HIGHLIGHT OF THE MONTH

First Principles Simulations of Nanoindentation and Atomic Force Microscopy on Silicon Surfaces

Ruben Perez* and Michael C. Payne
Theory of Condensed Matter, Cavendish Laboratory
University of Cambridge
Madingley Road, Cambridge CB3 0HE, U.K.

Ivan Štich and Kiyoyuki Terakura
JRCAT, Angstrom Technology Partnership
1–1–4 Higashi, Tsukuba, Ibaraki 305, Japan

Introduction

One of the most exciting trends in Solid State Physics in the last two decades is the fabrication of structures where several of the characteristic dimensions approach the atomic limit. Characterization techniques have evolved in the same direction, probing the properties of matter on progressively shorter scales. This is the case of indentation experiments, where the use of an atomic force microscope with a sharp indenter has made it possible to reduce the depth of the indentation from the several microns achieved in the 70's to just one nanometer[1]. At the same time, new techniques like the Scanning Tunneling Microscope (STM)[2] and the Atomic Force Microscope (AFM)[3] have been specifically devised to operate in the atomic scale. The development of these techniques and, especially, the interpretation of the images obtained with the tip–based microscopes should rely on a detailed knowledge of the interaction between the tip and the surface. The purpose of this paper is to show what we can learn about that interaction by the quantum mechanical simulation of the problem. The methods used, Total–Energy Pseudopotential calculations, were well established in the mid 80's, but their application have received a great impulse by the use of iterative minimization techniques, a work pioneered by Car and Parrinello[4], and the development of massively parallel computing. Nowadays, we are able to handle systems with hundreds of atoms, and tackle problems which are relevant to Material Science. In particular, we will show the ability of the technique in describing two extreme regimes in the tip–surface interaction problem: The nanoindentation regime, where tip and sample are in close mechanical contact, and the operation of the AFM in the attractive non-contact regime.

Atomic scale simulations using simplified interatomic potentials have already been applied to the problem of tip–surface interaction[5, 6, 7]. These studies have provided significant information about the process, but we consider that the explicit consideration of the electronic degrees of freedom provided by *ab initio* methods[8] is crucial to describe the breaking and remaking of bonds involved in the plastic response, and the weak interaction

between the dangling bonds of the adatoms and the apex atom in the tip which provides the contrast mechanism in the non–contact AFM experiments.

Nanoindentation

In indentation experiments, an indenter –made out of a hard material– is pressed against the surface of a sample. From the deformation of the surface we can learn about the mechanical properties of the sample. For small indentations, there is a perfect elastic recovery of the surface, but in general plastic deformations are induced during the loading process, and they remain after the indenter is removed from the surface.

The typical output of these experiments is a relation between the load applied and the indentation depth. While the elastic contribution to these load–indentation depth curves can be handled in an exact way, using the classical solutions of Hertz, very little is known about the plastic deformation, and empirical relations for the yield strength of the material are used in their modelling. As we have mentioned before, the technique has moved in the direction of smaller indentations, and indentation depths of the order of nanometers are now feasible. Due to the higher loads per atom reached in these nanoindentations, plasticity effects are expected to become particularly important. Our simulations will describe the mechanical response of the system for these nanometer–size indentation and explore the microscopic mechanisms involved in the onset and development of plasticity and the active role of the bonding sites of the tip in this process. These results are relevant to the description of this particular regime, as well as to the modelling of the indentation tests used in materials characterization.

The system that we have considered is a supercell containing a hard metallic indenter, a Si (111) slab and a vacuum region. The indenter is a sharp tetrahedral tip built up of twenty Al atoms stacked in four (111)-FCC planes. The relative positions of the atoms in the indenter are held fixed during the whole indentation process. This model retains the most important characteristics of a hard metallic tip, providing an accurate description of the bonding between the atoms in the surface and those on the tip. The Si slab contains 240 atoms and is built up of eight (111) planes in an almost square two-dimensional $5\times3\ \ \overline{3}$ unit cell each containing thirty atoms. The upper surface of the slab in contact with the tip has a 2×1 Pandey reconstruction[9]. The bottom surface is kept unreconstructed by saturating each dangling bond with an H atom. The total dimensions of the supercell, including the vacuum region of 6.26 Å, are $19.2 \times 19.9 \times 28.15$ Å. Initially the indenter is placed on top of the uppermost buckled atom in one of the π–bonded chains. The orientation of the tip is symmetric with respect to the (110) planes of the slab.

The indentation process was simulated in a stepwise, quasi–static manner by making small movements of the tip normal to the slab. At each step the atoms in the slab were allowed to relax to their equilibrium positions (zero forces) for that particular position of the tip[10]. Optimized non–local pseudopotentials[11, 12] including only s and p components

were used to describe the Al and Si ion cores[13]. The pseudopotentials were applied in the Kleinman–Bylander form [14] using the real space projection technique [15]. The electronic states were expanded at the Γ point of the Brillouin zone. A cutoff for the plane wave basis set of 7 Ry was used[16].

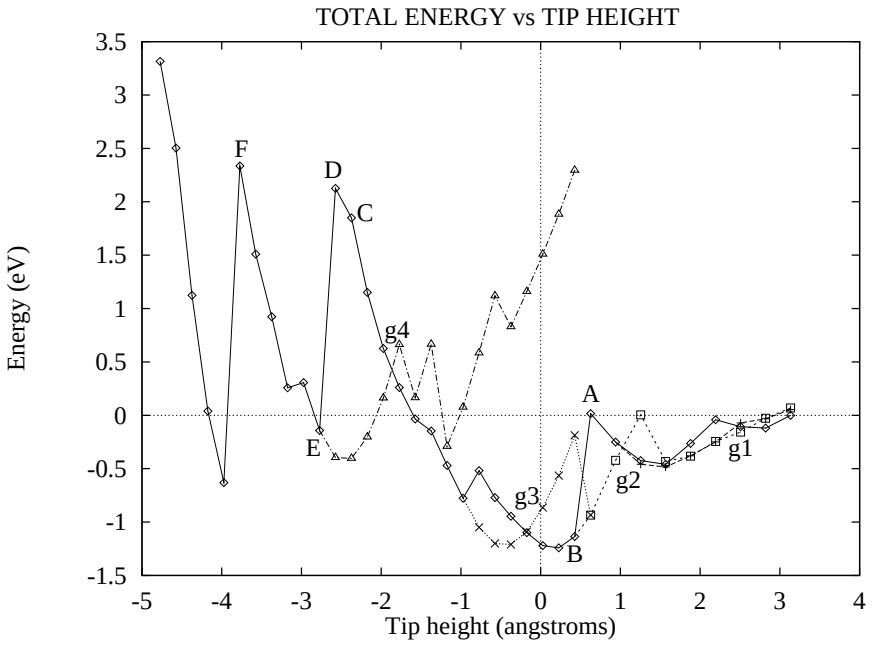

Figure 1: Total–energy of the system as a function of the the tip height. The continuous line corresponds to the indentation process while the different dashed lines refer to the retraction of the tip from different stages of the indentation process. The tip height is referred to the position of the last Si layer in an ideal Si (111) surface.

The total energy of the system for the different positions of the tip during the indentation process (black continuous line) and the retraction of the tip from different stages of the indentation (dashed lines) is shown in Fig. 1. We will consider first the indentation process. The results show a pattern of elastic deformations followed by plastic deformations seen as discontinuous jumps in the energy. These plastic deformations are associated with the breaking of few bonds where the stress has acumulated during the elastic regime. A few atoms in the structure undergo large displacements (of the order of 1 Å) to form a new stable bonding configuration in which the stress has been released. This new structure is elastically deformed upon further loading. The other remarkable feature in the total–energy curve is that it shows work hardening of the material: the stress necessary to produce those plastic deformations increases with the deformation.

As concerned the simulations pulling the tip back from different stages of the indentation process, the total–energy (dashed lines in Fig. 1) shows hysteresis, with several different stable structures for a given value of the tip height, and the system following different paths in the different backward displacements. This hysteric behaviour is related to the

breaking of the bonds already formed between the tip and the sample during the loading process. More importantly, the plot shows a transition from perfect recovery of the original structure of the Si slab for small indentations to permanent damage when we unload from a point beyond the step where plastic flow in the second double layer has taken place.

A detailed account of the atomistic processes involved in plastic flow for the different stages of the indentation will be published elsewhere[17]. Let us just mention that two different regimes can be identified. The first one (A–C, in Fig. 1) is characterized by an increasing coordination of the atoms in the first Si surface double layer with atoms in the walls of the tip, and by elastic deformations induced in the Si second double layer. The competition between the bonds formed with the atoms in the tip and the elastic deformation induced in the Si tetrahedral network determines that, upon further indentation, those atoms undergo a stick–slip motion along the walls of the tip, recovering positions in the direction of indentation which are very close to the ones in the original undeformed structure. This stick–slip motion is responsible for the friction, and hence the dissipation observed in earlier simulations[7].

After step C the system can no longer resist the induced stress with a pure elastic deformation, and plastic flow in the second double layer is induced. The transition to this new regime is mediated by the *spontaneous* breaking of the symmetry of the displacement pattern in the [110] direction. The mechanisms which dominate the plastic response are the plastic flow to the interstitials, inside the slab, and the shear flow around the indenter in the surface. As indentation proceeds (E–F), the plastically deformed areas in the surface and below the indenter make contact and we reach the full plastic flow regime, with plastic deformation inside the slab becoming easier because of the upward displacements of atoms in the planes above[17]. The evolution of plastic flow observed in the simulation supports the model used in some continuum elastic theories of the indentation contact to include plasticity effects[18].

Finite temperature and strain rate effects, not considered in our quasi-static, T= 0 simulations are known to influence the indentation process. An estimate of their effect can be obtained by calculating the energy barriers at the crossing points between the different stable structures found in the total–energy diagram. These barriers have been determined for four different crossings (labelled g1–g4 in Fig. 1). In the first three cases (g1–g3), the barriers are quite low: 0.10, 0.12 and 0.30 eV. These processes can be thermally activated easily and the system would follow the lowest energy structure available, avoiding states of high stress like A. On the contrary, the last transformation (g4) has a high barrier of 1.20 eV. A more detailed account of the influence of temperature and strain rate will be given elsewhere[17].

In conclusion, we have presented the first quantum mechanical simulation of a nanoindentation process. The atomic mechanisms for plastic deformation in this regime have been determined: plastic flow of atoms towards interstitial positions inside the slab, and extrusion of material towards the tip at the surface, induced by the non–uniform volume

strain and stabilized by the adhesive interactions with the tip. These adhesive interactions, disregarded in many continuum approaches to the problem, are also shown to be responsible for the friction through the induced stick–slip motion of Si atoms along the walls of the tip, the hysteric behaviour observed in the simulations pulling the tip out from different stages of the indentation process, and the recovery of plastic strains during unloading. The plastic deformations are triggered by the delocalization of the charge density induced by the stress in the elastically compressed structure. Finally, the onset of plastic irreversible deformation of the sample is related to the plastic deformation of the second double layer of the slab.

Atomic resolution in non–contact AFM

The Atomic Force Microscope (AFM) was developed as a tool capable of resolving surface structures of both conductors and insulators on the atomic scale, extending to a new whole class of materials the capabilities of the Scanning Tunneling Microscope (STM). The AFM probes the interatomic forces between a tip and a surface and their spatial variations. In the usual operational mode, the so–called constant force mode, the structure of the surface is determined by measuring the tip height which leads to a constant preset force between the surface and the tip as it scans the surface.

In principle, AFM images should be much simpler to interpret than the STM images since atomic forces reflect the total charge density (not the local density of states at the Fermi level), and so they are free of band-structure effects which prevent a direct interpretation of the STM images in terms of atomic structure. However, progress in force microscopy towards true atomic resolution has been slower, and there is a long–standing debate about the possibility of achieving true atomic resolution with the AFM. Many atomic-scale AFM images of crystalline solids have been published, but those images are taken in the contact regime, where the strong repulsive interaction between the tip and the sample provides the constrast mechanism. In this regime, local deformations of the sample and friction effects due to the adhesive interaction between tip and sample are known to affect the recorded topography. Most of the reported data show either perfectly ordered atomic structures or defects much larger than atomic–scale defects. On the other hand, images with point defects are routinely obtained with STM. These facts have raised the question whether the AFM is really a microscope like STM with true lateral resolution.

In the last two years several authors[19, 20] have claimed that it is possible to operate the AFM with atomic resolution, provided you work in the non–contact regime where the attractive Van der Waals forces acting between tip and sample are detected. Those forces are considerably weaker than the forces used in the contact mode, and, therefore, difficult to measure.

Recently Giessibl[20] has shown that atomic resolution is feasible in non–contact UHV conditions using a Si tip scanning a reactive surface, the reconstructed Si(111) 7×7 surface,

and a novel force detection scheme that senses the force gradient, instead of the force itself, through frequency modulation[21]. In few words, the tip is subject to a positive feedback, so it oscillates at its eigenfrequency. As the sample approaches the tip, the force gradient of the tip–surface interaction alters the effective force constant and the frequency changes. A scan of the surface at a constant frequency shift creates a map of a constant average force gradient. The lower part of Giessibl's experimental image shows a multitip image of the surface, then the quality deteriorates, and suddenly, for the width of a unit cell, the characteristic protrusions associated with the 12 adatoms on the top layer of the reconstruction according to the DAS model [22] can be clearly seen. After this, the resolution deteriorates again, and no image is obtained for the rest of the scan. It should be notice that the motion of the tip during imaging is very complicated, with the tip mean position much further away from the surface than the expected range of the interaction. According to ref. [20] the distance between the tip and the surface at the closest point is 5 Å and the estimated attractive force, at that point, is -0.14 nN.

The work presented here involves the understanding of the mechanisms of image formation in non–contact AFM. In particular, the possibility of achieving atomic resolution with the Van der Waals interactions, which are not dominated by the interaction between the two closest atoms, and the changes in the tip structure which are related to the true atomic resolution observed in a small area of the experimental image.

We have studied these questions in two menbers of the series of Takayanagi reconstructions [22], the 3×3 and the 5×5, which contain the same features present in the 7×7, particularly the presence of adatoms. The system that we have considered is a supercell containing a Si(111) slab, two tips (one on each side of the Si slab) and a vacuum region. Inversion symmetry has been impossed on the supercell. The Si slab is built up of eight (111) planes with 68 (200) atoms in the 3×3 (5×5) case, in which the central two layers are kept fixed to simulate the bulk crystal termination of the surface. The lenght of the unit cell normal to the surface is 11 times the the double layer spacing in the [111] direction of the bulk crystal. The tips used in the experiment are etched out of single-crystalline Si. As the natural cleavage planes of Si are (111) planes is reasonable to consider that the very end of the tip is bounded by those planes. According to that, we have considered sharp tetrahedral tips with 4 and 10 Si atoms stacked in two (three) (111)-Silicon planes. We have saturated the dangling bonds of the atoms in the base of these tips with Hydrogen atoms, except in one case for reasons to be discussed below. The Si atoms in the base of the tip and the H atoms attached to them are held fixed during the scanning process.

The operation of the microscope was simulated in a stepwise, quasi–static manner by making small movements of the tip parallel to the slab at a constant height above the surface. At each step the atoms in the slab were allowed to relax to their equilibrium positions (zero forces) for that particular position of the tip. In the electronic structure part of the calculation we used the Generalized Gradient Approximation (GGA)[23] of density functional theory, following the prescription of White and Bird[24]. The technical

details of the calculation are the same that in the nanoindentation simulations.

We want to explore other contrast mechanisms apart from the Van der Waals interaction. One of the likely candidates to provide a "bonding" attraction, bridging the distance between the tip and the surface, is the interaction between the dangling bonds in the adatoms and a dangling bond pointing out of the apex atom of the tip. In order to analyze this effect, scans in the 3×3 reconstruction have been performed using two different tips. They both have four Si atoms arranged in the same tetrahedral structure, the only difference being the presence or not of H atoms saturating the dangling bonds of the atoms in the base. This saturation changes the hybridization of the Si atoms in the base to a state close to the sp^3 of the bulk, and charge flows from the centre of the tetrahedra to the bonds with the hydrogen atoms and the Si-Si bonds, and the apex atom is left with three strong bonds with the other Si atoms and a dangling bond pointing out in the [111] direction.

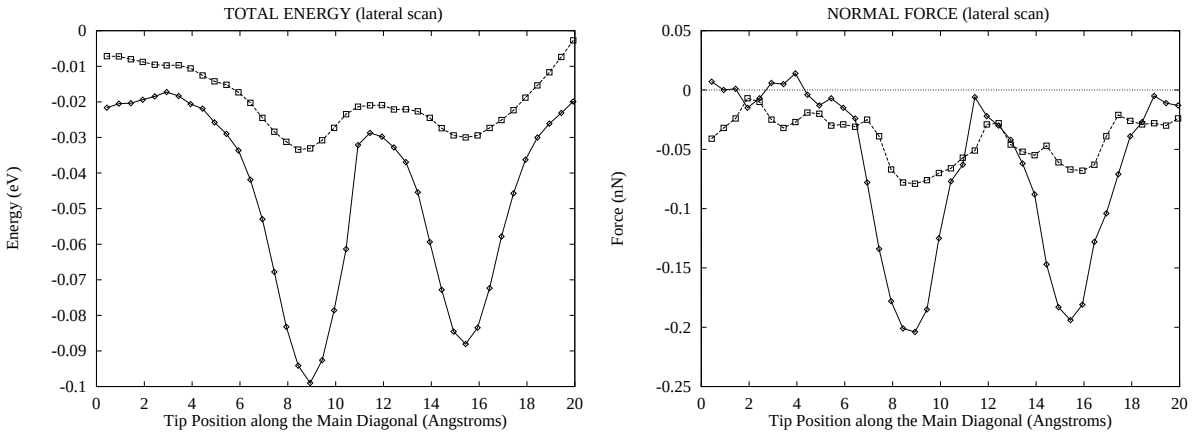

Figure 2: Total energy (in eV) and normal force (in nN) for the scan along the main diagonal of the 3×3 reconstruction at a constant height of 5 Å above the adatoms. The continuous (dashed) line corresponds to the tip with (without) a dangling bond. Both tips show minima at the position of the two adatoms in this reconstruction.

The total energy of the system and the normal force for the different positions of the two tips scanning along the main diagonal of the 3×3 reconstruction at a constant height of 5 Å above the adatoms are shown in Fig. 2. We will consider first the total energy. In each case, the zero of the energy corresponds to the sum of the total energy calculated independently for the slab and the tip using the same unit cell. Both tips show minima at the position of the two adatoms, the one in the stacking–faulted half being deeper, but there is a clear enhancement of the binding energy and, specially, of the contrast between the adatoms and other positions on the surface, when the scan is performed with the tip with the dangling bond pointing out from the apex atom towards the surface. Similar results are obtained in the case of the normal force, where the tip with the dangling bond shows a larger contrast between the position of the adatoms and those in between them or in the corner hole (left side of the figure). It should be noticed that the value of -0.2

nN obtained for the normal force when the tip is on top of one of the adatoms is close to the estimated experimental value.

It is interesting to compare those results with an estimation of the Van der Waals interaction of those tips with a Si surface[25]. Considering just the attractive part of the potential, a typical value for the interaction constant (the same for both Si and H), and a normal distance of 5 Å we obtain values of -0.03 eV and -0.07 eV for the two tips. While the first value is quite close to the value determined for the position on top of the adatom, the second one is clearly smaller than the value obtained in the first–principles calculation, suggesting that there is another contribution, apart from the Van der Waals interaction, to the binding energy of the system. In fact, the results are showing the onset of the bonding between the dangling bonds in the adatoms and the one of the apex atom in the tip. Clear evidence of this process comes from the comparison of the charge densities for the two positions in which the tip is on top of one of the adatoms: Charge accumulates in the region between the adatom and the apex of the tip, depleted from the backbond and from the dangling bond of the other adatom. The normal displacements of the adatoms during the tip scan are also consistent with this picture, with the adatom closer to the tip moving upwards, up to 0.05 Å and taking charge, while the other adatom moves downwards, and charge is removed from its dangling bond. The results for the constant–height scans on the 5×5 reconstruction, using a tip with ten Si atoms, where the dangling bonds of the atoms in the base have been saturated, confirm the picture obtained in the 3×3 case. Constant–force scans are currently being performed[26].

These results suggest an explanation for the sudden switch of the tip to atomic resolution observed in the experiments in terms of changes in the structure of the tip. If the tip picks up an atom from the surface, or it loses a contaminant, a dangling bond pointing towards the surface appears and the contrast is dramatically increased until that dangling bond is saturated again.

In summary, our simulations provide a clear understanding of the imaging process, showing that the range of operation, the values of the force, and the images determined experimentally, can be understood as a result of the interaction of the tip with the adatoms in the surface. Atomic resolution contrast is clearly enhanced by the interaction between the dangling bonds of the adatoms and the apex atom in the tip. The atomic resolution of the tip should be adscribed to the presence of a dangling bond in the apex of the tip, an esentially unstable situation. The contrast mechanism is related to the coupling between the tip and the charge transfer modes among the different dangling bonds in the surface, which have been observed in other experiments involving tip–surface interaction[27].

Acknowledgements

The calculations on nanoindentation were performed as a part of the "Grand Challenge" collaborative project, coordinated by Prof. M. J. Gillan, on the Cray T3D at the Edin-

burgh Parallel Computing Centre. Computer time for the AFM simulations was provided on the Connection Machine CM5E at JRCAT in Tsukuba. R. P. acknowledges the financial support of the Human Capital and Mobility Network of the European Union on *"Ab-initio (from electronic structure) calculation of complex processes in materials" (contract: ERBCHRXCT930369).*

* Electronic address: rp10013@phy.cam.ac.uk

References

[1] B. Bhushan and V. N. Koinkar, Appl. Phys. Lett. **64**, 1653 (1994).

[2] G. Binnig, H. Rohrer, Ch. Gerber and E. Weibel, Phys. Rev. Lett. **50**, 120 (1983).

[3] G. Binnig, C. F. Quate and Ch. Gerber, Phys. Rev. Lett. **56**, 930 (1986).

[4] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

[5] J. B. Pethica and A. P. Sutton, J. Vac. Sci. Technol. A **6**, 2494 (1988); A. P. Sutton *et al*, in *Electron Theory in Alloy Design*, edited by D. G. Pettifor and A. H. Cottrell (The Institute of Materials, London, 1992), p. 191–233.

[6] U. Landman, W. D. Luedtke, N. A. Burnham and R. J. Colton, Science **248**, 454 (1990); U. Landman *et al*, in *Scanning Tunneling Microscopy* III, edited by R. Wiesendanger and H.–J. Güntherodt (Springer, Berlin, 1993), p. 207–259.

[7] J. S. Kallman *et al.* , Phys. Rev. B **47**, 7705 (1993).

[8] M. C. Payne *et al.* , Rev. Mod. Phys. **64**, 1045 (1992).

[9] K. C. Pandey, Phys. Rev. Lett. **60**, 2156 (1988); F. Ancilotto *et al.* , *ibid.* **65**, 3148 (1990).

[10] The total energy was converged to within less than three-thousands of eV per supercell and the forces in the atoms to less than 0.03 eV/Å at each step.

[11] A. Rappe, *et al.* , Phys. Rev. B **41**, 1227 (1990).

[12] J. S. Lin, *et al.* , Phys. Rev. B **47**, 4174 (1993).

[13] The core radii used to generate the pseudopotentials are 1 Å for Si and 1.26 Å for Al. During our simulations, the minimum distance between the atoms, even in the highly compressed stages of the simulation, remains large enough for the cores not to overlap.

[14] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1980).

[15] R. D. King–Smith, M. C. Payne, and J–S. Li, Phys. Rev. B **44**, 13063 (1991).

[16] The details of the Si(111) 2×1 reconstruction seem to be very sensitive to the cutoff used, in particular the buckling between atoms in the chains. The optimized pseudopotentials that we use produce a well converged structure at the low cutoff of 7 Ry employed.

[17] R. Perez and M. C. Payne, accepted for publication in Phys. Rev. Lett. (1995).

[18] J. S. Field and M. V. Swain, J. Mater. Res. **8**, 297 (1993).

[19] F. Ohnesorge and G. Binnig, Science, **260**, 1451 (1993).

[20] F. J. Giessibl, Science, **267**, 68, (1995).

[21] T. R. Albrecht *et al*, J. Appl. Phys. **69**, 668 (1991).

[22] K. Takayanagi *et al*, J. Vac. Sci. Technol. **A 3**, 1502 (1985); Surf. Sci. **64**. 367 (1985).

[23] J. P. Perdew, in *Electronic Structure of Solids '91*, edited by P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991).

[24] J. A. White and D. M. Bird, Phys. Rev. B (RC), **50**, 4954 (1994).

[25] J. Israelachvili, *Intermolecular & Surface Forces*, (Academic Press, London, 1992)

[26] R. Perez, M. C. Payne, I. Štich and K. Terakura, to be published.

[27] R. Wolkow and Ph. Avouris, Phys. Rev. Lett. **60**, 1049 (1988).

Towards a "Quick and Dirty" Fully Variational Density Functional Theory Calculations

Volker Heine and Roger Haydock*
Cavendish Laboratory, University of Cambridge
Madingley Road, Cambridge CB3 0HE, U.K.

* Sabbatical visitor 1995-6 from University of Oregon

This is to publicise some ideas floating around, perhaps to stimulate some discussion: actually three ideas.

The first is a comment by Eric Wimmer at a workshop (by Mike Finnis as a *Network* precursor) discussing 'order-N' methods for large systems. He said *"You know, it isn't larger system that is our main problem. There is plenty of good science that can be done with systems of the present size. The killer is more usually the size of the parameter space to be explored, and to do that requires faster codes".*
Consider an atom coming down on a surface. Two parameters are needed to specify where in the unit cell it hits, and three more for the direction and the magnitude of the incident velocity: and of course more parameters if it is an incident molecule. The same applies if one thinks about sliding at grain boundaries or crystal growth, in fact, the same is true of most applications to real problems in materials science, surface reactions, etc.. And of course one would really like to do them dynamically at finite temperature.

The second item is that there exists a further generalisation of the Harris-Foulkes functional which treats also the Kohn-Sham potential V_{KS} as an adjustible variable. One expands the wave function Ψ (to be precise the one-electron orbitals ϕ_i) in terms of some chosen set of functions, one puts into the generalised functional $G(\Psi, \rho, V_{KS})$ some suitably <u>*chosen*</u> charge density ρ, and solves the Kohn-Sham equations with a suitably <u>*chosen*</u> V_{KS}, where Ψ, ρ and V_{KS} are not made self-consistent. If everything is self-consistent, then G gives the usual density functional theory (DFT) energy; but if Ψ, ρ, V_{KS} are not self-consistent, then the error is second order in $\delta\Psi$, $\delta\rho$, δV_{KS}. It is no longer an absolute minimum. In fact it is a maximum under variation of V_{KS} alone, and minimal under $\delta\Psi$, $\delta\rho$, so that the situation is controllable if one wants to adjust V_{KS} etc. (which on the whole I think one does not want to do). This was shown by Michael Methfessel who derived the precise form of the second order error terms. We will give and prove the

functional below. The point is that G gives an honest evaluation of our usual functional, with only second order errors.

The third question is whether one can now achieve what one may call the Foulkes objective (see his Ph. D. thesis, 1988), namely to develop tight-binding into a quantitatively useable tool, approximate but based on the variational property of the functional without arbitrary fudges. It would be a quick 'one-shot' calculation without self-consistency. This search caused him to rediscover the Harris functional and the present generalisation was already implicit in his work. Let us look at the three ingredients of the generalised functional. The basic purpose of writing down a wave function and solving the Kohn-Sham equation is to evaluate the kinetic energy. Here a minimal tight-binding basis set is known to give a good account of 'conformational' energy differences, i.e. depending on bond angles, as evidenced by the success of Hückel etc. theory and the critique of 'chemical pseudopotentials' by P.W. Anderson and Dave Bullett (see e.g. the discussion by Heine 1980). From chemistry one knows that 'contracted orbitals' are good, i.e. atomic orbitals pulled slightly inwards due to the energy lowering from the bonding. Similarly the charge density is to a good approximation a sum of slightly contracted atomic densities (see e.g. Robertson et al. 1991). Finally a good approximation to V_{KS} seems to be the Wigner-Seitz trick: when in atomic cell n, the electron sees the potential of a singly ionised ion. Namely, one assumes the exchange and correlation hole (containing correctly one electron) extends over the atomic cell n. Perhaps this is a bit too crude in highly ionic situations, but even there it is not bad: the more ionic one takes, say, the metal ion, the more attractive that potential becomes, but the more repulsive the Madelung addition to that potential from the surrounding negative ions. There is scope here for using variational power of the functional.

Otto Sankey (Demkov et al. 1995) already does something like this, though not quite, for various calculations involving SiO_2. Jose Ortega has used it to explore the migration of a Si ion from out of its tetrahedral cage of oxygens, and Sandro De Vita carried out full DFT calculations for several selected configurations. The two sets of energies tracked each other very well, with a consistent 30% difference in energy differences (as yet unpublished). The Sankey code ran 100 times faster than CASTEP (the Cambridge Serial Total Energy Package), a useful factor. Andrew Gibson (Gibson et al. 1993) has done calculation on a vacancy in MgO using the generalised functional and the above prescription for the basis set, ρ, and V_{KS}. Note that having chosen the basis, ρ, and V_{KS}, one must evaluate the functional accurately without further uncontrolled approximations which would destroy the variational nature. In particular, it seems all three-centre integrals in the tight-binding formulation have to be evaluated properly. I have heard this disputed, but Matthew Foulkes found (Ph. D. thesis 1988) that one could choose local orbitals such that three-centre contributions became small at some particular interatomic

spacing, but that they did not remain small as one varied the spacing. The biggest problem lies in the representation of the V_{KS}. The straight application of the Wigner-Seitz Ansatz, as done by Gibson, results in very complicated shapes for the atomic cells and hence difficulties in handling the various integrals. Even in a perfect crystal the Ansatz is a mess, with each spherically symmetric ionic potential chopped off to zero at the planar faces of the Wigner-Seitz cell. What is to be done? The strict Wigner-Seitz Ansatz is a good approximation but one should be able to modify it to make it better for calculations without making the approximation worse.

Finally it remains to specify and prove the generalised functional. It is

$$G(\Psi, \rho, V_{KS}) = V_{ion-ion} + < \Psi | T + V_{ext} + V_{KS} | \Psi > - \int V_{KS} \rho d^3\mathbf{r} + E_e(\rho) \qquad (1)$$

where $E_e(\rho)$ is the sum of the Hartree, exchange and correlation energies, as is V_{KS} in its way. Variation of G with respect to Ψ, or rather with respect to one-electron orbitals ϕ_i, and subject to the usual normalisation, gives the usual Kohn-Sham equation

$$(T + V_{ext} + V_{KS})\,\phi_i = \epsilon_i \phi_i, \qquad (2)$$

where T is the kinetic energy. Variation with respect to ρ gives

$$V_{KS} = \partial E_e(\rho)/\partial\rho \qquad (3)$$

as the usual definition of V_{KS}, and variation with respect to V_{KS} gives

$$\rho = |\Psi|^2 = \sum_i |\phi_i(\mathbf{r})|^2. \qquad (4)$$

Thus we have proved that G is stationary with respect to variations of Ψ, ρ and V_{KS}, and we see from Eq. (1) that the correct DFT energy is obtained if the self-consistency equations (3), (4) are satisfied. That's it.

References

[1] A.A. Demkov, J. Ortega, O.F. Sankey and M.P. Grumbach, Phys. Rev. B **52**, 1618 (1995).

[2] M. Foulkes and R. Haydock, Phys. Rev. B. **39**, 12520-36 (1989).

[3] M. Foulkes, Ph. D. thesis presented to the University of Cambridge (1988).

[4] A. Gibson, R.Haydock and J.P. LaFemina, Phys. Rev. B **47**, 9229-37 (1993).

[5] V. Heine, Solid State Phys. **35**, 1 (1980); D. Bullet, Solid State Phys. **35**, 128 (1980).

[6] M. Methfessel, (e-mail: msm@th4.ihp-ffo.de; preprint (1995)).

[7] I.J. Robertson, M.C. Payne and V.Heine, J. Phys. Condens. Matter **3**, 8351 (1991).