Surface Science 224 (1989) 515–524
North-Holland, Amsterdam

ON THE POSSIBILITY OF TWO-DIMENSIONAL GROWTH OF GaAs ON ATOMICALLY FLAT Si(100) SURFACES

Efthimios KAXIRAS

IBM Research Division, T.J. Watson Research Center, Yorktown Heights, NY 10598, USA

and

J.D. JOANNOPOULOS

Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 13 July 1989; accepted for publication 28 August 1989

The possibility of growing GaAs on atomically flat Si(100) surfaces is investigated at monolayer and bilayer coverages. Based on ab-initio total-energy calculations at zero-temperature we compare the stability of separate domains of Ga or As versus mixed monolayers, the stability of pure GaAs bilayers versus mixed bilayers and the formation of GaAs bilayer islands versus wetting of the surface. The results predict that two-dimensional growth of bulk GaAs on atomically flat regions (terraces) of Si(100) is severely inhibited at the initial stages, due to chemical and rehybridization reactions of the Ga and As atoms on the Si surface.

Heterostructures of semiconductors, elemental or compound, have recently been the subject of intense study [1] as technologically promising artificial materials. A system of particular interest, not only from the technological but also from the theoretical point of view, is GaAs on Si [2]. Usually GaAs growth is initiated on Si substrates oriented along the (100) direction. The quality of the GaAs overlayers is far from ideal, with very high density of defects and dislocations present presumably due to lattice mismatch and polarity differences. However, it is not known to what degree other interface factors, such as surface reactions or substrate morphology, play an important role in growth inhibition or initiation. These issues are at the core of current growth models of GaAs on Si(100) [3,4].

In this paper we investigate, from an ab-initio theoretical point of view, the possibility of layered growth of GaAs on atomically flat regions (terraces) of the Si surface. We shall focus on the initial stages of growth by considering very low coverages (one to two monolayers) where a number of important issues can be isolated and addressed in detail. For example, what is the structure of the initial monolayer? Are the Ga and As atoms locally mixed or

0039-6028/89/$03.50

do they separate into domains of each species? Is the Si surface covered by Ga and As atoms to the largest possible extent at monolayer thickness or do these atoms form GaAs bilayer islands on part of the surface? What is the structure of the initial bilayer of GaAs? Does it resemble that of bulk GaAs with alternating planes of Ga and As or does it have a novel structure?

The main conclusion which emerges from answers to the above questions is that two-dimensional growth of bulk GaAs on Si(100) terraces is severely inhibited at the initial stages, due to chemical and rehybridization reactions of the Ga and As atoms on the Si surface. The implication of this prediction is that if other growth channels exist, which can nucleate the bulk phase of GaAs, they will prevail in the growth process. We emphasize that here we direct our attention to the very beginning of the growth process, where *local* equilibrium is possible but *global* equilibrium with mass transport at large distances is not achievable. The latter case was recently examined by Northrup [4] who showed that formation of three-dimensional GaAs islands would be energetically favorable under *global* equilibrium conditions.

To address the possibility of two-dimensional growth growth on Si terraces we consider a slab of Si bounded by (100) surfaces on which we place Ga and As atoms. The issue of interest is what is the lowest-energy local structure and how does this structure affect the prospects for growth. There exist two possibilities: the lowest-energy local structure may be the same as the structure of bulk GaAs, in which case two-dimensional growth can proceed, or it may be different, in which case layered growth of zincblende GaAs will be inhibited. The easiest way to envision formation of the lowest-energy structure is to consider a Si surface covered by a monolayer of one species (corresponding to pre-deposition, for example, of Ga or As) and introduce some amount of the opposite species. One then must consider the possible local atomic configurations and identify the lowest-energy structure. In this sense the exact relative abundance (chemical potential) of Ga and As atoms is not important, but the presence of both species is crucial. Temperature and kinetic effects, which play an important role in subsequent growth, are not addressed here, under the assumption that growth is initiated at low temperature and local-equilibrium conditions. Note that, at the monolayer and bilayer coverages to be discussed, strain due to lattice mismatch between GaAs and Si is relieved by vertical relaxation of the Ga and As overlayers. The experimental lattice mismatch is 3.9% whereas the theoretical value obtained by our calculations is 2.5%. Thus the strain problem is underestimated here, but due to relaxation, it does not enter in physical considerations at the coverages studied. Only at much higher coverages where vertical relaxation is not possible, strain becomes an important issue and may give rise to dislocations and other defects.

Finally, the role of intermixing between the Ga and As overlayers with the Si substrate has to be considered. Experiments appear to indicate that when deposited separately, Ga and As remain on the surface at various coverages

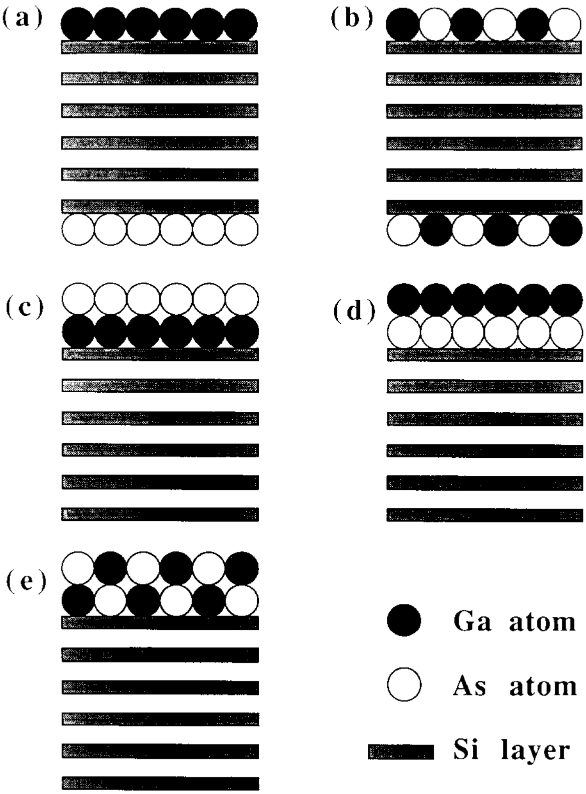

Fig. 1. Schematic representation of overlayer arrangements. (a) Pure monolayer coverage, (Ga)[Si](As). (b) Mixed monolayer coverage, (GaAs)[Si](AsGa). (c) Bilayer coverage with pure layers and As-exposed surface, (As)(Ga)[Si]. (d) Same as (c) with Ga-exposed surface, (Ga)(As)[Si}. (e) Bilayer coverage with mixed layers, (GaAs)(AsGa)[Si}.

without intermixing with the Si substrate [5–7]. Theory, on the other hand, shows that intermixing tends to lower the energy of the system by a small amount (of order 0.1 eV per (2×1) surface unit cell [4], in the following denoted as $/(2 \times 1)$ SUC). In the present work we shall consider one particular mode of intermixing, namely intermixing of entire layers of Si with Ga and As layers, analogous to a recent proposal for growth along the (111) direction [8].

Each surface layer of the slab is taken to be reconstructed, for simplicity, in the standard (2×1) dimer pattern (this choice will be justified below). The Ga and As atoms are placed on the slab in five different ways, shown in figs. 1a–1e:

(a) One surface of the slab is covered by Ga atoms while the other surface is covered by As atoms exclusively. We denote this by (Ga)[Si](As), where ( )

represents a single overlayer and [ ] represents the substrate slab. This situation models the complete coverage of a Si surface by separate domains of Ga and As atoms at monolayer thickness.

(b) Each surface of the slab is covered by GaAs pairs (one GaAs pair in each (2 × 1) surface unit cell). We denote this by (GaAs)[Si](AsGa). This corresponds to complete coverage of a Si surface by a uniform, mixed monolayer. This geometry would obstruct GaAs growth along the (100) direction.

(c) Only one of the two surfaces of the slab is covered by Ga and As atoms, with the Ga atoms in a single layer bonded to the surface Si layer and the As atoms in a single layer above the Ga atoms. This is denoted by (As)(Ga)[Si] and corresponds to a situation in which the available Ga and As atoms do not cover a Si surface completely but rather form a domain of zincblende GaAs along the (100) direction, with Ga bonded to the Si surface and As exposed.

(d) An arrangement similar to case (c) is formed, but with the two layers of GaAs interchanged. This is denoted by (Ga)(As)[Si].

(e) A bilayer of Ga and As atoms is formed on one side of the slab only, as in cases (c) and (d). However, this bilayer is not formed by two successive layers of pure composition (each layer composed either of Ga or As atoms exclusively) but by two mixed layers. Both the layer adjacent to the Si surface, as well as the exposed layer, contain equal numbers of Ga and As atoms. This is denoted by (GaAs)(AsGa)[Si] and, as in case (b), would obstruct the growth of GaAs.

Although the set of configurations described above is limited, it spans all the possibilities of qualitatively different *local* atomic arrangements for the stoichiometry and periodicity considered. In the real system the atomic configuration will in general be a mixture of such local structures. If energy differences between configurations (a) through (e) are large on the scale of growth temperatures, then the lowest-energy structure will dominate. As our calculations show, this is indeed the case, with energy differences being of order 1 eV/(2 × 1) SUC, whereas typical growth temperatures are of order 0.1 eV.

In order to obtain the relative energies of the configurations described above we have performed ab-initio total-energy calculations in the context of local-density-functional theory [9] using norm-conserving non-local pseudopotentials [10] to represent the atomic cores and a plane-wave basis. The Si slab consists of six atomic layers in the (100) direction, on which the two layers of Ga and As atoms are placed. A vacuum region equivalent to four layers along the (100) direction separates the GaAs-covered slabs. Since a wide variety of bonding arrangements and layer structures will be compared, serious consideration has to he given to convergence criteria. We have performed convergence studies of the number of sampling points in the surface Brillouin zone (SBZ), the energy-cutoff for the plane-wave basis, the need for

Table 1
Relative energy differences of the various overlayer configurations as depicted in fig. 1 in eV/(2×1) SUC

Fig.	Configuration	Relative energy difference
1a	(Ga)[Si](As)	0.8
1b	(GaAs)[Si](AsGa)	0.0
1c	(As)(Ga)[Si]	2.2
1d	(Ga)(As)[Si]	2.8
1e	(GaAs)(AsGa)[Si]	1.5

inversion symmetry to eliminate spurious vacuum fields in the polar configurations [11] and the number of relaxed layers. The final results were obtained using 32 sampling points in the SBZ and a plane-wave energy-cutoff of 10 Ryd. Since the exact structure of dimers on each surface is not known a priori, the three outermost layers of each side of the slab were allowed to relax fully by following the calculated Hellmann–Feynman forces. Thus for each atomic arrangement the lowest-energy equilibrium structure is obtained. The relative energies of the different configurations are given in table 1 (the error-bar in these numbers is estimated from the convergence study to be 0.2 eV/(2 × 1) SUC). The magnitude of these energy differences justifies the choice of the standard (2 × 1) dimer pattern mentioned earlier: Although other dimer patterns are possible (e.g. p(2 × 2), c(2 × 4), etc.), for a given layer and dimer composition the *local* atomic geometry and bonding configuration is the same for all patterns. Energy differences due to changes in the dimer patterns are typically in the meV range [12], two orders of magnitude smaller than the energy differences shown in table 1, which are due to variations in the layer and dimer composition, within a given pattern.

Based on the energies of the different configurations a number of conclusions can be drawn. We address the physics of the problem in two steps: first we compare configurations of constant coverage to establish the lowest energy geometry, and then we compare monolayer versus bilayer coverage by means of the lowest energy structure in each case.

In the case of monolayer coverage it is seen that segregation of the Ga and As atoms into domains of each species is unfavorable. A mixed layer with Ga and As atoms uniformly distributed is lower in energy than separate domains by 0.8 eV/(2 × 1) SUC (see table 1, energies of (Ga)[Si](As) and (GaAs)[Si](AsGa) configurations). The local geometries for the two surfaces of the (Ga)[Si](As) configuration (Ga covered and As covered) are shown in figs. 2a and 2b. Both the Ga and the As atoms form dimers in the separate domains. The Si dimers of the clean surface are broken and the surface Si atoms are four-fold coordinated with two As or two Ga nearest neighbors

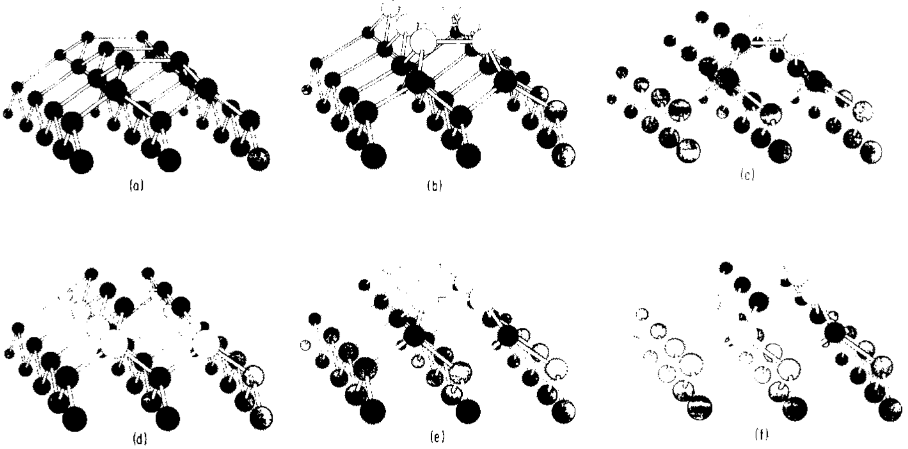

Fig. 2. Local bonding configurations for the different overlayer arrangements. (a) Ga dimers on one surface of (Ga)[Si](As). (b) As dimers on the other surface of (Ga)[Si](As). (c) Ga–As dimers on the surface of (GaAs)[Si](AsGa). (d) Ga dimers on the surface of (Ga)(As)[Si]. (e) As dimers on the surface of (As)(Ga)[Si]. (f) Ga–As dimers on the surface of (GaAs)(AsGa)[Si].

each. In the dimerized configuration As atoms tend to be p^3-bonded, whereas Ga atoms tend to be sp^2-bonded. Each surface Ga and As atom contributes three electrons to covalent bonding, with As atoms retaining a lone-pair of electrons. This arrangement renders the Ga and As surfaces electronically passive, since all the available electronic charge is in filled states. However, neither species is in ideal bonding configuration: the average bond-angle for the Ga atom is 111° rather than 120° as in ideal sp^2-bonding; the average bond-angle for the As atom is 104° compared to 97° of the three-fold coordinated elemental bulk As. In the mixed monolayer coverage ((GaAs)[Si](AsGa), fig. 2c), each surface atom is again three-fold coordinated providing an electronically passive surface. The average bond-angles are 117° for Ga and 99° for As, a considerable improvement over the bonding in the (Ga)[Si](As) configuration.

We consider next partial coverage of the Si surface by GaAs bilayer islands. In this case there are three possibilities as described earlier (cases (Ga)(As)[Si], (As)(Ga)[Si] and (GaAs)(AsGa)[Si]). These exhaust all possible growth sequences under stoichiometric conditions, that is initial deposition of As followed by Ga, the reverse sequence, and simultaneous deposition of Ga and As. However, under *equilibrium* conditions, the structure of the bilayer will be determined by the lowest-energy configuration, irrespective of the growth sequence. The lowest-energy configuration is the mixed bilayer (GaAs)(AsGa)[Si], which lies 0.7 eV/(2 × 1) SUC lower than the pure-bilayer with As exposed surface (As)(Ga)[Si] and 1.3 eV/(2 × 1) SUC lower than the pure-bilayer with Ga exposed surface (Ga)(As)[Si], as seen from table 1. The mixed layers are different from ordinary GaAs layers in the (100) direction, see figs. 2d–2f. The reason for the large energy gain upon mixing of the layers is attributable mainly to two reasons. First, the mixed layers eliminate the polar character of the pure GaAs bilayer. Thus, no dipole moment is created on the surface due to the spatial distribution of the atoms [13]. Second, the mixed layers allow for optimal relaxation (see fig. 2f) and hence rehybridization of the chemical bonds, which considerably lowers the energy of the system and passivates the surface. Thus, a stable structure of low energy and semiconducting character is obtained with mixed bilayers.

Finally, since mixed layers proved to be the lowest-energy configuration in either monolayer or bilayer coverage, we compare the energy of a mixed monolayer with that of a mixed bilayer plus exposed Si surface. This comparison will establish whether, at the coverage considered here, the available amount of Ga and As atoms wets the surface to the largest extent possible or forms islands of higher thickness on part of the surface area. The energy of the mixed monolayer is lower than that of the mixed bilayer by 1.5 eV/(2 × 1) SUC (see table 1, energies of (GaAs)[Si](AsGa) and (GaAs)(AsGa)[Si] configurations), which indicates that wetting is actually favored. This is not surprising since the amount of surface left uncovered during bilayer formation consists of Si dimers. These dimers are energetically less stable than the GaAs pairs in the mixed monolayer which have almost ideal geometric structure and are electronically passive. The Si dimers involve a weak second Si–Si bond which in the GaAs surface pair is replaced by a lone-pair filled As s-like state and an empty Ga p_z-like state, improving the stability of the surface layer and lowering the energy. These findings are consistent with some experimental evidence suggesting that at very low coverage island formation on the Si(100) surface is suppressed [6].

With the insight provided by the above discussion we can infer the merits of a particular mode of intermixing, namely exchanging entire Si layers with the Ga and As overlayers. This layer-exchange has been considered in connection with GaAs growth on Si along the (111) direction [8] and was shown to be advantageous in that case. More complicated intermixing patterns can occur but they cannot be analyzed on the basis of our present results. The possible configurations which we considered are shown in fig. 3. In configurations a–c of fig. 3 a Si layer has been displaced from the substrate all the way to the surface, embedding the Ga and As layers, whereas in d–f the Si layer has been displaced only partially toward the surface and is contained between layers of Ga and As. Configurations a–c of fig. 3 are structurally similar to configurations c–e of fig. 1 respectively, except for the termination of the surface by a Si layer. It was shown above that exposed layers of Si are energetically unfavorable in the presence of Ga and As, and consequently the structures a–c of fig. 3 must be higher in energy than the corresponding structures c–e of fig. 1. Thus, embedding the Ga and As layers under Si is not energetically feasible. In the

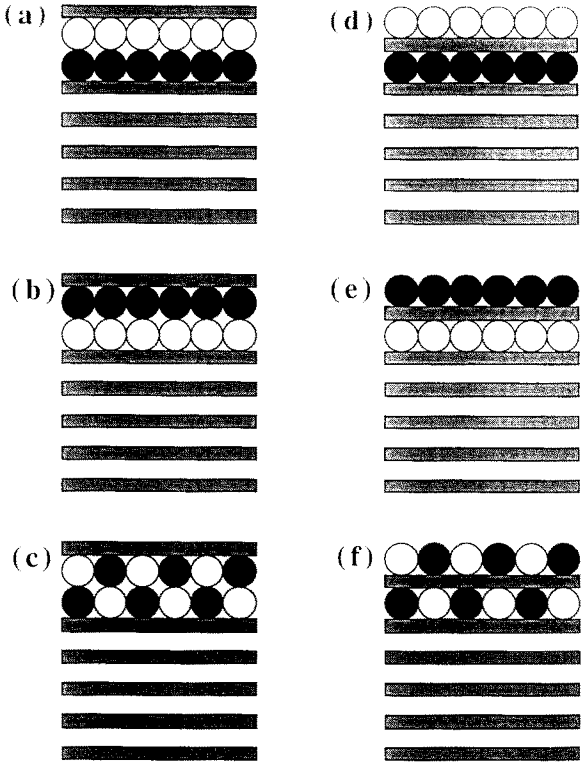

Fig. 3. Possible configurations obtained by intermixing entire layers of Ga and As with the Si substrate. Symbols are the same as in fig. 1.

case of partial displacement of the top Si layer, the dipole fields associated with configurations d and e of fig. 3 will be higher than in the corresponding configurations c and d of fig. 1, due to the larger spatial separation of the pure Ga and As layers. The larger dipole fields make the configurations d and e of fig. 3 energetically unfavorable compared to those already considered in fig. 1. The only structure which might be energetically comparable to the lowest energy bilayer coverage fig. 1e, is that shown in fig. 3f, with two mixed GaAs layers on either side of the displaced Si layer. This structure, however, has an exposed layer which is identical to the surface layer of (GaAs)(AsGa)[Si] and need not be considered separately as far as growth of bulk GaAs is concerned.

The following conclusions emerge from the discussion of our results:

(1) When the available amount of Ga and As atoms is less than a full monolayer, under equilibrium *stoichiometric* conditions, it wets the Si surface and formation of thicker GaAs islands is inhibited.

(2) At monolayer coverage the two species of atoms do not separate into pure Ga or pure As domains, but are locally mixed on the Si surface.

(3) When GaAs is available in stoichiometric amounts sufficient for bilayer coverage, the equilibrium structure of the bilayer is not a succession of pure Ga and As layers as in bulk GaAs along the (100) direction. Rather, chemical and rehybridization reactions are energetically very favorable, leading to a passivating mixed overlayer configuration. Once this mixed layer is formed it could persist in a disordered phase even under non-stoichiometric conditions because of its considerable stability. The mixed layer resembles locally a wurtzite structure under large uniaxial stress. A structural barrier (requiring creation of extensive defects at large energy cost) is encountered in reverting from the mixed layer structure to the zincblende structure of bulk GaAs.

This conclusion has important implications: since the mixed layers have very different structure from the thermodynamically stable zincblende phase of bulk GaAs, the growth of mixed layers on terraces will be undermined by any surface topology which could lead to nucleation of zincblende GaAs. Thus, nucleation centers of zincblende GaAs at specific surface topologies will *prevail* in the growth process. Experimentally it is possible to grow bulk GaAs on vicinal Si(100) surfaces [1], that is, surfaces with large density of steps. It is likely that surface steps are precisely the topology that nucleates zincblende GaAs, and growth occurs directly on the steps. Indeed, experimental evidence appears to support this view [14,15]. A possible microscopic mechanism for the *initiation* of this kind of growth, which involves formation of zincblende GaAs seeds at double-layer steps on the Si surface, was recently proposed by Kaxiras et al. [3]. These seeds can nucleate three-dimensional pyramidal islands of GaAs. Compelling evidence of why growth will *continue* in a three-dimensional island-like mode, once the seeds have grown to macroscopic size, was provided by the theoretical analysis of Northrup [4]. Ideally, the seeds should be stoichiometric for bulk-like growth to proceed. However, initial deposition conditions (e.g. excess Ga versus As) may affect the composition of the seeds. A non-stoichiometric initial composition may actually prove beneficial in certain cases (e.g. predominantly single-stepped surfaces, relief of strain, etc.). These more complicated issues are currently under investigation.

We should like to thank G.W. Turner for introducing us to the subject and helpful discussions. This work was supported in part by Joint Services Electronics Program, Contract No. DAAL 03-86-K-0002.

References

[1] See e.g. articles in: Heteroepitaxy in Silicon, Eds. J.C.C. Fan and J.M. Poate, Mater. Res. Soc. Symp. Proc. 67 (1986);
Heteroepitaxy in Silicon II, Eds. J.C.C. Fan and J.M. Phillips and B.-Y. Tsaur, Mater. Res. Soc. Symp. Proc. 91 (1987).
[2] For a review see: H. Kroemer, J. Crystal Growth 81 (1987) 193.

[3] E. Kaxiras, O.L. Alerhand, J.D. Joannopoulos and G.W. Turner, Phys. Rev. Letters 62 (1989) 2484.
[4] J.E. Northrup, Phys. Rev. Letters 62 (1989) 2487.
[5] R.I.G. Uhrberg, R.D. Bringans, R.Z. Bachrach and J.E. Northrup, Phys. Rev. Letters 56 (1986) 520.
[6] R.D. Bringans, M.A. Olmstead, R.I.G. Uhrberg and R.Z. Bachrach, Phys. Rev. B 36 (1987) 9569.
[7] J. Nogami, S. Park and C.F. Quate, Appl. Phys. Letters 53 (1988)2086;
B. Borguignon, K.L. Carleton, S.R. Leone, Surface Sci. 20 (1988) 455.
[8] J.E. Northrup, R.D. Bringans, R.I. Uhrberg, M.A. Olmstead and R.Z. Bachrach, Phys. Rev. Letters 61 (1988) 2957.
[9] P. Hohenberg and W. Kohn, Phys. Rev. 136 (1964) B864;
W. Kohn and C.J. Sham, Phys. Rev. 140 (1965) A1133;
We used the exchange–correlation energy functional of: D.M. Ceperley and B.J. Alder, Phys. Rev. Letters 45 (1980) 566.
[10] G.B. Bachelet, D.R. Hamann and M. Schlüter, Phys. Rev. B 26 (1982) 4199.
[11] For the origin of vacuum fields in such calculations see: E. Kaxiras, Y. Bar-Yam, J.D. Joannopoulos and K.C. Pandey, Phys. Rev. B 35 (1987) 9625.
[12] J. Ihm, D.H. Lee, J.D. Joannopoulos and J.J. Xiong, Phys. Rev. Letters 51 (1983) 1872.
[13] For a discussion of interface dipole moment due to the spatial distribution of atoms see: W.A. Harrison, E.A. Kraut, J.R. Waldrop and R.W. Grant, Phys. Rev. B 18 (1978) 4402.
[14] D.K. Biegelsen, F.A. Ponce, A.J. Smith and J.C. Tramontana, J. Appl. Phys. 61 (1987) 1856.
[15] R. Hull and A. Fischer-Colbrie, Appl. Phys. Letters 50 (1987) 851.