Multiscale modelling of nanomechanics and micromechanics: an overview

NASR M. GHONIEM†

Mechanical and Aerospace Engineering Department, University of California,
Los Angeles, California 90095-1597, USA

ESTEBAN P. BUSSO

Department of Mechanical Engineering, Imperial College, London, UK

NICHOLAS KIOUSSIS

Department of Physics, California State University Northridge,
Northridge, California 91330-8268, USA

HANCHEN HUANG

Department of Mechanical, Aerospace and Nuclear Engineering,
Rensselaer Polytechnic Institute, Troy, New York 12180-3590, USA

ABSTRACT

Recent advances in analytical and computational modelling frameworks to describe the mechanics of materials on scales ranging from the atomistic, through the microstructure or transitional, and up to the continuum are reviewed. It is shown that multiscale modelling of materials approaches relies on a systematic reduction in the degrees of freedom on the natural length scales that can be identified in the material. Connections between such scales are currently achieved either by a parametrization or by a 'zoom-out' or 'coarse-graining' procedure. Issues related to the links between the atomistic scale, nanoscale, microscale and macroscale are discussed, and the parameters required for the information to be transferred between one scale and an upper scale are identified. It is also shown that seamless coupling between length scales has not yet been achieved as a result of two main challenges: firstly, the computational complexity of seamlessly coupled simulations via the coarse-graining approach and, secondly, the inherent difficulty in dealing with system evolution stemming from time scaling, which does not permit coarse graining over temporal events. Starting from the Born–Oppenheimer adiabatic approximation, the problem of solving quantum mechanics equations of motion is first reviewed, with successful applications in the mechanics of nanosystems. Atomic simulation methods (e.g. molecular dynamics, Langevin dynamics and the kinetic Monte Carlo method) and their applications at the nanoscale are then discussed. The role played by dislocation dynamics and statistical mechanics methods in understanding microstructure self-organization, heterogeneous plastic deformation, material instabilities and failure phenomena is also discussed.

† Email: ghoniem@ucla.edu.

Finally, we review the main continuum-mechanics-based framework used today to describe the nonlinear deformation behaviour of materials at the local (e.g. single phase or grain level) and macroscopic (e.g. polycrystal) scales. Emphasis is placed on recent progress made in crystal plasticity, strain gradient plasticity and homogenization techniques to link deformation phenomena simultaneously occurring at different scales in the material microstructure with its macroscopic behaviour. In view of this wide range of descriptions of material phenomena involved, the main theoretical and computational difficulties and challenges are critically assessed.

§1. INTRODUCTION

Computational modelling of materials behaviour is becoming a reliable tool to underpin scientific investigations and to complement traditional theoretical and experimental approaches. In cases where an understanding of the dual nature of the structure of matter (continuous when viewed at large length scales and discrete when viewed at an atomic scale) and its interdependences are crucial, multiscale materials modelling (MMM) approaches are required to complement continuum and atomistic analyses methods. On transitional (or microstructure) scales (*in between* continuum and atomistic), continuum approaches begin to break down, and atomistic methods reach inherent time and length-scale limitations. Transitional theoretical frameworks and modelling techniques are being developed to bridge the gap between length-scale extremes. The power of analytical theories lies in their ability to reduce the complex collective behaviour of the basic ingredients of a solid (e.g. electrons, atoms, lattice defects, and single-crystal grains) into insightful relationships between cause and effect. For example, the description of deformation beyond the elastic regime is usually described by appropriate *constitutive equations*, and the implementation of such relationships within continuum mechanics generally relies on the inherent assumption that material properties vary continuously throughout the solid. However, certain heterogeneities linked to either the microstructure or the deformation *per se* cannot be readily described within the framework provided by continuum mechanics: dislocation patterns, bifurcation phenomena, crack nucleation in fatigue, some non-local phenomena, etc. Some examples will be discussed next to illustrate the role of MMM in nanomechanics and micromechanics research.

Recent interest in *nanotechnology* is challenging the scientific community to analyse, develop and design nanometre to micrometre-sized devices for applications in new generations of computers, electronics, photonics and drug delivery systems. These new exciting application areas require novel and sophisticated, physically based approaches for design and performance prediction. Thus, theory and modelling are playing an ever-increasing role in these areas to reduce development costs and manufacturing times. An important problem which concerns the microelectronic industry is the reliable operation of integrated circuits, where the lifetime is limited by the failure of interconnect wires in between submicrometre semiconducting chips. In some cases, the nucleation and growth of even a single nanovoid can cause interconnect failure. Statistical mechanics cannot adequately address this situation. Future electronic and optoelectronic devices are expected to be even smaller, with *nanowires* connecting a nano-sized memory and information storage and retrieval nanostructures. Understanding the mechanics of such nano-engineered devices will

enable high levels of reliability and useful lifetimes to be achieved. Undoubtedly, defects are expected to play a major role in these nanosystems and microsystems owing to the crucial impact on the physical and mechanical performance.

The potential of MMM approaches for computational materials design is also great. Such possibility was recently illustrated on a six-component Ti-based alloy, with a composition predetermined by electronic properties, which was shown to exhibit an entirely new twin- and dislocation-free deformation mechanism, leading to 'unexpected new properties, such as high ductility and strength, and Invar and Elinvar behaviour' (Saito *et al.* 2003). Such an alloy would be unlikely to be found by trial and error. This points to a paradigm shift in modelling, away from reproducing known properties of known materials and towards simulating the behaviour of possible alloys as a forerunner to finding real materials with these properties.

In high-pay-off, high-risk technologies, such as the design of large structures in the aerospace and nuclear industries, the effects of ageing and environment on failure mechanisms cannot be left to conservative approaches. Increasing efforts are now focused on developing MMM approaches to design new alloys and material systems in these areas. An illustration of an MMM based strategy for the development of large components surrounding the plasma core of a fusion energy system is shown in figure 1. The development of ultrastrong and yet ductile materials by combining nanolayers with different microstructures also requires detailed understanding of their mechanical properties. Such materials, if properly designed, may be candidates for many demanding applications (e.g. microelectronics, optoelectronics, laser mirrors, aircraft structures, rocket engines and fuel cells).

Appropriate validation experiments are also crucial to verify that the models predict the correct behaviour at each length scale, ensuring that the linkages between approaches are directly enforced. However, current nanoscale and microscale mechanical experiments have been mostly limited to indentation (Loubet *et al.* 1986, Wu *et al.* 1993) and bulge tests (Baker 1993, Small *et al.* 1994), and to non-contact tests such as X-ray residual stress measurements (James 1980,
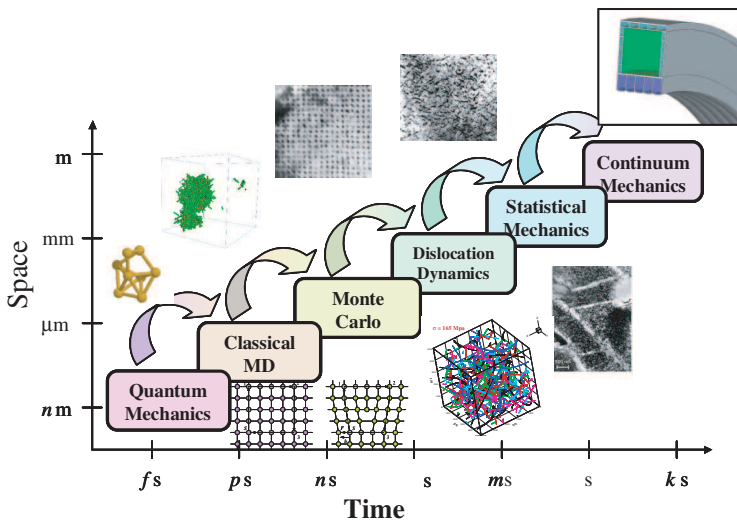

Figure 1. Illustration of a MMM approach for the design of radiation-resistant materials for fusion energy structures: MD, molecular dynamics.

Segmueller and Murakami 1988). Multiscale interconnected approaches will need to be developed to interpret new and highly specialized nanomechanical and micro-mechanical tests. One of the advantages of these approaches is that, at each stage, physically meaningful parameters are predicted and used in subsequent models, avoiding the use of empiricism and fitting parameters.

As the material dimensions become smaller, its resistance to deformation is increasingly determined by internal or external discontinuities (e.g. surfaces, grain boundaries and dislocation cell walls). The Hall–Petch relationship has been widely used to explain grain-size effects, although the basis of the relationship is strictly related to dislocation pile-ups at grain boundaries. Recent experimental observations on nanocrystalline materials with grains of the order of 10–20 nm indicate that the material is weaker than would be expected from the Hall–Petch relationship (Erb *et al.* 1996, Campbell *et al.* 1998). Thus, the interplay between interfacial or grain-boundary effects and slip mechanisms within a single-crystal grain may result in either strength or weakness, depending on their relative sizes. Although experimental observations of plastic deformation heterogeneities are not new, the significance of these observations has not been addressed until very recently. In the same metallic alloys, regular patterns of highly localized deformation zones, surrounded by vast material volumes which contain little or no deformation, are frequently seen (Mughrabi 1983, 1987, Amodeo and Ghoniem 1988). The length scale associated with these patterns (e.g. typically the size of dislocation cells, the ladder spacing in persistent slip bands (PSBs) or the spacing between coarse shear bands) controls the material strength and ductility. As it may not be possible to homogenize such types of microstructure in an average sense using either atomistic simulations or continuum theories, new intermediate approaches will be needed.

The issues discussed above, in addition to the ever-increasingly powerful and sophisticated computer hardware and software available, are driving the development of MMM approaches in nanomechanics and micromechanics. It is expected that, within the next decade, new concepts, theories and computational tools will be developed to make truly seamless multiscale modelling a reality. In this overview article, we briefly outline the status of research in each component that makes up the MMM paradigm for modelling nanosystems and microsystems: quantum mechanics (QM), molecular dynamics (MD), Monte Carlo (MC) method, dislocation dynamics (DD), statistical mechanics (SM) and, finally, continuum mechanics (CM).

§2. COMPUTATIONAL QUANTUM MECHANICS

2.1. *Essential concepts*

There is little doubt that most of the low-energy physics, chemistry, materials science and biology can be explained by the QM of electrons and ions and that, in many cases, the properties and behaviour of materials derive from the quantum-mechanical description of events at the atomic scale. Although there are many examples in materials science of significant progress having been made without any need for quantum-mechanical modelling, this progress is often limited as one pushes forward and encounters the atomic world. For instance, the understanding of the properties of dislocations comes from classical elasticity theory, but even in these cases, until recently, very little was known about the core of a dislocation or the effect of chemistry on the core, precisely because this part of the dislocation requires

detailed quantum-mechanical modelling. The ability of QM to predict the total energy and the atomic structure of a system of electrons and nuclei enables one to reap a tremendous benefit from a quantum-mechanical calculation. Many methods have been developed for solving the Schrödinger equation that can be used to calculate a wide range of physical properties of materials, which require only a specification of the ions present (by their atomic number). These methods are usually referred to as *ab initio* methods.

In considering the motion of electrons in condensed matter, we are dealing with the problem of describing the motion of an enormous number of electrons and nuclei (about 10^{23}) obeying the laws of QM. Prediction of the electronic and geometric structure of a solid requires calculation of the quantum-mechanical total energy of the system and subsequent minimization of that energy with respect to the electronic and nuclear coordinates. Because of the large difference between the masses of electrons and nuclei and the fact that the forces on the particles are the same, the electrons respond essentially instantaneously to the motion of the nuclei. Thus, the nuclei can be treated adiabatically, leading to a separation of the electronic and nuclear coordinates in the many-body wave function

$$\psi(\{\mathbf{R}_I, \mathbf{r}_n\}) = \psi_{el}(\{\mathbf{r}_n\}; \{\mathbf{R}_I\})\psi_{nuc}(\{\mathbf{R}_I\}), \tag{1}$$

which is the so-called Born–Oppenheimer approximation (Parr and Yang 1989, Dreizler and Gross 1990). The *adiabatic principle* reduces the many-body problem in the solution of the dynamics of the electrons in some frozen-in configuration $\{\mathbf{R}_I\}$ of the nuclei; this is an example of the philosophy behind systematic degree-of-freedom elimination which is itself at the heart of multiscale modelling.

Even with this simplification, the many-body problem remains formidable. The Hamiltonian for the N-electron system moving in condensed matter with fixed nuclei is (in atomic units)

$$H = \sum_{i=1}^{N} -\tfrac{1}{2}\boldsymbol{\nabla}_i^2 + \sum_{i=1}^{N} v_{\sigma_i}(\mathbf{r}_i) + \tfrac{1}{2}\sum_{i}\sum_{j\neq i}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \tag{2}$$

where the first term is the kinetic energy operator, $v_\sigma(\mathbf{r})$ is the one-electron spin-dependent external potential (e.g. the electron–nucleus interaction), and the third term represents the effect of the electron–electron interactions which poses the most difficult problem in any electronic structure calculation.

Density functional theory (DFT) (Hohenberg and Kohn 1964, Kohn and Sham 1965, Jones and Gunnarsson 1989, Parr and Yang 1989, Dreizler and Gross 1990) has been proven to be a very powerful quantum-mechanical method in investigating the electronic structure of atoms, molecules and solids. Here, the electron density $\rho(\mathbf{r})$, or spin density $\rho_\sigma(\mathbf{r})$, is the fundamental quantity, rather than the total wave function employed for example in Hartree–Fock theory (Ashcroft and Mermin 1976). The DFT makes successful predictions of ground-state properties for moderately correlated electronic systems (Perdew 1999). DFT can provide accurate ground-state properties for real materials (such as total energies and energy differences, cohesive energies of solids and atomization energies of molecules, surface energies, energy barriers, atomic structure, and magnetic moments) and provides a scheme for calculating them (Jones and Gunnarsson 1989, Fulde 1993, Perdew 1999).

We next give a brief discussion of the essential concepts of DFT. Hohenberg and Kohn (1964) proved that the total energy, including exchange and correlation, of an electron gas (even in the presence of a static external potential) is a unique functional of the electron density. The minimum value of the total-energy functional is the ground-state energy of the system, and the density that yields this minimum value is the exact single-particle ground-state density. Kohn and Sham (1965) then showed how it is possible, formally, to replace the many-electron problem by an exactly equivalent set of self-consistent one-electron equations

$$\left(-\tfrac{1}{2}\boldsymbol{\nabla}^2 + v_\sigma(\mathbf{r}) + \int \mathrm{d}^3 r' \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + v_{\mathrm{xc}}^\sigma(\mathbf{r})\right)\psi_{\mathbf{k}\sigma}(\mathbf{r}) = \epsilon_{\mathbf{k}\sigma}\psi_{\mathbf{k}\sigma}(\mathbf{r}), \tag{3}$$

where $\psi_{\mathbf{k}\sigma}(\mathbf{r})$ are the Kohn–Sham orbitals, and the spin-dependent exchange–correlation potential is defined as

$$v_{\mathrm{xc}}^\sigma(\mathbf{r}) \equiv \frac{\delta E_{\mathrm{xc}}[\rho_\uparrow, \rho_\downarrow]}{\delta \rho_\sigma(\mathbf{r})}. \tag{4}$$

Here, $E_{\mathrm{xc}}[\rho_\uparrow, \rho_\downarrow]$ is the exchange–correlation energy and the density $\rho_\sigma(\mathbf{r})$ of electrons of spin $\sigma(=\uparrow, \downarrow)$ is found by summing the squares of the occupied orbitals:

$$\rho_\sigma(\mathbf{r}) = \sum_{\mathbf{k}} |\psi_{\mathbf{k}\sigma}(\mathbf{r})|^2 \theta(\mu - \epsilon_{\mathbf{k}\sigma}), \tag{5}$$

where $\theta(x)$ is the step function and μ is the chemical potential. The exchange–correlation energy is 'nature's glue'. It is largely responsible for the binding of atoms to form molecules and solids. The total electronic density is

$$\rho(\mathbf{r}) = \rho_\uparrow + \rho_\downarrow(\mathbf{r}). \tag{6}$$

Aside from the nucleus–nucleus repulsion energy, the total energy is

$$E = \sum_{\mathbf{k}\sigma}\langle\psi_{\mathbf{k}\sigma}| -\tfrac{1}{2}\boldsymbol{\nabla}^2 |\psi_{\mathbf{k}\sigma}\rangle\theta(\mu - \epsilon_{\mathbf{k}\sigma}) + \sum_\sigma \int \mathrm{d}^3 r\, v_\sigma(\mathbf{r})\rho_\sigma(\mathbf{r}) + U[\rho] + E_{\mathrm{xc}}[\rho_\uparrow, \rho_\downarrow], \tag{7}$$

where

$$U[\rho] = \tfrac{1}{2}\int \mathrm{d}^3 r \int \mathrm{d}^3 r' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{8}$$

is the Hartree self-repulsion of the electron density.

The Kohn–Sham equations represent a mapping of the interacting many-electron system on to a system of non-interacting electrons moving in an effective non-local potential owing to all the other electrons. The Kohn–Sham equations must be solved self-consistently so that the occupied electron states generate a charge density that produces the electronic potential that was used to construct the equations. New iterative diagonalization approaches can be used to minimize the total-energy functional (Car and Parrinello 1985, Gilan 1989, Payne *et al.* 1992). These are much more efficient than the traditional diagonalization methods. If the exchange–correlation energy functional were known exactly, then taking the functional derivative with respect to the density would produce an exchange–correlation potential that included the effects of exchange and correlation exactly.

The complexity of the real many-body problem is contained in the unknown exchange–correlation potential $v_{\mathrm{xc}}^\sigma(\mathbf{r})$. Nevertheless, making simple approximations, we can hope to circumvent the complexity of the problem. Indeed the simplest

possible approximation, that is the local-density (LDA) or local-spin-density (LSDA) approximation, have been proven (Kohn and Sham 1965, von Barth and Hedin 1972) very successful.

2.2. *The local density approximation*

This simplest method of describing the exchange–correlation energy of an electron system has been the workhorse of condensed-matter physics for almost 30 years. In the LSDA the exchange–correlation energy of an electron system is constructed by assuming that the exchange–correlation energy ϵ_{xc} per electron at a point $\mathbf{r}$ in the electron gas is equal to the exchange–correlation energy per electron in an electron gas of *uniform* spin densities $\rho_\uparrow$, $\rho_\downarrow$, namely,

$$E_{xc}^{\text{LSDA}}[\rho_\uparrow, \rho_\downarrow] = \int \mathrm{d}^3 r \, \rho(\mathbf{r}) \epsilon_{xc}^{\text{unif}}(\rho_\uparrow(\mathbf{r})\rho_\downarrow(\mathbf{r})). \tag{9}$$

The LSDA assumes that the exchange–correlation energy functional is purely local. Several parametrizations exist for the exchange–correlation energy of a homogenous electron gas all of which lead to total-energy results that are very similar (Kohn and Sham 1965, Hedin and Lundqvist 1971, Vosko *et al.* 1980, Perdew and Zunger 1981). These parametrizations use interpolation formulae to link exact results for the exchange–correlation energy of high-density electron gases and calculations of the exchange–correlation energy of intermediate and low-density electron gases. The LDA or LSDA, in principle, ignores corrections to the exchange–correlation energy at a point $\mathbf{r}$ due to nearby inhomogeneities in the electron gas. Considering the inexact nature of the approximation, it is remarkable that calculations performed using the LDA have been so successful. Recent work has shown that this success can be partially attributed to the fact that the LDA gives the correct sum rules for the exchange–correlation hole (Hariss and Jones 1974, Gunnarsson and Lundqvist 1976, Langreth and Perdew 1977).

A number of attempts to improve the LDA or LSDA use gradient expansions of the charge or spin density. In this way, non-local information about the charge or spin density could be provided from the gradient terms. These type of approximation are in general referred to as *gradient expansion approximations* (Perdew 1999). The most popular, the Perdew–Wang generalized gradient approximation (GGA) (Perdew 1991, 1999, Perdew *et al.* 1996)

$$E_{xc}^{\text{GGA}}[\rho_\uparrow, \rho_\downarrow] = \int \mathrm{d}^3 r \, f(\rho_\uparrow, \rho_\downarrow, \boldsymbol{\nabla}\rho_\uparrow, \boldsymbol{\nabla}\rho_\downarrow) \tag{10}$$

starts from the second-order gradient expansion of the exchange–correlation hole density and then cuts off the spurious long-range ($|\mathbf{r} - \mathbf{r}'| \to \infty$) parts to restore the sum rules. The GGA is typically more accurate than LSD, especially for the rapidly varying densities of atoms and molecules. In recent years, GGAs have made DFT popular in quantum chemistry.

2.3. *Basis functions*

In the discussion of the Kohn–Sham equations earlier, nothing has been said about the way that $\psi_{\mathbf{k}}(\mathbf{r})$ itself will be determined. Many methods share the common feature that the wave function is presumed to exist as a linear combination of some set of basis functions which might be written generically as

$$\psi(\mathbf{r}) = \sum_n \alpha_n \phi_n(\mathbf{r}), \tag{11}$$

where α_n is the weight associated with the *n*th basis function $\phi_n(\mathbf{r})$. The solution of the Kohn–Sham equations in turn becomes a search for unknown coefficients rather than unknown functions. Part of the variety associated with the many different implementations of DFT codes is associated with the choice of basis functions and the form of the effective crystal potential. In deciding on a particular set of basis functions, a compromise will have to be made between high-accuracy results, which require a large basis set, and computational costs, which favour small basis sets. Similar arguments hold with respect to the functional forms of the $\phi_n(\mathbf{r})$. There are several powerful techniques in electronic structure calculations. The various methods may be divided into those which express the wave functions as linear combinations of some *fixed* basis functions, say linear combination of atomic orbitals localized about each atomic site (Eschrig 1989) or those which use a free-electron-like basis, the so-called plane-wave basis set (Payne *et al.* 1992, Denteneer and van Haeringen 1985). Alternatively, there are those methods which employ *matching* of partial waves, such as the full-potential linear-augmented-plane-wave method (Wimmer *et al.* 1981), the full-potential linear muffin-tin-orbital method (Price and Cooper 1989, Price *et al.* 1992) and the full-potential Korringa–Kohn–Rostoker method (Papanikolaou *et al.* 2002).

2.4. *Nanomechanics applications*

Electronic structure calculations based on DFT also can be applied for studies of non-periodic systems, such as those containing point, planar or line defects, or quantum dots if a periodic supercell is used. The supercell contains the defect surrounded by a region of bulk crystal or vacuum for the case of a surface. Periodic boundary conditions are applied to the supercell so that the supercell is reproduced throughout space. It is essential to include enough bulk solid (or vacuum) in the supercell to prevent the defects in neighbouring cells from interacting appreciably with each other. The independence of defects in neighbouring cells can be checked by increasing the volume of the supercell until the calculated defect energy is converged. Using the supercell geometry, one can study even molecules (Rappe *et al.* 1992), provided that the supercell is large enough that the interactions between the molecules are negligible.

The implementation of the DFT within the LDA or the GGA for the exchange–correlation energy, the development of new linearized methods for solving the single-particle Schrödinger equations and the use of powerful computers allow for a highly efficient treatment of up to several hundreds of atoms and has led to an outburst of theoretical work in condensed-matter physics and materials physics. Representative successful applications of DFT for a wide variety of materials ranging from metals to semiconductors to ceramics include the electronic, structural and magnetic properties of surfaces (LaBella *et al.* 1999, Ruberto *et al.* 2002), interfaces (Batirev *et al.* 1999), grain boundaries (Wu *et al.* 1994, Lu *et al.* 1999, Lu and Kioussis 2001), alloys (Kent and Zunger 2001, Janotti *et al.* 2002), chemisorption of molecules on surfaces (Dürr *et al.* 2001), C nanotubes (Chan *et al.* 2001, Yildrim *et al.* 2001) and quantum dots (Wang *et al.* 1999). For example, recent *ab initio* calculations predict that two rows of hydrogen atoms chemisorbed on selective sites *exterior* to an armchair carbon nanotube catalyses the breaking of the nearest-neighbour C–C bond of the nanotube through the concerted formation of C–H bonds, leading to the unzipping of the nanotube wall (Scudder *et al.* 2003). This remarkable hydrogen-induced unzipping mechanism lends strong support to the recent experimental observations

(Nikolaev *et al.* 1997) for the coalescence of single-walled nanotubes in the presence of atomic H.

Finally, *ab initio* calculations have been recently applied to study the dislocation core properties of isolated ⟨111⟩ screw dislocations in bcc molybdenum and tantalum (Woodward and Rao 2002). Alternatively, the Peierls–Nabarro model has come to serve as a link between atomistic and continuum approaches, by providing a means to incorporate information obtained from *ab initio* calculations directly into continuum models (Lu *et al.* 2000, 2001, 2002). The resultant approach can then be applied to problems associated with dislocation core structure and the cross slip process, which neither atomistic nor conventional continuum models could handle separately. This approach represents a combination of an *ab initio* treatment of the interactions across the slip plane and an elastic treatment of the continua on either side of the slip plane. Therefore, this approach is particularly useful for studying the interaction of impurities with dislocations when empirical potentials are either not available or not reliable to deal with such multielement systems. Furthermore, it allows us to study general trends in dislocation core properties and to correlate them with specific features of the underlying electronic structure. One of the many successful applications of this approach was the elucidation of the atomistic mechanism responsible for the hydrogen-enhanced local plasticity in aluminium (Lu *et al.* 2001). The γ surface for both the pure aluminium and the aluminium + hydrogen system, shown in figures 2 (*a*) and (*b*), respectively, indicate an overall reduction in energy in the presence of hydrogen. This, in turn, not only facilitates dislocation emission from the crack tip but also enhances the dislocation mobility dramatically.

2.5. *Limitations of the density functional theory*

The DFT within the LSDA and GGA generally works well for ground-state properties of moderately correlated electron systems. In some systems, however, the correlations among the electrons are stronger than might be expected from the local spin densities. This occurs in systems where the electrons partially preserve their localized atomic-like nature, as in the case for 4f (5f) states in rare-earth (actinide) atoms and sometimes for 3d states of transition-metal atoms. Typical examples include rare-earth metals, where LDA predicts a strong peak of 4f-orbital in the density of states at the Fermi level, which is in disagreement with experiment, and certain insulating transition-metal oxides, which LDA predicts to be metallic. Other systems with anomalous electronic and magnetic properties are heavy-fermion compounds, copper-oxide-based superconductors, colossal magnetoresistance manganites, etc. In the past few years, there has been a growing sense that realistic LDA-like approaches may be generalized to include the most significant correlation effects for strongly correlated electron systems of the Mott–Hubbard and charge-transfer variety (Anisimov 2000). All these approaches are based on LDA as a starting point and introduce additional terms intended to treat Coulomb correlations between electrons. Among these are the LDA + U method (Anisimov *et al.* 1991, Lichtenstein *et al.* 1995), the self-interaction correction method (Perdew and Zunger 1981, Svane and Gunnarson 1990), the LDA + dynamical mean-field theory method (Georges *et al.* 1996) and the optimized effective potential method (Gross *et al.* 1998).

In addition, only limited information about spectroscopic properties is available from such ground-state calculations. Quasiparticle excitations, as they
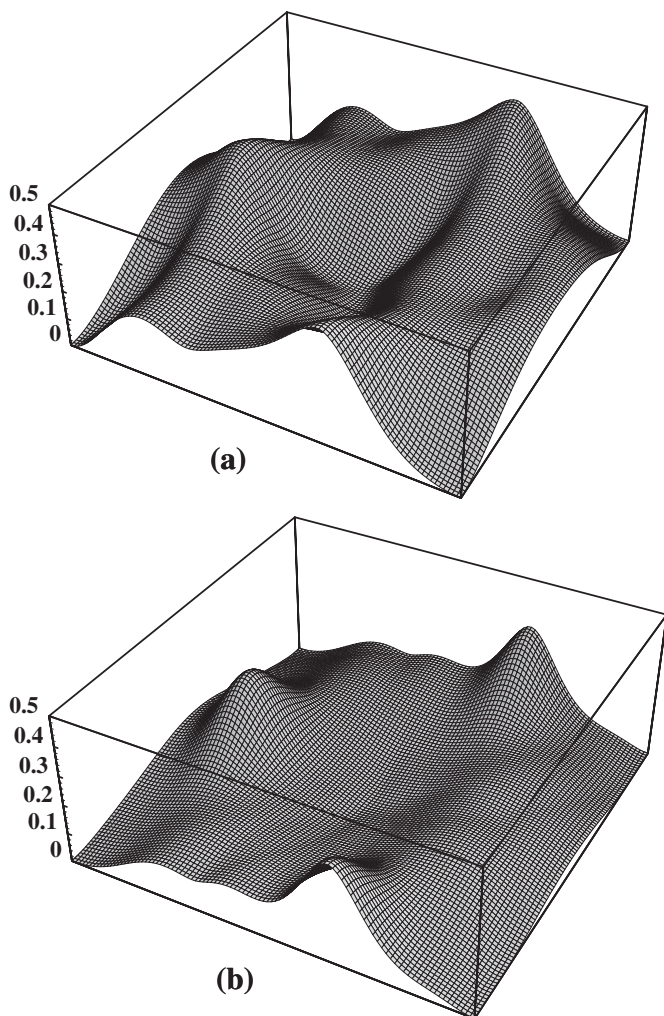
Figure 2.   The γ (J m^{-2}) surfaces for displacements along a (111) plane for (*a*) pure aluminium and (*b*) aluminium + hydrogen systems. (After Lu *et al.* (2001).)

occur in photoemission and tunnelling experiments, are not fully described by the Kohn–Sham eigenvalues. In fact, the band structures given by the LDA are often in distinct disagreement from experimental data, showing systematic deviations of band dispersions and band gaps. In addition to the failures of the band structure, two-particle excitations are not obtained with satisfying accuracy either. In particular, optical spectra, which correspond to electron–hole excitations in the electron system, cannot be described by straightforward use of DFT or other ground-state theories. A time-dependent extension of DFT (Gross *et al.* 1998) can treat excitation properties, but so far the method is limited to finite systems. Excitation properties have also been studied using the configuration interaction method (Dykstra 1998) but, similarly, this method is very much restricted to small systems such as small molecules or clusters. A major achievement in the field was given by the GW approximation (where G stands for the Green's function and W for the screened self energy

correction) (Hedin and Lundqvist 1969) for the self-energy, which allows for a very accurate evaluation of the self-energy on the basis of results from a preceding DFT calculation. Highly accurate band structures for real materials have been obtained by this method, including bulk semiconductors, insulators, metals, semiconductor surfaces, atoms, defects and clusters, thus making the GWA a standard tool in predicting the electron quasiparticle spectrum of moderately correlated electron systems (Louie 1996).

2.6. *Connections to interatomic potentials*

The direct use of DFT methods to perform full quantum MD simulations on real materials is limited to a hundred or so atoms and a simulation time of a few picoseconds. Thus, the next step of coarse graining the problem is to remove the electronic degrees of freedom by imagining the atoms to be held together by some sort of glue or interatomic potential, thereby allowing large-scale atomistic simulations for millions of atoms and a simulation time of nanoseconds. Such simulations, employing either the embedded-atom method (EAM) type (Daw and Baskes 1983, 1984) or the Finnis–Sinclair (1984) type in metals, and the Stillinger–Weber (1985) type (Ackland and Vitek 1990) or the Tersoff (1986) type in covalent materials, have been extremely useful in investigating generic phenomena in simple systems. Empirical potentials involve the fitting of parameters to a predetermined experimental or *ab initio* database, which includes physical quantities such as the lattice constant, the elastic constants, the vacancy formation energy and the surface energy. However, at the same time, they may not provide the desired physical accuracy for many real complex materials of interest. For example, reliable interatomic potentials usually are not available for multielement materials and for systems containing substitutional or interstitial alloying impurities.

There is consequently growing need to develop more accurate interatomic potentials, *derived from QM*, that can be applied to large-scale atomistic simulations. This is especially so for directionally bonded systems, such as transition metals, and for chemically or structurally complex systems, such as intermetallics and alloys. The tight-binding (TB) (Cohen *et al.* 1994, Mehl and Papaconstantopoulos 1996) and the self-consistent-charge density functional TB MD approach is becoming widespread in the atomistic simulation community, because it allows one to evaluate both ionic and electronic properties (Frauenheim *et al.* 1998). The success of TB MD stands on a good balance between the accuracy of the physical representation of the atomic interactions and the resulting computational cost. TB MD implements an empirical parametrization of the bonding interactions based on the expansion of the electronic wave functions on a very simple basis set. Recently, novel analytic bond-order potentials have been derived for atomistic simulations by coarse graining the electronic structure within the orthogonal two-centre TB representation (Pettifor 1989, Pettifor and Oleinik 1999, 2000). Quantum-based interatomic potentials for transition metals that contain explicit angular-force contributions have been developed from first-principles, DFT-based generalized pseudopotential theory (Moriarty 1988, Moriarty and Widom 1997, Moriarty *et al.* 2002).

2.7. *Linear scaling electronic structure methods*

A wealth of efficient methods, the so-called order-N (O(N)) methods, have recently been developed (Yang 1991, Li *et al.* 1993, Ordejon *et al.* 1993, Goedecker 1999) for

calculating the electronic properties of an extended system that require an amount of computation that scales linearly with the size N of the system. On the contrary, the widely used Carr–Parinello (1985) MD methods and various minimization techniques (Payne *et al.* 1992) are limited because the computational time required scales as N^3, where N can be defined to be the number of atoms or the volume of the system. Thus, these methods are an essential tool for the calculation of the electronic structure of large systems containing many atoms. Most $O(N)$ algorithms are built around the density matrix or its representation in terms of Wannier functions (Marzari and Vanderbilt 1997) and take advantage of its decay properties. To obtain linear scaling, one has to cut off the exponentially decaying quantities when they are small enough. This introduces the concept of a localization region. Only inside this localization region is the quantity calculated; outside it is assumed to vanish. For simplicity the localization region is usually taken to be a sphere, even though the optimal shape might be different.

$O(N)$ methods have become as essential part of most large-scale atomistic simulations based on either TB or semiempirical methods. The use of $O(N)$ methods within DFT methods is not yet widespread. All the algorithms that would allow us to treat very large basis sets within DFT have certain shortcomings. An illustration of the wide range of areas where $O(N)$ methods have made possible the study of systems that were too large to be studied with conventional methods include the 90° partial dislocation in silicon (Nunes *et al.* 1996), the surface reconstruction properties of silicon nanobars (Ismail-Beigi and Arias 1999), the molecular crash simulation of C_{60} fullerenes colliding with a diamond surface (Galli and Mauri 1994), the irradiation-mediated knockout of carbon atoms from carbon nanotubes (Ajayan *et al.* 1998) and the geometric structure of large bimolecules (Lewis *et al.* 1997).

§3. Large-scale atomistic simulations

The *ab initio* methods presented in §2 are rigorous but limited by present-day computers to systems containing a few hundred atoms at most. Such methods serve two important purposes. Firstly, they provide direct information on the response of materials to external environments (e.g. force and temperature). Secondly, they also generate a database of properties that can be used to construct effective (*empirical*) interatomic potentials. To determine the properties of an ensemble of atoms larger than can be handled by computational QM, the description of the atomic interactions must be approximated. The MD method is developed to enable studies of the properties of material volumes containing millions to billions of atoms with effective interatomic potentials. The basic idea is to eliminate all electronic degrees of freedom and to assume that the electrons are *glued* to the nuclei. Thus, the interaction between two atoms is represented by a *potential* function that depends on the atomic configuration (i.e. relative displacement) and the local environment (i.e. electrons). Based on the electronic structure database, or alternatively using experimental measurements of specific properties, approximate effective potentials can be constructed. According to classical Newtonian mechanics, the dynamic evolution of all atoms can be fully determined by numerical integration. In principle, once the positions and velocities of atoms in the finite ensemble within the simulation cell are known, all thermodynamic properties can be readily extracted. The implementation and practice of MD simulations are more involved than the conceptual description alluded to here. A successful simulation depends on three

major factors:

  (i) the computational implementation of the MD method;
 (ii) the construction of accurate interatomic potentials; and
(iii) the analysis of massive data resulting from computer simulations.

In the following, we briefly present key concepts in the numerical implementation of MD methods and then discuss salient aspects of the last two topics.

3.1. *The molecular dynamics method*

The MD method is about the simultaneous motion and interaction of atoms (or molecules). The following presentation outlines the basic equations and principles of the MD method, before proceeding to its numerical implementation. The dynamic evolution of the system is governed by classical Newtonian mechanics, where for each atom *i*, the equation of motion is given by

$$M_i \frac{\mathrm{d}^2 \mathbf{R}_i}{\mathrm{d}t^2} = \mathbf{F}_i = -\nabla_{\mathbf{R}_i} \Phi, \tag{12}$$

which is derived from the classical Hamiltonian of the system:

$$\mathcal{H} = \sum \frac{M_i V_i^2}{2} + \Phi. \tag{13}$$

An atom of mass M_i moves as a rigid particle at the velocity V_i in the effective potential $\Phi(\mathbf{R}_i)$ of other particles. The atomic force $\mathbf{F}_i$ is obtained as the negative gradient of the effective potential: $\mathbf{F}_i = -\nabla_{\mathbf{R}_i} \Phi$. Solving these second-order ordinary differential equations for all the atoms in a simulation cell may appear simplistic. A physical simulation involves proper selection of numerical integration scheme, employment of appropriate boundary conditions, and stress and temperature control to mimic physically meaningful thermodynamic ensembles.

A proper numerical integration scheme should be both numerically stable and computationally efficient. The numerical integration of the equations of motion is performed by either explicit or implicit methods. The simple Euler scheme is not appropriate for MD simulations because it lacks numerical stability. In the explicit Verlet's leapfrog method, the equation of particle motion is split into two first-order equations:

$$\frac{\mathrm{d}\mathbf{R}_i}{\mathrm{d}t} = \mathbf{V}_i, \quad \frac{\mathrm{d}\mathbf{V}_i}{\mathrm{d}t} = \frac{\mathbf{F}_i}{M_i}. \tag{14}$$

Based on the half-step leapfrog method, this set of equations is converted for small time increment δt to

$$\mathbf{R}_i(t + \delta t) = \mathbf{R}_i(t) + \delta t \mathbf{V}_i(t + \tfrac{1}{2}\delta t), \tag{15}$$

$$\mathbf{V}_i(t + \tfrac{1}{2}\delta t) = \mathbf{V}_i(t - \tfrac{1}{2}\delta t) + \delta t \frac{\mathbf{F}_i}{M_i}. \tag{16}$$

The Verlet algorithm is very popular in MD simulations because it is stable and memory-efficient and allows for a reasonably large time-step. Another popular implicit integration method for MD simulations is the predictor-corrector scheme, in particular the Gear (1971) algorithm.

Similar to the numerical integration scheme, proper use of boundary conditions is crucial to obtaining a physically meaningful MD simulation. The boundary of a simulation cell is usually within close proximity to each simulated atom, because of limited computational power. Simulations for 10^9 atoms represent the upper limit of computation today with simple interatomic potentials (for examples see Lennard-Jones (1924)). Total simulation times are typically less than 10 ns as a result of short integration time steps in the femtosecond range. The boundary of a simulation cell may not be a real physical interface, rather it is merely the interface of the simulation cell and its surroundings. Various boundary conditions are used in mechanics simulations. Since dislocations dominate the linking of nanomechanics and micromechanics, the following presentation focuses on those boundary conditions relevant to the long-range strain fields of dislocations. These include the rigid boundary condition (Kuramoto *et al.* 2000), the periodic boundary condition (Kido *et al.* 2000), the direct linking of atomistic and continuum regions (Oritz and Phillips 1999), and the flexible (Green's function) boundary condition (Sinclair 1971, Hoagland 1976, Woo and Puls 1976, Rao *et al.* 1998).

The rigid boundary condition is probably the simplest. According to this condition, boundary atoms are fixed during MD simulations. For simple dislocation configurations, the initial strain field, say of an infinitely long straight dislocation, is imposed on all atoms in the simulation cell. During subsequent simulations, boundary atoms are not allowed to relax. This rigidity of the boundary condition naturally leads to inaccuracy in the simulation results. The inaccuracy also exists in the application of the periodic boundary condition. Here, the simulation cell is repeated periodically to fill the entire space. When one dislocation is included in the simulation cell, an infinite number of its images are included in the entire space because of the infinite repetition of the simulation cell. Modifications have been made to include dipoles in a simulation cell, and to subtract the image effects. However, it is difficult to extend such modifications to general dislocation configurations, which are more complex.

In contrast with the two approximate boundary conditions, the direct linking and flexible boundary conditions can be rigorous. The direct linking of atomistic and continuum region (Qrtiz and Phillips 1999) aims at seamless bridging of two different scales. This idea has been extended to include direct linking of atomistic and electronic structure regions (Abraham *et al.* 1987). Once numerically robust and efficient, the direct linking scheme should be the most reliable and desirable boundary condition. At present, the other rigorous scheme, namely the flexible boundary condition, is more commonly used in dislocation simulations. This scheme has gone through the implementations in two and three dimensions. In the 1970s, the two-dimensional version of this boundary condition was implemented (Sinclair 1971, Hoagland *et al.* 1976, Woo and Puls 1976), allowing the simulation of a single straight dislocation. In these simulations, periodic boundary conditions are applied along the dislocation line. On the plane normal to the dislocation line, the simulation cell is divided into three regions. The innermost region is the MD region, in which atoms follow Newtonian dynamics according to the MD method outlined earlier. The intermediate region is the flexible region (or Green's function region), in which the force on each atom is calculated and then used to generate displacements of all atoms in the simulation cell according to the Green's function of displacement. Since a periodic boundary condition is applied along the dislocation line, line forces are calculated in the flexible region, making the displacement two dimensional. The

outermost region contains atoms that serve as background for force calculations in the flexible region. As a result of renewed interest in DD, flexible boundary conditions have been extended to three dimensions and are often referred to as the Green's function boundary conditions (Rao *et al*. 1998). Flexible or Green's function boundary conditions have clear advantages in allowing for full relaxation of one or a few dislocations in a simulation cell, without suffering from the artefacts of image dislocations. However, Green's function calculations are time-consuming limiting applications of the Green's function boundary condition. Recently, a tabulation and interpolation scheme (Golubov *et al*. 2001) has been developed, improving the computational efficiency by two orders of magnitude and yet maintaining the accuracy of linear elasticity. With this improvement, the Green's function boundary conditions work well for static simulations of dislocations. However, they are not applicable to truly dynamic simulations. During dislocation motion, elastic waves are emitted and, when they interact with the simulation cell boundaries or borders, they are reflected and can lead to interference or even resonance in the simulation cell. Approximate approaches (Ohsawa and Kuramoto 1999, Cai *et al*. 2000) have been proposed to damp the waves at the boundary, but a fully satisfactory solution is still not available.

In addition to the numerical integration and boundary condition issues, successful MD simulations rely on the proper control of thermodynamic variables: stress and temperature. When the internal stress, as derived from interatomic interactions, is not balanced by the external stress, the simulation cell deforms accordingly. At low stress levels, this response is consistent with linear elasticity. Under general stress conditions, the formulation of Parrinello and Rahman (1981) provides a mechanism to track the deformation. In this formulation, the three vectors describing the shape of simulation cell are equated to position vectors of three imaginary atoms. The driving force of the imaginary atoms is the imbalance of internal and external stresses, and the movements of those atoms are governed by Newton's equations of motion. This formulation works well when an equilibrated stress state is sought. However, large fluctuations in the deformation are unavoidable according to the Parrinello–Rahman method; the lattice constant may fluctuate by about 0.5% around its equilibrium value. This large fluctuation may introduce artefacts in the simulation of kinetic process, and caution should be exercised. The other thermodynamic variable, temperature, is controlled through the kinetic energy or velocities of all atoms. Using frictional forces to add or subtract heat, the Nose (1984)–Hoover (1985) method provides a mechanism of controlling the temperature of a simulation cell. When the temperature is uniform in space and constant in time, little difficulty exists. When the temperature changes, however, the strength of the frictional force dictates the transition time. A fast transition is usually necessary because of the short time reachable in MD simulations. However, such a transition is generally too fast compared with realistic processes. This transition problem exists in other temperature control mechanisms, such as velocity scaling or introduction of random forces.

3.2. *Interatomic potentials*

A proper MD method is a necessary condition for physically meaningful simulations. However, the method says nothing about how simulated atoms interact with each other. The latter aspect is solely determined by prescribed interatomic potentials and is more crucial in obtaining physically meaningful results. In general, there is a compromise between the potential rigorousness and the computational

efficiency. For high computational efficiency, pair potentials, such as the Lennard-Jones (1924) and the Morse (1929) potentials are used. With the increasing demand on accuracy and available computational power, many-body potentials such as the Finnis–Sinclair (1984) potential and the EAM (Daw and Baskes 1983, 1984) have been commonly used. In the same category are effective-medium and glue models (Ercolessi *et al.* 1986). Angle-dependent potentials, which are also many-body, include the well-known Stillinger–Weber (1985) potential and the Tersoff (1986) potential. Potentials that have a QM basis, such as the generalized pseudopotential (Moriarty 1988, 1994), the bond order potential (Pettifor 1989) and the inversion method (Chen 1990, Zhang *et al.* 1997) have also been used for greater accuracy.

In general, interatomic potentials are empirical or semiempirical and thereby have fitting parameters. Simpler potentials, like the Lennard-Jones potential, have very few fitting parameters, which can be easily determined from crystal properties. On the other hand, these potentials suffer from non-transferability. Since these potentials are fitted to only a few perfect crystal properties, their applicability in studying defects is by default questionable. The more sophisticated potentials, like the EAM, particularly the force matching approach (Ercolessi and Adams 1994), use several or many defect properties to determine the EAM functions, in either analytical or tabular form. Materials properties, of either the perfect crystal or the defective structures, are obtained from *ab initio* calculations and reliable experiments. It is worth pointing out that these interatomic potentials apply for specific classes of materials. Strictly speaking, pair potentials are applicable to simulations of rare-gas behaviour. However, applications to simple metal systems can also provide qualitative guidance. The EAM type of potential has proven to be a good choice for simple metals. Applications to metals such as aluminium, copper, silver and iron (except its magnetic properties) have been very successful. The radial function form of the EAM is computationally advantageous. However, this advantage is accompanied by the inability to describe covalent systems, in which angular dependence dominates. In studying silicon, diamond, carbon and other covalent systems, angular dependent potentials, particularly the Stillinger–Weber (1985) potential, the Tersoff (1986) potential, and the bond-order potential (Pettifor 1989) are among the leading candidates. A modification of the EAM has been proposed to include the angular dependence by Baskes (1987), and applied to SiC by Huang *et al.* (1995); the modification by Pasianot *et al.* (1991) is similar.

3.3. *Applications of atomistic simulations in mechanics*

This section focuses on the applications of MD simulations to mechanics problems at the nanoscale. Furthermore, the presentation is focused on nanomechanics of point defects, dislocations and interfaces (free surfaces and grain boundaries) under applied stress. Mechanics of nanotubes have also received much attention (Dumitrica *et al.* 2003) but will not be treated here.

We start with point defects. Under stress, the normal diffusion process governed by point defect motion is polarized by the action of the stress field. The diffusion process of a point defect has been investigated by a combination of molecular statics (MS) and MD simulations. Diffusion anisotropy, which is intrinsic in low-symmetry crystals, can be enhanced by the action of stress (Woo *et al.* 2002). The stress effect on a point defect, in particular its formation and migration energies, is crucial to the analysis of kinetic processes, but this is usually not a first-order effect in nanomechanics.

The second type of defect, namely dislocations, has received much attention in nanomechanics. In the following, we analyse the nanomechanics of dislocations in terms of statics and dynamics. In dislocation statics, one is primarily concerned with the final core structure of one or a few dislocations. The discovery of threefold symmetry of a screw dislocation in bcc metals is a beautiful example of applications of dislocation statics (Vitek 1974, Xu and Moriarty 1996, Duesbery and Vitek 1998, Rao and Woodward 2001, Wang *et al.* 2001a). During dislocation motion, many processes are involved; dislocation damping, dislocation intersection and dislocation clustering, just to name a few. MD simulations have provided much insight to details of kink–kink (Bulatov *et al.* 1995), dislocation–dislocation (Xu and Moriarty 1998, Zhou *et al.* 1998, Rodney and Phillips 1999, Rodney and Martin 2000), dislocation–point-defect (Justo *et al.* 2001) and dislocation–grain boundary (Ortiz and Phillips 1999) interaction processes. In the transonic velocity range, MD simulations have revealed new mechanisms and possibilities. Gumbsch and Gao (1999) demonstrated that transonic dislocations are possible. Later, it was shown that a transonic dislocation can cross the sound barrier back and forth (Shi *et al.* 2002, Li and Shi 2002). The availability of fast dislocations also enabled the study of dislocation dipole stability at the atomic level (Wang *et al.* 2001b), to serve as direct confirmation of the elasticity analysis (Huang *et al.* 1999). As shown in figure 3, the simulations further reveal that a dislocation dipole can be stabilized even with an overshoot; which is a direct result of the finite speed of wave propagation.

In addition to the static and dynamic behaviours, the nucleation of dislocations is another important area of nanomechanics. MD simulations have been applied in studies of dislocation emission from crack tips (Bulatov *et al.* 1998), nucleation of dislocations from surfaces (Liu *et al.* 2002a, b) and dislocation nucleation in nanocrystals (Van Swygenhoven *et al.* 2001, Cleri *et al.* 1997). One of the challenging
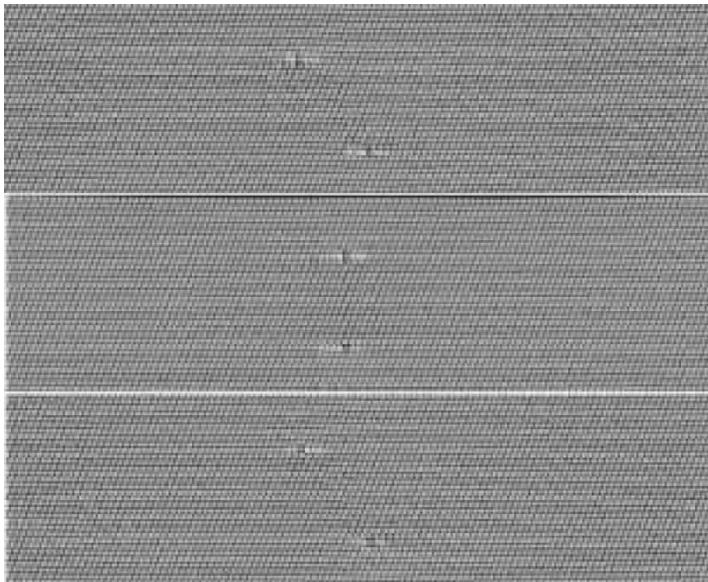

Figure 3.   Three snapshots of a pair of edge dislocations in a dipole configuration, showing the stabilization of a dipole despite the dynamic overshoot. (After Wang *et al.* (2001b).)

problems that remain is treating the interaction of boundaries with elastic waves emitted by a moving dislocation. It is usually necessary to separate this interaction from the true dynamics of a moving dislocation. Several recent studies have attempted to damp out boundary effects (Ohsawa and Kuramoto 1999, Cai *et al.* 2000, Wang *et al.* 2001b), but the problem is still not satisfactorily solved.

   Similar to point defects and dislocations, interfaces (e.g. surfaces and grain boundaries) respond to mechanical loading. In connection with dislocations, surfaces are nucleation sites. Although grain boundaries also have this function, they are more crucial in facilitating or blocking other deformation mechanisms. Using a multiscale approach, Ortiz and Phillips (1999) have studied the interaction of a dislocation with a grain boundary. It is well established that grain boundaries serve as barriers to dislocation motion and give rise to the Hall–Petch effect, which shows that the strength of materials is inversely proportional to the square root of its grain size. However, for materials with grain sizes below a few nano-metres, materials strength begins to decline when the grain size decreases (Van Swygenhoven 2002). On the nanoscale, grain boundaries facilitate mechanical deformation through grain-boundary sliding. At the same time, the grain structure evolves as well, and this evolution in turn affects the mechanical deformation. Taking thin slices of polycrystalline nanograins, Schonfelder *et al.* (1997, 1999) have investigated mechanisms of grain-boundary interactions. As shown in figure 4, high-angle grain boundaries shrink, and quadruple junctions are not stable and are reduced to triple junctions.
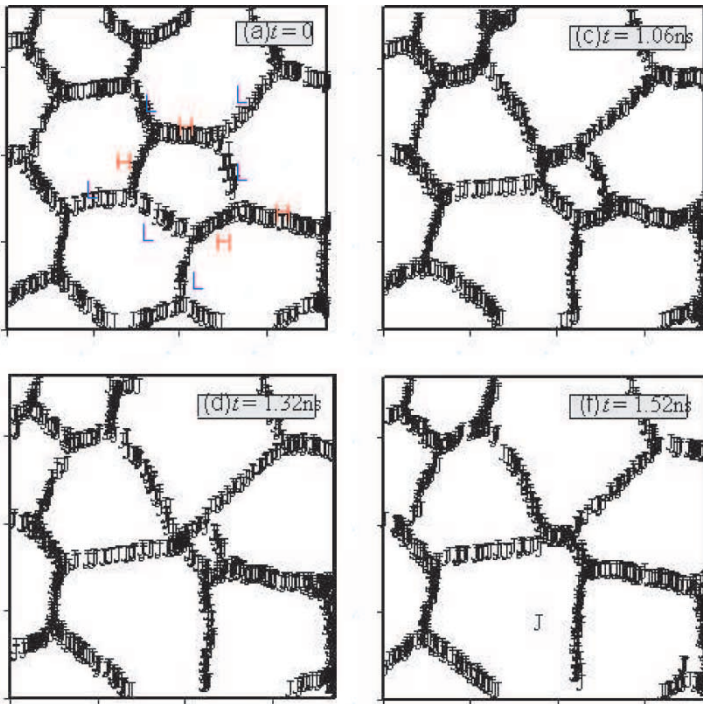

Figure 4.   MD simulations of nanoscale grain-boundary evolution under externally applied stress. Time is labelled in each frame. H indicates a high-angle grain boundary, and L a low-angle grain boundary. (After Schonfelder *et al.* (1999), courtesy of D. Wolf.)

The presentation in this section has so far focused on physical aspect of nanomechanics. This physical understanding relies on identification and analysis of point defects, dislocations and interfaces. Among the three types of defects, dislocations are probably the most difficult to characterize or identify. The identification is straightforward if the dislocation configuration is simple. One may rely on constructing the Burgers circuit, using disregistry functions (such as the centro-symmetry parameter), or tracking the atomic level stress or energy (Hamilton 1997, Hamilton *et al*. 1999, Chang *et al*. 2001). When complex dislocation configurations are involved, unambiguous identification of dislocation cores may require a combination of several of the methods discussed above. There is still not a universal characterization method that fits all purposes.

Before closing this section, let us examine the type of information that can be passed on to larger length scales. Here, we consider three types of defects: point defects, dislocations, and interfaces (particularly grain boundaries). Information on defect energetics and diffusion tensors obtained by MD or MS simulations provide parameters to rate equations and MC methods describing microstructure evolutions. Dislocation energetics (e.g. kink-pair formation energy and dislocation core energy), Peierls stress and dislocation mobility are used as input parameters in dislocation dynamics simulations at the mesoscopic level (Bulatov *et al*. 1998). Finally, dislocation nucleation conditions and interaction mechanisms with other defects (dislocations or grain boundaries or other interfaces) from MD simulations are critical to DD simulations.

§4. Mesomechanics and the role of defects

Studies of the mechanical behaviour of materials on a length scale larger than what can be handled by direct atomistic simulations and smaller than what is possible with continuum mechanics represent particular difficulties. Two complementary approaches have been advanced to model the mechanical behaviour in this *mesoscopic* length scale. The first approach, that of DD, was initially motivated by the need to understand the origins of heterogeneous plasticity and pattern formation. An early variant of this approach (the cellular automata) was first developed by Lepinoux and Kubin (1987), and that was followed by the proposal of DD (Ghoniem and Amodeo 1988a, b, Amodeo and Ghoniem 1990a, b). In these early efforts, dislocation ensembles were modelled as infinitely long and straight in an isotropic infinite elastic medium. The method was further expanded by a number of researchers (Guluoglu *et al*. 1989, Lubarda and Needleman 1993, Barts and Carlsson 1995, 1998, Wang and LeSar 1995), with applications demonstrating simplified features of deformation microstructure.

The second approach to mechanical models on the *mesoscale* has been based on SM methods (Walgraef and Aifantis 1985, Gregor and Kratochvil 1990, Kratochvil and Saxlova 1992, Saxlova *et al*. 1997, Hähner *et al*. 1998, Zaiser *et al*. 1998, El-Azab 2000, Thomson *et al*. 2002). In these developments, evolution equations for statistical averages (and possibly for higher moments) are to be solved for a complete description of the deformation problem. The main challenge in this regard is that, unlike the situation encountered in the development of the kinetic theory of gases, the topology of interacting dislocations within the system must be included (El-Azab 2000). In the following, we review the main theoretical and computational advances in mesomechanics.

4.1 *Defect mechanics on the mesoscale*

So far, we have dealt with simulations of the atomistic degrees of freedom, where various levels of approximations have been introduced to give tractable and yet rigorous solutions of the equations of motion. At present, MD simulations in the nanomechanics area have been limited to sizes less than 100 nm, and time scales less than tens of nanoseconds. Accurate descriptions of material volumes in the micrometre range where continuum descriptions break down are not yet attainable with MD techniques. Fortunately, however, the mechanical behaviour of materials is primarily determined by topological defects, ranging in size from atomic dimensions (i.e. vacancies and interstitials), to line defects (e.g. dislocations) and surface defects (e.g. voids, bubbles, cracks and grain boundaries). The mechanics problem can thus be greatly simplified if all atomic degrees of freedom were adiabatically eliminated (similar to the Born–Oppenheimer approach), and only those associated with defects are retained. Because the motion of all atoms in the material is not relevant, and only atoms around defects determine the mechanical properties, one can just follow material regions around defects. Since the density of defects is many orders of magnitude smaller than the atomic density, two useful results emerge. Firstly, defect interactions can be accurately described by long-range elastic forces transmitted through the atomic lattice. Secondly, the number of degrees of freedom required to describe their topological evolution is many orders of magnitude smaller than those associated with atoms. These observations have been instrumental in the emergence of mesomechanics on the basis of defect interactions, which has its roots in the pioneering work of Eshelby (1957), Kröner (1958a, b), Mura (1968, 1982) and Kossevich (1999). Because of the many computational advances during the past two decades, the field has steadily moved from conceptual theory to practical applications. While early research in defect mechanics focused on the nature of the elastic field arising from defects in materials, recent computational modelling has shifted the emphasis to defect ensemble evolution. We review here some of the more popular methods of computational mesomechanics that have been developed during the last two decades. These are the MC, the discrete DD and the SM methods.

4.2. *The kinetic Monte Carlo method*

The MC technique is a statistical method for solving deterministic or probabilistic problems, by sampling from random distributions utilizing concepts of probability theory. A simple method for generation of random numbers according to a given distribution function is the inversion method (James 1990). In this approach, if the distribution function is normalized to obtain a probability density function (PDF) $p(x)$ we can determine the probability that the random variable x' is less than an arbitrary x by integrating the PDF from the minimum value to x. The integral of the PDF is called the cumulative distribution function (CDF) $C(x)$. When the CDF is equated to a uniformly distributed random number ξ, that is $C(x) = \xi$, the resulting solution for x gives the desired distribution function. If the PDF $p(x)$ cannot be easily inverted analytically, sampling can be performed by the Von Neumann rejection technique. Another method to achieve the same result is known as importance sampling, and is a combination of the previous two methods. Here, we replace the original distribution function $p(x)$ by an approximate form $\tilde{p}(x)$ for which the inversion method can be applied. Then we obtain trial values for x with the inversion technique following $p'(x)$. Finally, we accept trial

values with the probability proportional to the weight w, given by $w = p(x)/\tilde{p}(x)$. The rejection technique has been shown to be a special case of importance sampling, where $p'(x)$ is a constant (James 1980).

When the configurational space is discrete, and all rates at which they occur can be determined, we can choose and execute a single change to the system from the list of all possible changes at each MC step. This is the general idea of the kinetic Monte Carlo (KMC) method (Doran 1970, Beeler 1982, Heinisch 1995). The main advantage of this approach is that it can take into account simultaneous different microscopic mechanisms and can cover very different time scales.

First, we tabulate the rates at which each event will occur anywhere in the system: r_i. The probability of event occurrence is defined as the rate at which the event takes place relative to the sum of all possible event rates. Once an event is selected, the system is changed, and the list of events that can occur at the next step is updated. Hence, one event denoted by m is randomly chosen from all of the M events that can possibly occur, as follows:

$$\sum_{i=0}^{m-1} r_i \bigg/ \sum_{i=0}^{M} r_i < \xi < \sum_{i=0}^{m} r_i \bigg/ \sum_{i=0}^{M} r_i, \qquad (17)$$

where r_i is the rate at which event i occurs ($r_0 = 0$) and ξ is a random number uniformly distributed in the range [$\in (0, 1)$].

After a particular event is selected, the table of all possible events is updated (Battaile and Srolovitz 1997). In the Metropolis *et al.* (1953) scheme, a fixed time increment is chosen such that at most one event occurs during a time step. However, this approach is inefficient since, in many time steps, no events will happen. An alternative technique, known as the *n*-fold way algorithm and introduced by Bortz *et al.* (1975), ensures that one event occurs somewhere in the system. Thus, the time increment itself, dt, corresponding to each step is variable, because it depends on the corresponding event probability, as

$$\Delta t = -\ln(\xi) \bigg/ \sum_{i=1}^{M} r_i. \qquad (18)$$

This method is particularly useful in cases where the events occur on very different time scales, and the fastest events are only possible in certain rare situations. For example, under high-energy ion or neutron irradiation conditions, self-interstitial atom (SIA) clusters are produced by atomic collisions in the crystal. They can be represented as rigid small prismatic dislocation loops that can migrate randomly in the drift field of dislocations (Osetsky *et al.* 2000). The results of KMC simulations shown in figure 5 (Ghoniem *et al.* 2002), illustrate the stages of SIA motion and clustering in the stress field of dislocations.

4.3. *Dislocation dynamics*

Since it was first introduced in the mid-1980s independently by Lepinoux and Kubin (1987) and by Ghoniem and Amodeo (1988a), DD has now become an important computer simulation tool for the description of plastic deformation on the microscale and mesoscale (i.e. the size range of a fraction of a micrometre to tens of micrometres). The method is based on a hierarchy of approximations that enable the solution of relevant problems with today's computational resources.
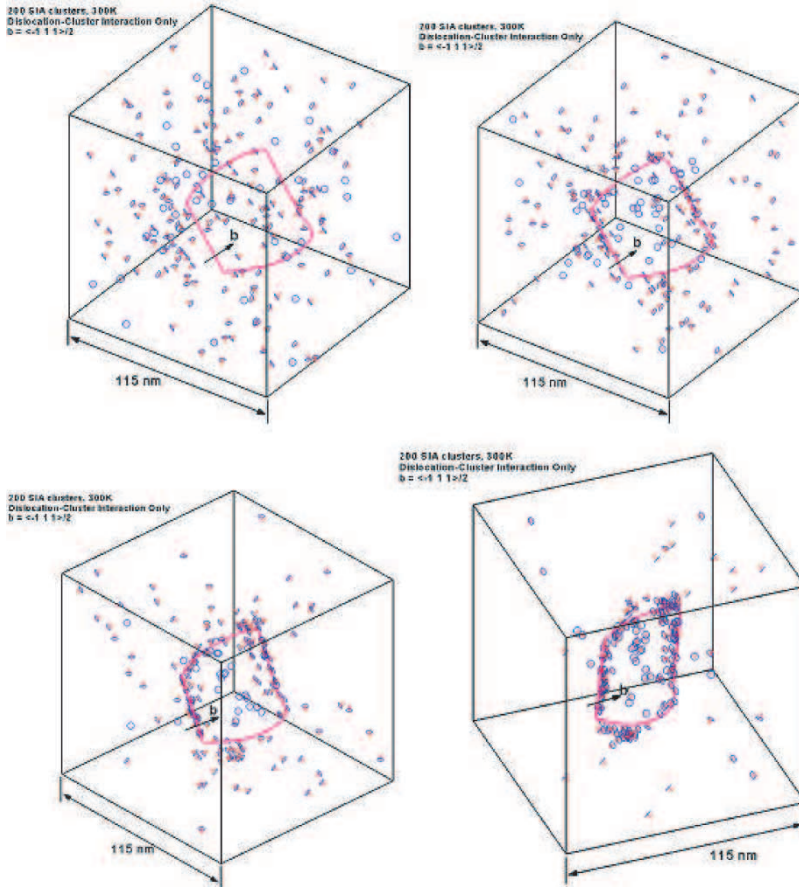
Figure 5.   Results of KMC simulations for SIA cluster agglomeration and interaction near dislocation segments. The sequence of figures are for $t = 0$, 0.043, 0.189 and 0.416 ns.

In its early versions, the collective behaviour of dislocation ensembles was determined by direct numerical simulations of the interactions between infinitely long straight dislocations (Lepinoux and Kubin 1987, Amodeo and Ghoniem 1988, 1990a, b, 1991, Ghoniem and Amodeo 1988a, 1990, Guluoglu *et al.* 1989, Ghoniem 1992, Groma and Pawley 1993, Lubarda and Needleman 1993, Barts and Carlsson 1995, 1998, Wang and LeSar 1995). Recently, several research groups extended the DD methodology to the more physical, yet considerably more complex three-dimensional simulations. The method can be traced back to the concepts of internal stress fields and configurational forces. The more recent development of three-dimensional lattice DD by Kubin and co-workers (Canova *et al.* 1992, DeVincre and Condat 1992, DeVincre *et al.* 1992, Kubin and Canova 1992, Kubin *et al.* 1992, Kubin 1993, DeVincre and Kubin 1994) has resulted in greater confidence in the ability of DD to simulate more complex deformation microstructure. More rigorous formulations of three-dimensional DD have contributed to its rapid development and applications in many systems (Hirth *et al.* 1996, Schwarz and Tersoff 1996, Schwartz 1997, Schwarz and Le Goues 1997, Rhee *et al.* 1998, Zbib *et al.* 1998, Ghoniem and Sun 1999, Ghoniem 1999, Ghoniem *et al.* 2000, 2001).

When the Green's functions are known, the elastic field of a dislocation loop can be constructed by a surface integration. The starting point in this calculation is the displacement field in a crystal containing a dislocation loop, which can be expressed as (Volterra 1907, Mura 1968)

$$u_i(\mathbf{x}) = -\int_S C_{jlmn} b_m \frac{\partial}{\partial x'_l} G_{ij}(\mathbf{x}, \mathbf{x}') n_n \, \mathrm{d}S(\mathbf{x}'), \qquad (19)$$

where C_{jlmn} is the elastic constants tensor, $G_{ij}(\mathbf{x}, \mathbf{x}')$ are the Green's functions at $\mathbf{x}$ due to a point force applied at $\mathbf{x}'$, S is the surface capping the loop, n_n is a unit normal to S and b_m is the Burgers vector. The elastic distortion tensor $u_{i,j}$ can be obtained from equation (19) by differentiation. The symmetric part of $u_{i,j}$ (elastic strain tensor) is then used in the stress–strain relationship to find the stress tensor at a field point $\mathbf{x}$, caused by the dislocation loop.

In an elastically isotropic and infinite medium, a more efficient form of equation (19) can be expressed as a line integral performed over the closed dislocation loop (deWit 1960):

$$u_i = -\frac{b_i}{4\pi} \oint_C A_k \mathrm{d}l_k + \frac{1}{8\pi} \oint_C \left(\epsilon_{iki} b_i R_{,pp} + \frac{1}{1-\nu} \epsilon_{kmn} b_n R_{,mi} \right) \mathrm{d}l_k, \qquad (20)$$

where μ and ν are the shear modulus and Poisson's ratio respectively, $\mathbf{b}$ is the Burgers vector with the Cartesian components b_i, and the vector potential $A_k(\mathbf{R}) = \epsilon_{ijk} X_i s_j / R(R + \mathbf{R} \cdot \mathbf{s})$ satisfies the differential equation $\epsilon_{pik} A_{k,p}(\mathbf{R}) = X_i R^{-3}$, where $\mathbf{s}$ is an arbitrary unit vector. The radius vector $\mathbf{R}$ connects a source point on the loop to a field point, as shown in figure 6, with Cartesian components R_i, successive partial derivatives $R_{,ijk....}$ and magnitude R. The line integrals are carried along the closed contour C defining the dislocation loop, of differential arc length $\mathrm{d}\mathbf{l}$ with components $\mathrm{d}l_k$.

Consider the virtual motion of the dislocation loop. The mechanical power during the virtual motion is composed of two parts:

(i) the change in the elastic energy stored in the medium upon loop motion under the influence of its own stress (i.e. the change in the loop self-energy):

(ii) the work done on moving the loop as a result of the action of external and internal stresses, excluding the stress contribution of the loop itself.
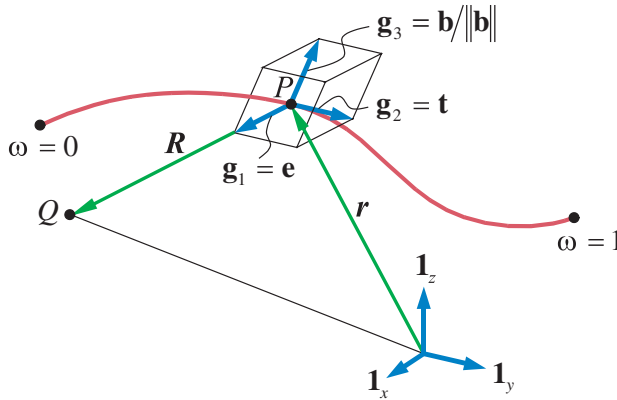

Figure 6. Parametric representation of dislocation segments. (After Ghoniem *et al.* (2001).)

These two components constitute the Peach–Koehler (1950) work . The main idea of DD is to derive approximate equations of motion from the principle of virtual power dissipation of the second law of thermodynamics (Ghoniem *et al.* 2001), by finding virtual Peach–Kohler forces that would result in the simultaneous displacement of all dislocation loops in the crystal. A major simplification is that this many-body problem is reduced to the *single-loop problem*. In this simplification, instead of moving all the loops simultaneously, they are moved sequentially, with the motion of each one against the collective field of all other loops. The approach is reminiscent of the single-electron simplification of the many-electron problem in QM.

In finite systems, Green's functions are not invariant under $\mathbf{x} \to \mathbf{x}'$ transformations, and they become functions of both source and field points (i.e. not functions of the radius vector alone any more). This does not allow the reduction of the surface integral of equation (20) to a simpler line integral by application of the Stokes theorem. In addition, closed-form solutions for Green's functions and their spatial derivatives are available only for elastically isotropic materials. Therefore, rigorous three-dimensional DD in finite anisotropic systems (e.g. thin films and quantum dots) are very demanding and have not yet been implemented.

If the material is assumed to be elastically isotropic and infinite, a great reduction in the level of required computations ensues. Firstly, surface integrals can be replaced by line integrals along the dislocation. Secondly, Green's functions and their derivatives have analytical solutions. Thus, the starting point in most DD simulations so far is a description of the elastic field of dislocation loops of arbitrary shapes by line integrals of the form proposed by deWit (1960) :

$$
\sigma_{ij} = \frac{\mu}{4\pi} \oint_C \left(\frac{1}{2} R_{,mpp}(\epsilon_{jmn}\mathrm{d}l_i + \epsilon_{imn}\mathrm{d}l_j) + \frac{1}{1-\nu}\epsilon_{kmn}(R_{,ijm} - \delta_{ij}R_{,ppm})\mathrm{d}l_k \right), \tag{21}
$$

where μ and ν are the shear modulus and Poisson's ratio respectively. The line integral is discretized, and the stress field of dislocation ensembles is obtained by a summation process over line segments.

Recently, Ghoniem and Sun (1999) and Ghoniem *et al.* (2001) have shown that, if dislocation loops are discretized into curved parametric segments, one can obtain the field by numerical integration over the scalar parameter that represents the segment. One of these segments is described by a parameter ω that varies, for example, from 0 to 1 at end nodes of the segment. The segment is fully determined as an affine mapping on the scalar interval $\in [0, 1]$, if we introduce the tangent vector $\mathbf{T}$, the unit tangent vector $\mathbf{t}$ and the unit radius vector $\mathbf{e}$, as follows: $\mathbf{T} = \mathrm{d}\mathbf{l}/\mathrm{d}\omega$, $\mathbf{t} = \mathbf{T}/|\mathbf{T}|$, $\mathbf{e} = \mathbf{R}/R$. Let the Cartesian orthonormal basis set be denoted by $\mathbf{1} \equiv \{\mathbf{1}_x, \mathbf{1}_y, \mathbf{1}_z\}$, $\mathbf{I} = \mathbf{1} \otimes \mathbf{1}$ as the second-order unit tensor, and $\otimes$ denotes the tensor product. Now define the three vectors ($\mathbf{g}_1 = \mathbf{e}$, $\mathbf{g}_2 = \mathbf{t}$, $\mathbf{g}_3 = \mathbf{b}/|\mathbf{b}|$) as a covariant basis set for the curvilinear segment, and their contravariant reciprocals as $\mathbf{g}^i \cdot \mathbf{g}_j = \delta_j^i$, where δ_j^i is the mixed Kronecker delta and $V = (\mathbf{g}_1 \times \mathbf{g}_2) \cdot \mathbf{g}_3$ the volume spanned by the vector basis, as shown in figure 6.

The differential stress field is given by

$$
\frac{\mathrm{d}\boldsymbol{\sigma}}{\mathrm{d}\omega} = \frac{\mu V|\mathbf{T}|}{4\pi(1-\nu)R^2} \left[(\mathbf{g}^1 \otimes \mathbf{g}_1 + \mathbf{g}_1 \otimes \mathbf{g}^1) + (1-\nu)(\mathbf{g}^2 \otimes \mathbf{g}_2 + \mathbf{g}_2 \otimes \mathbf{g}^2) - (3\mathbf{g}_1 \otimes \mathbf{g}_1 + \mathbf{I}) \right].
$$

$$
\tag{22}
$$

Once the parametric curve for the dislocation segment is mapped on to the scalar interval $\{\omega \in [0, 1]\}$, the stress field everywhere is obtained as a fast numerical quadrature sum (Ghoniem and Sun 1999). The Peach–Kohler (1950) force exerted on any other dislocation segment can be obtained from the total stress field (external and internal) at the segment as

$$\mathbf{F}_{PK} = \sigma \mathbf{b} \times \mathbf{t}. \tag{23}$$

The total self-energy of the dislocation loop is determined by double line integrals. However, Gavazza (1976) have shown that the first variation in the self-energy of the loop can be written as a single line integral, and that the majority of the contribution is governed by the local line curvature. Based on these methods for evaluations of the interaction and self-forces, the weak variational form of the governing equation of motion of a single dislocation loop was developed by Ghoniem *et al.* (2000):

$$\int_{\Gamma} \left(F_k^t - B_{\alpha k} V_\alpha \right) \delta r_k \, |\mathrm{d}\mathbf{s}| = 0. \tag{24}$$

Here, F_k^t are the components of the resultant force, consisting of the Peach–Koehler force $\mathbf{F}_{PK}$ (generated by the sum of the external and internal stress fields), the self-force $\mathbf{F}_s$ and the osmotic force $\mathbf{F}_O$ (in case climb is also considered (Ghoniem *et al.* 2000). The resistivity matrix (inverse mobility) is $B_{\alpha k}$, V_α are the velocity vector components, and the line integral is carried along the arc length of the dislocation $\mathrm{d}\mathbf{s}$. To simplify the problem, let us define the following dimensionless parameters:

$$\mathbf{r}^* = \frac{\mathbf{r}}{a}, \quad \mathbf{f}^* = \frac{\mathbf{F}}{\mu a}, \quad t^* = \frac{\mu t}{B}.$$

Here, a is lattice constant, μ the shear modulus and t is time. Hence equation (24) can be rewritten in dimensionless matrix form as

$$\int_{\Gamma^*} \delta \mathbf{r}^{*\mathrm{T}} \cdot \left(\mathbf{f}^* - \frac{\mathrm{d}\mathbf{r}^*}{\mathrm{d}t^*} \right) |\mathrm{d}\mathbf{s}^*| = 0. \tag{25}$$

Here, $\mathbf{f}^* = [f_1^*, f_2^*, f_3^*]^{\mathrm{T}}$, and $\mathbf{r}^* = [r_1^*, r_2^*, r_3^*]^{\mathrm{T}}$, which are all dependent on the dimensionless time t^*. Following Ghoniem *et al.* (2000), a closed dislocation loop can be divided into N_s segments. In each segment j, we can choose a set of generalized coordinates q_m at the two ends, thus allowing parametrization of the form:

$$\mathbf{r}^* = \mathbf{CQ}. \tag{26}$$

Here, $\mathbf{C} = [C_1(\omega), C_2(\omega), \dots, C_m(\omega)]$, $C_i(\omega)$, $(i = 1, 2, \dots, m)$ are shape functions dependent on the parameter $(0 \leqslant \omega \leqslant 1)$, and $\mathbf{Q} = [q_1, q_2, \dots, q_m]^{\mathrm{T}}$, where q_i are a set of generalized coordinates. Substituting equation (26) into equation (25) we obtain

$$\sum_{j=1}^{N_s} \int_{\Gamma_j} \delta \mathbf{Q}^{\mathrm{T}} \cdot \left(\mathbf{C}^{\mathrm{T}} \mathbf{f}^* - \mathbf{C}^{\mathrm{T}} \mathbf{C} \frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}t^*} \right) |\mathrm{d}\mathbf{s}| = 0 \tag{27}$$

Let

$$\mathbf{f}_j = \int_{\Gamma_j} \mathbf{C}^{\mathrm{T}} \mathbf{f}^* |\mathrm{d}\mathbf{s}|, \quad \mathbf{k}_j = \int_{\Gamma_j} \mathbf{C}^{\mathrm{T}} \mathbf{C} |\mathrm{d}s|.$$

Following a similar procedure to the finite-element method, we assemble the equation of motion for all contiguous segments in global matrices and vectors as

$$\mathbf{F} = \sum_{j=1}^{N_s} \mathbf{f}_j, \quad \mathbf{K} = \sum_{j=1}^{N_s} \mathbf{k}_j;$$

then, from equation (27), we obtain

$$\mathbf{K}\frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}t^*} = \mathbf{F}. \tag{28}$$

The solution of the set of ordinary differential equations (28) describes the motion of an ensemble of dislocation loops as an evolutionary dynamic system. However, additional *protocols* or algorithms are used to treat

  (i) strong dislocation interactions (e.g. junctions or tight dipoles),
 (ii) dislocation generation and annihilation and
(iii) adaptive meshing as dictated by large curvature variations (Ghoniem *et al.* 2000).

In the parametric method (Kukta and Freund 1998, Ghoniem 1999, Ghoniem *et al.* 2000, 2001) presented above, the dislocation loop can be geometrically represented as a continuous (to second derivative) composite space curve. This has two advantages: firstly, there is no abrupt variation or singularities associated with the self-force at the joining nodes in between segments; secondly, very drastic variations in dislocation curvature can be easily handled without excessive remeshing. Other approximation methods have been developed by a number of groups. These approaches differ mainly in the representation of dislocation loop geometry, the manner by which the elastic field and self-energies are calculated, and some additional details related to how boundary and interface conditions are handled. The suitability of each method is determined by the required level of accuracy and resolution in a given application.

Generally, coarse resolution is obtained by the lattice method, developed by Kubin *et al.* (1992) and Moulin *et al.* (1997), where straight dislocation segments (either pure screw or edge in the earliest versions, or of a mixed character in more recent versions) are allowed to jump on specific lattice sites and orientations. Straight dislocation segments of mixed character in the the force method, developed by Hirth *et al.* (1996) and Zbib *et al.* (1998) are moved in a rigid-body fashion along the normal to their midpoints, but they are not tied to an underlying spatial lattice or grid. The advantage of this method is that the explicit information on the elastic field is not necessary, since closed-form solutions for the interaction forces are directly used. The differential stress method developed by Schwarz and Tersoff (1996) and Schwarz (1997) is based on calculations of the stress field of a differential straight line element on the dislocation. Using numerical integration, Peach–Koehler forces on all other segments are determined . The Brown (1967) procedure is then utilized to remove the singularities associated with the self-force calculation. The method of the phase field microelasticity (Khachaturyan 2000, Wang *et al.* 2000, 2001c) is of a different nature. It is based on the Khachaturyan–Shatalov reciprocal space theory of the strain in an arbitrary elastically homogeneous system of misfitting coherent inclusions embedded into the parent phase. Thus, consideration of individual segments of all dislocation lines is not required. Instead, the temporal and
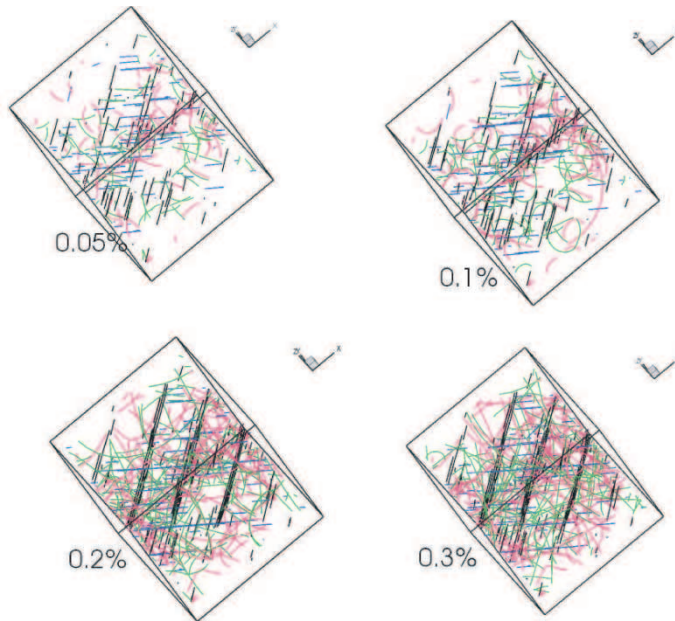
Figure 7. Results of computer simulations for dislocation microstructure deformation in copper deformed to increasing levels of strain (shown next to each microstructure). (After Wang *et al.* (2003a).)

spatial evolution of several density function profiles (fields) are obtained by solving continuum equations in Fourier space.

In the parametric DD method, the shape of loop ensembles is evolved using equations of motion for generalized coordinates representing the position, tangent and normal vectors of nodes on each loop. Figure 7 shows the results of such computations for simulation of plastic deformation in single-crystal copper at a constant strain rate of $100\,\mathrm{s}^{-1}$. The initial dislocation density of $\rho = 2 \times 10^{13}\,\mathrm{m}^{-2}$ has been divided into 300 complete loops. Each loop contains a random number of initially straight glide and superjog segments. When a generated or expanding loop intersects the simulation volume of $2.2\,\mu\mathrm{m}$ side length, the segments that lie outside the simulation boundary are periodically mapped inside the simulation volume to preserve translational strain invariance, without loss of dislocation lines. The number of nodes on each loop starts at 5 and is then increased adaptively proportional to the loop length, with a maximum number of 20 nodes per loop. The total number of degrees of freedom starts at 6000 and is increased to 24 000 by the end of the calculation. However, the number of *interacting* degrees of freedom is determined by a nearest-neighbour criterion, within a distance of $400a$ (where a is the lattice constant), and is based on a binary tree search. The dislocation microstructure is shown in figure 7 at different total strain (Wang *et al.* 2003a). It is observed that fine slip lines that nucleate at low strains evolve into more pronounced slip bundles at higher strains. The slip bundles are well separated in space, forming a regular pattern with a wavelength of approximately $1\,\mu\mathrm{m}$. Conjugate slip is also observed, leading to the formation of dislocation junction bundles and what appears to a stabilization of a cellular structure.

4.4. *Statistical mechanics*

Two of the most fascinating features of microscale plasticity are the spontaneous formation of dislocation patterns, and the highly intermittent and spatially localized nature of plastic flow. Dislocation patterns consist of alternating dislocation-rich and dislocation-poor regions usually in the micrometre range (e.g. dislocation cells, subgrains, bundles, veins, walls and channels). On the other hand, the local values of strain rates associated with intermittent dislocation avalanches are estimated to be of the order of $(1–10)\times10^6$ times greater than externally imposed strain rates (Neuhäuser 1983, Zaiser 2001). Understanding such collective phenomena is of paramount importance in two respects.

 (i) It will allow materials design approaches that are more fundamentally based to mitigate the detrimental effects of plastic deformation, fatigue and fracture on material failure.
 (ii) It will shed light on the intriguing physics of both self-organization phenomena and the behaviour of critical-state systems (e.g. avalanches and percolation).

Because of the high density of dislocations and the strong nature of interactions, direct computer simulations of inhomogeneous plastic deformation and dislocation patterns are still unattainable. Even if direct computer simulations are successful in the description of these collective phenomena, it is very desirable to obtain global kinetic or thermodynamic principles for understanding the self-organization associated with plasticity on the microscale. To attain these objectives, and to enable a direct link with continuum deformation theory, the SM approach has been advanced (Neumann 1971, Walgraef and Aifantis 1985, Ghoniem *et al.* 1990, Gregor and Kratochvil 1990, Kratochvil and Saxlova 1992, Hähner 1996, Saxlova *et al.* 1997, Zaiser and Hähner 1997, Hähner *et al.* 1998, Zaiser *et al.* 1998, El-Azab 2000, Thomson *et al.* 2002, LeSar and Richman 2003). The fundamental difficulty here is that dislocations, unlike particles, are linear objects of considerable topological complexity. Hence, when concepts of statistical mechanics and the theory of rate processes are used, some level of phenomenological description is unavoidable.

Statistical models of dislocation cells, PSBs, shear bands or other organized dislocation structures attempt to develop continuum equations for dislocation densities that contain spatial gradients that drive instabilities and self-organization. Several models have been recently advanced, and will be reviewed briefly here.

The Walgraef–Aifantis (1985) statistical model has gained interest, because of its ability to predict dislocation pattern formation under cyclic deformation. The model is formulated as a set of reaction–diffusion equations with spatial derivatives that lead to inhomogeneous dislocation distributions and will thus be considered first here. In this model, the static dislocation density, formed by the immobilized dislocations of the forest, subgrain walls or boundaries, is defined as ρ_s, and the mobile dislocation density for dislocations gliding between obstacles is defined as ρ_m. For simplicity, consider systems oriented for single slip. Hence, the mobile dislocation density ρ_m is divided into two subfamily densities representing dislocations gliding in the direction of the Burgers vector (ρ_m^+) or in the opposite direction (ρ_m^-) (with $\rho_m = \rho_m^+ + \rho_m^-$). These dislocation densities are related to the strain rate via the Orowan relation $\dot{\epsilon} = b\rho_m v_g$, where b is the length of Burgers vector, ρ_m the total mobile dislocation density and v_g the glide velocity in the primary

slip plane. Moreover, the dislocation densities are related to the internal stress by the relation

$$\sigma_i = \frac{\mu b}{2\pi\lambda} + \xi\mu b\,\rho_s^{1/2} \tag{29}$$

where μ is the shear modulus and ξ is a constant. In the last equation, the first contribution comes from obstacles such as precipitates or pre-existing walls separated by an effective spacing λ and, the second part is the contribution from the static dislocation population which also opposes dislocation motion. The internal stress σ_i reduces the effective stress σ_e defined as $\sigma_e = \sigma_a - \sigma_i$, with σ_a representing the applied stress. Finally, the glide velocity is related to the effective stress via appropriate phenomenological relations expressing the fact that individual dislocation motion is initiated when the effective stress acting on a dislocation exceeds the yield stress, in the form $v_g \propto (\sigma_e/\sigma_0)^m$. This can be put more explicitly as

$$v_g = v_0 \exp\left[-\frac{\mu}{kT}\left(\frac{\sigma_e}{\sigma_0}\right)^{-m}\right], \tag{30}$$

where σ_0 is the yield stress and $m > 1$. The essential features of DD are their mobility, which includes thermal diffusion and climb, and their mutual interaction processes. By taking into account these mechanisms, the resulting dynamic system can be written as (Walgraef and Aifantis 1985)

$$\partial_t\rho_s = -\boldsymbol{\nabla}\cdot\mathbf{J}_s + v_g\rho_m\rho_s^{1/2} - v_s d\rho_s^2 - v_g\delta\rho_m\rho_s - \beta\rho_s + v_g G(\rho_s)\rho_m,$$

$$\partial_t\rho_m^+ = -\boldsymbol{\nabla}\cdot\mathbf{J}_+ + \frac{\beta}{2}\rho_s - v_g G(\rho_s)\rho_m^+ - v_g\delta\rho_m^+(\rho_s + \rho_m^-), \tag{31}$$

$$\partial_t\rho_m^- = -\boldsymbol{\nabla}\cdot\mathbf{J}_- + \frac{\beta}{2}\rho_s - v_g G(\rho_s)\rho_m^- - v_g\delta\rho_m^-(\rho_s + \rho_m^+),$$

where δ is the characteristic separation length between dislocations for spontaneous annihilation (Essmann and Mughrabi 1979), d is the characteristic length of spontaneous dipole collapse, β is the frequency of a dislocation freeing from the forest and is proportional to $v_g/\bar{d}$ where $\bar{d}$ is the characteristic dipole destabilization length which is inversely proportional to the effective stress, and $\beta = \beta_0 v_g\sigma_e$. The different characteristic lengths introduced here, or at least their order of magnitude, may in principle be evaluated from microscopic analysis (Neumann 1971, Essmann and Mughrabi 1979). Because of mutual interactions, thermal activation and climb, the forest dislocations mobility is represented by a diffusive current $\mathbf{J} = -D_s\boldsymbol{\nabla}\rho_s$, which represents the effective diffusion within the forest. The current of mobile dislocations is taken as $\mathbf{J}_\pm = \pm\mathbf{v}_g\rho_m^\pm$ and represents the flux caused by gliding dislocations, in the present case, it is the flux caused by their edge component. Stability and numerical analyses of the previous set of equations have provided information on the conditions for formation of PSBs in fatigued specimens. It is shown that PSB formation is triggered by the clustering of dislocations or dislocation dipoles, which become finally immobile and arrange themselves in regularly spaced walls of high dislocation density (Walgraef and Aifantis 1985).

Another class of models is based on the evolution of dipolar loops, triggered by their interaction with gliding dislocations (Kratochvil and Saxlova 1992, Saxlova *et al.* 1997, Kubin and Kratochvil 2000). The proposed statistical models are of the reaction–transport type and focus on the feedback between the evolution of glide dislocations and the dipole density, as conceptualized earlier by Mughrabi (1981).

The result is the sweeping of dipole loops by screw dislocations, which initiates the formation of dislocation walls. In this approach, dipole generation and interactions play a secondary role and are introduced in an *ad hoc* and qualitative way.

To illustrate briefly the statistical approach of Kratochvil and co-workers, let us consider N loops of one type (i.e. either interstitial or vacancy), $i = 1, \ldots, N$. For the ith loop, the equation of motion can be then expressed as

$$\frac{\mathrm{d}\mathbf{r}_i}{\mathrm{d}t} = B\left(\sum_{j\neq i}^{N} \mathbf{f}_{\mathrm{int}}(\mathbf{r}_i - \mathbf{r}_j) + \mathbf{f}_{\mathrm{ext}}\right), \tag{32}$$

where B is the loop mobility, $\mathbf{f}_{\mathrm{ext}}$ is the force caused by stress gradients and glide dislocations and $\mathbf{f}_{\mathrm{int}}$ is the interaction force between dipoles. In order to obtain a continuous field description, equation (32) is multiplied by the delta function $\delta(\mathbf{r} - \mathbf{r}_i)$, and the result is differentiated with respect to $\mathbf{r}$; thus we have

$$\frac{\mathrm{d}}{\mathrm{d}\mathbf{r}}\left(\frac{\mathrm{d}\mathbf{r}_i}{\mathrm{d}t}\delta(\mathbf{r} - \mathbf{r}_i)\right) = B\frac{\mathrm{d}}{\mathrm{d}\mathbf{r}}\left[\left(\sum_{i=1}^{N} \mathbf{f}_{\mathrm{int}}(\mathbf{r}_i - \mathbf{r}_j) + \mathbf{f}_{\mathrm{ext}}\right)\delta(\mathbf{r} - \mathbf{r}_i)\right]. \tag{33}$$

By introduction of the discrete density function $\varrho(\mathbf{r}, \mathbf{r}_1, \ldots, \mathbf{r}_N) = \sum_{j\neq i}^{N} \delta(\mathbf{r} - \mathbf{r}_i)$, and taking the sum on the right-hand side of equation (33), it can be replaced by a weighted integral. Summing the above equation for all loops, one obtains

$$\frac{\mathrm{d}}{\mathrm{d}t}\varrho(\mathbf{r}) + \frac{\mathrm{d}}{\mathrm{d}\mathbf{r}}[B\mathbf{f}(\mathbf{r})\varrho(\mathbf{r})] = 0, \tag{34}$$

where

$$\mathbf{f}(\mathbf{r}) \equiv \int \mathbf{f}_{\mathrm{int}}(\mathbf{s} - \mathbf{r})\varrho(\mathbf{s})\,\mathrm{d}\mathbf{s} + \mathbf{f}_{\mathrm{ext}}. \tag{35}$$

As a first approximation it can be assumed that ϱ in equation (34) and definition (35) represents the local averaged loop density $\varrho(\mathbf{r})$ independent of detailed loop positions $\mathbf{r}_i$. Then, equation (34) represents the balance equation for the continuous distribution of loops described by the density $\varrho(\mathbf{r})$. Since the interaction force between loops f_{int} is short range, Kratochvil and co-workers utilized a Taylor series expansion of $\varrho(\mathbf{r})$ around $\mathbf{r}$ to obtain finally

$$\mathbf{f}(\mathbf{r}) = -\frac{\mathrm{d}\varrho(\mathbf{r})}{\mathrm{d}\mathbf{r}}\int(\mathbf{s} - \mathbf{r})f_{\mathrm{int}}(\mathbf{s} - \mathbf{r})\mathrm{d}\mathbf{s} + \mathbf{f}_{\mathrm{ext}}. \tag{36}$$

From the balance equation (34) and approximation (36), the problem of loop–loop interaction was shown to be represented as a diffusion equation in a drift field:

$$\frac{\mathrm{d}}{\mathrm{d}t}\varrho(\mathbf{r}) = \frac{\mathrm{d}}{\mathrm{d}\mathbf{r}}\left(\mathcal{D}\frac{\mathrm{d}\varrho(\mathbf{r})}{\mathrm{d}\mathbf{r}}\right) + \frac{\mathrm{d}}{\mathrm{d}\mathbf{r}}\mathbf{f}_{\mathrm{ext}}, \tag{37}$$

where $\mathcal{D}$ is an effective diffusion coefficient, given by

$$\mathcal{D} = B\varrho(\mathbf{r})\int \mathbf{s}\cdot\mathbf{f}_{\mathrm{int}}(\mathbf{s})\,\mathrm{d}\mathbf{s}. \tag{38}$$

When the motion of screw dislocations is coupled with equation (37), self-organized patterns of dipolar loops may be attainable, as conjectured by Mughrabi (1981).

In an attempt to generalize the statistical description of dislocations, El-Azab (2000) described the dislocation content of a slip system by introducing the distribution $\phi^{(i)}(\mathbf{x}, \mathbf{v}, \mathbf{t}, t)$. Considering only glide motion, the tangent vector is described

by a single scalar parameter: $\mathbf{t} = \mathbf{t}(\theta)$. The distribution function $\phi^{(i)}$ is defined in a generalized phase space, such that $\phi^{(i)}(\mathbf{x}, \mathbf{v}, \theta, t)\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{v}\,\mathrm{d}\theta$ is the dislocation line length contained in the phase space volume $\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{v}\,\mathrm{d}\theta$ at time t on the ith slip system. The approach of El-Azab (2000) is to consider the conservation equation of the dislocation density tensor $\boldsymbol{\alpha}$, for which dislocations on the ith slip system contribute:

$$\boldsymbol{\alpha}^{(i)}(\mathbf{x}, t) = \int_{\mathbf{v}} \int_{\theta} \mathbf{t} \otimes \mathbf{b}^{(i)} \phi^{(i)}(\mathbf{x}, \mathbf{v}, \theta, t)\,\mathrm{d}\mathbf{v}\,\mathrm{d}\theta. \tag{39}$$

The condition of compatibility of the total distortion field in the crystal is written as $\boldsymbol{\alpha} + \boldsymbol{\nabla} \times \boldsymbol{\beta}^{\mathrm{P}} = 0$ (Kröner 1981). Considering a volume Ω, this can be extended to

$$\mathcal{A} = \int_{\Omega} \boldsymbol{\alpha}\,\mathrm{d}\Omega + \int_{\Omega} \boldsymbol{\nabla} \times \boldsymbol{\beta}^{\mathrm{P}}\,\mathrm{d}\Omega = \mathbf{0}. \tag{40}$$

$\boldsymbol{\beta}^{\mathrm{P}}$ is the plastic distortion tensor. Equation (40) corresponds to the principle of invariance of the total Burgers vector. For interacting dislocations, however, El-Azab (2000) included additional source terms to the conservation equations. Thus, for the ith slip system we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{\Omega} \boldsymbol{\alpha}^{(i)}\,\mathrm{d}\Omega\right) = \int_{\Omega} \sum \mathbf{S}^{(i)}\,\mathrm{d}\Omega, \quad i = 1, N. \tag{41}$$

$\sum \mathbf{S}^{(i)}$ includes all possible tensorial sources $\mathbf{S}^{(i)}$ resulting from Burgers vector reactions and cross-slip. The time rate of variation in the dislocation density tensor in terms of the distribution function has been worked out by El-Azab (2000) as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{\mathbf{x}} \boldsymbol{\alpha}^{(i)}(\mathbf{x}, t)\,\mathrm{d}\mathbf{x}\right) = \int_{\mathbf{x}} \int_{\mathbf{v}} \int_{\theta} \mathbf{t} \otimes \mathbf{b}^{(i)}\left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \boldsymbol{\nabla} + \dot{\mathbf{v}} \cdot \boldsymbol{\nabla}_v\right)\phi^{(i)}\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{v}\,\mathrm{d}\theta, \tag{42}$$

where the terms containing $\dot{\theta}\,\partial/\partial\theta$ cancel each other. The right-hand side of equation (41) can also be represented by a phase space integral of scalar source functions

$$\int_{\Omega} \sum \mathbf{S}^{(i)}\,\mathrm{d}\Omega = \int_{\mathbf{x}} \int_{\mathbf{v}} \int_{\theta} \mathbf{t} \otimes \mathbf{b}^{(i)} \sum S^{(i)}\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{v}\,\mathrm{d}\theta. \tag{43}$$

From the last two equations, El-Azab (2000) finally obtained a set of kinetic equations, of the form

$$\left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \boldsymbol{\nabla} + \dot{\mathbf{v}} \cdot \boldsymbol{\nabla}_v\right)\phi^{(i)} = \sum S^{(i)}, \quad i = 1, N. \tag{44}$$

It is also shown that this generalized approach can be reduced to the more simplified formulation of Groma (1997) in the special case of a system of parallel edge dislocations in a single-slip configuration. The linear stability analysis of Groma (1997) shows that a homogeneous stationary solution is unstable and that density perturbations grow, leading to pattern formation.

In an effort to connect DD simulations to continuum theories, Hartley (2003) introduced the concept of a dislocation distribution vector, defined as

$$\boldsymbol{\rho}_\beta(\mathbf{t}) \equiv \mathbf{t}\rho_\beta(\mathbf{t}) \tag{45}$$

so that the magnitude of $\boldsymbol{\rho}_\beta(\mathbf{t})$ is

$$|\boldsymbol{\rho}_\beta(\mathbf{t})| = \rho_\beta(\mathbf{t}), \tag{46}$$

which is the total length per unit volume of dislocation lines with tangents parallel to $\mathbf{t}$, and with Burgers vector $\mathbf{b}_\beta$. The dislocation distribution vector is related to the Nye tensor $\boldsymbol{\alpha}$, by

$$\boldsymbol{\alpha} = \sum_\beta \mathbf{b}_\beta \otimes \boldsymbol{\rho}_\beta, \qquad (47)$$

where the sum is carried over all Burgers vectors in all slip systems. The curvature tensor of the deformed volume is given by

$$\boldsymbol{\kappa} = \tfrac{1}{2}\mathrm{Tr}\,(\boldsymbol{\alpha})\,\mathbf{I} - \boldsymbol{\alpha}. \qquad (48)$$

Thus, it may be possible to compute the volumetric distortions associated with the movement of dislocations within a simulation volume, and then connect these distortions to continuum theories.

Zaiser (2001) noted that, because of the highly intermittent nature of plastic flow, the instantaneous active slip volume in fcc metals is of the order of 10^{-6} only and that, at typical strain rates of $10^{-4}\,\mathrm{s}^{-1}$, the time between neighbouring avalanches exceeds the lifetime of an individual avalanche by four to six orders of magnitude. On that basis, he concluded that plastic deformation can be characterized as a slowly driven non-equilibrium system exhibiting large fluctuations. To estimate the magnitude of such fluctuations, Hahner (1996) proposed that, if one neglects dislocation multiplication and annihilation, the average work done by the internal stress must be zero, and that the work of the external stress is dissipated into heat. Thus, one can write

$$\left\langle \tau_{\mathrm{int}}^\beta v^\beta \right\rangle = 0, \qquad (49)$$

where $\tau_{\mathrm{int}}^\beta$ and v^β are internal stresses and dislocation velocities on slip system β. Recognizing that the effective stress τ_{eff} has an external τ_{ext} plus an internal τ_{int} component, and upon separating the stress and velocity into average (represented by $\langle \cdots \rangle$) and fluctuating (represented by $\delta \cdots$) components, Zaiser (2001) arrives at the following expressions for the autocorrelation functions of the effective stress $\tau_{\mathrm{eff}}^\beta$, and dislocation velocity v^β fluctuations:

$$\left\langle (\delta\tau_{\mathrm{eff}}^\beta)^2 \right\rangle = \left\langle \tau_{\mathrm{int}}^\beta \right\rangle \left\langle \tau_{\mathrm{eff}}^\beta \right\rangle,$$

$$\frac{\left\langle (\delta v^\beta)^2 \right\rangle}{\left\langle v^\beta \right\rangle^2} = \frac{\left\langle \tau_{\mathrm{int}}^\beta \right\rangle}{\left\langle \tau_{\mathrm{eff}}^\beta \right\rangle} = \frac{b^2 \rho^\beta \left\langle \tau_{\mathrm{int}}^\beta \right\rangle}{B \left\langle \dot{\gamma}^\beta \right\rangle}, \qquad (50)$$

where B is the mobility, and $\dot{\gamma}^\beta$ is the average shear strain rate. Using typical values for copper in the previous equations, the relative velocity fluctuation is found to be of the order of 10^7 Zaiser (2001) in accordance with the experimental measurements of Neuhäuser (1983).

Recently, detailed measurements of the misorientation angles across dislocation walls have been made by Hughes *et al.* (1998). Zaiser (2001) formulated a velocity-averaged stochastic version of equation (44) for the distribution function of *surplus dislocation segments* or geometrically necessary dislocations to show that the probability distribution function for misorientation angles is in good agreement with the experimental measurements of Hughes *et al.* (1998).

A percolation-type stochastic dislocation model has been advanced by Thomson and co-workers (Thomson and Levine 1998, Shim *et al.* 2001, Thomson *et al.* 2002). In this model, which is limited to stage III deformation, the weakest cell in a band undergoes a burst of strain, which initiates a cluster of newly strained cells. Slip transmission to neighbouring cells is assumed to be of the form $s^* = as$, where s is the number of dislocations in the strained cell, a is a pile-up amplification factor and s^* is the number of dislocations in the previously unstrained cell. The key parameter a in such a model is assumed to depend on two types of mechanisms: firstly, a random distribution of sources within the cell walls and/or, secondly a distribution of locks holding up dislocations within the cell walls. Their analysis shows that percolating slip clusters conform to the universality class of percolation theory (Stauffer and Aharony 1992). Furthermore, it can be effectively used to describe slip transmission in stage III, and that a well-defined percolation threshold can be identified.

The power of SM can be brought to bear on one of the most important challenges for the multiscale modelling approach. The large disparity in time and length scales during plastic deformation limit the reach of direct DD simulations. Hence, it is of interest to determine the necessary coarse-graining strategies for both temporal and spatial variables. LeSar and co-workers (Lesar *et al.* 1989, LeSar and Rickman 2002, 2003, Rickman *et al.* 2003) have begun to tackle this issue by investigations of a number of specific problems, namely the combined motion of dislocations and solutes, and the problem of the long-range interactions of blocks of dislocation ensembles.

In the limit where the solute mobility is high relative to that of dislocations, dislocation–solute atmosphere pairs can be treated as quasi-particles, and the degrees of freedom associated with individual solute atoms can be eliminated (Rickman *et al.* 2003). Thus, an effective mobility can be derived and can be related to the details of such interaction. However, in the other extreme limit of low solute mobility compared with dislocations, the solute atoms function essentially as static traps for dislocations. In this case, LeSar and co-workers modelled slip as a discrete one-dimensional 'birth-and-death' stochastic process (Van Kampen 1992). LeSar and Rickman (2003) found that the dislocation diffusivity is modified by $\Delta D(c, \epsilon/k_B T)$:

$$\frac{\Delta D}{D'} = -c \, \frac{\tau_t/\tau - 1}{c \, (\tau_t/\tau - 1) + 1}, \tag{51}$$

where $\tau_t/\tau = \exp(-\epsilon/k_B T)$ is the mean residence time in a trapped state, D' is the diffusivity when there are no traps, $c \equiv N_t/N$ is the trap (i.e. solute) concentration. These examples show that it may be possible to average out fast time variables in DD simulations.

The issue of spatial averaging has been recently dealt with by LeSar and Rickman (2002) and Wang *et al.* (2003b). Starting from the work of Kossevich (1979) on the interaction energy of systems of dislocations, an energy expression in terms of the dislocation density tensor (Rickman *et al.* 2003) is derived. The basic idea in this approach is to divide space into small averaging volumes to calculate the local multipole moments of the dislocation microstructure and then to write the energy (LeSar and Rickman 2002) or the force (Wang *et al.* 2003b) as an expansion over the multipoles. Numerical implementation of the multipole expansion for the Peach–Kohler force on a test dislocation separated from a volume containing a dense dislocation aggregate revealed that such a technique is highly accurate and
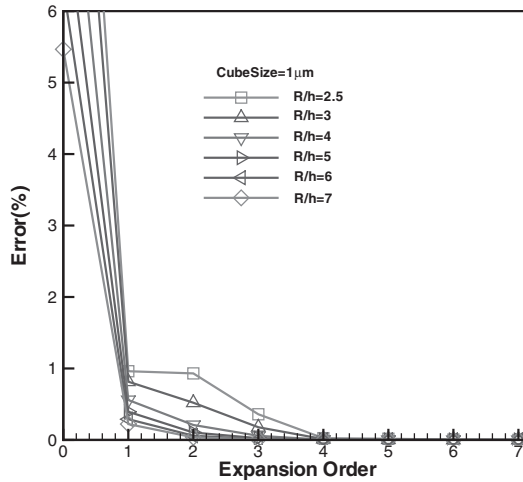
Figure 8. Relative error versus the expansion order for different R/h (volume size, 1 μm).

can result in significant coarse graining in current DD simulations. The relative error in the multipole expansion, expressed as $|\sigma_{\text{multipole}} - \sigma_{\text{accurate}}|/\sigma_{\text{accurate}}$ has been tested by Wang *et al.* (2003b). These numerical tests are performed for different sizes (1 μm, 5 μm and 10 μm), different orders of the expansion and different R/h ratios (R being the distance from the centre of the volume to the test dislocation and h being the volume size). A representative result is shown in figure 8. It is clear that the multipole approximate solution converges very rapidly, that the third-order expansion (quadropole) gives a relative error less than 1% and that the fourth-order expansion gives a relative error of less than 0.05%.

§5. Continuum mechanics methods

In this section, an overview will first be given of the main CM-based framework used today to describe the nonlinear deformation behaviour of materials at the local (e.g. single-phase or grain level) and macroscopic (e.g. polycrystal level) scales. Emphasis will be placed on recent progress made in crystal plasticity, strain gradient plasticity, and homogenization techniques to link deformation phenomena simultaneously occurring at different scales in the material microstructure with its macroscopic behaviour.

Standard tensorial notation will be used throughout. Vectors will be described by bold lower case letters, second-order tensors by bold capital letters and fourth-order tensors by bold script capital letters. Also, $\mathbf{a} \cdot \mathbf{b} = a_i b_i$, $\mathbf{Ab} = A_{ij} b_j$, $\mathbf{A} : \mathbf{B} = A_{ij} B_{ij}$, $\mathbf{AB} = A_{ij} B_{jk}$, $\mathcal{L} : \mathbf{A} = L_{ijkl} A_{kl}$ and $(\mathbf{a} \otimes \mathbf{b})_{ij} = a_i b_j$, where Einstein summation applies for repeated indices.

5.1. *Continuum discretization of a boundary value problem*

In a generic boundary value problem (BVP), the deformation of a body subjected to external forces and prescribed displacements is governed by the

   (i) equilibrium equations,
  (ii) constitutive equations,
 (iii) boundary conditions and
 (iv) initial conditions.

The 'weak' form of the boundary value problem is obtained when the equilibrium equations and the boundary conditions are combined into the 'principle of virtual work'. This 'weak form' constitutes the basis for obtaining a numerical solution of the deformation problem via, for example, the finite-element method. Thus, in a CM Lagrangian formulation of a quasistatic BVP, the principle of virtual work is the vehicle by which the global equilibrium equations are obtained (for example, Zienkiewicz and Taylor (1994)). The basic features of a generic Galerkin-type discretization framework are given next.

Consider a structure occupying a domain V in the deformed configuration which is subjected to external forces and displacements on its boundary, $\varGamma$. In the absence of body forces and inertial effects, the principle of virtual work for the structure, in its rate form, satisfies the following equation:

$$\int_V \boldsymbol{\sigma} : \delta\dot{\boldsymbol{\epsilon}}\, dV - \int_\varGamma \mathbf{t} \cdot \delta\mathbf{v}\, d\varGamma = 0, \tag{52}$$

for any arbitrary virtual velocity vector field $\delta\mathbf{v}$ compatible with all kinematics constraints. In the above equation, $\mathbf{t} = \boldsymbol{\sigma}\mathbf{n}$ represents the boundary traction forces, $\boldsymbol{\sigma}$ the Cauchy stress, $\mathbf{n}$ the normal to the surface on which the tractions act, and $\delta\dot{\boldsymbol{\epsilon}}$ the virtual strain rate associated with the velocity field $\delta\mathbf{v}$.

To solve a complex BVP numerically, the discretization of the principle of virtual work is generally performed using the finite-element method. Let $\mathbf{v}$ be approximated at a material point within an element by

$$\mathbf{v} = \sum_{i=1}^{N_{\max}} N^i\, \hat{\mathbf{v}}^i \equiv \mathbf{N}\,\hat{\mathbf{v}}, \tag{53}$$

where $\hat{\mathbf{v}}$ denotes the nodal values of the element velocity field and $\mathbf{N}$ are the isoparametric shape functions. Substituting equation (53) into equation (52) leads to the discretized version of the principle of virtual work on the finite element V_{e}:

$$\mathbf{r}\{\hat{\mathbf{v}}\} \equiv \mathbf{f}^{\mathrm{I}} - \mathbf{f}^{\mathrm{E}} = 0, \tag{54}$$

where

$$\mathbf{f}^{\mathrm{I}} = \int_{V_{\mathrm{e}}} \mathbf{B}^{\mathrm{T}}\boldsymbol{\sigma}\, dV_{\mathrm{e}}, \quad \mathbf{f}^{\mathrm{E}} = \int_{\varGamma_{\mathrm{e}}} \mathbf{N}^{\mathrm{T}}\mathbf{t}\, d\varGamma_{\mathrm{e}}, \tag{55}$$

are the internal and external global force vectors, respectively, and $\mathbf{B}$ relates the symmetric strain rate tensor with $\hat{\mathbf{v}}$. The global equilibrium relations (equation (54)) represent a set of implicit nonlinear equations which may be solved incrementally using a Newton-type algorithm. In a Newton–Raphson iterative scheme, the nonlinear system (equation (54)) is typically expanded using Taylor series in the neighbourhood of $\hat{\mathbf{v}}$:

$$\mathbf{r}\{\hat{\mathbf{v}}^k + \delta\hat{\mathbf{v}}^k\} = \mathbf{r}\{\hat{\mathbf{v}}^k\} + \frac{\partial\mathbf{r}\{\hat{\mathbf{v}}^k\}}{\partial\hat{\mathbf{v}}^k}\, \delta\hat{\mathbf{v}}^k + \mathrm{O}\{\delta\hat{\mathbf{v}}^{k^2}\}, \tag{56}$$

where k represents a generic iteration and $\partial\mathbf{r}/\partial\hat{\mathbf{v}}$ is the global tangent stiffness or Jacobian matrix of the nonlinear system of equations. The formulation of accurate estimates of the global Jacobian is at the heart of most numerical schemes developed to provide robust algorithms for the use of complex constitutive models with continuum approaches (for example, Crisfield (1997), Esche *et al.* (1997) and Busso *et al.* (2000)).

5.2. *Continuum approaches for single-crystal plasticity*

Constitutive models developed to predict the anisotropic behaviour of single-crystal materials generally follow either a Hill-type or a crystallographic approach. As a common feature, they treat the material as a continuum in order to describe properly the plastic or viscoplastic effects. Hill-type approaches (for example, Nouailhas and Freed (1992) and Schubert *et al.* (2000)) are based on a generalization of the von Mises yield criterion proposed by Hill (1950) to account for the non-smooth yield or flow potential surface required to describe the anisotropic flow stress behaviour of single crystals. In constitutive formulations based on crystallographic slip, the macroscopic stress state is resolved on to each slip system following Schmid's law. Internal state variables are generally introduced in both formulations to represent the evolution of the microstructural state during the deformation process. Although recent developments in these two approaches have now reached an advanced stage, the major improvements have been made by crystallographic models due to their ability to incorporate complex micromechanisms of slip within the flow and evolutionary equations of the single crystal models. These include the effects of dislocation interactions (for example, Meric and Cailletaud (1991) and Meric *et al.* (1991)), hardening and strain gradient phenomena (for example, Gurtin (2000), Meissonnier *et al.* (2001) and Stainer *et al.* (2002)) and general anisotropic plastic and viscoplastic behaviour (for example Jordan and Walker (1992), Anand and Kothari (1996) and Busso and McClintock (1996)). A brief outline of the salient features of local and non-local crystal plasticity approaches are given below.

5.3. *Generic local crystallographic framework*

A generic internal-variable-based crystallographic framework is said to be local when the evolution of its internal variables can be fully determined by the local microstructural state at the material point. The description of the kinematics of most crystal plasticity theories follows that originally proposed by Asaro and Rice (1977), which has been widely reported in the computational mechanics literature (for example Pierce *et al.* (1983), Asaro and Needleman (1985), Cuitiño and Oritz (1992), Kalidindi *et al.* (1992) and Busso *et al.* (2000)). It relies on the multiplicative decomposition of the total deformation gradient $\mathbf{F}$ into an inelastic component $\mathbf{F}^p$ and an elastic component $\mathbf{F}^e$ (Lee 1969). Thus, under isothermal conditions,

$$\mathbf{F} = \mathbf{F}^e \mathbf{F}^p. \tag{57}$$

Although single-crystal laws can be formulated in a corotational frame, that is the stress evolution is computed on axes which rotate with the crystallographic lattice, the most widely used approach is to assume that the material response is hyperelastic, that is that its behaviour can be derived from a potential (i.e. free energy). Such a potential may be expressed in terms of the elastic Green–Lagrange tensorial strain measure

$$\mathbf{E}^e = \tfrac{1}{2}(\mathbf{F}^{e\mathrm{T}} \mathbf{F}^e - \mathbf{1}), \tag{58}$$

and the corresponding objective work conjugate (symmetric) stress (Hill 1978) or second Piola–Kirchhoff stress $\mathbf{T}$. Note that the Cauchy stress is related to $\mathbf{T}$

by $\boldsymbol{\sigma} = \det(\mathbf{F}^e)^{-1} \mathbf{F}^e \mathbf{T} \mathbf{F}^{eT}$. The hyperelastic response of the single crystal is governed by

$$\mathbf{T} = \frac{\partial \Phi\{\mathbf{E}^e\}}{\partial \mathbf{E}^e}, \tag{59}$$

where $\Phi\{\mathbf{E}^e\}$ represents the Helmholtz potential energy of the lattice per unit reference volume. Differentiation of equation (59), and the assumption of small elastic stretches yield

$$\mathbf{T} \approx \mathcal{L} : \mathbf{E}^e. \tag{60}$$

where $\mathcal{L}$ is the anisotropic linear elastic moduli.

In rate-dependent formulations, the time rate of change in the inelastic deformation gradient is related to the slip rates $\dot{\gamma}^\alpha$ on each slip system (Asaro and Rice 1977), as

$$\dot{\mathbf{F}}^p = \left(\sum_{\alpha=1}^{n_\alpha} \dot{\gamma}^\alpha \mathbf{P}^\alpha \right) \mathbf{F}^p, \quad \text{with } \mathbf{P}^\alpha \equiv \mathbf{m}^\alpha \otimes \mathbf{n}^\alpha. \tag{61}$$

Here, $\mathbf{m}^\alpha$ and $\mathbf{n}^\alpha$ are unit vectors defining the slip direction and the slip plane normal on the slip system.

In rate-independent formulations, in contrast, flow rules are based on the well-known Schmid law and a critical resolved shear stress τ_c^α, whereby the rate of slip is related to the time rate of change in the resolved shear stress $\tau^\alpha (= \mathbf{T} : \mathbf{P}^\alpha)$. Then,

$$\dot{\tau}^\alpha = \dot{\tau}_c^\alpha = \sum_{\alpha=1}^{n_\alpha} h^{\alpha\beta} \dot{\gamma}^\alpha, \quad \text{if } \dot{\gamma}^\alpha > 0. \tag{62}$$

In the above equation, $h^{\alpha\beta}$, the slip-hardening rates, is the slip interaction matrix which accounts for latent hardening effects. Owing to the severe restrictions placed on material properties, such as latent hardening, to ensure uniqueness in the mode of slip (for example Anand and Kothari (1996) and Busso and Cailletaud (2003)), and the associated difficulties in its numerical implementation, the use of rate-independent formulations has been somehow restricted and is much more limited than rate-dependent models. This has been compounded by the fact that, by calibrating their strain-rate sensitivity response accordingly, rate-dependent models have been successfully used in quasirate-independent regimes. Thus, henceforth the focus of the discussions will be on rate-independent approaches.

The slip rate in equation (61) can functionally be expressed as

$$\dot{\gamma}^\alpha = \widehat{\dot{\gamma}}^\alpha \{ \tau^\alpha, S_1^\alpha, \ldots, S_{m_S}^\alpha, \theta \}, \tag{63}$$

where S_i^α for $i = 1, \ldots, m_S$ denotes a set of internal state variables for the slip system α, and θ is the absolute temperature. A useful and generic expression for the overall flow stress in the slip system can be conveniently found by inverting equation (63) with $m_S = 2$:

$$\tau^\alpha = \pm \hat{f}_v^\alpha \{ \dot{\gamma}^\alpha, S_2^\alpha, \theta \} \pm c_S S_1^\alpha, \tag{64}$$

where c_S is a scaling parameter, and S_1^α and S_2^α represent additive and multiplicative slip resistances respectively. Here the distinction between a multiplicative slip resistance S_2^α and an additive slip resistance S_1^α is motivated by the additive and multiplicative use of non-directional hardening variables rather than by mechanistic

considerations. By expressing the flow stress in the slip system in the way shown in equation (64), the contributions from viscous effects (first term in equation (64)), and hardening mechanisms (second term) can be clearly identified. The majority of formulations relied on power-law functions for equation (63), which results in $S_2 \neq 0$ and $S_1 = 0$ in equation (64) (for example Pierce *et al.* (1983)). This introduces a coupling between the viscous term and microstructure which is inconsistent with most strengthening mechanisms. Recently, work by Meric and Cailletaud (1991) and Busso *et al.* (2000) have proposed flow stress relations with $S_1 \neq 0$ and $S_2 = 0$, which allows a more physically meaningful interpretation of strengthening phenomena. For a more detailed discussion in these issues, see Busso and Cailletaud (2003).

The crystallographic formulation is completed with the evolutionary relations for the S_i^α internal slip system variables. The time rate of change in each internal slip system variable $\dot{S}_i^\alpha$ is, in its most general form, expressed as

$$\dot{S}_i^\alpha = \widehat{\dot{S}}_i^\alpha \left\{ S_1^\alpha, \ldots, S_{m_S}^\alpha, \dot{\gamma}^1, \ldots, \dot{\gamma}^\alpha, \ldots, \dot{\gamma}^{n_\alpha}, \theta \right\}. \tag{65}$$

Note that the dependence of equation (65) on the slip rates on all systems enables cross-hardening effects to be accounted for.

5.4. *Non-local approaches*

The study of experimentally observed size effects in a wide range of mechanics and materials problems has received much attention recently. Most continuum approaches and formulations dealing with these problems are based on strain gradient concepts and are known as non-local theories since the material behaviour at a given material point depends not only on the local state but also on the deformation of neighbouring regions. Examples of such phenomena include the particle size effects on composite behaviour (for example Nan and Clarke (1996)), the precipitate phase size in two-phase single-crystal materials (for example Busso *et al.* (2000)), the increase in measured microhardness with decreasing indenter size (for example Swadener *et al.* (2002)), and the effects of decreasing film thickness (for example Huber and Tsakmakis (1999)), among others.

The dependence of mechanical properties on length scales can in most cases be linked to features of the microstructure, the boundary conditions or the type of loading, which give rise to localized strain gradients. In general, the local material flow stress is controlled by the actual gradients of strain when the dominant geometric or microstructural length scales force the deformation to develop within regions of less than approximately 5–10 μm wide in polycrystalline materials, and of the order of 0.1–1 μm in single-crystal materials (Busso *et al.* (2000)). Thus, gradient-dependent behaviour is expected to become important once the length scale associated with the local deformation gradients becomes sufficiently large when compared with the controlling microstructural feature (e.g. average grain size in polycrystal materials). In such cases, the conventional crystallographic formulations discussed in the previous section will be unable to predict properly the evolution of the local material flow stress. To accommodate these strain gradients, generation of geometrically necessary dislocations is required in these regions of incompatibility (for example see Arsenlis and Parks (2001), Busso *et al.* (2000) and Gao and Huang (2003)). The introduction of these geometrically necessary dislocations, in addition to those stored in a random way (so-called 'statistically stored' dislocations), is what causes the additional strengthening of the material.

One of the first non-local theories was that proposed by Aifantis (1987) and Zbib and Aifantis (1988) to describe the formation of shear bands. This type of formulations rely on first and second derivatives of strain linked to the flow rule to describe strain gradient effects without the need to use higher-order stresses. It requires additional boundary conditions, is relatively easy to implement numerically into the finite-element method but is limited in describing strain gradient problems that involve only one material length scale. Furthermore, by the nature of the formulation, it provides a limited mechanistic insight into the non-local phenomena.

A more physically intuitive continuum approach to describing strain gradient effects is that followed by constitutive theories such as those developed by Acharya and Beaudoin (2000), Busso *et al.* (2000), Arsenlis and Parks (2001) and Bassani (2001). They rely on internal state variables to describe the evolution of the obstacle or dislocation network within the material and generally introduce the strain gradient effects directly in the evolutionary laws of the slip system internal variables without the need for higher-order stresses. Thus, in some of these formulations, the functional dependence of the evolutionary laws of the slip system internal variables, such as the general form given for the slip resistance in equation (65), will now include an additional dependence in the gradient $\nabla \dot{\gamma}^\alpha$ of the slip rates. Then,

$$\dot{S}_i^\alpha = \widehat{\dot{S}}_i^\alpha \left\{ S_1^\alpha, \ldots, S_{m_\mathrm{S}}^\alpha, \dot{\gamma}^1, \ldots, \dot{\gamma}^\alpha, \ldots, \dot{\gamma}^{n_\alpha}, \nabla \dot{\gamma}^1, \ldots, \nabla \dot{\gamma}^\alpha, \ldots, \nabla \dot{\gamma}^{n_\alpha}, \theta \right\}, \qquad (66)$$

This class of theories has been shown capable of providing considerable physical insight into the effects of microstructure on the observed macroscopic phenomena, including rate-independent plastic deformation and viscoplasticity in both single-crystal and polycrystalline materials (for example Acharya and Beaudoin (2000), Busso *et al.* (2000) and Arsenlis and Parks (2001)). One additional attractive aspect of the these theories is that they are relatively easy to implement numerically and do not require higher-order stresses or additional boundary conditions. However, one limitation of these types of theory is that they are unable to describe problems which may require non-standard boundary conditions such as the boundary layer problem modelled by Shu *et al.* (2001).

A significant amount of work has been based on the treatment of the solid as a Cosserat continuum (for example Mühlhaus (1989) and Forest (1998)), where the material flow stress is assumed to be controlled not just by the rate of slip but also by the material curvature. However, even though Cosserat-type models have shown to be well suited to predicting localization phenomena, they are analytically complex and phenomenological in nature, thus making it difficult to gain a direct insight into the controlling physical phenomena at the microstructural level.

Another class of non-local theories with higher-order fields is that proposed recently by Fleck and Hutchinson (2001) and by Gurtin (2003). They contain higher-order stresses and are extensions of the original theory proposed earlier by Fleck and Hutchinson (1997). In such non-local formulations, the balance laws are based on a principle of virtual work where additional field variables, namely the work conjugates of the slip and slip rate gradients, are required. Consider a crystallographic approach and let $\mathbf{q}$ be a higher-order stress vector with components q_i, let the work conjugate to the slip rate gradients be $\nabla \dot{\gamma}^\alpha$, and let π^α be the work

conjugate to the slip rate $\dot{\gamma}^\alpha$. Then,

$$\int_V \left(\boldsymbol{\sigma} : \delta\dot{\boldsymbol{\epsilon}} + \sum_\alpha (\pi^\alpha - \tau^\alpha)\delta\dot{\gamma}^\alpha + \sum_\alpha \mathbf{q}^\alpha \cdot \nabla\delta\dot{\gamma}^\alpha \right) dV$$

$$- \int_\Gamma \left(\boldsymbol{\sigma}\,\mathbf{n} \cdot \delta\mathbf{v} + \sum_\alpha \mathbf{q}^\alpha \cdot \mathbf{n}\,\delta\dot{\gamma}^\alpha \right) d\Gamma = 0. \tag{67}$$

By integrating equation (67) by parts, it can be shown that, in addition to the usual equilibrium relation in V given by

$$\operatorname{div}\boldsymbol{\sigma} = 0, \tag{68}$$

the following microforce balance is obtained:

$$\pi^\alpha = \tau^\alpha + \nabla\mathbf{q} : \mathbf{1}, \tag{69}$$

where $\mathbf{1}$ is the second-order unit tensor and τ^α is the resolved shear stress previously defined. Furthermore, equation (67) also implies that the following relations be satisfied at the boundary Γ:

$$\mathbf{t} = \boldsymbol{\sigma}\mathbf{n} \text{ or } \mathbf{v}, \quad \text{and} \quad \mathbf{q}^\alpha \cdot \mathbf{n}, \text{ or } \dot{\gamma}^\alpha. \tag{70}$$

It can be seen that, when no slip rate gradients are present, equation (68) gives $\pi^\alpha = \tau^\alpha$ and equation (67) resolves to the local version of the principle of virtual work given by equation (52).

The numerical implementation of these typically highly nonlinear theories require the development of complex numerical algorithms. However, the introduction of higher-order stresses and the required boundary conditions to satisfy the additional higher-order field equations introduced by some non-local theories often precludes and limits the use of these formulations to in-house programmes (e.g. see implementation of a coupled stress theory by Shu *et al.* 1999)).

5.5. *Homogenization approaches for heterogeneous microstructures*

The bridging between the mechanical behaviour of heterogeneous materials and that of their individual constituents remains a topic of major interest and is at the heart of homogenization schemes developed to predict the behaviour of materials at different scales. Such schemes are based on the assumption that the mechanical behaviour of individual constituents can lead to the description of the mechanical response of a macroscopic aggregate through either suitable interaction laws or a numerical averaging process.

When distinct heterogeneities exist at different microstructural levels, it is always possible to identify at each level the smallest possible representative volume element (RVE) of the microstructure which contains all the information concerning the distribution and morphology of the material's heterogeneities. Irrespective of the homogenization schemes, once the relevant scales in the heterogeneous microstructure are identified, it is then necessary to select a RVE of the microstructure at the level of interest. Typically, the representative length scales in the microstructure are the average size D of the largest heterogeneity at that particular level and the size L of the RVE. They must satisfy

$$D \ll L. \tag{71}$$

The value of D can be, for instance, the average grain size in a polycrystal with randomly oriented grains or can be defined by the size of the precipitates relative to their mean spacing.

Homogenization schemes require not only constitutive models for the individual constituents but also appropriate rules to make the transition between scales. The cases to be discussed here are those where loading in the RVE is homogeneous. If the loads applied on the RVE were inhomogeneous, then the homogenized equivalent medium is said to be generalized, and special kinematics and equilibrium equations would apply (for example Besson *et al.* (2002)). Some of the most widely used approaches to link local fields with macroscale phenomena, such as the large deformation and texture evolution of polycrystals, are Taylor-type (for example Kalidindi *et al.* (1992)) or Sachs-type (for example Leffers (2001)) models. The former assumes strain uniformity and can only fulfil compatibility at grain boundaries, but not equilibrium. In contrast, Sachs-type models assume homogeneous stresses and ignore local compatibility at grain boundaries. Among the most successful recent methods proposed to overcome the limitations of the assumed plastic strain or stress uniformity are those based on the relaxed constrained method (for example van Houtte, Delannay and Samajdar (1999)) and on mean-field approaches, such as self-consistent (for example Hill (1965)) and variational (for example Ponte Castañeda (1991)) methods. Here, compatibility and equilibrium between grains are satisfied at both the local and the macroscopic levels.

In the self-consistent averaging approach, the interaction between a single-crystal grain and its neighbours is treated as that between an inclusion with the same properties as those of the grain, embedded in an homogeneous equivalent medium (HEM) which has the same (unknown) properties as those of the macroscopic aggregate (figure 9 (*a*)). A critical aspect of the self-consistent method is that the strain distribution within the inclusion is assumed uniform when in reality, in cases of low strain-rate sensitivity and large property mismatch, it is seldom the case.

One of the simplest self-consistent frameworks is that proposed by Hill (1965), where the stress-rate tensor $\dot{\mathbf{T}}$ and strain-rate tensor $\dot{\mathbf{E}}$ in each phase or grain are related to those of the HEM, $\dot{\boldsymbol{\sigma}}$ and $\dot{\boldsymbol{\epsilon}}$, through an elastic accommodation tensor $\mathcal{L}_e$

$$\dot{\mathbf{T}} - \dot{\boldsymbol{\sigma}} = \mathcal{L}_e : (\dot{\boldsymbol{\epsilon}} - \dot{\mathbf{E}}). \qquad (72)$$

The determination of suitable interaction relations between the inclusion and HEM is at the centre of most self-consistent approaches.

In equation (72), an elastic interaction between the grain and the polycrystal aggregate is implicitly assumed; thus a high constraint is imposed on the inclusion
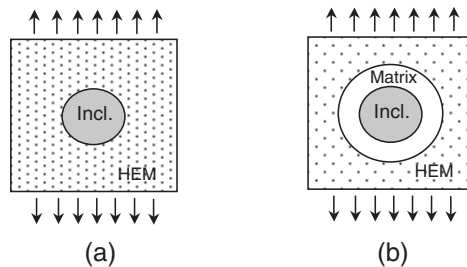

Figure 9. Comparison between the assumptions behind (*a*) standard and (*b*) generalized self-consistent approaches.

by the surrounding elastic aggregate. In reality, such a high constraint is partially relaxed by the plastic deformation of the polycrystalline aggregate. The work of Berveiller amd Zaoui (1979) addressed this problem by introducing a plastic accommodation factor $\mathcal{L}_p$. Thus, for the nonlinear case, equation (72) becomes

$$\dot{\mathbf{T}} - \dot{\boldsymbol{\sigma}} = \mathcal{L}_p : \left(\dot{\boldsymbol{\epsilon}}^p - \dot{\mathbf{E}}^p \right), \tag{73}$$

where the global and local strain-rate tensors are now the plastic tensors. Similarly, the tangent approach method proposed by Molinari (2002) enables approximate solutions for nonlinear material behaviour problems to be obtained while preserving the structure of the Eshelby linear inclusion solution. Generally, self-consistent schemes are well suited to plastic or viscoplastic aggregates which can be treated as elastically rigid. The incorporation of elastic effects within a self-consistent framework is more difficult and clear solutions remain elusive. As self-consistent schemes are generally implicit, their numerical implementation require Newton-type iterative procedures to solve the highly nonlinear systems of equations. A more elaborate self-consistent approach (for example Herve and Zaoui (1993)), referred to as a generalized or a three-phase scheme, is illustrated in figure 9 (*b*). Here, a composite sphere made up of two phases (i.e. matrix and inclusion) is embedded in an infinite matrix which is the HEM of unknown properties. The additional requirement in this case is that the average strain in the two-phase composite sphere must be the same as the strain prescribed in the far field. Herve and Zaoui (1993) showed that such scheme provides a framework for statistical analyses and demonstrated this by determining the effective behaviour of a random assembly of arbitrary-size spheres. Pitakthapanaphong and Busso (2002) used a similar generalized self-consistent approach to determine the elastoplastic properties of functionally gradient materials.

An alternative homogenization method to the classical self-consistent approach is that derived from a variational procedure. The variational formulation proposed by Ponte Castañeda (1991) relies on the effective modulus tensor of 'linear elastic comparison composites', whereby the effective stress potentials of nonlinear composites are expressed in terms of the corresponding potentials for linear composites with similar microstructural distributions, to generate the corresponding bounds. Ponte Castañeda's variational principles have been used successfully to derive elastoplastic relations for metal matrix composites, among other applications. For further details, see the references given herein.

As a result of the ever-increasing computer power available, it is now becoming possible to replace the approximate mean-field methods for more accurate techniques based on numerical homogenization. One of the most powerful is that based on periodic unit-cell concepts, whereby a microstructural 'window' of the constituents is periodically arranged and subjected to homogeneous far-field loading. By defining appropriate periodic boundary conditions on the smallest window or RVE which contains all the information about the local heterogeneities, the average stress and strain responses of the unit cell can be obtained numerically. Even though these methods are computationally intensive, they offer the capability for digitized images of typical microstructures to be easily superimposed on to a regular finite-element mesh, thus enabling a precise model of the microstructure to be made and a framework to predict accurately the inhomogeneity of the deformation at the level of the individual phases. Alternatively, the finite-element mesh can be designed so that the element boundaries correspond to phase boundaries. This method has the advantage
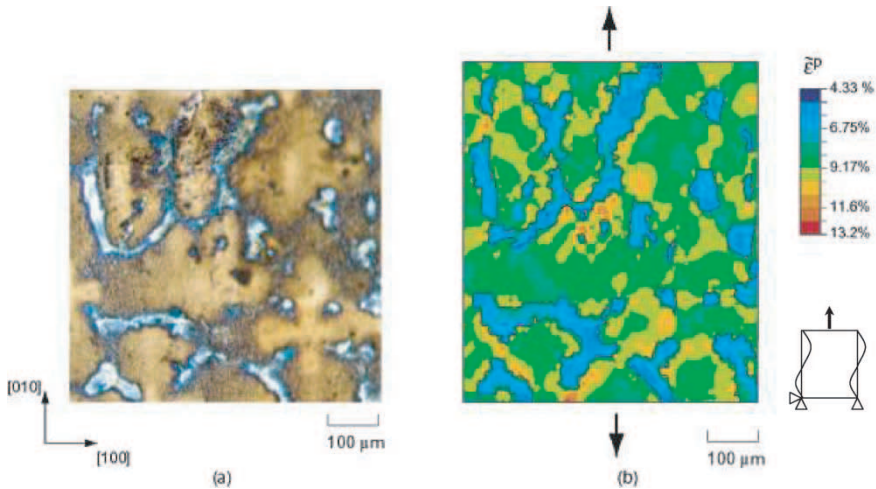
Figure 10. Homogenization of a typical heterogeneous single-crystal superalloy: (*a*) RVE; and (*b*) predicted contours of accumulated inelastic strain under [010] uniaxial loading (Regino *et al.* 2002).

that almost any heterogeneity can be modelled, but the modelling effort required is generally impractical for very complex microstructures.

An example of a periodic unit cell of a typical heterogeneous single-crystal superalloy RVE is presented in figure 10 (Regino *et al.* 2002). Figure 10 (*a*) shows a scanning electron micrograph of the microstructure's RVE at an intermediate scale, and figure 10 (*b*) the predicted RVE contour plot of uniaxially equivalent accumulated inelastic strain, after a 10% straining along the [010] orientation at a rate of $10^{-3}\,\text{s}^{-1}$ at 950°C. It can be seen that the localization of inelastic strain occurs in the vicinity of the stronger (eutectic) region. For complex three-dimensional microstructures, such as those of polycrystal aggregates with randomly oriented grains, realistic finite-element meshes can be built based on Voronoi tesselations (for example Ghosh *et al.* (2002) and Barbe *et al.* (2002)).

§6. Outstanding issues and future prospects

In this review, we have discussed the different modelling approaches which address specific phenomena at different length scales and have highlighted the rich variety of physical, computational and technological issues within the broad area of nanomechanics and micromechanics which have been successfully addressed. We conclude this article by briefly summarizing the current level of understanding in these areas and discuss our expectations of forthcoming progress.

6.1. *Bridging the length and time scales*

Even though recent advances in computing power have led to new and vastly improved simulation techniques at the atomistic and continuum levels, there has not yet been a concerted effort to develop explicit links between atomistic and CM models. When properly reinforced, links between length scales should bring the overall field of material modelling one step closer to predict, *ab initio*, the final properties of a proposed material.

A natural sequence to produce a fully integrated multiscale modelling framework will require the following:

(i) the prediction of crystal energies, structures and the derivation of interatomic potentials from *ab initio* calculations;

(ii) the prediction of complex microstructural heterogeneities and morphologies resulting from solidification from the surface energies and kinetics predicted *ab initio*;

(iii) the identification of the crystal structures, slip systems and interface structures of the phases and the kinetics of phase transformations during subsequent thermal or thermomechanical processes using atomistic, DD and topological modelling techniques;

(iv) the formulation of crystallographic models for each phase to describe the dominant deformation processes and the development of homogenization schemes to obtain the macroscopic mechanical response of the material.

Figure 11 illustrates the structure and the links between the modelling paradigms, together with the outputs and inputs for each subarea. Here, the physical properties of the alloy are predicted at the atomistic scale and these are linked to the continuum level via appropriate constitutive and kinetic models. The properties determined by *ab initio* calculations (e.g. crystal energies and structures, and interatomic potentials) needs to be related to the range of compositions predicted after solidification. The *ab initio* results can then be used in MD calculations of deformation mechanisms. The resulting defect structures and interface kinetics are then fed into the continuum formulations to provide the crystallographic description and slip systems of the individual phases present in the heterogeneous (multiphase) microstructures. The overall behaviour of RVEs of such microstructures can finally be determined using homogenization techniques.

While progress on linking length scales has just started, linking time scales remains an outstanding problem. As pointed out earlier, systematic and rigorous
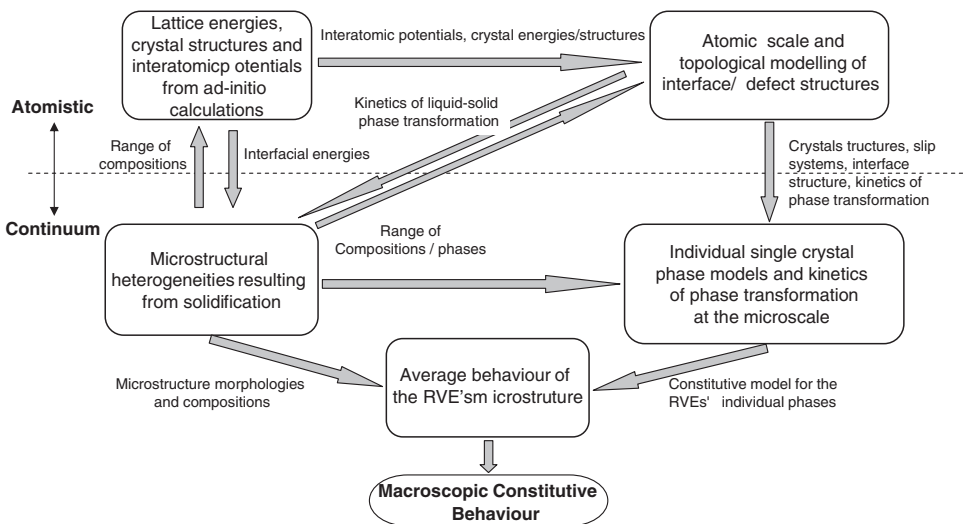

Figure 11.	Flow chart showing the links between the different modelling approaches.

reduction in the degrees of freedom that describe material evolution will lead to self-consistent length scale linking, and hence confident predictions of mechanical properties. On the other hand, methods for self-consistent linking of time scales are still lacking. Events at the atomic scale are often in the picosecond to nanosecond range, while microstructure evolution takes place on much longer time scales: seconds to years. Since the time evolution of the microstructure is path dependent, events that occur in the picosecond to nanosecond time scale (e.g. atomic jumps and nucleation) may have profound effects on microstructure evolution. In addition, relaxation time scales in atomic models cannot be matched by continuum relaxation time scales. Thus, when a MD model is directly linked to the continuum, matching can be accomplished for static or quasistatic problems, while it has not been shown for fully dynamic problems. Progress in this area is needed.

6.2. *Atomistic scales*

Atomistic simulations have mainly supplemented experimentally obtained information until now. Nevertheless, future developments are expected in four major directions:

  (i) higher accuracy;
  (ii) larger systems;
 (iii) computationally faster methods; and
 (iv) more general approaches.

Density functional approaches have evolved rather rapidly during the past decade to address accuracy. The introduction of generalized gradient corrections has improved the accuracy of binding energies, surface energies and energy barriers. While the standard local spin density and generalized gradient approximations for the exchange–correlation energy can work for certain cases of strong correlation, that is where electrons partially preserve their localized atomic-like nature, they fail dramatically for others. Novel functionals, such as the recently developed self-correlation-free meta-GGA and self-interaction-free hyper-GGA, might yield a more reliable description of strong correlations. It is also possible that the use of quantum MC methods for benchmark calculations could provide a fruitful path to assess and improve density functionals. Concurrent with the above efforts in developing more accurate functionals within the spirit of DFT, several other methods have been recently developed to treat the ground-state properties of condensed systems that exhibit strong correlations. As alluded to in §2, these include the self-interaction correction method, the LDA+U method, the LDA+dynamic mean-field theory method, and the optimized effective potential method. In addition, the time-dependent extension of DFT and the GWA will provide a way to treat the excitation properties of moderately correlated electron systems.

Two other areas of developments, perhaps the most important ones, are faster methods, to enable dynamic processes to be simulated over sufficiently long time scales, and O(N) methods. Recasting the electronic structure calculations in an O(N) form which scales linearly with N will have important conceptual and practical implications for the treatment of sufficiently large systems. If the full calculation on a large system could be carried out in a time of O(N), then the properties of a small region of the large system could be calculated in a time independent of the system size. Such *quantum* O(N) methods will have exactly the same scaling as classical empirical potential methods. Furthermore the independence between the

different regions will allow them to be readily adapted to parallel computations. This would enable calculations on large systems of great interest in areas which are beyond current capabilities, such as materials science and biology, to be performed.

A single method is unlikely to meet all goals in multiscale modelling of nano-mechanics and micromechanics. Even with $O(N)$ algorithms, it will not be possible in the foreseeable future to treat systems containing millions of atoms at a highly accurate DFT level using large basis sets, as would be necessary for certain materials science applications. Such problems can only be approached if one succeeds in *linking* methods of different accuracies, such as DFT methods with classical force fields, and applying the high-accuracy method only to regions where the low-accuracy method is expected to fail. Hybrid methods of this type will certainly be based on the same notions of locality as $O(N)$ methods and will employ similar techniques.

In atomistic MD simulations, two areas are worth exploring. The first is the direct linking between atomistic and mesomechanics or microstructure–mechanics simulations. Approaches linking directly atomistic and continuum methods do exist; however, a two-way connection between atomistic and DD simulations has not yet been achieved. Atomistic simulations are capable of providing interaction rules or mechanisms for a finite number of dislocation configurations, most of these being of high symmetry. In DD simulations, a few dislocations may interact at close proximity, forming a low-symmetry configuration with respect to the simulation cell boundaries. It would be desirable to take such a configuration, to describe it in MD simulations and then to feed the resulting configuration back to DD simulations in a seamless fashion. The second area is dislocation interaction with interfaces. A dislocation behaves differently near a surface or at a grain boundary, when compared with bulk dislocations. A fair understanding of how a dislocation nucleates at a surface or grain boundary (i.e. dislocation absorption or blockage at interfaces) exists. The interaction of a dislocation with an interface depends very much on the nature of the interface. Different grain-boundary structures (e.g. low-angle versus high-angle grain boundaries) and compositions (e.g. pure versus segregated grain boundaries) lead to a large number of interaction mechanisms. Many of these mechanisms remain to be explored; thus a comprehensive description of dislocation–grain-boundary interaction needs to be developed.

In addition to the areas discussed above, it is worth pointing out that the mechanics of nanotubes (and related nanocomposites) is within the reach of MD simulations today. There are common features between the mechanics of nanotubes and that of crystalline solids. For instance, the 5–7 structure in a nanotube glides in a similar fashion to a dislocation in a crystal, thus making the study of the mechanics of a single nanotube relatively easy. However, when the interface of a nanotube and a matrix (e.g. polymer) is involved, the representation of atomic interactions can be a challenging issue.

6.3. *Transitional-continuum scales*

The significant role that defects play in determining the mechanical properties of metallic materials and the vast progress recently made in computational techniques has led to the emergence of the field of mesomechanics, which focuses on the behaviour of defects rather than that of atoms. Thus mesomechanics describes the mechanics of heterogeneities and irregularities in materials, including topological

defects (point, line, surface and volume), and compositional and structural heterogeneities. One of the most powerful mesomechanics methods is DD, where considerable progress has been made during the past two decades owing to a variety of conceptual and computational developments. It has moved from a curious proposal to a full and powerful computational method. In its present stage of development, DD have already addressed complex problems, and quantitative predictions have been validated experimentally. Progress in three-dimensional DD has contributed to a better understanding of the physical origins of plastic flow and has provided tools capable of quantitatively describing experimental observations at the nanoscale and microscale, such as the properties of thin films, nanolayered structures, microelectronic components and micromechanical elements. Advances in areas such as the mechanical behaviour of very small volumes and dislocation–interphase interactions are expected to continue in the near future.

One of the most difficult future challenges is the development of explicit links between the transitional-microstructure scales addressed by mesomechanics approaches and the continuum level. As previously discussed, mesomechanics approaches are needed to complement atomistic methods and to provide information about defect interaction and the kinetics of slip and interphase motion (see for example figure 11). Such fundamental information can then be transferred to the continuum level to underpin the formulation of flow and evolutionary behaviour of CM-based constitutive equations. Crystallographic approaches for single-crystal behaviour which rely on internal slip system variables will continue to provide the most powerful framework to incorporate basic mechanistic understanding in continuum models. Numerical and analytical homogenization techniques at the continuum level will be relied upon to a much greater extent than at present to model the behaviour of complex multiphase and polycrystalline microstructures. This will enable the resulting constitutive models to incorporate explicit links between features of the microstructure at different levels and the macroscopic behaviour. However, further development of this type of multiscale material design capability will require a few challenges to be overcome.

New and efficient computational techniques for processing and visualizing the enormous amount of data generated in mesomechanical and continuum multiscale simulations must be developed. Then, the issue of computational efficiency must be addressed so that truly large-scale simulations on thousands of processors can be effectively performed. Another example of the severe computational complexities that can arise with deformation is given by the highly intermittent nature of plastic slip, as it may require to be dynamically described as quick *avalanche* events separated by long time intervals, thus imposing severe computational limitations. Theoretical methodologies that enable the proper *coarse graining* of space should be pursued, such as multipolar expansion techniques, continuum averaging of slip gradients, grains and heterogeneous microstructures, and extraction of average lattice curvatures. The increasing complexity of non-local formulations to predict size effects in multiphase materials and composites will require improved and more robust numerical schemes to be developed, especially when a more physical description of dislocation interaction with themselves and with grain boundaries or other obstacles is required. Methods for a more direct coupling between DD simulations and continuum methods would also be required to improve the ability of non-local CM approaches to predict complex deformation phenomena, such as size effects. For instance, details obtained by DD simulations can provide information about the

evolution of the material topology, and the dislocation density tensor which can lead to precise descriptions of lattice curvatures.

The integration of the approaches discussed in this section is expected to lead to more physically based multiscale formulations and establish *materials by design* as an approach to optimize the mechanical properties of materials with complex microstructures, and to develop exciting new materials with extraordinary properties.

REFERENCES

ABRAHAM, F. F., BROUGHTON, J. Q., BERNSTEIN, N., and KAXIRAS, E., 1987, *Comput. Phys.*, **12**, 538.

ACHARYA, A., and BEAUDOIN, A., 2000, *J. Mech. Phys. Solids*, **48**, 2213.

ACKLAND, G. J., and VITEK, V., 1990, *Phys. Rev.* B, **41**, 10 324.

AIFANTIS, E., 1987, *Int. J. Plast.*, **95**, 211.

AJAYAN, P. M., RAVIKUMAR, V., and CHARLIER, J.-C., 1998, *Phys. Rev. Lett.*, **81**, 1437.

AMODEO, R., and GHONIEM, N. M., 1988, *Res Mech.*, **23**, 137; 1990a, *Phys. Rev.*, **41**, 6958; 1990b, *ibid.*, **41**, 6968; 1991, *Modeling of Deformation of Crystalline Solids*, edited by P. F. T. Lowe, T. Rollett and G. Daehn (Warrendale, Pennsylvania: Metallurgical Society of AIME), p. 125.

ANAND, L., and KOTHARI, M., 1996, *J. Mech. Phys. Solids*, **44**, 525.

ANISIMOV, V., 2000, *Strong Coulomb Correlations in Electronic Structure Calculations: Beyond the Local Density Approximation* (New York: Gordon and Breach).

ANISIMOV, V. I., ZAANEN, J., and ANDERSEN, O., 1991, *Phys. Rev.* B, **44**, 943.

ARSENLIS, A., and PARKS, D., 2001, *J. Mech. Phys. Solids*, **50**, 1979.

ASARO, R. J., and NEEDLEMAN, A., 1985, *Acta metall.*, **33**, 309.

ASARO, R. J., and RICE, J. R., 1977, *J. Mech. Phys. Solids*, **25**, 309.

ASHCROFT, N. W., and MERMIN, N. D., 1976, *Solid State Physics* (Philadelphia, Pennsylvania: Saunders College).

BAKER, S., 1993, *Mater. Res. Soc. Symp. Proc.*, **308**, 209.

BARBE, F., DECKER, I., JEULIN, D., and CAILLETAUD, G., 2002, *Int. J. Plast.*, **17**, 513.

BARTS, D. B., and CARLSSON, A. E., 1995, *Phys. Rev.* E, **52**, 3195; 1998, *Phil. Mag.* A, **77**, 751.

BASKES, M. I., 1987, *Phys. Rev. Lett.*, **59**, 2666; 2001, *J. Mech. Phys. Solids*, **49**, 1983.

BATIREV, I. G., ALAVI, A., FINNIS, M. W., and DEUTSCH, T., 1999, *Phys. Rev. Lett.*, **82**, 1510.

BATTAILE, C. C., and SROLOVITZ, D. J., 1997, *J. appl. Phys.*, **82**, 6293.

BEELER, J. R., 1982, *Radiation Effects Computer Experiments* (Amsterdam: North-Holland).

BERVEILLER, B., and ZAOUI, A., 1979, *J. Mech. Phys. Solids*, **26**, 325.

BESSON, J., CAILLETAUD, G., CHABOCHE, J.-L., and FOREST, S., 2002, *Mecanique Non-Linéaire des Matériaux*, first edition (Paris: Hermes Science).

BORTZ, A. B., KALOS, M. H., and LEBOWITZ, J. L., 1975, *J. Comput. Phys.*, **17**, 10.

BROWN, L. M., 1967, *Phil. Mag.*, **15**, 363.

BULATOV, V. V., ABRAHAM, F. F., KUBIN, L., DEVINCRE, B., and YIP, S., 1998, *Nature*, **391**, 669.

BULATOV, V. V., YIP, S., and ARGON, A. S., 1995, *Phil. Mag.* A, **72**, 453.

BUSSO, E., and CAILLETAUD, G., 2003, *J. Phys. Paris, IV*, **105**, 255.

BUSSO, E., and MCCLINTOCK, F., 1996, *Int. J. Plast.*, **12**, 1.

BUSSO, E. P., MEISSONNIER, F. T., and O'DOWD, N. P., 2000, *J. Mech. Phys. Solids*, **48**, 2333.

CAI, W., DE KONING, M., BULATOV, V. V., and YIP, S., 2000, *Phys. Rev. Lett.*, **85**, 3213.

CAMPBELL, G., FOILES, S., HUANG, H., HUGHES, D., KING, W., LASSILA, D., NIKKEL, D., DIAZ DE LA RUBIA, T., SHU, J. Y., and SMYSHLYAEV, V. P., 1998, *Mater. Sci. Engng*, **A251**, 1.

CANOVA, G., BRECHET, Y., and KUBIN, L. P., 1992, *Modeling of Plastic Deformation and its Engineering Applications*, edited by S. I. Anderson, J. B. Bilde-Sorensen, N. Hansen, D. Juul Jensen, T. Leffers, H. Lilholt, T. Lorentzen, O. B. Pedersen, and B. Ralph (Risø, Roskilde: Risø National Laboratory).

CANOVA, G., BRECHET, Y., KUBIN, L. P., DEVINCRE, B., PONTIKIS, V., and CONDAT, M., 1993, *Microstructures and Physical Properties*, edited by J. Rabiet (Geneva: CH-Transtech).

CAR, R., and PARRINELLO, M., 1985, *Phys. Rev. Lett.*, **55**, 2471.

CHAN, S.-P., CHEN, G., GONG, X., and LIU, Z., 2001, *Phys. Rev. Lett.*, **87**, 205 502.

CHANG, J. P., CAI, W., BULATOV, V. V., and YIP, S., 2001, *Mater. Sci. Engng*, **A309**, 160.

CHEN, N. X., 1990, *Phys. Rev. Lett.*, **64**, 1193.

CLERI, F., WOLF, D., YIP, S., and PHILLPOT, S. R., 1997, *Acta mater.*, **45**, 4993.

COHEN, R. E., MEHL, M., and PAPACONSTANTOPOULOS, D. A., 1994, *Phys. Rev. B*, **50**, 14 694.

CRISFIELD, M. A., 1997, *Non-linear Finite Element Analysis of Solids and Structures*, Vols 1 and 2, fourth edition (New York: Wiley).

CUITIÑO, A. M., and ORTIZ, M., 1992, *Modelling Simulation Mater. Sci. Engng*, **1**, 225.

DAW, M., and BASKES, M., 1983, *Phys. Rev. Lett.*, **50**, 1285; 1984, *Phys. Rev. B*, **29**, 6443.

DENTENEER, P. J., and VAN HAERINGEN, W., 1985, *J. Phys. C*, **18**, 4127.

DEVINCRE, B., and CONDAT, M., 1992, *Acta metall.*, **40**, 2629.

DEVINCRE, B., and KUBIN, L. P., 1994, *Modelling Simulation Mater. Sci. Engng*, **2**, 559.

DEVINCRE, B., PONTIKIS, V., BRECHET, Y., CANOVA, G., CONDAT, M., and KUBIN, L. P., 1992, *Microscopic Simulations of Complex Hydrodynamic Phenomena*, edited by M. Mareschal and B. L. Lolian (New York: Plenum), p. 413.

DEWIT, R., 1960, *Solid. St. Phys.*, **10**, 249.

DORAN, D. G., 1970, *Radiat. Effects*, **2**, 249.

DREIZLER, R. M., and GROSS, E. K. U., 1990, *Density Functional Theory* (Berlin: Springer).

DUESBERY, M. S., and VITEK, V., 1998, *Acta mater.*, **46**, 1481.

DUMITRICA, T., BELYTSCKO, T., and YAKOBSON, B. I., 2003, *J. chem. Phys.*, **118**, 9485.

DÜRR, M., RASCHKE, M. B., PEHLKE, E., and HÖFER, U., 2001, *Phys. Rev. Lett.*, **86**, 123.

DYKSTRA, C. E., 1988, *Ab initio Calculation of the Structure and Properties of Molecules* (Amsterdam: Elsevier).

El-AZAB, A., 2000, *Phys. Rev. B*, **61**, 11956.

ERB, U., PALUMBO, G., ZUGIC, R., and AUST, K., 1996, *Treatise on Materials Science and Engineering* (Pittsburgh, Pennsylvania: The Minerals, Metals and Materials Research Society), p. 93.

ERCOLESSI, F., and ADAMS, J., 1994, *Europhys. Lett.*, **26**, 583.

ERCOLESSI, F., TOSATTI, E., and PARRINELLO, M., 1986, *Phys. Rev. Lett.*, **57**, 719.

ESCHE, S. K., KINZEL, G. L., and ALTAN, T., 1997, *Int. J. numer. Meth. Engng*, **40**, 4577.

ESCHRIG, H., 1989, *Optimized LCAO Method and the Electronic Structure of Extended Systems* (Berlin: Springer).

ESHELBY, J. D., 1957, *Proc. R. Soc. A*, **241**, 376.

ESSMANN, U., and MUGHRABI, H., 1979, *Phil. Mag. A*, **40**, 731.

FINNIS, M., and SINCLAIR, J., 1984, *Phil. Mag. A*, **50**, 45.

FLECK, N. A., and HUTCHINSON, J., 1997, *Adv. appl. Mech.*, **33**, 295; 2001, *J. Mech. Phys. Solids*, **49**, 2245.

FOREST, S., 1998, *Acta mater.*, **46**, 3265.

FRAUENHEIM, T., POREZAG, D., ELSTNER, M., JUNGNICKEL, G., ELSNER, J., HAUGK, M., and SIECK, A., 1998, *Mater. Res. Soc. Symp. Proc.*, **491**, 91.

FULDE, P., 1993, *Electron Correlations in Molecules and Solids* (Berlin: Springer).

GALLI, G., and MAURI, F., 1994, *Phys. Rev. Lett.*, **73**, 3471.

GAO, H., and HUANG, Y., 2003, *Scripta metall.*, **48**, 113.

GAVAZZA, S., and BARNETT, D., 1976, *J. Mech. Phys. Solids*, **24**, 171.

GEAR, C., 1971, *Numerical Initial Value Problems in Ordinary Differential Equations* (Englewood Cliffs, New Jersey: Prentice-Hall).

GEORGES, A., KOTLIAR, G., KRAUTH, W., and ROZENBERG, M., 1996, *Rev. mod. Phys.*, **68**, 13.

GHONIEM, N. M., 1992, *Non-Linear Phenomena in Material Science II*, edited by L. Kubin and G. Martin (Dordrecht: Kluwer); 1999, *J. Engng Mater. Technol.*, **121**, 136.

GHONIEM, N. M., and AMODEO, R. J., 1988a, *Solid St. Phenomena*, **3–4**, 377; 1988b, *Int J. Engng Sci.*, **26**, 653; 1990, *Patterns, Defects and Material Instabilities*, edited by D. Walgreaf and N. Ghoniem (Dordrecht: Kluwer), p. 303.

GHONIEM, N. M., HUANG, J., and WANG, Z., 2001, *Phil. Mag. Lett.*, **82**, 55.

GHONIEM, N. M., MATTHEWS, J. R., and AMODEO, R. J., 1990, *Res Mech.*, **29**, 197.

GHONIEM, N. M., and SUN, L. Z., 1999, *Phys. Rev.* B, **60**, 128.

GHONIEM, N. M., TONG, S., HUANG, J., SINGH, B., and WEN, M., 2002, *J. nucl. Mater.*, **307–311**, 843.

GHONIEM, N. M., TONG, S.-H., and SUN, L. Z., 2000, *Phys. Rev.*, **61**, 913.

GHOSH, S., NOWAK, Z., and LEE, K., 2002, *Acta mater.*, **45**, 2215.

GILLAN, M. J., 1989, *J. Phys.: condens. Matter*, **1**, 689.

GOEDECKER, S., 1999, *Rev. of mod. Phys.*, **71**, 1085.

GOLUBOV, S. I., LIU, X. L., HUANG, H. C., and WOO, C. H., 2001, *Comput. Phys. Commun.*, **137**, 312.

GREGOR, V., and KRATOCHVIL, J., 1990, *Res. Mech.*, **29**, 197.

GROMA, I., 1997, *Phys. Rev.* B, **56**, 5807.

GROMA, I., and PAWLEY, G. S., 1993, *Phil. Mag.* A, **67**, 1459.

GROSS, E. K. U., KREIBICH, T., LEIN, M., and PETERSILKA, M., 1998, *Electron Correlations and Materials Properties*, edited by N. K. A. Gonis and M. Ciftan (New York: Kluwer-Plenum), p. 393.

GULUOGLU, A. N., SROLOVITZ, D. J., LeSAR, R., and LOMDAHL, R. S., 1989, *Scripta metall.*, **23**, 1347.

GUMBSCH, P., and GAO, H., 1999, *Science*, **283**, 965.

GUNNARSSON, O., and LUNDQVIST, B. I., 1976, *Phys. Rev.* B, **13**, 4274.

GURTIN, M. E., 2000, *J. Mech. Phys. Solids*, **48**, 989; 2003, *Int. J. Plast.*, **19**, 47.

HÄHNER, P., 1996, *Appl. Phys.* A, **62**, 473.

HÄHNER, P., BAY, K., and ZAISER, M., 1998, *Phys. Rev. Lett.*, **81**, 2470.

HAMILTON, J. C., 1997, *Phys. Rev.* B, **55**, R7402.

HAMILTON, J. C., STUMPF, R., BROMANN, K., GIOVANNINI, M., KERN, K., and BRUNE, H., 1999, *Phys. Rev. Lett.*, **82**, 4488.

HARRIS, J., and JONES, R. O., 1974, *J. Phys.* F, **4**, 1170.

HARTLEY, C., 2003, *Proceedings of the International Conference on Plasticity*, Quebec City, Canada. *Phil. Mag.*, this issue.

HEDIN, L., and LUNDQVIST, S., 1969, *Solid St. Phys.*, **23**, 1; 1971, *J. Phys.* C, **4**, 2064.

HEINISCH, H. L., 1995, *Nucl. Instrum. Meth.* B, **102**, 47.

HERVE, E., and ZAOUI, A., 1993, *Int. J. Engng Sci.*, **31**, 1.

HILL, R., 1950, *The Mathematical Theory of Plasticity*, fourth edition (Oxford: Clarendon); 1965, *J. Mech. Phys. Solids*, **14**, 89.

Hirth, J. P., Rhee, M., and Zbib, H., 1996, *J. Comput.-Aided Mater. Des.*, **3**, 164.

HOAGLAND, R. G., HIRTH, J. P., and GEHLEN, P. C., 1976, *Phil. Mag.* A, **34**, 413.

HOHENBERG, P., and KOHN, W., 1964, *Phys. Rev.*, **136**, B864.

HOOVER, W. G., 1985, *Phys. Rev.* A, **31**, 1695.

HUANG, H. C., GHONIEM, N. M., DIAZ DE LA RUBIA, T., RHEE, M., ZBIB, H., and HIRTH, J. P., 1999, *J. Engng Mater. Technol.*, **121**, 143.

HUANG, H. C., GHONIEM, N. M., WONG, J. K., and BASKES, M. I., 1995, *Modelling Simulation Mater. Sci. Engng*, **3**, 615.

HUBER, N., and TSAKMAKIS, C., 1999, *J. Mech. Phys. Solids*, **47**, 1589.

HUGHES, D., CHRZAN, D., LIU, Q., and HANSEN, N., 1998, *Phys. Rev. Lett.*, **81**, 4664.

ISMAIL-BEIGI, S., and ARIAS, T. A., 1999, *Phys. Rev. Lett.*, **82**, 2127.

JAMES, F., 1980, *Rep. Prog. Phys.*, **43**, 1145; 1990, *Comput. Phys. Commun.*, **60**, 329.

JANOTTI, A., ZHANG, S. B., WEI, S.-H. and VAN DE WALLE, C. G., 2002, *Phys. Rev. Lett.*, **89**, 086 403.

JONES, R. O., and GUNNARSSON, O., 1989, *Rev. mod. Phys.*, **61**, 689.

JORDAN, E., and WALKER, K., 1992, *J. Engng Mater. Technol.*, **114**, 19.

JUSTO, J. F., DE KONING, M., CAI, W., and BULATOV, V. V., 2001, *Mater. Sci. Engng*, **A309**, 129.

KALIDINDI, S., BRONKHORST, C., and ANAND, L., 1992, *J. Mech. Phys. Solids*, **40**, 537.

KENT, P., and ZUNGER, A., 2001, *Phys. Rev. Lett.*, **86**, 2613.

KHACHATURYAN, A. G., 2000, *Proceedings of a Symposium, Science of Alloys for the 21st Century: A Hume–Rothey Symposium Celebration*, edited by E. Turchi and A. G. A. Shull (Warrendale, Pennsylvania: Metallurgical Society of AIME), p. 293.

KIDO, T., MARUYAMA, T., NIITA, K., and CHIBA, S., 2000, *Nucl. Phys.* A, **663**, 877c.

KOHN, W., and SHAM, L., 1965, *Phys. Rev.*, **140**, A1133.

KOSSEVICH, A. M., 1979, *Dislocations in Solids*, Vol. 1, edited by F. Nabarro (New York: North-Holland), p. 33; 1999, *Crystal Lattice Phonons, Solitons, and Dislocations* (New York: Wiley).

KRATOCHVIL, J., and SAXLOVÀ, N., 1992, *Scripta metall. mater.*, **26**, 113.

KRÖNER, E., 1958a, *Kontinuumstheorie der Versetzungen und Eigenspannungen* (Berlin: Springer); 1958b, *Erg. angew. Math.*, **5**, 1; 1981, *Physics of Defects*, *Les Houches Session XXXV* (Amsterdam: North-Holland), p. 217.

KUBIN, L. P., 1993, *Phys. Stat. sol.* (a), **135**, 433.

KUBIN, L. P., and CANOVA, G., 1992, *Scripta metall. mater.*, **27**, 957.

KUBIN, L. P., CANOVA, G., CONDAT, M., DEVINCRE, B., PONTIKIS, V., and BRECHET, Y., 1992, *Diffusion Defect Data—Solid St. Data*, B, **23–24**, 455.

KUBIN, L. P., and KRATOCHVIL, J., 2000, *Phil. Mag.* A, **80**, 201.

KUKTA, R. V., and FREUND, L. B., 1998, *Multiscale Modelling of Materials*, edited by V. Bulatov, T. Diaz de la Rubia, R. Phillips, E. Kaxiras and N. Ghoniem (Boston, Massachusetts: Materials Research Society), p. 99.

KURAMOTO, E., OHSAWA, K., and TSUTSUMI, T., 2000, *J. Comput.-Aided Mater. Des.*, **7**, 89.

LaBELLA, V. P., YANG, H., BULLOCK, D. W., THIBADO, P. M., KRATZER, P., and SCHEFFER, M., 1999, *Phys. Rev. Lett.*, **83**, 2989.

LANGRETH, D. C., and PERDEW, J. P., 1977, *Phys. Rev.* B, **15**, 2884.

LEE, E. H., 1969, *J. appl. Mech.*, **36**, 1.

LEFFERS, T., 2001, *Acta mater.*, **17**, 469.

LENNARD-JONES, J. E., 1924, *Proc. R. Soc.* A, **106**, 463.

LEPINOUX, J., and KUBIN, L. P., 1987, *Scripta metall.*, **21**, 833.

LeSAR, R., and RICKMAN, J. M., 2002, *Phys. Rev.* B, **65**, 144 110; 2003, *Phil Mag.*, this issue.

LeSAR, R., NAJAFABADI, R., and SROLOVITZ, D. J., 1989, *Phys. Rev. Lett.*, **63**, 624.

LEWIS, J. P., ORDEJON, P., and SANKEY, O., 1997, *Phys. Rev.* B, **55**, 6880.

LI, Q. K., and SHI, S. Q., 2002, *Appl. Phys. Lett.*, **80**, 3069.

LI, X. P., NUNES, R. W., and VANDERBILT, D., 1993, *Phys. Rev.* B, **47**, 10 891.

LICHTENSTEIN, A. I., ZAANEN, J., and ANISIMOV, V. I., 1995, *Phys. Rev.* B, **52**, R5467.

LIU, W. C., SHI, S. Q., HUANG, H. C., and WOO, C. H., 2002a, *Comput. Mater. Sci.*, **23**, 155.

LIU, W. C., SHI, S. Q., WOO, C. H., and HUANG, H. C., 2002b, *Comput. Modelling Engng Sci.*, **3**, 213.

LOUBET, J., GEORGES, J., and MEILLE, G., 1986, *Microindentation Techniques in Materials Science and Engineering*, ASTM Special Technical Publication 889, edited by P. Blau and B. Lawn (Philadelphia, Pennsylvania: American Society for Testing and Materials), p. 72.

LOUIE, S. G., 1996, *Quantum Theory of Real Materials*, edited by J. R. Chelikowsky and S. G. Louie (Boston, Massachusetts Kluwer), p. 83.

LU, G., BULATOV, V., and KIOUSSIS, N., 2002, *Phys. Rev.* B, **66**, 144 103.

LU, G., and KIOUSSIS, N., 2001, *Phys. Rev.* B, **64**, 024 101.

LU, G., KIOUSSIS, N., BULATOV, V., and KAXIRAS, E., 2000, *Phys. Rev.* B, **62**, 3099.

LU, G., KIOUSSIS, N., WU, R., and CIFTAN, M., 1999, *Phys. Rev.* B, **59**, 891.

LU, G., ZHANG, Q., KIOUSSIS, N., and KAXIRAS, E., 2001, *Phys. Rev. Lett.*, **87**, 095 501.

LUBARDA, V. A., BLUME, B. J. A., and NEEDLEMAN, A., 1993, *Acta metall. mater.*, **41**, 625.

MARZARI, N., and VANDERBILT, D., 1997, *Phys. Rev.* B, **56**, 12 847.

MEHL, M., and PAPACONSTANTOPOULOS, D. A., 1996, *Phys. Rev.* B, **54**, 4519.

MEISSONNIER, F., BUSSO, E., and O'DOWD, N., 2001, *Int. J. Plast.*, **17**, 601.

MERIC, L., and CAILLETAUD, G., 1991, *J. Engng Mater. Technol.*, **113**, 171.

MERIC, L., POUBANNE, P., and CAILLETAUD, G., 1991, *J. Engng Mater. Technol.*, **162–170**, 947.

METROPOLIS, N., ROSENBLUTH, A. W., ROSENBLUTH, M. N., TELLER, A. H., and TELLER, E., 1953, *J. chem. Phys.*, **21**, 1087.

MOLINARI, A., 2002, *J. Engng Mater. and Technol.*, **124**, 62.

MORIARTY, J. A., 1988, *Phys. Rev.* B, **38**, 3199; 1994, *ibid.*, **49**, 12 431.

MORIARTY, J. A., BELAK, J. F., RUDD, R. E., SÖDERLIND, P., STREITZ, F. H., and YANG, L. H., 2002, *J. Phys.: condens. Matter*, **14**, 2825.

MORIARTY, J. A., and WIDOM, M., 1997, *Phys. Rev.* B, **56**, 7905.

MORSE, P. M., 1929, *Phys. Rev.*, **34**, 57.

MOULIN, A., CONDAT, M., and KUBIN, L. P., 1997, *Acta mater.*, **45**, 2339.

MUGHRABI, H., 1981, *Proceedings of the Fourth International Conference on Continuum Models of Discrete Systems*, edited by O. Brulin and R. K. T. Hsi (Amsterdam: North Holland); 1983, *Acta metall.*, **31**, 1367; 1987, *Mater. Sci. Engng*, **85**, 15.

MÜHLHAUS, H. B., 1989, *Int. Arch.*, **59**, 124.

MURA, T., 1968, *The Continuum Theory of Dislocations, Advances in Materials Research 3* (New York: Interscience); 1982, *Micromechanics of Defects in Solids* (Dordrecht: Martinus Nijhoff).

NAN, C.-W., and CLARKE, D., 1996, *Acta mater.*, **44**, 3801.

NEUHÄUSER, H., 1983, *Dislocations in Solids*, Vol. 6, edited by F. Nabarro (Amsterdam: North-Holland).

NEUMANN, P., 1971, *Acta metall.*, **19**, 1233.

NIKOLAEV, P., THESS, A., RINZLER, A. G., COLBERT, D. T., and SMALLEY, R., 1997, *Chem. Phys. Lett.*, **266**, 422.

NOSE, S., 1984, *Mol. Phys.*, **52**, 255.

NOUAILHAS, D., and FREED, A., 1992, *J. Engng Mater. Techol.*, **114**, 97.

NUNES, R. W., BENNETTO, J., and VANDERBILT, D., 1996, *Phys. Rev. Lett.*, **77**, 1516.

OHSAWA, K., and KURAMOTO, E., 1999, *J. appl. phys.*, **86**, 179.

ORDEJON, P., DRABOLD, D. A., GRUMBACH, M. P., and MARTIN, R. M., 1993, *Phys. Rev.* B, **48**, 14 646.

ORTIZ, M., and PHILLIPS, R., 1999, *Adv. appl. Mech.*, **36**, 1.

OSETSKY, Y. N., BACON, D., SERRA, A., SINGH, B., and GOLUBOV, S., 2000, *J. nucl. Mater.*, **276**, 65.

PAPANIKOLAOU, N., ZELLER, R., and DEDERICHS, P., 2002, *J. Phys.: condens. Matter*, **14**, 2799.

PARR, R. G., and YANG, W., 1989, *Density Functional Theory of Atoms and Molecules* (Oxford University Press).

PARRINELLO, M., and RAHMAN, A., 1981, *J. appl. Phys.*, **52**, 7182.

PASIANOT, R., FARKAS, D., and SAVINO, E. J., 1991, *Phys. Rev.* B, **43**, 6952.

PAYNE, M. C., TETER, M. P., ALLAN, D. C., ARIAS, T. A., and JOANNOPOULOS, J. D., 1992, *Rev. mod. Phys.*, **64**, 1045.

PEACH, M. O., and KOEHLER, J. S., 1950, *Phys. Rev.*, **80**, 436.

PERDEW, J. P., 1991, *Electronic Structure of Solids '91*, edited by P. Ziesche and H. Eschrig (Berlin: Akademic), p. 11; 1999, *Electron Correlations and Materials Properties*, edited by A. Gonis, N. Kioussis and M. Ciftan (New York: Kluwer and Plenum), p. 287.

PERDEW, J. P., BURKE, K., and WANG, Y., 1996, *Phys. Rev.* B, **54**, 16 533.

PERDEW, J. P., and ZUNGER, A., 1981, *Phys. Rev.* B, **23**, 5048.

PETTIFOR, D. G., 1989, *Phys. Rev. Lett.*, **63**, 2480.

PETTIFOR, D. G., and OLEINIK, I., 1999, *Phys. Rev.* B, **59**, 8487; 2000, *Phys. Rev. Lett.*, **84**, 4124.

PIERCE, D., ASARO, R. J., and NEEDLEMAN, A., 1983, *Acta metall.*, **31**, 31 951.

PITAKTHAPANAPHONG, S., and BUSSO, E., 2002, *J. Mech. Phys. Solids*, **50**, 695.

PONTE CASTAÑEDA, P., 1991, *J. Mech. Phys. Solids*, **39**, 45.

PRICE, D., and COOPER, B., 1989, *Phys. Rev.* B, **39**, 4945.

PRICE, D., WILLS, J., and COOPER, B., 1992, *Phys. Rev.* B, **46**, 11 368.

RAO, S., HERNANDEZ, C., SIMMONS, J. P., PARTHASARATHY, T. A., and WOODWARD, C., 1998, *Phil. Mag.* A, **77**, 231.

RAO, S. I., and WOODWARD, C., 2001, *Phil. Mag.* A, **81**, 1317.

RAPPE, A. M., JOANNOPOULOS, J. D., and BASH, P. A., 1992, *J. Am. chem. Soc.*, **114**, 6466.

REGINO, M., BUSSO, E., O'DOWD, N., and ALLEN, D., 2002, *Materials for Advanced Power Engineering, Part I*, edited by J. Lecomte-Beckers, M. Carton, F. Schubert and P. J. Ennis (Forschungszentrum Julich), p. 283.

RHEE, M., ZBIB, H. M., HIRTH, J. P., HUANG, H., and DE LA RUBIA, T., 1998, *Modelling Simulation Mater. Sci. Engng*, **6**, 467.

RICKMAN, J., LESAR, R., and SROLOVITZ, D., 2003, *Acta mater.*, **51**, 1199.

RODNEY, D., and MARTIN, G., 2000, *Phys. Rev.* B, **61**, 8714.

RODNEY, D., and PHILLIPS, R., 1999, *Phys. Rev. Lett.*, **82**, 1704.

RUBERTO, C., YOURDSHAHYAN, Y., and LUNDQVIST, B. I., 2002, *Phys. Rev. Lett.*, **88**, 226 101.

SAITO, T., FURUTA, T., HWANG, J., KURAMOTO, S., NISHINO, K., SUZUKI, N., CHEN, R., YAMADA, A., ITO, K., SENO, Y., NONAKA, N., IKEHATA, H., NAGASAKO, N., IMAMOTO, C., IKUHARA, Y., and SAKUMA, T., 2003, *Science*, **300**, 464.

SAXLOVÀ, N., KRATOCHVIL, J., and ZATLOUKAL, J., 1997, *Mater. Sci. Engng*, **A234**, 205.

SCHONFELDER, B., KEBLINSKI, P., WOLF, D., and PHILLPOT, S. R., 1999, *Mater. Sci. Forum*, **294**, 9.

SCHONFELDER, B., WOLF, D., PHILLPOT, S. R., and FURTKAMP, M., 1997, *Interface Sci.*, **5**, 245.

SCHUBERT, F., FLEURY, G., and STEINHAUS, T., 2000, *Modelling Simulation Sci. Engng*, **8**, 947.

SCHWARZ, K. V., and TERSOFF, J., 1996, *Appl. Phys. Lett.*, **69**, 1220.

SCHWARZ, K. W., 1997, *Phy. Rev. Lett.*, **78**, 4785.

SCHWARZ, K. W., and LEGOUES, F. K., 1997, *Phys. Rev. Lett.*, **78**, 1877.

SCUDDER, H., LU, G., and KIOUSSIS, N., 2003, *Phys. Rev.* B (to be published).

SEGMUELLER, A., and MURAKAMI, M., 1988, *Treatise Mater. Sci. Engng*, **27**, 143.

SHI, S. Q., HUANG, H. C., and WOO, C. H., 2002, *Comput. Mater. Sci.*, **23**, 95.

SHIM, Y., LEVINE, L., and THOMSON, R., 2001, *Mater. Sci. Engng*, **A309–A310**, 340.

SHU, J., FLECK, N., VAN DER GIESSEN, E., and NEEDLEMAN, A., 2001, *J. Mech. Phys. Solids*, **49**, 1361.

SHU, J., KING, W. E., and FLECK, N. A., 1999, *Int. J. Numer. Meth. Engng*, **44**, 373.

SINCLAIR, J. E., 1971, *J. appl. Phys.*, **42**, 5321.

SMALL, M., DANIELS, B., CLEMENS, B., and NIX, W., 1994, *J. Mater. Res.*, **9**, 25.

STAINER, L., CUTINO, A., and ORTIZ, M., 2002, *J. Mech. Phys. Solids*, **50**, 1511.

STAUFFER, D., and AHARONY, A., 1992, *Introduction to Percolation Theory* (London: Taylor & Francis).

STILLINGER, F. H., and WEBER, T. A., 1985, *Phys. Rev.*, B, **31**, 5262.

SVANE, A., and GUNNARSSON, O., 1990, *Phys. Rev. Lett.*, **65**, 1148.

SWADENER, J., MISRA, A., HOAGLAND, R., and NASTASI, M., 2002, *Scripta metall.*, **47**, 343.

TERSOFF, J., 1986, *Phys. Rev. Lett.*, **56**, 632.

THOMSON, R., and LEVINE, L., 1998, *Phys. Rev. Lett.*, **81**, 3884.

THOMSON, R., LEVINE, L. E., SHIM, Y., and SAVAGE, M. F., 2002, *Comput. Modelling Engng Sci.*, **3**, 245.

VAN HOUTTE, P., DELANNAY, L., and SAMAJDAR, I., 1999, *Textures Microstruct.*, **31**, 109.

VAN KAMPEN, N., 1992, *Stochastic Processes in Physics and Chemistry* (Amsterdam: North-Holland).

VAN SWYGENHOVEN, H., 2002, *Science*, **296**, 66.

VAN SWYGENHOVEN, H., CARO, A., and FARKAS, D., 2001, *Mater. Sci. Engng*, **309**, 440.

VITEK, V., 1974, *Crystal Lattice Defects*, **5**, 1.

VOLTERRA, V., 1907, *Annal. Sci. École Norm. Super.*, Paris, **24**, 401.

VON BARTH, U., and HEDIN, L., 1972, *J. Phys.* C, **5**, 1629.

VOSKO, S. H., WILK, L., and NUSAIR, M., 1980, *Can. J. Phys.*, **58**, 1200.

WALGRAEF, D., and AIFANTIS, C., 1985, *Int. J. Engng Sci.*, **23**, 1351.

WANG, G. F., STRACHAN, A., CAGIN, T., and GODDARD, W. A., 2001a, *Mater. Sci. Engng*, **A309**, 133.

WANG, H. Y., and LESAR, R., 1995, *Phil. Mag.* A, **71**, 149.

WANG, J., WOO, C. H., and HUANG, H. C., 2001b, *Appl. Phys. Lett.*, **79**, 101.

WANG, L.-W., KIM, J., and ZUNGER, A., 1999, *Phys. Rev.* B, **59**, 5678.

WANG, Y., JIN, Y. M., CUITINO, A. M., and KHACHATURYAN, A. G., 2000, *Proceedings of the International Conference on Dislocations 2000* (Gaithersburg, Maryland: National Institute of Standards and Technology), p. 107; 2001c, *Acta mater.*, **49**, 1847.

WANG, Z., GHONIEM, N., LESAR, R., and SRIRAM, S., 2003a, *Modelling Simulation Mater. Sci. Engng* (submitted).

WANG, Z., LESAR, R., and GHONIEM, N., 2003b (to be published).

WIMMER, E., KRAKAUER, H., WEINERT, M., and FREEMAN, A., 1981, *Phys. Rev.* B, **24**, 864.

WOO, C. H., HUANG, H. C., and ZHU, W. J., 2002, *Appl. Phys.* A, **76**, 101.

WOO, C. H., and PULS, M. P., 1976, *J. Phys.* C, **9**, L27.

WOODWARD, C., and RAO, S. I., 2002, *Phys. Rev. Lett.*, **88**, 216 402.

WU, R., FREEMAN, A., and OLSON, G., 1994, *Science*, **265**, 376.

WU, T., LEE, C.-K., and WANG, R., 1993, *Mater. Res. Soc. Symp. Proc.*, **308**, 133.

XU, W., and MORIARTY, J. A., 1996, *Phys. Rev.* B, **54**, 6941; 1998, *Comput. Mater. Sci.*, **9**, 348.

YANG, W., 1991, *Phys. Rev. Lett.*, **66**, 1438.

YILDIRIM, T., GÜLSEREN, O., and CIRACI, S., 2001, *Phys. Rev.*, B, **64**, 075 404.

ZAISER, M., 2001, *Mater. Science Engng*, **A309–A310**, 304.

ZAISER, M., AVLONITIS, M., and AIFANTIS, E. C., 1998, *Acta mater.*, **46**, 4143.

ZAISER, M., and HÄHNER, P., 1997, *Phys. Stat. sol.* (b), **199**, 267.

ZBIB, H., and AIFANTIS, E., 1988, *Res. Mech.*, **23**, 261.

ZBIB, R. M., RHEE, M., and HIRTH, J. P., 1998, *Int. J. Mech. Sci.*, **40**, 113.

ZHANG, W. Q., XIE, Q., GE, X. J., and CHEN, N. X., 1997, *J. appl. Phys.*, **82**, 578.

ZHOU, S. J., PRESTON, D. L., LOMDAHL, P. S., and BEAZLEY, D. M., 1998, *Science*, **279**, 1525.

ZIENKIEWICZ, O. C., and TAYLOR, R. L., 1994, *The Finite Element Method*, fourth edition (New York: McGraw-Hill).