Ab initio Hartree–Fock calculations for periodic compounds: application to semiconductors

R Orlando†, R Dovesi†, C Roetti† and V R Saunders‡

† Department of Inorganic, Physical and Material Chemistry, University of Torino, via Giuria 5, I-10125 Torino, Italy
‡ Daresbury Laboratory, Science and Engineering Research Council, Daresbury, Warrington WA4 4AD, UK

Abstract. The *ab initio* Hartree–Fock crystalline orbital program CRYSTAL is applied to diamond, silicon, BN, BP, SiC and AlP. The effects of the computational parameters controlling the accuracy of the infinite Coulomb and exchange series are analysed; the performances of five standard (but re-optimised in the valence part) molecular basis sets (STO-3G; 3-21G; 3-21G*; 6-21G; 6-21G*) are documented with reference to equilibrium binding energy, lattice parameter and bulk modulus. The analysis is then extended, with the largest basis set, to transverse optical phonon frequencies, band-structure and charge-density data. The results show trends similar to those expected from molecular calculations; typically, the mean lattice parameter and bulk modulus errors obtained at a 6-21G* level are about +1% and +10%, respectively.

1. Introduction

The Hartree–Fock(HF)-LCAO technique is by far the most widely available computational scheme for the study of the electronic structure of molecules; standard programs have been available for at least 15 years [1].

In solid state physics 'first-principle' or *ab initio* calculations have become common only recently; most are based on density functional (DF) schemes [2], in conjunction with plane-wave (PW) basis sets and pseudopotential (PP) techniques [3]. There are many reasons for this choice.

(i) The DF-PW-PP approach is the natural evolution of the 'empirical pseudopotential' techniques implemented more than 20 years ago for the qualitative and quantitative study of the band structure of crystalline compounds [4].

(ii) The orthogonality of the basis functions and their simple analytical form drastically reduce the complexity of the integral evaluation; the corresponding computer programs are not more than few thousand lines of code long.

(iii) The use of the PP ensures in many cases (metals and small-unit-cell non-ionic compounds) a sufficiently rapid convergence of the total energy with respect to the basis set size.

(iv) Finally, correlation effects are taken into account to some extent, through the 'exchange-and-correlation' potential derived from the electron gas model.

The popularity of the DF-PW-PP method is only one of the reasons for the scant interest in solid state physics for the HF–crystalline orbital (CO)–LCAO method; one could add

(i) the prejudice against the HF method due to the qualitatively incorrect behaviour of the electron gas density of states at the Fermi energy [5],

(ii) the lack of correlation and

(iii) the much greater complexity of an *ab initio* HF–LCAO computational scheme with respect to DF-PW-PP methods, due to the non-orthogonality of the basis functions and to the triple-infinite summation over four-centre integrals ('the calculation and manipulation of horrendous numbers of many-centre integrals' [6]).

In spite of this situation, many arguments strongly suggest the need for *ab initio* HF–LCAO calculations for solids.

(i) In many situations, reference to chemical concepts, language, numerical tools is essential; typically, for molecular crystals, one must compare bulk and isolated-molecule results obtained with the same Hamiltonian, the same basis sets and possibly the same program, to obtain reasonable answers. Similar arguments hold for large classes of compounds, such as silicates.

(ii) For large-unit-cell systems a local basis is far more convenient than the PW one; in the HF–CO scheme to be discussed below the cost of the calculation increases roughly linearly with the basis set in the unit cell (no fourth-power 'explosion').

(iii) The considerable experience gained in molecular quantum chemistry in algorithms, computational devices and basis sets can be partially exploited in a HF–CO–LCAO scheme.

(iv) As regards correlation effects, *a posteriori* corrections to the HF energy and energy-dependent quantities through correlation-only electron density functionals are available [7, 8], and have been tested [9]. The present work extends these tests, with excellent results.

Alternatively, many-body perturbation theory methods or the closely related coupled-electron-pair models (see, e.g., [10]) should shortly be transferred from the molecular to the non-metallic crystalline context. Indeed, recent calculations based on CEPA-0 formalism have been carried out for diamond [11] and silicon [12]. In another interesting development, the quantum Monte Carlo method [13] has been applied to diamond [14] with some success. One may expect that the HF method will prove useful in supplying the Slater determinant component of the Jastrow–Slater trial function [14] commonly employed in the quantum Monte Carlo method.

In the past 20 years a few HF computational schemes for periodic systems have been proposed [15] but, as far as we know, the only general (i.e. applicable to realistic polymers, slabs and crystals) implementation of those formal schemes is CRYSTAL, the program written by the present authors and collaborators [16–19]. CRYSTAL is able to treat systems periodic in 3D (crystals), 2D (slabs), 1D (polymers) and, as a limiting case, 0D (molecules) of any symmetry (in particular, the 230 space groups and the 80 two-sided plane groups are available). It has been systematically improved and generalised in the past 10 years. Much effort has been devoted to the standardisation of the program and to the improvement of both speed and accuracy; a version of CRYSTAL is available from Quantum Chemistry Program Exchange (QCPE) [20], such that systems containing up to 150 AOs in the unit cell can be treated (more recent versions are able to treat larger unit cells). The implementation of the program has been accompanied by applications ranging from polymers [21] to slabs [22], from metals [23] to semiconductors [24], and

from ionics [25] to molecular crystals [26]. The studies on semiconductors (diamond and silicon, graphite and BN), in particular, refer to an early version of CRYSTAL, when the numerical accuracy of the program was much lower than now, and they were performed with a minimal STO-3G basis [27, 28].

The aim of the present paper is to document the quality of the results that one can obtain with the HF–LCAO–CO method as implemented in CRYSTAL for a relatively large set of semiconductors: diamond and silicon (in the diamond structure) and SiC, BN, BP, AlP (in the zincblende structure). Attention is mainly directed to energy-dependent quantities (equilibrium lattice parameter, bulk modulus, bulk modulus derivative, frozen phonon frequency and binding energy); charge distribution and electronic structure data are also reported, in order to characterise the bond structure of the investigated compounds.

In section 2 we discuss the influence of the computational parameters which control the truncation of the Coulomb and exchange infinite summations and the reciprocal space integrations on the accuracy of the results and on the cost. This requires some care, because at present total energy derivatives are evaluated numerically, an analytical gradient code (see, e.g., [29]) being unavailable, for the moment at least.

In section 3, basis set problems for periodic compounds are discussed; the packed nature of the compounds and the presence in condensed matter of electronic states quite at variance with respect to molecular situations (e.g. ions such as O^{2-} and N^{3-} are stabilised by the crystal field) prevent a straightforward use of the standard molecular basis sets. In particular, overdiffuse functions are not only not very useful (there are no 'tails' to describe) but also costly (the cost in the present approach is mainly driven by the exponent of the most diffuse functions) and problematic (risk of linear dependence) [19]. Five standard [27, 28, 30–33] molecular basis sets (STO-3G, 3-21G, 3-21G*, 6-21 and 6-21G*) are considered and re-optimised in the valence and polarisation part.

Section 4 is devoted to the energy-related data: the energy versus volume curve is evaluated point by point and fitted to an equation of state (the Murnaghan [34] equation), which provides the equilibrium lattice parameter, the minimum energy, the bulk modulus and its derivative. The transverse optical phonon frequency is also evaluated with a third-order fit. *A posteriori* correlation corrections to the HF total energy according to the Colle–Salvetti [7] or the Perdew [8] functional are reported.

Section 5 contains a brief comparison with density-functional results for binding energy, lattice parameter, bulk modulus and transverse optical phonon frequency.

Section 6 describes the electron distribution and the band structure of the six compounds, to highlight similarities and differences in covalency and ionicity among them, with section 7 presenting our conclusions.

2. Computational details

In this section the fundamental equations of the HF–CO–LCAO method are given and the main features of the method illustrated. The definition of the computational parameters involved in the approximations will serve the discussion of the accuracy of results attainable, since at the moment no standard parameter set can be chosen in a range where the numerical error is negligible in the determination of the observables of interest. For a deeper insight the reader may consult [19]. The periodic HF–LCAO scheme is similar to the molecular one from many points of view. The first step involves the evaluation of the mono- and bi-electronic integrals, which are required for the definition

of the Fock matrix **F** at each SCF iterative cycle:

$$F_{12}^g = H_{12}^g + B_{12}^g \tag{1}$$

where **H** and **B** respectively denote the mono- and bi-electronic contributions in the basis of the AO χ_i^g, g being a direct lattice vector ($\chi_i^g = \chi_i(r - s_i - g)$ and s_i the position in the reference cell of the atom to which χ_i belongs). In particular, B_{12}^g is the sum of the Coulomb and exchange terms as follows:

$$B_{12}^g = \sum_{3,4} \sum_l P_{34}^l \sum_h [(\chi_1^0 \chi_2^g | \chi_3^h \chi_4^{h+l}) - \tfrac{1}{2}(\chi_1^0 \chi_3^h | \chi_2^g \chi_4^{h+l})] \tag{2}$$

where **P** is the density matrix and the usual notation for the bi-electronic integrals has been adopted. The summations over h and l are over all the direct lattice vectors and thus extend to infinity; because of translational invariance the first AO can always be centred in the origin cell, denoted by the null vector **0**. The Fock matrix so defined and the overlap matrix S^g are then Fourier transformed to the reciprocal space, in the Bloch function (BF) basis; eigenvalues ε and eigenvectors A are obtained after solving for each reciprocal lattice k-vector, in the first Brillouin Zone (BZ); the matrix equation is

$$\mathbf{F}(k)\mathbf{A}(k) = \mathbf{S}(k)\mathbf{A}(k)\boldsymbol{\varepsilon}(k). \tag{3}$$

Once the eigenvectors are known, the new density matrix elements in the AO basis, are calculated by integration over the BZ volume:

$$P_{12}^g = \sum_n \int_{BZ} dk \exp(ik \cdot g) A_{1n}^*(k) A_{2n}(k) \theta[\varepsilon_F - \varepsilon_n(k)] \tag{4}$$

where ε_F and θ denote the Fermi energy and Heaviside function, respectively. At this point the Fock matrix in direct space (equation (1)) can be redefined and the whole process iterated until convergence is obtained.

The implementation of the above equations requires the solution of some formal and computational problems concerning the infinite series on the direct lattice vectors g, h and l and the integration in reciprocal space. Treatments of the Coulomb and exchange series, which for periodic systems need to be considered separately, and a technique of integration in reciprocal space had to be implemented. Here our computational scheme is briefly sketched, in order to define the computational parameters discussed below.

2.1. Coulomb series

It is easily seen in equation (2) that, when Gaussian functions are used, an exponential decay is observed in integral values by increasing the modulus squared of g and l; therefore, all terms, whose overlaps S_{12}^g and S_{34}^l are estimated to be less than a threshold 10^{-sc}, are neglected. The remaining terms, related to the infinite-h summation, are then classified into two sets.

(i) To the long-range set belong all the interactions between the charge distributions $\{\chi_1^0 \chi_2^g\}$ and the λth-shell distributions ρ_λ^h within cell h, when their penetration is smaller than a threshold 10^{-t_M}, ρ_λ^h being defined as

$$\rho_\lambda^h = \sum_{3 \in \lambda} \sum_4 \sum_l P_{34}^l \chi_3^h \chi_4^{h+l}. \tag{5}$$

These interactions are evaluated by an Ewald-type technique [19, 35], after a multipolar expansion of ρ_λ^h up to L_M order [17].

(ii) In the short-range interaction region, only when the penetration between the charge distributions $\{\chi_1^0\,\chi_2^g\}$ and $\{\chi_3^h\,\chi_4^{h+l}\}$ is greater than 10^{-bc}, integrals are calculated exactly; otherwise they are estimated by using a bipolar expansion of L_C maximum order (see [19], section II.4, and [36]).

2.2. Exchange series

As with Coulomb series, the exponential decay of integrals with increasing $|h|$ and $|h+l-g|$ allows neglect of all terms whose overlap between AOs couples, $\{\chi_1^0\,\chi_3^h\}$ and $\{\chi_2^g\,\chi_4^{h+l}\}$, respectively, is less than a threshold $10^{-s_{ex}}$; a bipolar expansion of I_{ex} maximum order is used for the calculation of integrals involving two distributions whose penetration is smaller than a threshold $10^{-b_{ex}}$. More critical is the truncation of the g and l summations, as the convergence rate of the series largely depends on the asymptotic behaviour of the density matrix [37]. For this purpose two other tolerances, p_{ex}^g and p_{ex}^l, are introduced and the truncation is based on an overlap criterion scaled on the expected density matrix behaviour (see [19], section II.5).

2.3. Integration in reciprocal space

The problem of the evaluation of the integral (4) at each SCF cycle is a simple task for insulators, since the integrand function is quite regular. The integral is reduced to a weighted average over a set of sampling k-points belonging to a commensurate net, which reproduces the reciprocal lattice with a suitable shrinking factor. For silicon, using shrinking factors of 4, 8 and 12 (corresponding to 8, 29 and 72 sampling k-points in the irreducible wedge of the BZ) yields the following total energy values (6-21G basis set, experimental lattice parameter): $-577.819\,686$ au, $-577.824\,285$ au and $-577.824\,320$ au, respectively.

As no analytical evaluation of the energy gradients is available in CRYSTAL yet, the equilibrium data (the energy E_0, the volume V_0 and the bulk modulus B) have been determined by interpolating on a certain number of energy points using the Murnaghan [34] equation of state:

$$E(V) = E_0 + (BV/B')[(V_0/V)^{B'}/(B'-1)+1] - V_0B/(B'-1) \qquad (6)$$

where the fourth parameter B' is the bulk modulus derivative. A careful check of the quality of the calculation, which is related to the values adopted for the computational parameters discussed above, is necessary, owing to the numerical nature of the evaluation of the energy derivatives through equation (6). This point requires some comment. All the computational parameters previously illustrated act essentially in the same way: they define two or more subspaces, where interactions are evaluated with differing approximations; for example, in the Coulomb series the g infinite set is divided into two subsets by s_C, one of which is completely neglected. Similarly, the b_{ex}-parameter divides the h infinite series into two subsets: in the first, integrals are calculated exactly; in the second they are evaluated by a much cheaper multipolar expansion.

A consequence of this 'direct space truncation scheme' of the infinite series is that the definition of the subsets of interactions evaluated at different levels of approximation depends on the cell geometry; for example, a contraction of silicon lattice parameter

Table 1. Total energy E for silicon at the experimental lattice parameter as a function of different approximation levels in the treatment of Coulomb and exchange infinite series. For the definition of computational parameters see text. A 6-21G basis set is adopted. N_C and N_{ex} are the number of bi-electronic Coulomb (and exchange) integrals evaluated. t_1 and t_2 are the cost for the integral evaluation and scf steps in IBM 3090 (scalar mode) seconds.

	Computational parameters												E
Case	s_C	b_C	t_M	L_M	s_{ex}	b_{ex}	p_{ex}^g	p_{ex}^l	N_C	N_{ex}	t_1	t_2	(Hartree)
I	4	5	5	3	4	3	5	9	29	19	96	42	−577.8301
II	5	6	6	4	5	4	6	11	66	45	199	44	−577.8244
III	6	9	9	6	6	7	7	13	211	93	471	52	−577.8253

from 5.43 to 5.03 Å implies an increase of about 75% in the number of bi-electronic integrals to be calculated exactly. This migration of integrals among differing precision zones generates numerical noise in the calculated total energy versus lattice parameter curve, which may affect the fitted parameters. The numerical noise can be reduced by using quite strict tolerances; greater accuracy implies higher cost of the calculation and a fair compromise must be found. Alternatively, the approximation scheme can be frozen at a given lattice parameter (usually the minimum considered), so that the partition of the interactions among the different zones is the same for the calculation of all energy points. This method will be referred to in the following as the 'geometry-independent' approximation scheme.

Table 1 shows the influence of the computational parameters on the total energy of silicon evaluated at the experimental geometry and on the cost of the calculation. In table 2 the four parameters of equation (6) are reported as obtained by 10 energy points calculated according to the conditions specified in table 1, and when using the 'geometry-independent' scheme or not. Table 1 indicates that in going from case I to case III conditions the cost increases by a factor of 5; however, table 2 shows that, when the 'geometry-independent' scheme is adopted, the differences resulting from case I and case III conditions for E_0, a_0 and B are quite small (negligible for total energy, less than 0.1% for the lattice parameter and about 3% for the bulk modulus). The usefulness of the 'geometry-independent' scheme is confirmed by the root mean square of the difference between calculated and interpolated energies. On the basis of above data,

Table 2. Calculated equilibrium parameters for silicon. Total energy E_0, lattice parameter a_0, bulk modulus B and bulk modulus derivative B' have been interpolated by the Murnaghan equation with the same basis set and within the approximation levels illustrated in table 1. σ is the root mean square between the calculated and the interpolated energies. The 'geometry-dependent' and 'geometry-independent' computational conditions are defined in text.

	'Geometry-dependent' conditions					'Geometry-independent' conditions				
Case	E_0 (Hartree)	a_0 (Å)	B (GPa)	B'	σ	E_0 (Hartree)	a_0 (Å)	B (GPa)	B'	σ
I	−577.8333	5.427	187	18.9	2.7×10^{-3}	−577.8283	5.559	97	4.1	1.3×10^{-4}
II	−577.8275	5.559	97	3.7	5.1×10^{-4}	−577.8280	5.561	94	4.1	1.7×10^{-4}
III	−577.8283	5.564	96	3.8	2.1×10^{-4}	−577.8278	5.565	94	3.9	8.9×10^{-5}

Table 3. The Murnaghan equation parameters for silicon as a function of the explored Δa lattice parameter range around the experimental value; n is the number of points used for the fitting, ΔE is the energy range in the explored interval. The other quantities are as in previous tables.

			Case I			Case II			Case III		
n	Δa (Å)	ΔE (mHartree)	a_0 (Å)	B (GPa)	B'	a_0 (Å)	B (GPa)	B'	a_0 (Å)	B (GPa)	B'
5	0.4	8	5.560	100	4.0	5.565	83	1.1	5.566	95	3.4
6	0.5	9	5.560	99	3.8	5.561	92	3.9	5.565	96	3.8
7	0.6	15	5.562	97	3.5	5.565	91	3.2	5.568	94	3.4
8	0.7	19	5.561	98	3.8	5.562	94	4.0	5.566	95	3.7
9	0.8	23	5.558	97	4.2	5.560	94	4.2	5.564	95	4.0
10	0.9	35	5.559	97	4.1	5.561	94	4.1	5.565	94	3.9

the calculations to be presented in the next section were performed at a level of approximation similar to case II.

The last point to be discussed here is the influence on the best-fit parameters of the explored lattice parameter range and of the number of energy points considered. It can be seen in table 3 that in the range from six to 10 points (i.e. from $\Delta a = 0.5$ Å to $\Delta a = 0.9$ Å) the a_0, B and B' variations, at fixed computational conditions, are smaller than 0.1%, 2% and 30%, respectively; E_0 is determined with better accuracy. The parameters obtained for a given Δa interval after doubling the number of points (i.e. halving the step) are very similar to the corresponding results in table 3. We conclude that the precision in the E_0, a_0 and B calculated values is acceptable; for B', higher numerical precision and, possibly, a more flexible interpolation curve would be necessary.

3. Basis set

The choice of the basis set is a crucial step of the calculation. As discussed in previous work [19], standard molecular and atomic basis sets must be modified to some extent for periodic systems. Present experience with LCAO periodic calculation has shown that, although for highly ionic crystalline systems such as alkali halides or metallic compounds the molecular basis sets are rather inadequate, often little modification is needed for covalent compounds. In the present paper this possibility is explored; standard molecular basis sets of the STO-3G [27, 28], N-21G and N-21G* [30–33] ($N = 3$ or 6) type are adopted. In the 'minimal-basis' STO-3G sets, each atomic function is represented by a Slater-type function, whose exponent has been optimised in an atomic calculation. The Slater functions are in turn expanded as a linear combination of three Gaussian-type functions (GTFs) according to the least-squares fitting criterion. In the N-21G sets the core functions are expanded in N GTFs, and there are two sets of valence functions, resulting from a contraction of two GTFs and one GTF, respectively. The Gaussian exponents and contraction coefficients have been determined variationally from atomic calculations. The addition of an asterisk to the basis set symbols denotes the inclusion of a shell of d polarisation functions. In the case of sp shells, the exponents of the s and p GTFs have been constrained to be equal for both the STO-3G and N-21G sets, in the interests of computational efficiency when evaluating the one- and two-electron integrals. For the

Table 4. Parameters of the valence basis sets used in the present calculation, η is the scale factor for the STO-3G basis, α and c are the exponents and the coefficients for the N-21G (or N-21G*) bases. The values in parentheses are the standard values; when only one set of data is reported, the unmodified standard basis set has been adopted for the crystalline calculation.

System		STO-3G η	N-21G*				
			α(sp)	c(s)	c(p)	α(sp)	α(d)
C		1.73	3.6645	−0.39590	0.23646	0.226	0.8
		(1.72)	0.77054	1.2158	0.86062	(0.19586)	
BN	B	1.46	2.2819	−0.36866	0.23115	0.197	0.8
		(1.50)	0.46525	1.1994	0.86676	(0.12433)	
	N	1.98	5.4252	−0.41330	0.23797	0.297	0.8
		(1.95)	1.1491	1.2244	0.85895	(0.28320)	
SiC	C	1.55	3.6650	−0.39590	0.23646	0.184	0.8
		(1.72)	0.77054	1.2158	0.86062	(0.19586)	
	Si	1.65	1.0791	−0.37611	0.067103	0.180	0.5
		(1.75)	0.30242	1.2516	0.95688	(0.093339)	(0.45)
BP	B	1.25	2.2819	−0.36866	0.23115	0.140	0.8
		(1.50)	0.46525	1.1994	0.86676	(0.12433)	
	P	1.90	1.2186	−0.37150	0.091582	0.160	0.43
			0.39555	1.2710	0.93492	(0.12281)	(0.55)
Si		1.58	1.0791	−0.37611	0.067103	0.130	0.5
		(1.75)	0.30242	1.2516	0.95688	(0.093339)	(0.45)
AlP	Al	1.50	1.164	−0.17574	−0.0021333	0.110	0.37
		(1.70)	0.268	0.25280	0.0099580	(0.063909)	(0.33)
			(0.94616)	(−0.32033)	(0.051938)		
			(0.20251)	(1.1841)	(0.97266)		
	P	1.70	1.2186	−0.37150	0.09158	0.151	0.48
		(1.90)	0.39555	1.2710	0.93492	(0.12281)	(0.55)

present study, the Slater exponent of the STO-3G valence shell has been re-optimised with respect to the total energy of each crystalline compound. In the 3-21G basis the only exponent of the outer valence shell has been re-optimised, which has also been used for the 6-21G and the N-21G* cases; for AlP the inner valence shell of aluminium had to be redefined, too, in order to reduce the overlap with the outer shell, whose exponent in the bulk is higher than in the isolated atom or ion. The two exponents and the contraction coefficients of this inner valence shell have been obtained by optimising the total energy of the Al^+ isolated ion. The results of the optimisation process are shown in table 4. As expected, the exponents of the valence shells in the bulk are higher than in the isolated atom, the exception being carbon in SiC, which is slightly expanded in order to be able to allocate the extra electrons provided by Si (SiC is the most ionic among the six considered compounds).

The relevance of a better description of the core was investigated by comparing the accuracy in the determination of the binding energy, lattice parameter and bulk modulus, obtained from calculations with the 3-21G and 6-21G basis sets, respectively. The effect of the inclusion of one polarisation set in both 3-21G and 6-21G sets was also analysed. Molecular polarisation function exponents for the first-row elements are known to require no significant adjustment in the description of crystalline atoms and the standard definition [30] can be safely used. On the other hand, a re-optimisation is often necessary for second-row-element d orbitals, whose role in the wavefunction description is more important than for first-row elements. The cost of the calculation for diamond and silicon

Table 5. Cost of the calculation for diamond and silicon as a function of the basis set (experimental lattice parameter). Of the N_C (and N_{ex}) bi-electronic Coulomb (and exchange) integrals evaluated, M_C (and M_{ex}) were approximated by using a bipolar expansion; N' is the number of symmetrised sums of integrals stored for the SCF cycle; all these numbers are in 10^6 units. Computational parameters correspond to case II in table 1. t_1 and t_2 are as in table 1. Eight to 10 iterations were necessary to gain a convergence of 1×10^{-8} au on total energy.

		STO-3G	3-21G	3-21G*	6-21G	6-21G*
C						
	N_C	46	116	289	117	291
	M_C	18	92	232	93	234
	N_{ex}	27	79	197	80	198
	M_{ex}	21	66	164	67	165
	N'	0.1	0.7	2.5	0.7	2.5
	t_1	184	276	645	288	672
	t_2	8	16	48	18	43
Si						
	N_C	22	63	143	66	149
	M_C	18	51	117	53	121
	N_{ex}	15	42	96	44	100
	M_{ex}	13	35	80	37	82
	N'	0.2	0.8	2.4	0.9	2.5
	t_1	95	167	382	199	473
	t_2	14	53	109	44	111

with the five adopted basis sets is reported in table 5. It must be noted that, as a consequence of the schemes used for the evaluation of the bi-electronic integrals (the expensive part of the calculation), the cost is dominated by those integrals involving valence (low-exponent) shells. This explains why little difference is observed in t_1 for the STO-3G and 3-21G calculations, if compared with the respective numbers of bi-electronic integrals to be evaluated; the cost-determining outer shell in the 3G basis is a contraction of three Gaussian functions, while in the 3-21G there is only one 'outer' Gaussian. So there seems to be apparently no real advantage in using the STO-3G basis set instead of the richer N-21G. It must, however, be remarked that the relative cost of the SCF part is higher, the larger the basis sets, so that for large-unit-cell systems the choice of a minimal basis set might be necessary; this is the reason why it was considered in the present analysis. Another consequence of the 'outer Gaussian effect' is that the cost difference between the 3-21G (3-21G*) and 6-21G (6-21G*) calculation is only about 4% for diamond and 20% for silicon, which has a larger core. Much higher is the additional cost due to polarisation functions; computational time and integral storage increase by a factor of about 3, as the integrals involving d-type shells are quite expensive. Evidently the N-21G and N-21G* basis sets are not at all the best basis sets which may be adopted; they, however, represent a reasonable compromise between accuracy and cost of the calculation. Improvements in the basis set might be found in the use of a larger split-valence N-111G set and in the consideration of a second set of polarisation functions. The addition of an extra diffuse sp shell (exponents between 0.1 and 0.01 au) is on the contrary not possible at present, both because of the cost of the calculation (the 'outer Gaussian effect' mentioned above) and for numerical reasons (risk of near linear dependence, due to the very large overlap between neighbouring atoms). However, such

Table 6. Calculated lattice parameters and percentage error (in parentheses) with respect to experimental values.

	Calculated lattice parameter (Å) (error (%))						Experimental
	STO 3G	3-21G	3-21G*	6-21G	6-21G*	LDA	
C	3.581	3.569	3.562	3.577	3.574	3.561[a]	3.560[b]
	(+0.60)	(+0.25)	(+0.06)	(+0.48)	(+0.39)	(+0.02)	
BN	3.601	3.630	3.607	3.637	3.619	3.606[c]	3.615[b]
	(−0.39)	(+0.41)	(−0.22)	(+0.61)	(+0.11)	(−0.24)	
SiC	4.413	4.371	4.311	4.438	4.390	4.361[a]	4.360[a]
	(+1.22)	(+0.25)	(−1.12)	(+1.79)	(+0.69)	(+0.02)	
BP	4.595	4.640	4.577	4.656	4.598	4.538[c]	4.538[b]
	(+1.26)	(+2.25)	(+0.86)	(+2.60)	(+1.32)	(0.00)	
Si	5.450	5.525	5.441	5.558	5.501	5.433[a]	5.431[b]
	(+0.35)	(+1.73)	(+0.18)	(+2.34)	(+1.29)	(+0.04)	
AlP	5.396	5.527	5.442	5.598	5.553	5.471[d]	5.451[b]
	(−1.01)	(+1.39)	(−0.16)	(+2.70)	(+1.87)	(+0.37)	

[a] [38].
[b] [39].
[c] [40].
[d] [41].

functions are expected to be much less useful in the present context than for molecules, because there are no 'tails' to describe in systems with such a packed structure.

4. Energy-related results

In table 6 the equilibrium lattice parameters obtained with the five basis sets mentioned above are compared with the experimental data. Our comments are as follows.

(i) As expected from molecular experience, the lattice parameters obtained with the re-optimised STO-3G basis set are reasonably good, the maximum error being smaller than 2%; on the contrary, the use of the standard [27, 28] scale factors produces errors as large as 5%.

(ii) Unlike molecules, where the 3-21G and 6-21G sets lead to very similar equilibrium geometries [31, 32], in periodic packed systems the accuracy in the description of the core is more important, particularly in compounds where a second row element is involved. The 3-21G(*) lattice parameters are always smaller than the 6-21G(*) values, the difference being in some cases more than 1%; the shorter equilibrium distances must be attributed to the 'improper' use of the valence functions of neighbouring atoms for the improvement of the core states. This hypothesis is confirmed by the much higher formation energies obtained with the 3-21G(*) bases than with the 6-21G(*) bases (see table 9).

(iii) The addition of polarisation functions always yields a reduction in lattice parameters which can be as large as 1.5% when second-row elements are involved; something analogous happens for molecules, where the inclusion of d functions generally results in shorter bond lengths [33].

Table 7. Calculated bulk moduli and percentage error (in parentheses) with respect to experimental values. The percentage is referred to the first entry, which for BN is obtained by an empirical relation, as explained in [42].

	Calculated bulk modulus (GPa) (error (%))						
	STO-3G	3-21G	3-21G*	6-21G	6-21G*	LDA	Experimental
C	543	461	491	460	476	438[a]	443[a]
	(+22.6)	(+4.1)	(+10.8)	(+3.8)	(+7.4)	(−1.1)	
BN	496	406	429	402	416	367[b]	367[b]
	(+35.1)	(+10.6)	(+16.9)	(+9.5)	(+13.3)	(0.0)	465[c]
SiC	240	218	244	219	238	212[a]	224[a]
	(+7.1)	(−2.7)	(+8.9)	(−2.2)	(+6.2)	(−5.4)	
BP	207	158	172	157	170	165[b]	173[b]
	(+19.7)	(−8.7)	(−0.6)	(−9.2)	(−1.7)	(−4.6)	190[c]
Si	128	93	108	94	107	92[a]	99[a]
	(+29.3)	(−6.1)	(+9.1)	(−5.1)	(+8.1)	(−7.1)	
AlP	137	86	96	84	92	85[d]	86[d]
	(+59.3)	(−0.2)	(+11.6)	(−2.3)	(+7.0)	(−1.2)	

[a] [38].
[b] [40].
[c] [43].
[d] [41].

(iv) With the 6-21G and 6-21G* basis sets regular trends are observed, similar to the ones resulting from molecular calculations. Note in particular that the error is larger for second-row-element systems; it is not clear at the moment whether this is due to larger correlation errors or to a different basis set efficiency.

(v) Note finally that the most accurate basis set (6-21G*) gives errors always smaller than 2%.

A coherent picture is also obtained for bulk moduli B, which are reported in table 7. Owing to its limited variational freedom, the STO-3G basis gives B-values too large, with errors as large as 50%. As regards split-valence calculations, while the core description has nearly no influence on this parameter, a conspicuous increase is induced by polar-isation functions, particularly when d-type orbitals are added to the 3-21G set. At the 6-21G* level a regular trend with respect to the experimental data is observed; the mean error is 7.5%. In two cases (BN and BP), table 7 reports two reference data, the percentage error being evaluated with respect to the first entry. For BN the first 'exper-imental' datum is in fact obtained from a simple empirical relation [42], which produces, however, very accurate B-values for the III–V semiconductors. The second entry, resulting from measurements of Cardona and co-workers [43], is affected by a declared experimental error larger than 10%, as is the second experimental value reported for BP [43]. Bulk modulus derivatives are not reported. Although values interpolated by equation (6) are all in the range between 3.4 and 4.4, which is quite reasonable (experimental results for diamond and silicon are 4.0 and 4.2 [38], respectively), numeri-cal error is such that fluctuations larger than 20% are observed with respect to different choices of energy points and no definite trend can be demonstrated.

Another quantity of interest is the energy of the transverse optical mode at $q = 0$. In the 'frozen-phonon' harmonic approximation it is given as

$$E_{TO} = (m_e k/\mu)^{1/2}$$

Table 8. Transverse optical frequencies and percentage error (in parentheses) with respect to experimental (mean, for SiC and BP) values. The 6-21G* basis set was used.

	Transverse optical frequency (THz) (error (%))					
	C	BN	SiC	BP	Si	AlP
Calculated[a]	44.60	33.85	25.65	27.56	17.02	13.92
	(+11.8)	(+7.0)	(+9.2)	(+13.6)	(+9.6)	(+5.7)
Calculated[b]		30.1[c]		23.0[c]	15.16[d]	
		(−4.8)		(−5.2)	(−2.4)	
Experimental	39.9[e]	31.62[e]	23.86[e]	23.9[c]	15.53[e]	13.17[e]
			23.11[e]	24.6[c]		

[a] This work.
[b] LDA calculations.
[c] [40].
[d] [45].
[e] [46].

where k is the second derivative of the energy (in Hartrees) with respect to anion–cation displacements (in Bohr radii) from equilibrium, m_e and μ are the electron mass and the reduced mass of the unit cell respectively (both in atomic mass units), so that E_{TO} is in Hartrees. To convert to the more commonly used units of terahertz, we use

$$\nu_{TO} = h^{-1} E_{TO} 10^{-12} = h^{-1} m_e^{1/2} (k/\mu)^{1/2} 10^{-12} = 154.1079 (k/\mu)^{1/2}$$

where we have used the most recently available [44] values for the fundamental constants. The force constant is obtained from a third-order polynomial best fit performed on a sample of five energy points corresponding to different sublattice displacements. This interpolation is less troublesome than the calculation of bulk modulus, as no alteration of the unit-cell volume takes place and consequently numerical noise is less. Transverse optical frequencies calculated with the most extended basis set used, the 6-21G*, are reported in table 8. Deviations from experimental data are all in the positive direction and of about the same order of magnitude as for bulk moduli; errors range from +5.7 to +13.6%.

Binding energies reported in table 9 are obtained as the difference between total energies of the bulk and of the isolated atoms calculated with the same basis set, the only difference being the exponent of the outer Gaussian, which has been optimised separately for the bulk and for the isolated atom. A comparison between the 3-21G and 6-21G columns shows that the first basis set describes more bound systems, because of the basis set superposition error, which favours the bulk situation when the core is poor. Polarisation functions supply appreciable contributions, of about 10%. The 6-21G* binding energies are smaller than the experimental values by about 30%, the main source of error being the lack of correlation. An attempt at a posteriori recovery of correlation contributions is possible by using density functionals of the correlation energy, which have been proposed recently [7, 8]. Two were tested; results are reported in table 10. Similarly to HF binding energy the correlation contribution to binding energy is evaluated as the difference between correlation energies of the bulk and of the atoms. Satisfactory agreement with experiment is gained this way and trends are regular; while the Colle–Salvetti functional always underestimates the correlation contributions, the Perdew functional gives values in good agreement with experiment. Preliminary tests seem to

Table 9. Calculated binding energies BE and percentage error (in parentheses) with respect to experimental values. $BE = E_{atom} - E_{bulk}$, E_{bulk} being the total energy (per atom pair) of the periodic systems and E_{atom} the sum of the corresponding atom HF energies calculated with the same basis sets. The experimental binding energies have been corrected for the zero-point vibration energies.

	STO-3G	3-21G	3-21G*	6-21G	6-21G*	Experimental
C						
E_{bulk}(Hartree)	−74.8769	−75.3439	−75.3794	−75.6803	−75.7139	
E_{atom}(Hartree)	−74.4573	−74.9621	−74.9621	−75.3179	−75.3179	
BE (Hartree/cell)	0.4196	0.3818	0.4173	0.3624	0.3960	0.555[a]
(error (%))	(−24.4)	(−31.2)	(−24.8)	(−34.7)	(−28.6)	
BN						
E_{bulk}(Hartree)	−78.3156	−78.8419	−78.8693	−79.1858	−79.2130	
E_{atom}(Hartree)	−77.9517	−78.4952	−78.4952	−78.8593	−78.8593	
BE (Hartree/cell)	0.3639	0.3467	0.3741	0.3265	0.3537	0.498[b]
(error (%))	(−26.9)	(−30.4)	(−24.9)	(−34.4)	(−29.0)	
SiC						
E_{bulk}(Hartree)	−323.1471	−325.1536	−325.1980	−326.7711	−326.8118	
E_{atom}(Hartree)	−322.7926	−324.8255	−324.8255	−326.4844	−326.4844	
BE (Hartree/cell)	0.3545	0.3281	0.3725	0.2867	0.3274	0.475[c]
(error (%))	(−25.4)	(−30.9)	(−21.6)	(−39.6)	(−31.1)	
BP						
E_{bulk}(Hartree)	−361.4786	−363.6197	−363.6646	−365.4036	−365.4380	
E_{atom}(Hartree)	−361.1768	−363.3898	−363.3898	−365.2009	−365.2009	
BE (Hartree/cell)	0.3018	0.2299	0.2748	0.2027	0.2371	0.383[b]
(error (%))	(−21.2)	(−40.0)	(−28.3)	(−47.1)	(−38.1)	
Si						
E_{bulk}(Hartree)	−571.4237	−574.9006	−574.9494	−577.8274	−577.8730	
E_{atom}(Hartree)	−571.1279	−574.6889	−574.6889	−577.6508	−577.6508	
BE (Hartree/cell)	0.2958	0.2117	0.2605	0.1766	0.2222	0.345[d]
(error (%))	(−14.3)	(−38.6)	(−24.5)	(−48.8)	(−35.6)	
AlP						
E_{bulk}(Hartree)	−576.2450	−579.7684	−579.8110	−582.7028	−582.7375	
E_{atom}(Hartree)	−575.9735	−579.5511	−579.5511	−582.5367	−582.5367	
BE (Hartree/cell)	0.2715	0.2173	0.2599	0.1661	0.2008	0.317[e]
(error (%))	(−14.4)	(−31.5)	(−18.0)	(−47.6)	(−36.7)	

[a] [11].
[b] [40].
[c] Data from [38], corrected for thermal motion (9.0 mHartree, obtained from a Debye temperature of 1270 K [47]).
[d] [48].
[e] Data from [49], corrected for thermal motion (4.2 mHartree, obtained from a Debye temperature of 588 K [47]).

indicate that those functionals give negligible corrections to the other parameters of the Murnaghan equation, and in particular to the lattice parameter and the bulk modulus. More sophisticated schemes are probably required to take into account correlation effects on those quantities correctly.

5. Comparison with density-functional results

Comparison with the results of DF schemes is difficult both because of the wide variety of functionals in use, usually a local-density approximation (LDA), and because of

Table 10. Correlation correction to HF binding energies obtained with the 6-21G* basis set and percentage error (in parentheses) with respect to experimental values. The Colle–Salvetti (CS) and the Perdew (P) correlation-only density functionals have been used.

	Correlation correction (Hartree/cell) (error (%))				
	HF	HF + CS	HF + P	LDA	Experimental
C	0.3960	0.5109	0.5488	0.584[a]	0.555
	(−28.6)	(−7.9)	(−1.1)	(+5.2)	
BN	0.3537	0.4649	0.4963	0.526[b]	0.498
	(−29.0)	(−6.6)	(−0.3)	(+5.6)	
SiC	0.3274	0.4320	0.4697	0.490[a]	0.475
	(−31.1)	(−9.1)	(−1.1)	(+3.2)	
BP	0.2371	0.3401	0.3743	0.423[b]	0.383
	(−38.1)	(−11.2)	(−2.3)	(+10.4)	
Si	0.2222	0.3128	0.3466	0.356[a]	0.345
	(−35.6)	(−9.3)	(+0.5)	(+3.2)	
AlP	0.2008	0.2882	0.3188	—	0.317
	(−36.7)	(−9.1)	(+0.6)		

[a] [38].
[b] [40].

uncertainty in the quality of the basis set used. The DF values quoted in tables 6–8 and 10 are all taken from the work of M L Cohen and his co-workers, using the DF-PP-PW scheme with the Wigner [50] interpolating functional. These studies appear to be the only ones which have systematically covered the same systems and properties as in the present work. We note that the calculated DF lattice constants (table 6) are in close agreement with experiment, while the bulk moduli are too small by typically 3–5% (table 7), the phonon frequencies too small by 3–5% (table 8), and the binding energies are too great by typically 5–10% (table 10). However, it is estimated that the DF binding energies would increase by a further 10% given a complete basis, further exaggerating the tendency of DF theory to over-bind. For example, the DF limit binding energy of diamond has been estimated to be 0.645 Hartree/cell [11], and a recent highly accurate study in a local (muffin tin orbital) basis using the von Barth and Hedin [51] functional gave a binding energy for silicon of 0.384 Hartree/cell [52], both considerably greater than the DF-PP-PW results [38] quoted in table 10. In any event, HF theory corrected by the Perdew [8] correlation-only functional gives by far the most accurate binding energies of any available theory.

Another complicating factor when comparison with DF theory is made is the question of the exact form of functional used. Thus Wentzcovitch *et al* [40] have shown that changing from the Wigner to the Hedin and Lundqvist [53] functional causes approximately a 1% decrease in lattice constant, thereby degrading the level of agreement with experiment. However, the bulk moduli increase by 3%, so improving the agreement with experiment. Yin and Cohen [45] have shown that the same change in functional causes the calculated phonon frequencies to increase by 1%. It is worth mentioning here that DF calculations using Gaussian basis sets (and thus perhaps most comparable with the present work) are well known. Thus the results of Harmon *et al* [54] for silicon are 5.488 Å, 87 GPa, 0.356 Hartree/cell and 15.0 THz for lattice constant, bulk modulus, binding energy and transverse optical phonon frequency, respectively. The status of DF
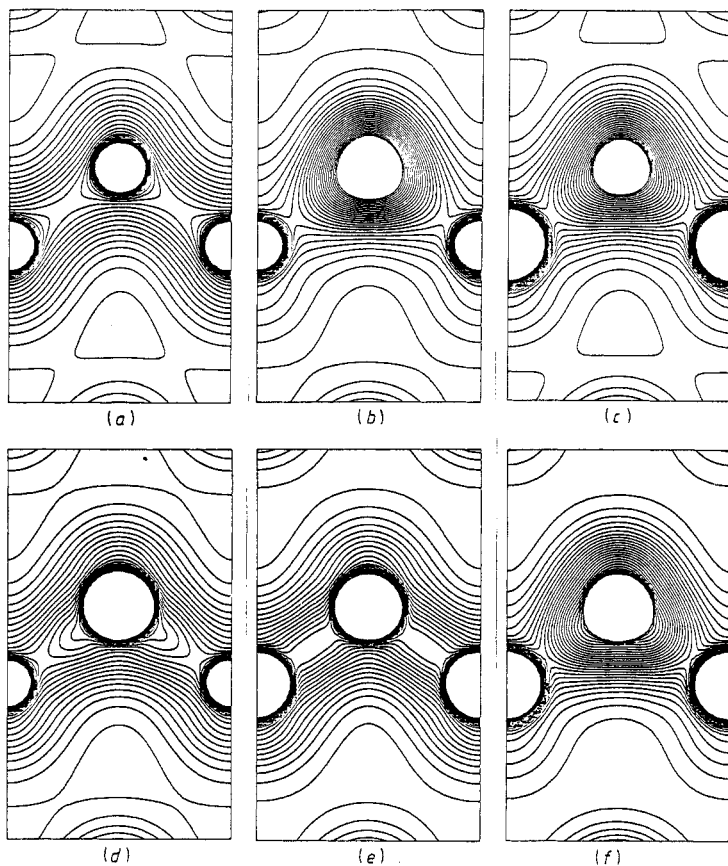
Figure 1. Electronic total charge isodensity plots for (*a*) diamond, (*b*) BN, (*c*) SiC, (*d*) BP, (*e*) silicon and (*f*) AlP in the (110) plane at the experimental value V_0 of the primitive cell volume. Charge density is in units of electrons per V_0, with a step of 1.5 electron per V_0 between contiguous curves. The atomic positions are as in figure 2.

calculations using local basis sets has been well stated by Chelikowsky [55] as follows: 'typically, one has confidence in the calculated cohesive energies to within 10%, the lattice parameter to within 2–3%, and the bulk modulus to within 5%'. The present work demonstrates rather similar uncertainties within the HF method.

6. Charge-density and band-structure data

The electron charge-density maps obtained with the 6-21G* basis are reported in figure 1, whereas figure 2 shows the difference between the crystalline charge density and the superposition of the spherical atomic densities. For the isolated atoms the standard atomic 6-21G basis sets [30–33] have been adopted. Complementary information is supplied in table 11, where Mulliken population data are reported. It must be stressed that population data for periodic compounds should be used in an even more qualitative sense than in molecular contexts; the large number of neighbours and the packed nature
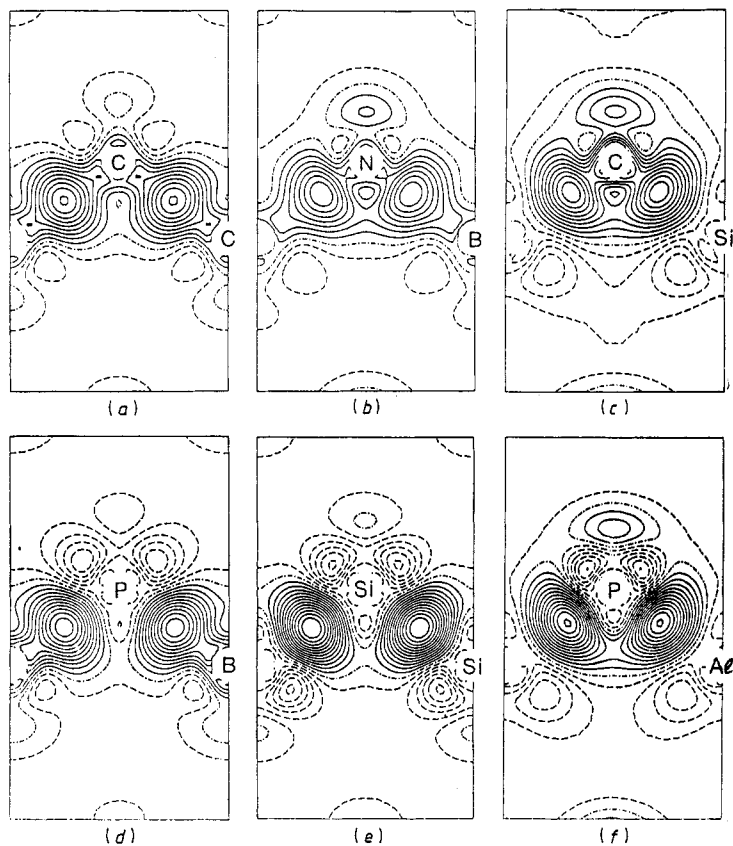
Figure 2. Electronic charge difference isodensity plots between the studied systems and the corresponding atomic superposition arrays. The plane and units are as in figure 1. The full and broken curves indicate positive and negative increases of 0.75 electron per V_{ti}, respectively; the chain curve line corresponds to zero charge density.

of those compounds is such that small changes in the basis or in the partition scheme of the charge can produce changes as large as 0.2–0.3 electrons.

Let us consider first the situation in the two homonuclear compounds. In both cases, as expected, there is a large build-up of electronic charge along the bonds; however, the deformation maps in figure 2 show that in the Si case the bond charge is much more diffuse than in diamond, where the lack of core p electrons allows the valence electrons to get nearer the nuclei. Figure 1 and table 11 also show that, whereas for diamond there is a monotonic decrease in the total charge density from the atomic positions to the point midway between the two atoms, for silicon there is a large plateau, of nearly constant density. In the case of silicon, the structure factors calculated from the HF charge density can be compared with the very accurate experimental data reported by Spackman [56]. The 6-21G* basis gives 0.208 for the 'forbidden' 222 reflection, to be compared with Spackman's value of 0.193. The agreement factor $R = \Sigma(|F_{exp}| - |F_{calc}|)/\Sigma|F_{exp}|$ evaluated from the 18 non-equivalent structure factors is 0.22, not too far from the experimental uncertainty ($R = 0.13$ in Spackman's refinement) and much better than that obtained from any previous calculation. In particular, the R-factors from LDA cal-

Table 11. Electronic distribution data for the six semiconductors. Q_A is the Mulliken gross charge of atom A, Q_{AB} is the bond population between atom A and B, and Q_{dA} is the contribution from d orbitals to Q_A. ρ_m is the minimum electron density along the A–B path, at a fractionary distance d_m from atom A. In the silicon case there are two symmetric minima, and a relative maximum midway the bond, as shown in figure 1. The calculations were performed with a 6-21G* basis set, at the experimental lattice parameter.

	q_A (electrons)	q_B (electrons)	q_{AB} (electrons)	q_{dA} (electrons)	q_{dB} (electrons)	ρ_m (electrons Bohr^{-3})	d_m
C	6.00	6.00	0.752	0.063	0.063	0.2901	0.50
BN	4.14	7.86	0.590	0.078	0.023	0.1527	0.31
SiC	12.19	7.81	0.576	0.215	0.013	0.1110	0.39
BP	5.34	14.66	0.700	0.028	0.128	0.1306	0.70
Si	14.00	14.00	0.678	0.087	0.087	0.0943	0.50
						0.0927	0.35
AlP	12.20	15.80	0.575	0.155	0.041	0.0592	0.35

culations are very large ($R = 1.12$ for the data in [45]), probably because the core correction to the pseudo-valence contribution has been carried out incorrectly [57]. Figures 1 and 2 reproduce well the general features of the corresponding experimental maps [56]. The electron density midway between neighbouring atoms is 0.094 (see table 11), to be compared with 0.085 from experiment, the difference being mainly due to basis set limitations. LDA calculations in PW basis sets give values of between 0.084 and 0.089 electron Bohr^{-3} [56].

As regards the heteronuclear compounds, the electronic charges suggest the following 'ionicity' order: SiC > AlP, BN > BP. Actually, the BP charge transfer is 'reversed', and boron appears to be negatively charged. The low ionicity of BP is confirmed by figure 1, which shows features along the atom–atom direction quite similar to the Si–Si ones; the bond charge in figure 2 appears to be localised roughly midway between the two atoms, whereas for SiC, AlP and BN it is nearer the 'anion'. Note that the bond population of BP is much higher than those of the three other compounds. The electronegativity decreases both going from the first to the second row (second to third period) and going from group V to group III. AlP and BN are 'in-row' III–V compounds and show then roughly the same 'ionicities'; BP is a 'cross-row' compound, and the two effects balance nearly exactly. The very high 'ionicity' resulting from our calculations for SiC indicates that the 'row' effect is at least as important as the 'group' effect shown by BN and AlP.

The 'ionicity scale' resulting from those data can be compared with the empirical electronegativity scales proposed by Pauling [58] and by Phillips [59]; the B minus A differences are 0.7, 1.0, 0.6, 0.1 and 1.1, 1.0, 0.6, −0.4 for SiC, BN, AlP and BP, respectively. In both cases the BP ionicity is quite small; on the Phillips scale it is 'reversed'. SiC is more ionic than BN and AlP on the Phillips scale but not on the Pauling scale. The 'ionicity scale' recently proposed by Christensen *et al* [60] on the basis of quantum mechanical LMTO [61] calculations gives very similar ionicities for SiC, BN and AlP, which are much larger than the BP value (nearly null), in qualitative agreement with the present results.

Table 12 reports the most relevant data of the band structure for the six compounds, which is shown in figure 3. It is well known that the eigenvalues of the one-electron Fock Hamiltonian are a poor approximation of the spectrum of quasi-particle energies of

Table 12. Band-structure main features.

	Valence band width (eV)			Minimum gap (eV)	
	Lower	Upper	Full	Indirect	Direct
C	11.6	18.5	29.9	12.1	13.8
BN	7.1	14.3	27.7	13.5	18.9
SiC	6.4	11.7	21.7	8.6	13.1
BP	6.9	13.0	22.0	7.6	10.0
Si	5.9	11.5	17.4	6.0	8.4
AlP	3.3	7.8	16.5	7.7	10.8
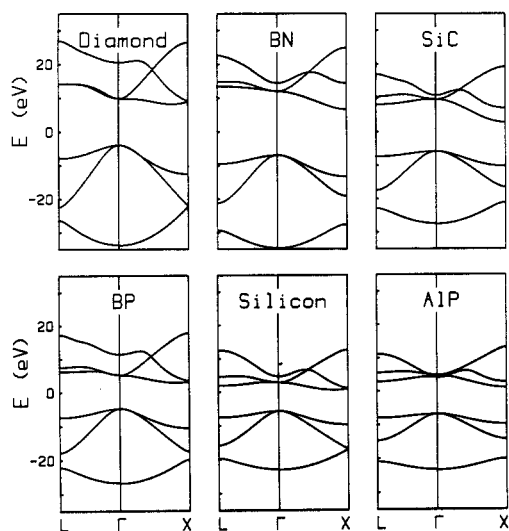
Figure 3. Electronic band structures at the experimental lattice parameters along the path L–Γ–X: valence and lower conduction bands.

crystalline compounds. Band gaps and band widths are systematically overestimated by a factor of 2 or more; for example the 'experimental' vertical band gaps at Γ for diamond and silicon are 7.3 eV and 3.4 eV [62], respectively, to be compared with 13.8 eV and 8.4 eV from the present calculations. On the contrary, the topology of the band structure and its main features are well reproduced. The minimum gap in the six cases is indirect (not vertical), as expected from experiment and in agreement with LDA calculations [40, 41, 63]. The lowest point in the conduction band structure is at point X for the heteronuclear compounds, and along the Γ–X line in diamond and silicon.

7. Conclusions

The HF–CO–LCAO program CRYSTAL has been applied to six semiconductors; the effects of the computational parameters and of the basis set on binding energy, equilibrium distance, bulk modulus and transverse optical frequencies have been discussed; band-structure and charge distribution data have also been reported.

On the basis of these data, the following conclusions can be drawn.

(i) The numerical accuracy of the computational scheme (treatment of the Coulomb and exchange series, reciprocal space integrations; more generally, stability and reliability of the code), although still far from that of molecular programs, is now sufficient to give quantitative answers not only for total energy, but also for geometry optimisations and vibrational frequencies.

(ii) For semiconductors, standard but not very large molecular basis sets can be used after the re-optimisation of the outer shell; on the contrary, more extended basis sets ('triple-zeta' type, or even larger) are to be discarded to avoid numerical instabilities and huge costs.

(iii) The errors with respect to experimental data for lattice parameters, bulk moduli and binding energies are, roughly speaking, quite similar to those obtained in molecular calculations with similar basis sets; probably the main difference is the greater import- ance of the core description quality in condensed systems.

(iv) The cost of such calculations is not very high (2–3 min on Cray-XMP machines per energy point); the availability of the program, from the present authors or from QCPE, should encourage more quantum chemists to carry out similar work.

(v) Limitations of the approach are evident from the previous discussion; a better numerical-accuracy-to-cost ratio, and an analytical energy gradient option would be useful; work is in progress in both directions.

Acknowledgments

References

[1] Hehre D J, Lathan W A, Newton M D, Ditchfield R and Pople J A 1973 GAUSSIAN 70 Program no. 236, Quantum Chemistry Program Exchange, Indiana University, Bloomington, IN
[2] Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864
 Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133
[3] Bachelet G B, Hamann D R and Schlüter M 1982 *Phys. Rev.* B **26** 4199
[4] Callaway J 1964 *Energy Band Theory* (New York: Academic)
[5] Ashcroft N W and Mermin N D 1976 *Solid State Physics* (Philadelphia, PA: Holt–Saunders) p 335
[6] Monkhorst H J 1979 *Phys. Rev.* B **20** 1504
[7] Colle R and Salvetti O 1975 *Theor. Chim. Acta* **37** 329
 Colle R, Fortunelli A and Salvetti O 1984 *J. Chem. Phys.* **80** 2654
[8] Perdew J P 1986 *Phys. Rev.* B **33** 8822; **34** 7406
[9] Causà M, Dovesi R, Pisani C, Colle R and Fortunelli A 1987 *Phys. Rev.* B **36** 891
 Causà M, Colle R, Fortunelli A, Dovesi R and Pisani C 1988 *Phys. Scr.* **38** 194
[10] Kutzelnigg W 1977 *Modern Theoretical Chemistry* vol 3, ed H F Schaefer III (New York: Plenum) p 129
[11] Stollhoff G and Bohnen K P 1988 *Phys. Rev.* B **37** 4678
[12] Ganduglia-Pirovano M V, Stollhoff G, Fulde P and Bohnen K P 1989 *Phys. Rev.* B **39** 5156
[13] McMillan W L 1965 *Phys. Rev.* **138** A442
 Ceperley D, Chester G V and Kalos M H 1977 *Phys. Rev.* B **16** 3087
[14] Fahy S, Wang X W and Louie S G 1988 *Phys. Rev. Lett.* **61** 1631
[15] André J M, Gouverneur L and Leroy G 1967 *Int. J. Quantum Chem.* **1** 451
 Del Re G, Ladik J and Biczo G 1967 *Phys. Rev.* **155** 997

Harris F E and Monkhorst H J 1969 *Phys. Rev. Lett.* **23** 1026

Euwema R N, Wilhite D L and Surratt G T 1973 *Phys. Rev.* B **7** 818

Stoll H and Preuss H 1973 *Phys. Status Solidi* b **60** 185

Kertész M, Koller J and Azman A 1977 *J. Chem. Phys.* **67** 1180

Brener N and Fry J L 1978 *Phys. Rev.* B **17** 506

Suhai S 1980 *J. Chem. Phys.* **73** 3843

Karpfen A 1981 *Int. J. Quantum Chem.* **19** 1207

Svane A and Andersen O K 1986 *Phys. Rev.* B **34** 5512

[16] Pisani C and Dovesi R 1980 *Int. J. Quantum Chem.* **17** 501

[17] Dovesi R, Pisani C, Roetti C and Saunders V R 1983 *Phys. Rev.* B **28** 5781

[18] Dovesi R 1986 *Int. J. Quantum Chem.* **29** 1755

[19] Pisani C, Dovesi R and Roetti C 1988 *Hartree–Fock Ab-initio Treatment of Crystalline Systems, Lecture Notes in Chemistry* (Heidelberg: Springer)

[20] Dovesi R, Pisani C, Roetti C, Causà M and Saunders V R 1989 CRYSTAL 88. Program no. 577, Quantum Chemistry Program Exchange, Indiana University, Bloomington, IN

[21] Dovesi R 1984 *Int. J. Quantum Chem.* **26** 197

[22] Causà M, Dovesi R, Pisani C and Roetti C 1986 *Surf. Sci.* **175** 551

[23] Dovesi R, Pisani C, Ricca F and Roetti C 1982 *Phys. Rev.* B **25** 3731

[24] Dovesi R, Pisani C and Roetti C 1980 *Int. J. Quantum Chem.* **17** 517

Dovesi R, Pisani C, Ricca F and Roetti C 1980 *Phys. Rev.* B **22** 5936

Dovesi R, Pisani C, Roetti C and Dellarole P 1981 *Phys. Rev.* B **24** 4170

[25] Dovesi R, Ermondi C, Ferrero E, Pisani C and Roetti C 1984 *Phys. Rev.* B **29** 3591

Causà M, Dovesi R, Pisani C and Roetti C 1986 *Phys. Rev.* B **33** 1308

[26] Dovesi R, Causà M, Orlando R, Roetti C and Saunders V R 1990 *J. Chem. Phys.* at press

[27] Hehre W J, Stewart R F and Pople J A 1969 *J. Chem. Phys.* **51** 2657

[28] Hehre W J, Ditchfield R, Stewart R F and Pople J A 1970 *J. Chem. Phys.* **52** 2769

[29] Pulay P 1977 *Modern Theoretical Chemistry* vol 4, ed H F Schaefer III (New York: Plenum) p 153

[30] Hariharan P C and Pople J A 1973 *Theor. Chim. Acta* **28** 213

[31] Binkley J S, Pople J A and Hehre W J 1980 *J. Am. Chem. Soc.* **102** 939

[32] Gordon M S, Binkley J S, Pople J A, Pietro W J and Hehre W J 1982 *J. Am. Chem. Soc.* **104** 2797

[33] Pietro W J, Francl M M, Hehre W J, DeFrees D J, Pople J A and Binkley J S 1982 *J. Am. Chem. Soc.* **104** 5039

[34] Murnaghan F D 1944 *Proc. Natl Acad. Sci. USA* **30** 244

[35] Tosi M P 1964 *Solid State Physics* vol 16, ed F Seitz and D Turnbull (New York: Academic) p 1

Ewald P P 1921 *Ann. Phys., Lpz.* **64** 253

[36] Carlson B G and Rushbrooke G S 1950 *Proc. Camb. Phil. Soc.* **46** 626

Buehler R J and Hirschfelder J O 1951 *Phys. Rev.* **83** 628

[37] Causà M, Dovesi R, Orlando R, Pisani C and Saunders V R 1988 *J. Phys. Chem.* **92** 909

[38] Chang K J and Cohen M L 1987 *Phys. Rev.* B **35** 8196

[39] Weast R C, Astle M J and Beyer W H 1987 *Handbook of Chemistry and Physics* 67th edn (West Palm Beach, FL: CRC)

[40] Wentzcovitch R M, Chang K J and Cohen M L 1986 *Phys. Rev.* B **34** 1071

[41] Zhang S B and Cohen M L 1987 *Phys. Rev.* B **35** 7604

[42] Cohen M L 1985 *Phys. Rev.* B **32** 7988

[43] Sanjurjo J A, Lopez-Cruz E, Vogl P and Cardona M 1983 *Phys. Rev.* B **28** 4579

[44] Cohen E R and Taylor B N 1987 *Rev. Mod. Phys.* **59** 1121

[45] Yin M T and Cohen M L 1982 *Phys. Rev.* B **26** 3259

[46] Weyrich K H, Brey L and Christensen N E 1988 *Phys. Rev.* B **37** 1392

[47] *Landolt–Börnstein New Series* 1982 vol 17a ed O Madelung (Berlin: Springer)

[48] Yin M T and Cohen M L 1982 *Phys. Rev.* B **26** 5668

[49] Harrison W A 1981 *Phys. Rev.* B **23** 5230

[50] Wigner E 1938 *Trans. Faraday Soc.* **34** 678

[51] von Barth U and Hedin L 1972 *J. Phys. C: Solid State Phys.* **5** 1629

[52] Methfessel M, Rodriguez C O and Andersen O K 1989 *Phys. Rev.* B **40** 2009

[53] Hedin L and Lundqvist B I 1971 *J. Phys. C: Solid State Phys.* **4** 2064

[54] Harmon B N, Weber W and Hamann D R 1982 *Phys. Rev.* B **25** 1109

[55] Chelikowsky J R 1987 *Phys. Rev.* B **35** 1174

[56] Spackman M A 1986 *Acta Crystallogr.* A **42** 271

[57] von Barth U and Pedroza A C 1985 *Phys. Scr.* **32** 353

[58] Pauling L 1960 *The Nature of the Chemical Bond* 3rd edn (Ithaca, NY: Cornell University Press)
[59] Phillips J C 1970 *Rev. Mod. Phys.* **42** 317
[60] Christensen N E, Satpathy S and Pawlowska Z 1987 *Phys. Rev.* B **36** 1032
[61] Skriver H L 1984 *The LMTO Method* (Berlin: Springer)
[62] Godby R W, Schluter M and Sham L J 1987 *Phys. Rev.* B **36** 6497
[63] Jansen R W and Sankey O F 1987 *Phys. Rev.* B **36** 6520