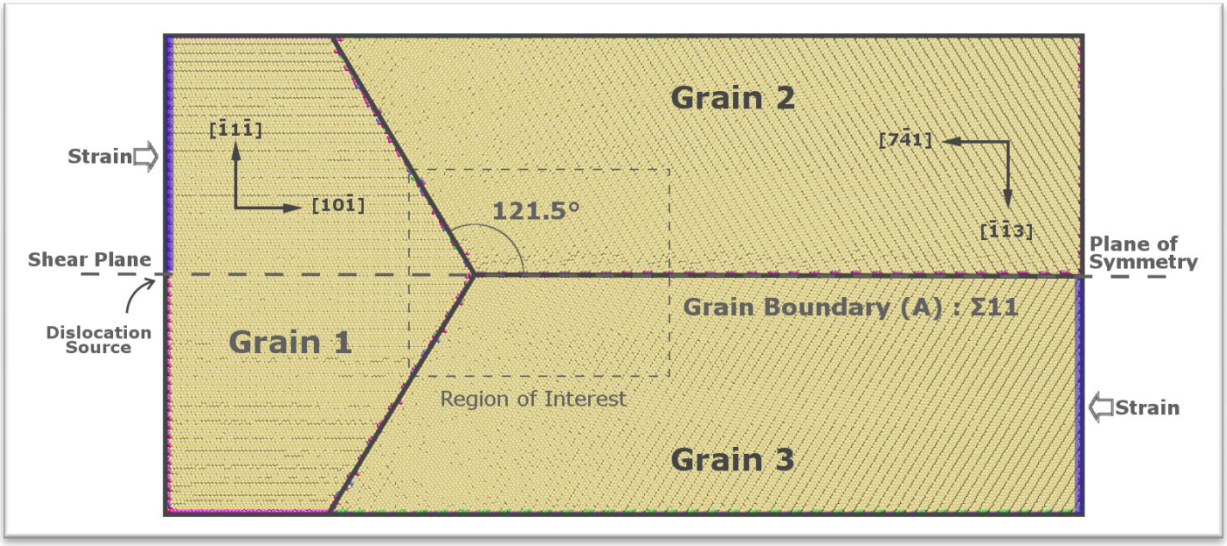


Figure 5.1: The configuration of a symmetrical triple junction system.

5.3 Discussion

5.3.1 Triple Junction Deformation

Shear strain introduced to the system generates a continuous stream of dislocations on the symmetry plane. At the early stage, these dislocations glide into the triple junction and emerge on the other side as some intrinsic grain boundary dislocations (GBD). These GBD move along grain boundary A and induce a net sliding motion between grain 2 and 3. Because the grain boundary has a tilt-symmetry and the symmetry plane aligns with the shear plane, grain boundary migration (normal motion) via shear coupling is kinetically favoured [11]. The coupling can be understood as a shifting motion driven by a shear stress, at which atoms next to a grain boundary adopt the crystalline periodicity of the neighbouring grain. This process causes a transfer of atoms between neighbouring grains in a layer-by-layer manner and induces a net grain boundary motion normal to the shear direction. As plastic deformation continues, the triple junction disfigures into a structure with two acute dihedral angles as shown in *figure 5.2*. The configuration is not dynamically stable for two reasons: 1) the dislocation pathway becomes discontinued. As such, dislocations can no longer cross the junction easily. 2) Grains do not favour acute dihedral angles for energetic reasons based on interfacial energy balance between grain boundaries. As a consequence to the

discontinuity in the dislocation pathway, dislocations from grain 1 arriving at the junction begin to accumulate and the junction becomes increasingly amorphized. The amorphized region extends along the grain boundary following the shear stress. As a result, the triple junction and the grain boundaries at the region become indiscernible and distorted. As shown in *figure 5.3*, the nucleation of an amorphized region removes the acute dihedral angles and smoothens the dislocation pathway.

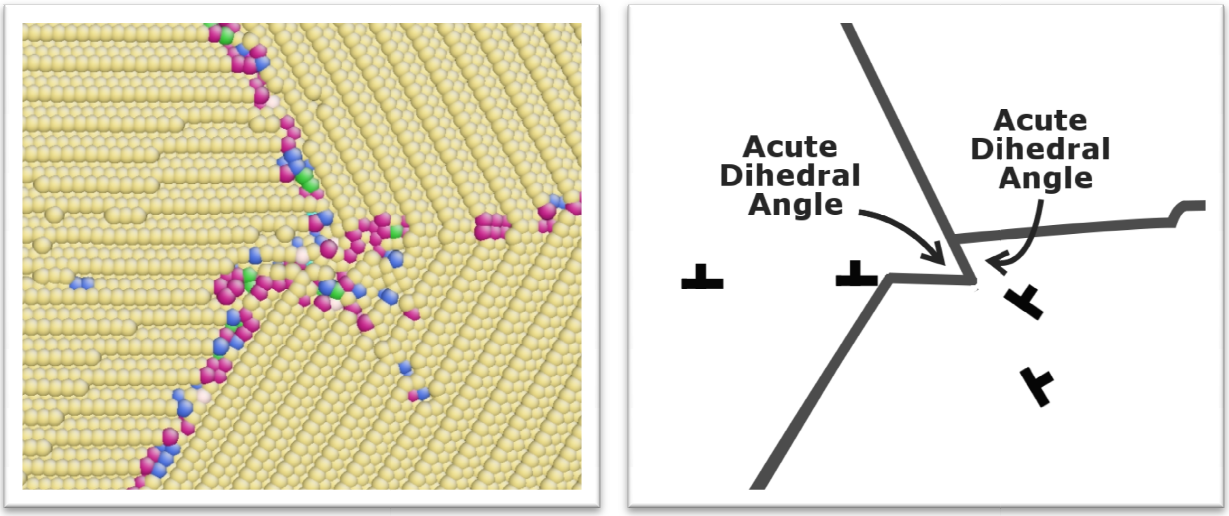


Figure 5.2: Visualization (in coordinate number coloring) and the corresponding schematic of a triple junction in its early stage of deformation. Two acute dihedral angles are formed.

As the deformation continues, it is found that the smoothened triple junction configuration becomes stable. *Stable* in this context means the angles between grain boundaries remain constant. It implies that migration of grain boundary is ceased near the triple junction, and that sliding is the sole form of grain boundary motion. The abeyance of migration can be considered as a pinning effect due to the triple junction. The section of grain boundary further away from the junction will experience little influence of the pinning and continues to migrate. However, migration motion is constrained in the section joining the pinned and the free grain boundary. In particular, grain boundary in that section cannot migrate beyond the pinned orientation because this configuration will obstruct the sliding motion. For this reason, the front of the newly pinned grain boundary will induces a pinning effect to the nearby migrating boundaries. As a result, the migrating boundary will continue to adopt the pinned orientation and be reduces in length. *Figure 5.4* illustrates the mechanism of such adaptation process. The consequence is a transformation of the grain boundary

action from migration to a pure sliding motion. Such process can be important for crack propagation. For instance, as oppose to a sliding, grain boundary migration is not associated with crack propagation. In a scenario where a short intergranular fatigue crack arrives at a triple junction, the shear stress at the front of its crack tip can initiate the dynamical process, transforming an initially migration-preferred grain boundary configuration into a sliding-preferred orientation before the crack continuous to propagate.

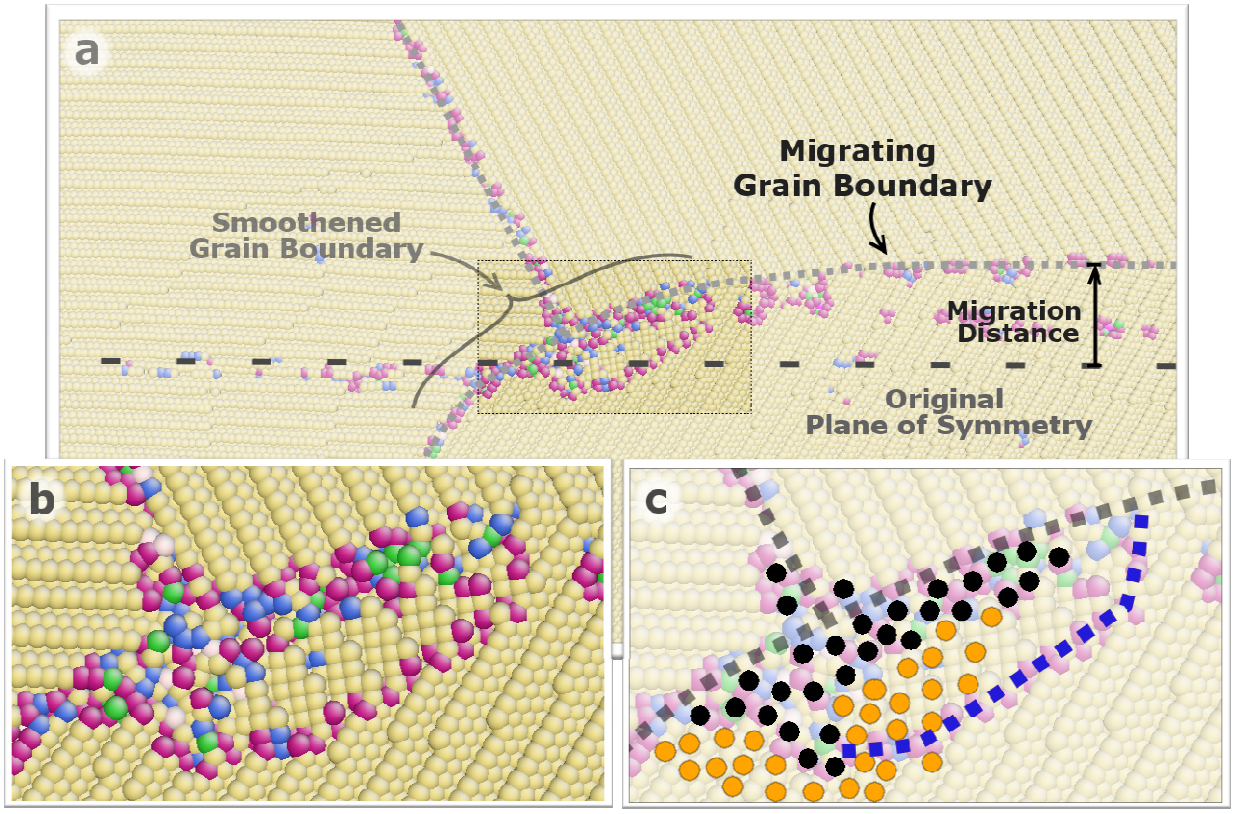


Figure 5.3: a) Snapshot of the deforming triple junction. b) Magnified view of a nano-grain formed from dynamic recrystallization. Its crystalline orientation is clearly different from all neighbouring grains. c) Schematic of the nano-grain. Black dot represent amorphized region visible in the picture. Orange dots represent amorphized region behind the visible atoms. Blue dash line represents the newly nucleated high angle grain boundary.

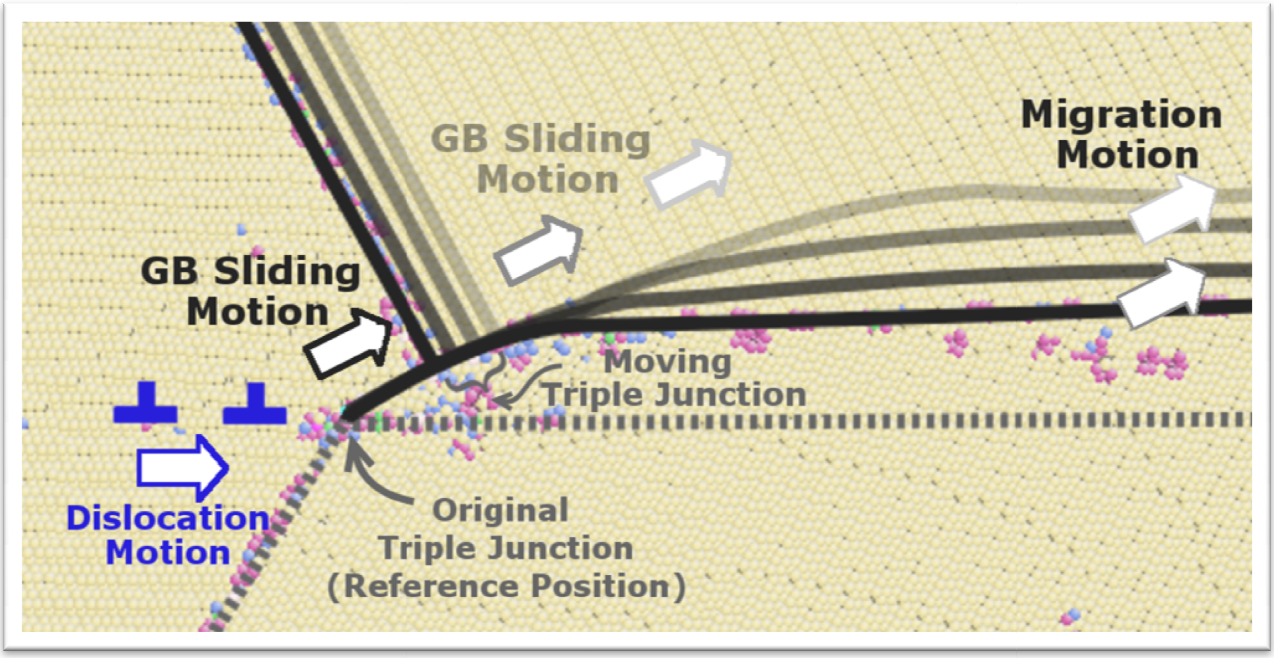


Figure 5.4: Time series of grain boundary motions when the triple junction is being deformed. The lines correspond to grain boundaries position at 60ps, 70ps, 80ps and 90ps (dark to light) from the beginning of the simulation. It illustrated the relationship between GB sliding and migration motion. As the triple junction deform, an increasing section of grain boundary adopted the pure sliding motion.

5.3.2 Dynamic Recrystallization

As the deformation continues, it is found that the amorphized region extends along grain boundary *A* following the plastic deformation. At an instant, some dozens of atoms in the region rearrange themselves into a new crystalline seed. The seed quickly rearrange disordered atoms nearby and extends its border into grain 3 by nucleating a high angle grain boundary. The nucleation process can be interpreted as the generation of a dense wall of dislocations from the amorphized region, which sweeps into the crystallinity of grain *3* and rotates the surrounded atoms into a new crystalline orientation. The aberrant nature of this dislocation dynamics is driven by the highly-localized contour of stress presented in a deforming triple junction. The contour of stress at a triple junction induced by the shear has a 'rotation' tendency and prefers to reorient small regions of atoms. In the simulation, two such recrystallized grains are found along the triple junction. They have an average size of around 32±6Å in span and share a similar flattened shape. Because of the nucleation mechanism, these grains are found to have one side attached to the amorphized region,

with an indistinct interface, while the other side is interfaced to grain 3 with a discerned boundary. Their sizes and shapes are found to remain approximately constant throughout the simulation. As illustrated in *figure 5.3*, a recrystallized region resembles a nano-sized grain attached to the grain boundary. These grains nucleate without going through the stage of subgrain formation, ie: nucleation of a low angle grain boundary. For this reason, the process is recognized as a discontinuous dynamic recrystallization (DDRX)[12]. Recrystallization near a triple junction has been reported previously, however, the scale at which it was observed is vastly different from the process discussed here. The presence of the amorphous region and the nano-grains deepened the crystalline incompatibility of the region. The deformed triple junction becomes a more predominated site for defect accumulation. It is believed that such aggregation process will eventually leads to the nucleation of a crack.

5.4 Conclusions

Simulation results suggest that a triple junction could deform via nucleating a small amorphized region nearby. The amorphized region is a result of the dislocations accumulation at the junction. The formation of the amorphized region is found to smoothen a grain boundary in the disfiguring junction. Consequentially, sliding motion is favoured on that section of the grain boundary. It is found that the length of that section increase as the grain boundary attach to it continues to migrate. Eventually the grain boundary is pinned from migrating. Some nano-sized grains are found to nucleate in the amorphized region and extend its border to the crystalline region. The event is identified as a discontinued dynamics recrystallization for it occurs without the formation of a low angle grain boundary.

Acknowledgement

This research is supported by DND-TIF project: A Coupled Atomic-Meso-Macroscopic Modeling of Cracks in Aircraft Components, and is made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET:www.sharcnet.ca) and Compute/Calcul Canada.

References

[1] Williams, J. a A theoretical derivation of the creep life of commercial materials failing by triple-point cracking. *Philosophical Magazine* **15**, 1289-1291(1967).

[2] Venkataraman, G., Chung, Y.W. & Mura, T. Application of minimum energy formalism in a multiple slip band model for fatigue – II crack nucleation and derivation of generalised Coffin-Manson law. *Acta metallurgica et materialia* **39**, 2631-2638(1991).

[3] Manonukul, A. & Dunne, F. High-and low-cycle fatigue crack initiation using polycrystal plasticity. *Proceedings: Mathematical, Physical and Engineering Sciences* **460**, 1881–1903(2004).

[4] Kobayashi, S., Tsurekawa, S. & Watanabe, T. Roles of structure-dependent hardening at grain boundaries and triple junctions in deformation and fracture of molybdenum polycrystals. *Materials Science and Engineering: A* **483-484**, 712-715(2008).

[5] Kobayashi, S., Inomata, T. & Kobayashi, H., S Effects of grain boundary-and triple junction-character on intergranular fatigue crack nucleation in polycrystalline aluminum. *Journal of Materials* 3792-3799(2008).

[6] Walter, J. & Cline, H. Grain Boundary Sliding, Migration, and Deformation in High-Purity Aluminum. *Trans. AIME* **242**, 1823–1830(1968).

[7] Ovid'ko, I. Triple junction nanocracks in deformed nanocrystalline materials. *Acta Materialia* **52**, 1201-1209(2004).

[8] Fedorov, A., Gutkin, M. & Ovid'ko, I. Transformations of grain boundary dislocation pile-ups in nano- and polycrystalline materials. *Acta Materialia* **51**, 887–898(2003).

[9] Liu, X.Y. et al. Anisotropic surface segregation in Al-Mg alloys. *Surface science* **373**, 357–370(1997).

[10] Allen, MP & Tildesley DJ. Computer Simulation of Liquids. *Oxford University Press* (1989)

[11] Cahn, J., Mishin, Y. & Suzuki, A. Coupling grain boundary motion to shear deformation. *Acta Materialia* **54**, 4953-4975(2006).

[12] Miura, H. Preferential nucleation of dynamic recrystallization at triple junctions. *Scripta Materialia* **50**, 65-69(2004).

[13] Poirier, J. & Nicolas, A. Deformation-induced recrystallization due to progressive misorientation of subgrains, with special reference to mantle peridotites. *The Journal of Geology* **83**, 707–720(1975).

[14] Bestmann, M. Intragranular dynamic recrystallization in naturally deformed calcite marble: diffusion accommodated grain boundary sliding as a result of subgrain rotation recrystallization. *Journal of Structural Geology* **25**, 1597-1613(2003).

[15] Shu-nong, Liu C.J. & Zhang, X. Continuous dynamic recrystallization and discontinuous dynamic recrystallization in 99.99% polycrystalline aluminum during hot compression. *Transactions of Nonferrous Metals Society of China* **15**, (2005).

[16] Li, J. AtomEye: an efficient atomistic configuration viewer. *Modelling and Simulation in Materials Science and Engineering* **11**, 173-177(2003).

Chapter 6

Conclusion

The presented thesis comprises of discussions on fatigue related dislocations dynamics. Numerous studies were carried out to address key issues in this topic. In particular, these studies have provided an understanding on the essential aspects of the dislocation dipoles disintegration and triple junction deformation mechanisms. These findings are of scientific values and are accredited by my supervisor, Dr. Chen, for publications. In conclusion, here presents a summary of the findings and scientific contribution of this thesis:

In regards to the disintegration mechanism of a group of dislocation dipoles piled-up on a grain boundary, it is found that the process is initiated by cross-slip of partial dislocations at the grain boundary. The cross-slip promote a rich set of dislocation dynamics which leads to the formation of Lomer-Cottrell locks, vacancy rolls, a batch of stacking fault tetrahedrons and vacancy clusters. Lomer-Cottrell locks are found to form via the union of multiple partial dislocations while tetrahedrons are formed via the union of multiple stacking fault plates. Stacking fault tetrahedrons eventually collapse and result in the production of a large amount of point defects closely packed in the width of the dipoles. To the best of the author's knowledge, the dynamics have not been reported previously. The finding provided a useful description, embracing the collective behavior of a series of dislocations.

As for the deformation mechanism of a triple junction, it is realized via atomistic simulation that the process can be initiated by nucleating a small amorphized region. The amorphized region assists the recrystallization of nano-grains next to the triple junction. The region is also found to be related to the smoothening of the disfigured triple junction. It is also found to assist grain boundary sliding motion near-by. Grain boundary migration and sliding were

observed in the simulation. Their relationship in the scope of triple junction deformation has also been discussed in detail.

A massively parallel molecular dynamics simulator was developed for the purpose of the studies presented in this thesis. The program illustrated superior performance and efficiency in some brief benchmark procedures. The technologies involved in the acceleration were discussed in detail. The code was written from scratch with the intention to serve both educational and research purposes for the scientific community. It is an ongoing effort to implement the latest computation technology which enables efficient and massive simulation in the field of molecular dynamics simulation. It is in hope the simulator can handle real-life sample sizes in the foreseeable future as computation technology mature.