# Symmetry-general *ab initio* computation of physical properties using quantum software integrated with crystal structure databases: results and perspectives

**Yvon Le Page,**[a]* **Paul W. Saxe**[b] **and John R. Rodgers**[c]

[a]ICPET, National Research Council of Canada, Ottawa, Canada K1A 0R6, [b]Materials Design Inc., PMB#176, 825 College Blvd, Oceanside, CA 92057, USA, and [c]Toth Information Systems Inc., 2045 Quincy Ave., Ottawa, Canada K1J 6B2

Correspondence e-mail: yvon.le_page@nrc.ca

The timely integration of crystal structure databases, such as CRYSTMET, ICSD *etc.*, with quantum software, like *VASP*, *OresteS*, *ElectrA etc.*, allows *ab initio* cell and structure optimization on existing pure-phase compounds to be performed seamlessly with just a few mouse clicks. Application to the optimization of rough structure models, and possibly new atomic arrangements, is detailed. The ability to reproduce observed cell data can lead to an assessment of the intrinsic plausibility of a structure model, even without a competing model. The accuracy of optimized atom positions is analogous to that from routine powder studies. Recently, the *ab initio* symmetry-general least-squares extraction of the coefficients of the elastic tensor for pure-phase materials using data from corresponding entries in crystal structure databases was automated. A selection of highly encouraging results is presented, stressing the complementarity of simulation and experiment. Additional physical properties also appear to be computable using existing quantum software under the guidance of an automation scheme designed following the above automation for the elastic tensor. This possibility creates the exciting perspective of mining crystal structure databases for new materials with combinations of physical properties that were never measured before. Crystal structure databases can accordingly be expected to become the cornerstone of materials science research within a very few years, adding immense practical value to the archived structure data.

## 1. Introduction

*Ab initio* calculations for materials have been surfing two phenomenal waves in the past 15 years. The first one is the wave of enthusiasm in the theoretical and computing communities that was initiated by the Car & Parrinello (1985) algorithm. This algorithm enables the *ab initio* cell and structure optimization of crystalline compounds. That wave of enthusiasm is still propagating, with numerous recent developments. The second wave is of course the 'factor 2 each 18-months' of Moore's law, which has packed into today's inexpensive off-the-shelf PCs number-crunching power that exceeds the power obtainable 10 years ago using central computers in respectable institutions. In simple terms, the three-year interval between IUCr meetings consistently corresponds to a fourfold increase in computing power. A month-long computation then becomes a week-long one at the next meeting and takes 2 d after two meetings. This wave shows no sign of slowing down for at least several more years. The result of the busy and bubbling past 15 years has accordingly been the birth of a new and now fully recognized

science called 'computational materials science' (see *e.g.* Bernholc, 1999), which is still far from maturity.

Over the years, the *ab initio* simulation of crystalline systems has come a long way. It used to be mostly a help in understanding the experimentally observed phenomena through quantum calculations on tractable simplified models. Nowadays, many crystalline compounds can be fully modelled to considerable accuracy on supercomputers or even PCs, using no greater simplifications than just the widespread assumptions that are common to most quantum-mechanical simulations of materials. A recent review of the topic was presented by Winkler (1999). *Ab initio* calculations have accordingly become, in some cases, an exploration and discovery tool capable of guiding the experiment. Physics journals such as *Physical Review* now quite regularly publish papers about materials where the 'experimental part' is entirely composed of *ab initio* computations.

From its achievements, and because of its sustained speedup through improved algorithms and faster clock speeds, it was already obvious a few years back that computational materials science was soon going to need a direct interface with the large number of published crystal structure results accumulated in existing databases. We reported on the topic, the need and the progress, at the Glasgow meeting of the IUCr (Rodgers *et al.*, 1999). We report here on the creation of that interface between crystal structure databases and quantum software, and on its exploitation for the point-and-click extraction of a first physical property, stiffness, together with its cluster of related thermo-mechanical properties. We also examine the perspectives opened by this success on the feasibility of computing additional physical properties.

## 2. Building the interface

The purpose of an interface between crystal structure databases and quantum software is to put the whole crystal structure literature at the fingertips of the quantum physicist or chemist. There is no reason why this data should be partitioned into organic, inorganic and intermetallic compounds. That separation, arising from the practicalities of the abstraction of printed data, is arbitrary to a degree and it has more to do with historical reasons than rational ones. That separation is completely absent from the quantum concepts. Through corresponding distribution contracts, we have accordingly brought together CRYSTMET (White *et al.*, 2002) and the ICSD (Behrens, 1999) crystal structure databases under two unified data exploitation frameworks with different purposes.

### 2.1. Materials Toolkit

Toth Information Systems' database exploitation framework, called *Materials Toolkit* (White *et al.*, 1999), is oriented towards the crystallographic manipulation of crystal structures. Among its current tools are the visualization of structures, allowing mostly crystal-chemically sensible default three-dimensional structure plots across the combined databases, fully searchable calculated powder data, flexible interactive powder pattern calculations, structure types, missing symmetry search *etc*. The tools presently included with this crystallographic type of user interface do not currently include quantum software, but such software could be included shortly if needed.

### 2.2. MedeA

Materials Design's software framework, called *MedeA* (Saxe, 1998), is more oriented towards computational materials science. Its principal basic tools are: (i) a flexible database search called *InfoMaticA* (Saxe & Le Page, 1999), which includes the CRYSTMET, ICSD and Crystal Data databases; (ii) a graphical crystal structure editor with extensive editing capabilities; (iii) a graphical user interface (GUI) for the seamless submission of quantum calculations to a variety of *ab initio* software, such as *VASP* (Kresse, 1993; Kresse & Hafner, 1993, 1994), *OresteS* (Methfessel & van Schilfgaarde, 1993) and *ElectrA* (Eyert, 2000).

Among the editing features, one can of course change cell data, atom type, atom coordinates within the Wyckoff site *etc*. One can also create supercells, lower the symmetry to space group *P*1, and add, subtract or move individual atoms. If desired, a conventional description of the altered structure can then be produced automatically for the purpose of efficient quantum computing. All of this is seamless and instantaneous because it is assisted by the symmetry-general tools required for implementing the effortless point-and-click guidance of the user from the database search to the quantum software submission.

## 3. Simple exploitation of the interface

The above functionality puts at the user's fingertips everything that is required to find and adapt existing crystal structures, and then rapidly submits the correct input files for *ab initio* jobs for the purpose of the problem. Those quantum jobs usually reduce to the application of some form or variant of the Car & Parrinello (1985) algorithm. The input comprises the space-group information combined with possibly crude cell data and with atomic positions that are at least qualitatively correct. The result is a 'relaxed structure' with minimum total energy.

Relaxed cell data usually agree within a percent or so with observed X-ray cell data. In itself, that is rarely an improvement over the X-ray results if the compound has been prepared before. It can, however, be an improvement if the material is in finely divided form, with corresponding widening of Bragg peaks, or if it has never been prepared before. However, failure to reproduce the experimental cell data within a percent or so is a serious indication of a deficient crystal structure model. Conversely, without being an absolute proof, successful quantum modelling of the cell data then greatly increases the probability that a hypothesized crystal structure, *e.g.* one that cannot be refined because of insufficient intensity data, is sound.

**Table 1**
Models for 'Type D52', $La_2O_3$.

(*a*) Early Zachariasen (1926) model (*Strukturbericht*, 1931, p. 242)

| Space group $P321$, $a = 3.94$, $c = 6.13$ Å | | | | | |
|---|---|---|---|---|---|
| La | $2c$ | $3..$ | 1/3 | 2/3 | 0.25 |
| O | $3e$ | $.2.$ | 0.75 | 0 | 0 |

After *ab initio* optimization with cell and structure relaxation:

| Space group $P321$, $a = 4.68$, $c = 5.02$ Å | | | | | |
|---|---|---|---|---|---|
| La | $2c$ | $3..$ | 1/3 | 2/3 | 0.26 |
| O | $3e$ | $.2.$ | 0.70 | 0 | 0 |

Cohesive energy: 2805 kJ mol$^{-1}$

(*b*) Later Pauling (1928) model (*Strukturbericht*, 1931, p. 744)

| Space group $P\bar{3}m1$, $a = 3.94$, $c = 6.13$ Å | | | | | |
|---|---|---|---|---|---|
| La | $2d$ | $3m.$ | 1/3 | 2/3 | 0.235 |
| O1 | $1a$ | $\bar{3}m.$ | 0 | 0 | 0 |
| O2 | $2d$ | $3m.$ | 1/3 | 2/3 | 0.63 |

After *ab initio* optimization with cell and structure relaxation:

| Space group $P\bar{3}m1$, $a = 3.92$, $c = 6.08$ Å | | | | | |
|---|---|---|---|---|---|
| La | $2d$ | $3m.$ | 1/3 | 2/3 | 0.245 |
| O1 | $1a$ | $\bar{3}m.$ | 0 | 0 | 0 |
| O2 | $2d$ | $3m.$ | 1/3 | 2/3 | 0.646 |

Cohesive energy: 3630 kJ mol$^{-1}$

(*c*) Neutron results of Koehler & Wollan (1953)

| Space group $P\bar{3}m1$, $a = 3.94$, $c = 6.13$ Å | | | | | |
|---|---|---|---|---|---|
| La | $2d$ | $3m.$ | 1/3 | 2/3 | 0.245 |
| O1 | $1a$ | $\bar{3}m.$ | 0 | 0 | 0 |
| O2 | $2d$ | $3m.$ | 1/3 | 2/3 | 0.645 |

On the other hand, the relaxed atomic coordinates will often represent an improvement over a qualitatively correct but poorly refined crystal structure. X-ray refinement of an oxide phase of a heavy metal, or powder refinement in the presence of an impurity phase, are both well known cases which often result in poorly refined structures. The correctness of the *ab initio* modelling of the distortions of distances and angles in a material can be astounding, especially when those distortions were not introduced in the first place in the structure model to be optimized. Therefore, a powder diffraction pattern computed using a combination of experimental X-ray cell data with *ab initio* optimized atomic coordinates can often turn out to be superior to one recalculated using coordinate data from printed powder work.

In other words, the interfaces described above allow the easy and seamless submission by non-specialists of correct *ab initio* jobs from entries in crystal structure databases for various purposes aimed at complementing the experiment or aiding in its interpretation. We now provide some illustrative examples.

### 3.1. Example 1: plausibility of the early D5$_2$ structure-type model

A first model for structure-type D5$_2$ (*Strukturbericht*, 1931, p. 242) was proposed by Zachariasen (1926). Table 1 reports the present-day description for that structure model with the formula $La_2O_3$ and cell data $a = 3.94$, $c = 6.13$ Å. *Ab initio* cell and structure optimization on Zachariasen's model gives the optimized cell data $a = 4.68$, $c = 5.02$ Å: a suggested change of nearly 20% in both $a$ and $c$. The inconsistency between the cell derived from the structure model and the observed cell data raises very serious doubts about the validity of the structure model. This conclusion is reached without having to resort to any comparison with another model. The model is internally inconsistent. The computing time for the cell and structure optimization was about 4.5 h.

In fact, a second structure model for $La_2O_3$ was proposed by Pauling (1928; *Strukturbericht*, 1931, p. 744). It is also reported in Table 1. *Ab initio* cell data calculated with structure optimization of Pauling's model are $a = 3.92$, $c = 6.08$ Å. They are in reasonable agreement with the experiment, indicating a plausible model. The computing time for the cell and structure optimization was about 2.2 h. Modern neutron results obtained by Koehler & Wollan (1953) are also quoted in Table 1 for comparison. They indicate full agreement of the optimized atom coordinates. Space group $P321$ is a subgroup of $P\bar{3}m1$, but the fact that Zachariasen's model was topologically different nevertheless inhibits its convergence onto Pauling's model. Zachariasen's model was derived from enlightened crystal-chemical intuition in 1926, but it is rated as internally inconsistent by modern quantum software.

In spite of the fact that it has been labelled '*falsch*' on p. XXII in *Strukturbericht* (1937), and is clearly superseded by Pauling's (1928) model, Zachariasen's (1926) model is listed under entry number 26864 in the ICSD, which bears no editorial remark warning the user that the entry corresponds to an incorrect structure model. Quantum software interfaced with the ICSD database in *MedeA* raises sufficient doubts about that entry for users to beware.

### 3.2. Example 2: improvement of approximate crystal structure results for quartz

Table 2 summarizes the crystal structure results for quartz taken from Table *c* in *Strukturbericht* (1931, p. 168). Atom coordinates in Table *c* correspond to the general and special positions in *Internationale Tabellen zur Bestimmung von Kristalstrukturen* (1935). An origin shift of 0, 0, 1/3 was introduced in the *International Tables for X-ray Crystallography* (Vol. I, 1952) for space group $P3_221$. It was retained by the *International Tables for Crystallography* (Vol. A, 1983) and it is taken into account here in Table 2. The results after structure optimization of atomic coordinates, a calculation of 3 h, deviate from modern refined low-temperature neutron results (Jorgensen *et al.*, 1982) by approximately 0.02 Å for Si and 0.04 Å for O, from a starting point that was 0.04 and 0.39 Å away. In other words, the accuracy of the optimized atomic coordinates is comparable with that of many powder X-ray refinements. This accuracy is quite sufficient to compute powder patterns with very recognizable low-angle intensities after optimization of an approximate starting model, *e.g.* derived from a rough structure sketch. Comparison with Table 1 of Donnay & Le Page (1978) shows that *Strukturbericht* was

**Table 2**
Improvement of the quartz crystal structure model from *Strukturbericht* (1931, p. 168).

Cell data as 4.90 and 5.39 Å from kX.

Space group $P3_221$, $a = 4.90$, $c = 5.39$ Å

| Si | 3a | .2. | 0.46 | 0 | 2/3 |
| O | 6c | 1 | 0.5 | 0.3 | 0.77 |

After a single cycle of *ab initio* optimization (3 h on a 1.4 GHz Pentium IV computer):

| Si | 3a | .2. | 0.464 | 0 | 2/3 |
| O | 6c | 1 | 0.410 | 0.276 | 0.778 |

13 K neutron results, adapted from Jorgensen *et al.* (1982) in order to correspond to the above setting:
Space group $P3_221$, $a = 4.9021$, $c = 5.3997$ Å

| Si | 3a | .2. | 0.468 | 0 | 2/3 |
| O | 6c | 1 | 0.412 | 0.271 | 0.784 |

correctly reporting quartz in the *z+* setting of Donnay & Le Page.

### 3.3. Example 3: computed powder patterns from 'cell and structure type' data entries

Many contributions about intermetallic compounds detail the structure for one composition and give cell and structure-type information for several analogous chemical compositions. Since its inception in the early 1960s at Los Alamos under the name Metals Data File or MDF, the CRYSTMET database has abstracted those cell, formula and structure-type entries in parallel with the full structure entries. The rationale was that, someday, those entries could be exploited on an equal footing with the structure entries. Quantum software now provides us with such an opportunity, as shown below.

Entry AL4132 is a 'cell and structure type' entry describing a material with composition $Ti_4As_3$ crystallizing in space group $I\bar{4}3d$ with cell parameter $a = 7.6795$ (2)Å and structure type $Th_3P_4$, but no full crystal structure for $Ti_4As_3$ is apparently available. The structure data for $Th_3P_4$ in entry AL4092 places Th in Wyckoff position 12(a) with $\bar{4}$ site symmetry and fixed coordinates 3/8, 0, 1/4. P is at Wyckoff position 16(c) on a threefold axis with coordinates $x, x, x$ ($x = 0.083$) in a cell with edge $a = 8.6170$ Å. We could attempt to compute a powder pattern with the edge of $Ti_4As_3$ and the fractional positions of $Th_3P_4$, but we would be taking chances by doing that. The cell volumes (453 and 640 Å$^3$) are quite different and the substitutions interchange the relative electronegativities of the sites, pointing to possibly major coordinate adjustments and inviting some caution. *Ab initio* cell and structure optimization using the cell and composition data from AL4132 and the fractional coordinates from AL4092 gave $a = 7.6779$ Å and $x$(Ti) = 0.0753. The match of observed and calculated cell data supports its assignment to the $Th_3P_4$ structure type by Berger (1977).

Examination of the structure shows that Ti is sandwiched between two layers with As triangles made of fixed-coordinate positions of the space group, in a site with odd angular distortion, halfway between an octahedron and a trigonal prism. Packing of spherical atoms within that structure type would then result in $x = 1/12$, even if this does not correspond to a fixed-position Wyckoff site. In fact, Meisel (1939) results do not report 0.083, but 1/12, because the structure was established before the days of least squares. This is in contrast with $Ti_4As_3$, for which *ab initio* methods predict that the Ti atom is displaced by 0.107 Å from this ideal position towards the origin. This seems to be in agreement with X-ray refinements on other arsenides with $Th_3P_4$ structure type, where arsenic occupies Wyckoff position 12(a), *e.g.* $Ce_4As_3$, $Pr_4As_3$ and $Yb_4As_3$ as reported by Ono *et al.* (1970). From X-ray data, they refined a value of $x = 0.069$, *i.e.* about 0.223 Å away from the equidistance to the As planes, in the sense indicated by our *ab initio* optimization. In fact, *ab initio* optimization of $Th_3P_4$ gives $x = 0.0792$, about 0.061 Å from the value of 1/12 printed by Meisel (1939). After this *ab initio* analysis, which is in support of the experimental cell for $Ti_4As_3$ and proposes only a slight adjustment to the free coordinate in structure-type data, we can proceed with the calculation of a credible powder pattern for $Ti_4As_3$, based on X-ray cell data and *ab initio* atom coordinates.

### 3.4. Usefulness and current limits of applicability

With inorganic or intermetallic materials, obtaining a reliable chemical analysis can often present some difficulties. It is therefore comforting to see the cell and structure results derived from the analysis of diffraction data supported within confidence limits by *ab initio* calculations, especially when the diffraction data are not of high quality. Without constituting a proof, such an agreement indicates at least that the structural model is plausible. In cases where several structure models are possible, comparison between observed and calculated cell parameters can constitute a very solid argument in deciding which is right.

Quantum simulations are time-consuming calculations. The corresponding input data can also be quite complex to figure out. Now that simple-to-use GUIs exist to link crystal structure databases and quantum software, the main remaining obstacle to the widespread use of *ab initio* software is the computing time. Roughly speaking, a current-day PC can optimize, to useful precision, the cell and structure of an insulating compound with 30–35 atoms per primitive cell in about a week. We feel that this is then about the current feasibility limit for economical calculations with a PC. Beowulf farms of PCs, where typically 70 PCs are computing in parallel, would be able to tackle similar problems with about three or four times that number of atoms, but for a significant cost. This is because the scaling factor for the time required by quantum computations is about the cube of the number of atoms per primitive cell. There are quite a number of known compounds with more atoms than that per primitive cell. However, there is plenty to be done now with structures of 30 atoms and less. Also, Moore's law, which is expected to be applicable for at least five more years, should let us tackle 60-atom structures on a PC, and 250-atom structures on a Beowulf farm by 2005.

This factor alone will put most compounds presently in crystal structure databases within the reach of *ab initio* cell and structure optimization, even if no improvement in algorithms occurs in the next few years. In other words, there will be plenty to do at all times with quantum software interfaced with crystal structure databases, even during the ramping-up period that is still ahead of us. Given this projected gradual lowering of its limitations, quantum modelling is clearly going to become an intrinsic part of crystal structure reporting within a very few years, at least in the difficult cases where the evidence from the diffraction data is not overwhelming.

It should be noted that at no point in the past was Moore's law credibly predicted to have that many extra years ahead of it. Sixty-four-bit chips, with breakthrough architectures and dedicated optimizing compilers, will all be ready to take the relay of mere clock speedup when that latter factor will have of necessity reached a physical limit. Just those three factors, already existing for a cost or clearly within reach, are probably sufficient to keep Moore's law going for five more years should higher clock speeds face a material impossibility for further improvement tomorrow.

## 4. Automated extraction of elasticity from *ab initio* simulations

Elasticity is one of the most basic and deceptively simple everyday physical properties of nearly all crystalline matter. This property is in fact deeply rooted in the chemical bonding between atoms. It corresponds to the reversible transformation of chemical bonding energy into macroscopic mechanical potential energy. *Ab initio* software, being able to perform accurate total energy calculations for single-crystal matter, is also in principle capable of computing its elastic coefficients. Depending on crystal symmetry, there are from 3 to 21 independent stiffness tensor coefficients (see *e.g.* Nye, 1964) measurable on the single-crystal material and calculable from its crystal structure using *ab initio* methods. Considering that the elasticity of single-crystal bars can be measured to an accuracy of about 1% by resonance methods, comparison between computed and measured elastic coefficients then constitutes a powerful objective test for quantum packages. If desired, it is then a simple matter to extract the equivalent properties for isotropic finely crystallized matter from those elastic coefficients computed on the single crystal. Such properties are widely known as Young's modulus, the shear modulus and the bulk modulus. The further extension to additional thermo-mechanical properties is detailed in §5 of this paper.

Since about 1990, elastic calculations have been fairly commonly performed on cubic and hexagonal close packed (h.c.p.) elements, and on a few cubic compounds (see *e.g.* references in Le Page & Saxe, 2001*a*). The elasticity of a few compounds with low symmetry has also been elucidated (see *e.g.* Winkler *et al.*, 2001), but overall, elasticity studies have remained a patient and time-consuming handicraft concen-trated on cubic and hexagonal materials and practiced mostly by quantum software developers. This is partly because of the considerable difficulties in generating input files for appropriate and efficient *ab initio* jobs and then analysing and merging the results into computed stiffness coefficients.

In two recent papers, we have developed and implemented two novel symmetry-general least-squares approaches to the extraction of precise values for symmetry-unrelated elastic tensor coefficients for pure-phase materials. The first (Le Page & Saxe, 2001*a*) is based on the least-squares analysis of efficient total-energy calculations for properly selected strained states of the material and produces from them satisfactory coefficients for the relaxed material, whatever its symmetry. The second method (Le Page & Saxe, 2002) achieves the same thing from the least-squares analysis of stress calculated *ab initio* for properly selected strained states of the material. The strained states for the stress-based approach are not the same as those for the energy-based approach. It produces satisfactory coefficients for materials in their relaxed configuration, under pressure or even undergoing a large general stress or strain. The difference in the capabilities of the two methods is mostly due to the fact that: (i) the formulation of the problem is truly linear in terms of stress; (ii) an energy is a scalar quantity while a stress is a symmetrical rank 2 tensor with six independent components. Point (i) leads to simpler equations, and therefore allows additional theoretical extensions without too much algebraic complexity. Point (ii) allows extraction of stiffness coefficients from fewer simulations because each symmetry-independent stress component is a least-squares observation, similar to a total energy 'observation'.

These two methods have been automated and implemented within *MedeA* (Saxe, 1998; Saxe & Le Page, 1999). For both methods, the quantum input files are produced, submitted and analysed transparently for the user, reducing the learning barrier to a higher-level point-and-click user guidance, *i.e.* a gross simplification by comparison with the numerical complexity and symmetry details of the resulting input files. The first method can be used to drive two total-energy quantum engines, *VASP* and *OresteS*. The second method only drives *VASP* because it is the only one of the engines in *MedeA* to produce stress results from *ab initio* calculations on strained structures.

Both methods are very efficient in terms of computing time because the symmetry aspects of the distorted structures are automatically taken into account, for all symmetries. They give, of course, very similar relaxed numbers when driving the same quantum engine and applied to the same materials. When applied to materials for which elasticity has been measured, the fit with the experimental data is usually sufficiently good for the calculated numbers to have been technically usable if no corresponding experimental data had existed.

As indicated above, we have used existing experimental stiffness data for elements and compounds as a test for the quantum engines and the analysis methods. Both passed those tests so convincingly that, with time, we have gained sufficient trust in the computational methods. We now use those
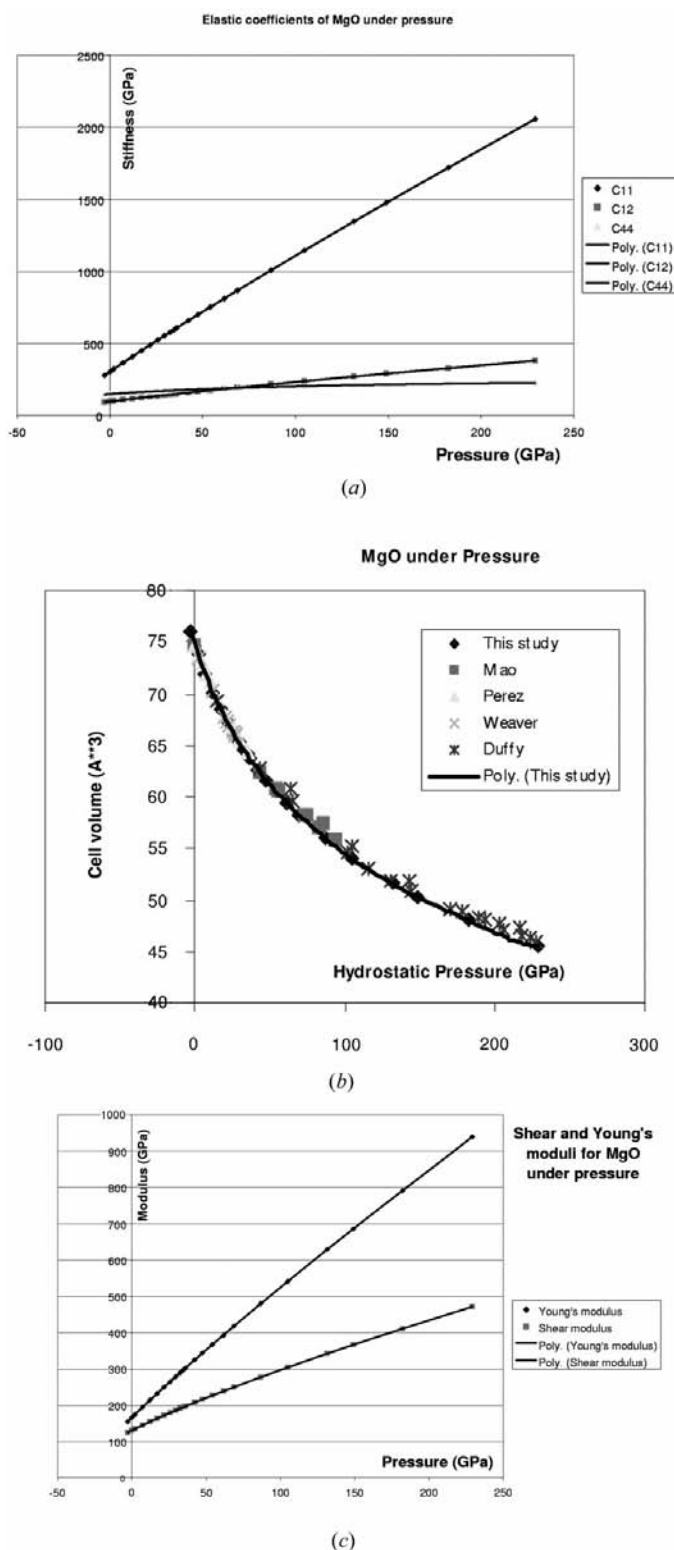
**Figure 1**
Compression of MgO. (*a*) Elastic coefficients of MgO under hydrostatic pressure. For a full analysis, with comparisons of first and second derivatives, see work by Le Page & Saxe (2002). (*b*) Equation of state for MgO. For full explanations and references to the experimental data, see work by Le Page & Saxe (2002). (*c*) Shear and Young's moduli for polycrystalline MgO up to 240 GPa, derived from stiffness coefficients computed by *ab initio* methods.

methods as a way to locate and clarify shortcomings in published experiments or their reports, as shown in the examples below.

### 4.1. Example 1: experimental crystal settings used for Ga and calcite

In order to explain an unconvincing agreement between observed and calculated elastic coefficients for gallium, we showed that it in fact arose from a poorly reported reference system for experimental measurements on Ga (Le Page & Saxe, 2001*b*). The experimental setting corresponded to space group *Abma* rather than the expected *Cmca*.

In the same vein, we were also able to show that the experimental setting for the measurement of physical properties of calcite had unknowingly been its reverse setting (Le Page *et al.*, 2002). That setting was propagated from experimental study to experimental study since at least 1890 (Voigt, 1890). Among half-a-dozen reports, we could find only one recent study (Kaga, 1968) providing sufficient information to ensure the correctness of the crystal setting.

### 4.2. Example 2: compression curves for materials

Elastic coefficients, as well as their first and second derivatives with respect to pressure, or equivalently Murnaghan (1951) higher-order elastic tensors, provide clues about second-order phase transformations under pressure (Born & Huang, 1954; Landau & Lifshitz, 1958; Karki *et al.*, 1997). Le Page & Saxe (2002) detail a computationally very simple and efficient method for applying a known large stress to a material through a succession of small strain steps, each computed from the elastic coefficients after the previous step. Small steps ensure the quasi-linearity of elasticity over each step. This simple development brings these otherwise complex calculations under general stress within easy reach for users.

Fig. 1(*a*) reports the elastic coefficients of MgO under hydrostatic pressure. Fig. 1(*b*) reports the equation of state for MgO up to 240 GPa. It shows that, for this type of information, computations may already be just as reliable as the difficult experiments for pressures exceeding the solidification pressure of argon, around 60–80 GPa. Fig. 2 reports the elastic coefficients of Si under uniaxial [001] stress, indicating a phase transformation at −11.7 GPa, in agreement with experiments on uniaxial loading. Note that Fig. 1 is about hydrostatic pressure, while Fig. 2 is about uniaxial [001] compression. When interpreting those two figures, it should be kept in mind that compressive stresses in general are negative by convention, but hydrostatic pressure is an exception to this and is always taken as positive.

Full experimental sets of accurate elastic coefficients under hydrostatic pressure come from Brillouin scattering of laser light in diamond-anvil compression experiments. They are then applicable to transparent materials only. Compression information on opaque materials is mostly limited to X-ray diffraction information, *i.e.* cell and some structure information. *Ab initio* computational approaches tested on transparent materials until they become routine can then be

extended to the derivation of full elastic properties of opaque materials, after checking that they confirm the available cell and structure information.

There is little doubt that *ab initio* methods are already sufficiently reliable to routinely produce stiffness and compliance data that are technically useable for materials on which such data were never measured. Time will tell how close the prediction was, but extensive tests based on elements and simple compounds with up to 10 atoms per primitive cell seem to indicate sufficient reliability in the calculations to disclose actual deficiencies in the experiment or the reporting.

## 5. Derivation of thermo-mechanical properties from calculated total energy and stiffness data

In tensor notation, the stiffness coefficients $C_{ijkl}$ relate the stress $s_{ij}$ and the strain $e_{kl}$ through $s_{ij} = C_{ijkl}e_{kl}$, with the summation over $k$ and $l$ that is implicit for tensor algebra. Conversely, the compliance coefficients $S_{ijkl}$ relate the strain $e_{ij}$ and the stress $s_{kl}$ through $e_{ij} = S_{ijkl}s_{kl}$. It follows that the compliance $S$ is the inverse of the stiffness $C$, making it easily calculable through matrix inversion.

Few engineering applications actually make use of the full tensor expression for the stiffness or compliance. This is because most materials used by the engineer are in polycrystalline form with more-or-less random orientation of the crystallites. The engineer can then describe the elastic behaviour of polycrystalline materials through the use of two numbers: Young's modulus and the shear modulus. A frequently used third number is the bulk modulus describing the volume compressibility of the material. The bulk modulus $K$ is related to Young's modulus $E$ and the shear modulus $G$ through $E = 9KG/(3K + G)$.

It is difficult to find out what truly happens in a polycrystalline material, but there are two limiting models. The first one assumes that the polycrystalline material is uniformly strained. This is Voigt's approximation, the maximum approximation, based on the use of stiffness. The other one assumes that the material is uniformly stressed. This is Reuss' approximation, the minimum approximation, based on the use of compliance. Under the assumption of random orientation of the crystallites, general triclinic expressions for $G$ and $K$ exist (see *e.g.* Appendix I in Anderson, 1963), giving maximum estimates $K_V$, $G_V$ and $E_V$, and minimum estimates $K_R$, $G_R$ and $E_R$ for the bulk, the shear and Young's modulus, respectively. The high and the low estimates are usually not very different, but Hill's estimate (Hill, 1952), namely the average between high and low estimates, usually corresponds quite closely to the engineering values.

Knowledge of the bulk, the shear and Young's modulus for the isotropic material as well as its density then gives immediately the speed of sound in the material for longitudinal and transverse acoustic waves. Cast or rolled polycrystalline materials usually have a porosity that can reach several percent. For good results in the comparison with actual values of the speed of sound, it is essential to enter the true density of the as-cast material in the formula giving the speed of sound, rather than the density of the single crystals constituting it, which is calculable crystallographically.

Anderson (1963) indicates and justifies a simple method for the calculation of the Debye temperature $\Theta_D$ of a solid from an average of the speed of sound for longitudinal and transverse acoustic waves in the material with zero porosity. Grüneisen's coefficient is also easily obtainable *ab initio* by fitting the equation of state of the material as a function of hydrostatic pressure. Again through simple calculations, the Debye temperature and Grüneisen's coefficient are combined to give the heat capacity and the thermal expansion coefficient of the isotropic material.

In other words, symmetry-unrelated stiffness coefficients for the single-crystal result from an elaborate analysis of time-consuming *ab initio* simulations for distorted states of a material. The corresponding files that are input to the quantum program can be fairly complex. But, as soon as a set of symmetry-unrelated stiffness coefficients is obtained, then 'back-of-an-envelope' calculations produce from them nearly a dozen other properties of scientific or technical interest, most of them being grouped under the label 'thermo-mechanical properties' (see, *e.g.*, Kittel, 1996). Those values may sometimes appear approximate when compared with measurements. Not all, but some of the differences arise from experimental problems in obtaining fine-grained and low-porosity materials made of randomly oriented crystallites. Very similar deviations are observed when the isotropic values are recalculated using the much more accurate experimental stiffness coefficients.

### 5.1. Shear and Young's modulus for MgO under pressure

Fig. 1(c) shows the shear and Young's modulus for MgO, as calculated in the way indicated above. These moduli in turn give access to the speed of longitudinal and transverse sound waves in polycrystalline MgO. Propagation of sound waves are the main means of investigation available for understanding the internal structure of the Earth; therefore, such computa-
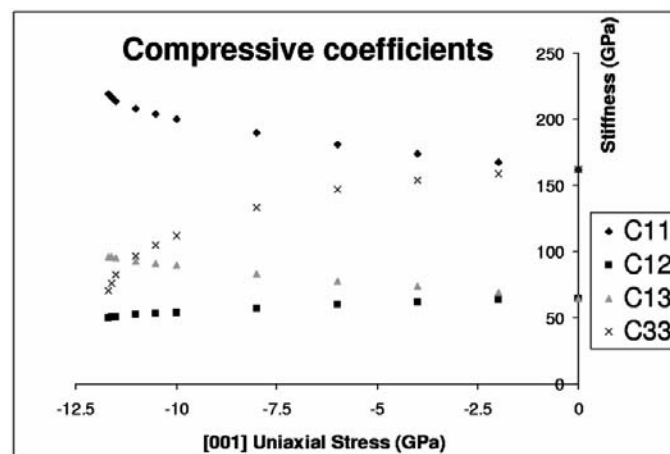


**Figure 2**
Uniaxial compression of Si along [001]. A phase transformation to the $\beta$-tin structure type is observed at $-11.7$ GPa, in agreement with experiment.

tions, which clarify the high-pressure behaviour of minerals at lower-mantle depths, are of significant geological importance. MgO being a transparent material, can be experimentally investigated in diamond-anvil cells, with comparisons between experimental and calculated data. As similar experimental data for opaque materials are much more difficult to obtain, reliance on elastic data derived from *ab initio* computations is accordingly going to become increasingly important for the understanding of minerals under pressure.

## 6. Perspectives opened by the automated computation of elasticity

The automated calculation of stiffness described above would not have been widely successful without an obviously correct decomposition of materials into their electronic states by the quantum software. In other words, the stiffness constitutes another powerful indication that the wavefunction derived by *ab initio* software is correct, pointing to the possible extraction of additional physical properties from the same wavefunction. Numerous such developments are indeed currently ongoing, with good to excellent results on other properties of specific compounds. For example, the literature has started to include impressive *ab initio* work on thermal expansion and related properties (Lichtenstein *et al.*, 2000). We have built on previous work by others to develop the novel general methods (Le Page & Saxe, 2001*a*, 2002) capable of extracting elasticity from entries with any symmetry. In the same way, the above-mentioned specific studies now need additional thought and additional work to generalize them, followed by extensive testing of the implementation for materials for which the property has been measured. The dielectric properties of matter (polarizability, optical, piezoelectric, pyroelectric, ferroelectric *etc.*) also appear to be already within reach. See for example the recent treatment of photoelasticity in quartz by Detraux & Gonze (2001).

Stiffness constitutes a 'pivot property' among the thermo-mechanical properties, in the sense that maybe a dozen other physical properties can be directly calculated from stiffness, but not the other way around. From an *ab initio* viewpoint, stiffness, being derivable from total energy, is an easy case. However, from a crystal-physics viewpoint (see *e.g.* Nye, 1964; Wooster, 1973), stiffness may be the most complex among the primary properties because it relates the stress and the strain, which are both rank 2 tensor quantities.

Extraction of each additional pivot property, for example dielectric polarizability, requires a specialized calculation on the *ab initio* wavefunction, such as a Berry phase summation (Berry, 1984) to extract a dipole moment. That may or may not require bringing additional quantum engines under our present computational frameworks. From a crystal-physics viewpoint, those properties listed above link a quantity that is a scalar, a vector or a rank 2 tensor, and either a scalar or a vector. It then seems that the crystal-physics theoretical and computational framework developed above for the relationship between two rank 2 tensors is more complex than that required for the above properties. The elasticity framework

could then be readily used as a blueprint and adapted to the automation of the calculation of many additional intrinsic physical properties of pure-phase materials.

Our point is that: firstly, quantum engines capable of extracting a number of such pivotal properties currently exist. They may still be in experimental form and they may require huge calculations, but they exist and they work. Secondly, from a crystal-physics viewpoint, the automation of elasticity, a property linking two rank 2 tensors, was probably more difficult than the automation of the other primary properties, which link a scalar, a vector or a rank 2 tensor to a scalar or a vector. In other words, as we were able to automate the calculation of elasticity, automating the calculation of several other properties now also seems to be feasible, requiring mostly time and appropriate collaborations.

## 7. Conclusions

Within a short period of time, several physical properties for pure-phase materials can be expected to become credibly computable from corresponding crystal structure database entries, both more quickly and more cheaply than they can be measured. The range of materials for which at least some form of property data exists can then be expected to increase considerably over the coming years. It goes without saying that the above developments and perspectives will increase considerably the value of the data stored in crystal structure databases.

Materials design is about finding materials with given combinations of properties. Past experience shows that modelling or experiment can both just as easily go wrong in exceptional cases. It is only the convergence of the two sets of numerical results for a property that can build full trust in the numbers. In the search for materials with a given combination of properties, we are not very far from the time when it will be more advantageous to use modelling rather than experiment as a fast and economical tool to eliminate, in the first instance, the poor prospects. Costly experimentation would follow on just the surviving best prospective materials. Materials observed to deviate significantly from the predictions should receive full attention. The field of computational materials design can then be expected to experience a considerable growth or even a quantum jump in the coming few years because of the current availability of easy-to-drive flexible quantum software interfaced with crystal structure databases, running on off-the-shelf, inexpensive, but nevertheless powerful computers.

All the timings reported in this manuscript correspond to computations performed in July and August 2001 on a single 1.4 GHz Pentium IV machine, built in December 2000, and running under Windows 2000. Moore's law predicts that, by the time this issue of *Acta Crystallographica* hits the shelves of libraries around July 2002, all calculation times printed in the present paper will have been approximately halved. The maximum number of atoms for a primitive cell to be accurately tractable in a week with a single PC will also have increased from 30 to *ca* 38 atoms with no increase in costs. This

assumes, of course, that the same calculations would be run with the same software and the same execution parameters on a July 2002 state-of-the-art, off-the-shelf, single PC.

## References

Anderson, O. L. (1963). *J. Phys. Chem. Solids.* **24**, 909–917.

Behrens, H. (1999). *XVIIIth International Congress and General Assembly of the IUCr*, Glasgow, Abstract No. C07.ID.001.

Berger, R. (1977). *Acta Chem. Scand. Ser. A*, **31**, 514.

Bernholc, J. (1999). *Phys. Today*, September issue, 30–35.

Berry, M. V. (1984). *Proc. R. Soc. London Ser. A*, **392**, 45–57.

Born, M. & Huang, K. (1954). *Dynamical Theory of Crystal Lattices*. Oxford: Clarendon Press.

Car, R. & Parrinello, M. (1985). *Phys. Rev. Lett.* **55**, 2471–2474.

Detraux, F. & Gonze, X. (2001). *Phys. Rev. B*, **63**, 115118.

Donnay, J. D. H. & Le Page, Y. (1978). *Acta Cryst.* A**34**, 584–594.

Eyert, V. (2000). *Int. J. Quant. Chem.* **77**, 1007–10031.

Hill, R. (1952). *Proc. Phys. Soc. Lond.* **65**, 350–354.

*Internationale Tabellen zur Bestimmung von Kristalstrukturen* (1935). Edited by C. Hermann. Berlin: Gebrüder Borntraeger.

*International Tables for X-ray Crystallography* (1952). Vol. I. Birmingham: Kynoch Press.

*International Tables for Crystallography* (1983). Vol. A. Dordrecht: D. Reidel.

Jorgensen, J. D., Rotella, F. J. & Lager, G. A. (1982). *J. Appl. Phys.* **53**, 6751–6756.

Kaga, H. (1968). *Phys. Rev. B*, **172**, 900–919.

Karki, B. B., Ackland, G. J. & Crain, J. (1997). *J. Phys. Condens. Matter*, **9**, 8579–8589.

Kittel, C. (1996). *Introduction to Solid State Physics*, 7th ed. Chichester: John Wiley.

Koehler, W. C. & Wollan, E. O. (1953). *Acta Cryst.* **6**, 741–742.

Kresse, G. (1993). PhD thesis. Technische Universität Wien.

Kresse, G. & Hafner, J. (1993). *Phys. Rev. B*, **48**, 13115–13118.

Kresse, G. & Hafner, J. (1994). *Phys. Rev. B*, **49**, 14251–14269.

Landau, L. D. & Lifshitz, E. M. (1958). *Statistical Physics*. London: Pergamon Press.

Le Page, Y. & Saxe, P. W. (2001*a*). *Phys. Rev. B*, **63**, 174103.

Le Page, Y. & Saxe, P. W. (2001*b*). *Physica B*, **307**, 191–193.

Le Page, Y. & Saxe, P. W. (2002). *Phys. Rev. B*, **65**, 104104.

Le Page, Y., Saxe, P. W. & Rodgers (2002). *Phys. Status Solidi B*, **229**, 1155–1161.

Lichtenstein, A. I., Jones, R. O., de Gironcoli, S. & Baroni, S. (2000). *Phys. Rev. B*, **62**, 11487–11492.

Meisel, K. (1939). *Z. Anorg. Allg. Chem.* **240**, 300–312.

Methfessel, M. & van Schilfgaarde, M. (1993). *Phys. Rev. B*, **48**, 4937–4940.

Murnaghan, F. D. (1951). *Finite Deformation of an Elastic Solid*. New York: Wiley.

Nye, J. F. (1964). *Physical Properties of Crystals; Their Representation by Tensors and Matrices*, 3rd ed. Oxford University Press.

Ono, S., Despault, G., Calvert, L. D. & Taylor, J. B. (1970). *J. Less-Common Met.* **22**, 51–59.

Pauling, L. (1928). *Z. Kristallogr.* **69**, 415–421.

Rodgers, J. R., Le Page, Y., White, P. S., Harris, J. & Saxe, P. (1999). *XVIIIth International Congress and General Assembly of the IUCr*, Glasgow, Abstract No. C07.ID.006.

Saxe, P. W. (1998). http://www.MaterialsDesign.com/.

Saxe, P. W. & Le Page, Y. (1999). http://www.MaterialsDesign.com/.

*Strukturbericht* (1931). Vol. 1, pp. 1913–1928, edited by P. P. Ewald & C. Hermann. Leipzig: Akademische Verlag.

*Strukturbericht* (1937). Vol. 2, pp. 1928–1932, edited by C. Hermann, O. Lohrmann & H. Philipp. Leipzig: Akademische Verlag.

Voigt, W. (1890). *Ann. Physik.* **31**, 474.

White, P. S., Rodgers, J. R. & Le Page, Y. (1999). *XVIIIth International Congress and General Assembly of the IUCr, Glasgow*, Abstract No. P07.23.003.

White, P. S., Rodgers, J. R. & Le Page, Y. (2002). *Acta Cryst.* B**58**, 343–348.

Winkler, B. (1999). *Z. Kristallogr.* **212**, 506–527.

Winkler, B., Hytha, M., Warren, M. C., Milman, V., Gale, J. D. & Schreuer, J. (2001). *Z. Kristallogr.* **216**, 67–70.

Wooster, W. A. (1973). *Tensors and Group Theory for the Physical Properties of Crystals*. Oxford: Clarendon Press.

Zachariasen, W. (1926). *Z. Phys. Chem. Leipzig*, **123**, 134–150.