

First-Principles Calculations for Insulators at Constant Polarization

Oswaldo Diéguez and David Vanderbilt

Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey 08854-8019, USA
(Received 29 November 2005; published 7 February 2006)

We develop an exact formalism for performing first-principles calculations for insulators at fixed electric polarization. As shown by Sai, Rabe, and Vanderbilt (SRV) [Phys. Rev. B **66**, 104108 (2002)], who designed an approximate method to tackle the same problem, such an approach allows one to map out the energy landscape as a function of polarization, providing a powerful tool for the theoretical investigation of polar materials. We apply our method to a system in which the ionic contribution to the polarization dominates (a broken-inversion-symmetry perovskite), one in which this is not the case (a III–V semiconductor), and one in which an additional degree of freedom plays an important role (a ferroelectric phase of KNO_3). We find that while the SRV method gives rather accurate results in the first case, the present approach provides important improvements to the physical description in the latter cases.

DOI: [10.1103/PhysRevLett.96.056401](https://doi.org/10.1103/PhysRevLett.96.056401)

PACS numbers: 71.15.–m, 77.22.–d, 77.80.–e

In 1993 King-Smith and Vanderbilt [1] introduced a theory for computing the electric polarization of an infinite solid, showing for the first time that this property was, indeed, a well-defined quantity for an insulating material. Their theory of the bulk polarization (TBP) is based on computing the electronic contribution to the polarization as a Berry phase of the valence-band Bloch wave functions transported across the Brillouin zone. A practical numerical scheme to compute it was given in the same paper [1], and it is now widely used in first-principles calculations. Later, Souza, Íñiguez, and Vanderbilt (SIV) [2] used the TBP as the cornerstone for a method to calculate the exact ground state of an insulator in the presence of an electric field by minimizing an electric enthalpy functional expressed in terms of occupied Bloch-like states on a uniform grid of points in reciprocal space. It is therefore possible to compute from first principles many interesting properties of materials related to their behavior under electric fields.

In this Letter we propose a method to do first-principles calculations not at constant electric field, but at constant electric polarization. In this way, it would become possible to map the energy E of an insulator as a function of its polarization \mathbf{P} . There are several reasons why this is useful. First, it allows for an exhaustive search for competing local minima and for saddle points in a system with a complicated energy surface. These features would otherwise be hidden in a standard optimization procedure, which would be likely to find only a single minimum of the energy. Second, knowing $E(\mathbf{P})$, we can calculate various properties related to derivatives of E with respect to \mathbf{P} , such as the linear and nonlinear dielectric susceptibilities. While methods of density-functional perturbation theory [3] can also be used to access these susceptibilities, our approach is advantageous in that it does not require much special-purpose programming and yields nonlinear susceptibilities with very little extra effort beyond that needed to obtain the linear ones. And third, Landau-Devonshire theories have historically provided an important avenue to the under-

standing of ferroelectric materials, based on an expansion of the energy in powers of polarization. The ability to compute $E(\mathbf{P})$ may open the door to the first-principles derivation of Landau-Devonshire descriptions, as opposed to the usual empirical formulations, for a wide range of ferroelectric materials.

The problem of performing simulations at constant polarization in the context of first-principles calculations has been addressed previously. In particular, Sai, Rabe, and Vanderbilt (SRV) [4] developed an *approximate* method to do this and applied it to obtain $E(\mathbf{P})$ for several perovskite systems of interest. A similar approach was used by Fu and Bellaïche to study the electromechanical response of solids under finite electric fields [5]. Our method relies on many of the ideas of the SRV one, but it is instead an *exact* method because it incorporates the SIV theory of finite electric fields [2].

Our goal is to find the minimum energy E of a system of atoms for which the electric polarization takes some target value \mathbf{P}_t . The system has both ionic and electronic degrees of freedom, and we use vectors \mathbf{X} and Ψ to represent them. Vector \mathbf{X} has $3N + 6$ components given by the Cartesian coordinates of the N atoms in the simulation cell and the 6 components of the strain tensor. On the other hand, Ψ contains the one-electron wave functions of the system. We are therefore facing a standard constrained optimization problem that can be solved by introducing a Lagrange multiplier λ and searching for the minimum of $E + \lambda \cdot (\mathbf{P} - \mathbf{P}_t)$ or, discarding the constant term, the minimum of $E + \lambda \cdot \mathbf{P}$. Defining $\lambda = -\Omega \mathcal{E}$, where Ω is the volume of the unit cell of the material under consideration, we have

$$\begin{aligned} \min_{\substack{\mathbf{X}, \Psi \\ \lambda = -\Omega \mathcal{E}}} \{E(\mathbf{X}, \Psi) + \lambda \cdot \mathbf{P}(\mathbf{X}, \Psi)\} &= \min_{\substack{\mathbf{X} \\ \mathcal{E} = \mathbf{P} = \mathbf{P}_t}} \{ \min_{\Psi} \{E(\mathbf{X}, \Psi) \\ &\quad - \Omega \mathcal{E} \cdot \mathbf{P}(\mathbf{X}, \Psi)\} \} \\ &= \min_{\substack{\mathbf{X} \\ \mathcal{E} = \mathbf{P} = \mathbf{P}_t}} H(\mathbf{X}, \mathcal{E}), \end{aligned} \quad (1)$$

where $\boldsymbol{\lambda} \rightarrow \mathbf{P} = \mathbf{P}_t$ indicates a minimization over the values of $\boldsymbol{\lambda}$ that impose the constraint $\mathbf{P} = \mathbf{P}_t$. The function $H(\mathbf{X}, \boldsymbol{\mathcal{E}})$ is the electric enthalpy that is minimized in the SIV method to find the electronic ground state of an insulator in the presence of an electric field $\boldsymbol{\mathcal{E}}$ [2].

Therefore, the problem of finding the minimum energy of a system that is constrained to have polarization \mathbf{P} is equivalent to finding the ground state of a system under an electric field $\boldsymbol{\mathcal{E}}$ that imposes polarization \mathbf{P} . However, the apparently straightforward approach of doing calculations at different values of the electric field to find $E(\mathbf{P})$ parametrically will not work in general. Figure 1 illustrates why this is so for a typical double-well potential characteristic of ferroelectric perovskites. In this case, several values of the polarization correspond to the *same* value of the electric field, since $\boldsymbol{\mathcal{E}} = \Omega^{-1}(dE/d\mathbf{P})$, and an optimization using the SIV theory will most likely find the point that has the lowest energy of the three, i.e., the global minimum of $H(P)$. With extra care it might be possible to find the secondary local minimum of $H(P)$, but not the third solution, which is a maximum (or, in 3D, a saddle point) of $H(P)$. It follows from this line of reasoning that one cannot map the region of $E(P)$ that has negative curvature, and it may also prove difficult to map nearby regions of positive curvature.

We now describe how to perform a minimization of the enthalpy H over the ionic degrees of freedom \mathbf{X} and the electric fields $\boldsymbol{\mathcal{E}}$ that produce a desired polarization \mathbf{P}_t in a way that parallels the presentation in Sec. III.A of Ref. [4]. From now on, we assume that the optimization is done while keeping the cell vectors fixed. (It is possible to remove this constraint, but it is necessary to take into account some technical subtleties related to the way in which the stress tensor is computed in the presence of an electric field. The details, together with examples, will be presented elsewhere [6].) We begin with a trial guess $(\mathbf{X}_0, \boldsymbol{\mathcal{E}}_0)$, and we expand $H(\mathbf{X}, \boldsymbol{\mathcal{E}})$ and $\mathbf{P}(\mathbf{X}, \boldsymbol{\mathcal{E}})$ as low-order Taylor series

$$H = H_0 - \mathbf{F}\delta\mathbf{X} + \frac{1}{2}\delta\mathbf{X}\mathbf{K}\delta\mathbf{X} - \Omega\mathbf{P}\delta\boldsymbol{\mathcal{E}}, \quad (2)$$

$$\mathbf{P} = \mathbf{P}_0 + \frac{1}{\Omega}\mathbf{Z}\delta\mathbf{X} + \frac{1}{4\pi}\boldsymbol{\chi}\delta\boldsymbol{\mathcal{E}}, \quad (3)$$

where $H_0 = H(\mathbf{X}_0, \boldsymbol{\mathcal{E}}_0)$ and $\mathbf{P}_0 = \mathbf{P}(\mathbf{X}_0, \boldsymbol{\mathcal{E}}_0)$ (atomic units

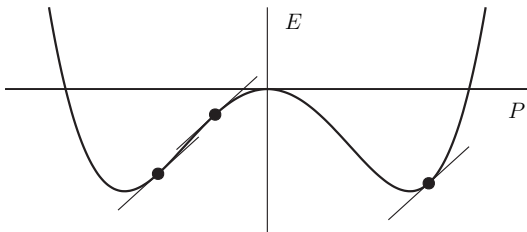


FIG. 1. Sketch of energy as a function of polarization for a double-well potential. The value of the electric field needed to impose the polarization is the same at the three indicated points, as their slopes are the same.

are used throughout unless stated otherwise). The $F_i = -(\partial H/\partial X_i)$ are the $3N$ components of the forces on the atoms at finite electric field [2], the $K_{ij} = (\partial^2 H/\partial X_i \partial X_j)$ are the $3N \times 3N$ force-constant matrix elements, the $Z_{i\alpha} = \Omega(\partial P_\alpha/\partial X_i)$ are the $3N \times 3$ Born effective charges, and the $\chi_{\alpha\beta} = 4\pi(\partial P_\beta/\partial \epsilon_\alpha)$ are the 3×3 elements of the dielectric susceptibility tensor. We can then predict that $\partial H/\partial X_i = 0$ and $\mathbf{P} = \mathbf{P}_t$ at an $\mathbf{X} = \mathbf{X}_0 + \delta\mathbf{X}$ and an $\boldsymbol{\mathcal{E}} = \boldsymbol{\mathcal{E}}_0 + \delta\boldsymbol{\mathcal{E}}$ given by

$$\begin{pmatrix} \mathbf{K} & -\mathbf{Z} \\ -\mathbf{Z} & -\frac{\Omega}{4\pi}\boldsymbol{\chi} \end{pmatrix} \begin{pmatrix} \delta\mathbf{X} \\ \delta\boldsymbol{\mathcal{E}} \end{pmatrix} = \begin{pmatrix} \mathbf{F} \\ \Omega\Delta\mathbf{P}_0 \end{pmatrix}, \quad (4)$$

where $\Delta\mathbf{P}_0 = \mathbf{P}_0 - \mathbf{P}_t$. This system of linear equations can be solved iteratively, refining $\delta\mathbf{X}$ and $\delta\boldsymbol{\mathcal{E}}$ until \mathbf{F} and $\Delta\mathbf{P}_0$ both vanish. As mentioned before, \mathbf{F} can be computed exactly in the presence of an electric field according to the SIV prescription [2]. The guiding tensors \mathbf{K} , \mathbf{Z} , and $\boldsymbol{\chi}$ do not need to be computed exactly; replacing these by approximate versions only results in a slower convergence to the correct solution, without shifting the solution itself. In particular, we normally find it sufficient to compute \mathbf{K} , \mathbf{Z} , and $\boldsymbol{\chi}$ at zero electric field.

The scheme described here has been implemented to work with the ABINIT [7] density-functional theory (DFT) [8] code. Instead of performing a direct iterative solution of Eq. (4), we have used a nested-loop algorithm for the sake of robustness. Starting with some guess for \mathbf{X} and $\boldsymbol{\mathcal{E}}$, we keep the atoms fixed and vary the field in the internal loop until the polarization is the target one, a problem that is well behaved. Once this is achieved, we solve Eq. (4) to get new values of \mathbf{X} and $\boldsymbol{\mathcal{E}}$, and iterate until convergence is achieved.

As a first example of how our method works, we apply it to a soft-mode system studied by Sai, Rabe, and Vanderbilt [4]. They have shown that breaking the inversion symmetry in perovskites by modulating their composition in a cyclic sequence of layers produces some interesting features in their energy landscape, apart from making them promising candidates for new materials with useful piezoelectric and dielectric properties. They presented results for $\text{Ba}(\text{Ti}-\delta, \text{Ti}, \text{Ti} + \delta)\text{O}_3$, where the two species that alternate with Ti on the *B* site are virtual atoms that differ from Ti in that their nucleus has a defect or excess of charge equal to δ . For different values of δ they find the $E(P)$ curves that we reproduce in Fig. 2 (dashed lines). As δ increases, the double-well potential becomes asymmetric (a feature not seen in normal perovskites) until one of the minima eventually disappears at around $\delta = 0.4$.

We have repeated their study using the same DFT methodology, but applying our exact method to compute $E(P)$ instead of their approximate one. We used the ABINIT [7] code to do our calculations, with the Ceperley-Alder form [9] of the local-density approximation (LDA) [10] to obtain the exchange-correlation term in DFT, a plane-wave cutoff of 35 Ha to define the basis set, a $4 \times 4 \times 4$ Monkhorst-Pack [11] grid for computations in reciprocal space, and Troullier-Martins [12] norm-conserving pseu-

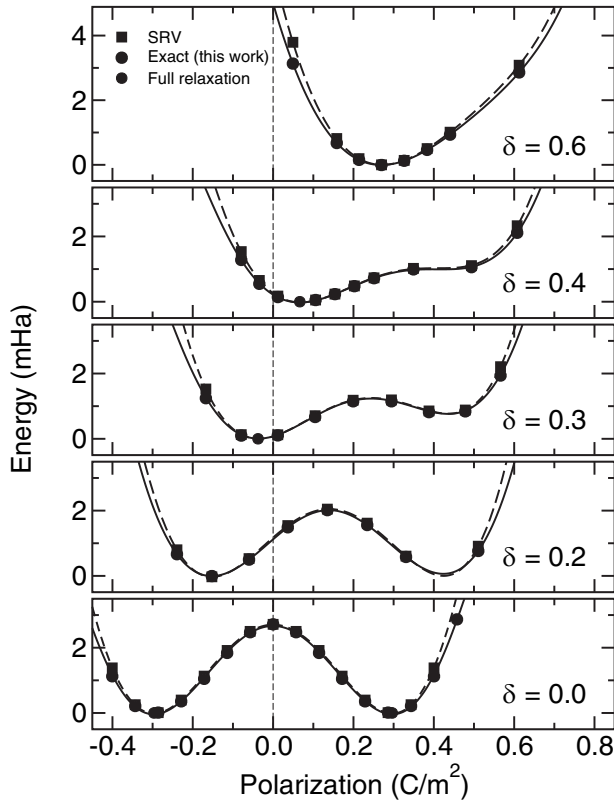


FIG. 2. Energy versus polarization along the c axis for the broken-inversion-symmetry perovskite described in the text, for different values of δ . The solid (dashed) line is a polynomial fit of the values obtained using our (the SRV) method.

dopotentials [13] to model the ion-electron interactions. The tetragonal unit cell vectors were kept fixed, with cell parameters $a = 7.547$ a.u. and $c/a = 3.036$. Here, our results (solid lines in Fig. 2) are very similar to the SRV ones; the curves differ noticeably only for large values of the polarization. This is not surprising, since the dielectric behavior is dominated by the ionic response for soft-mode systems like this one.

On the other hand, the electronic response is found to be profoundly important to the dielectric behavior in the case of a III-V semiconductor like AlAs, as can be seen in Fig. 3. Here we employed ABINIT [7] using the LDA [9], a plane-wave cutoff of 9 Ha, a $6 \times 6 \times 6$ reciprocal space grid [11], Troullier-Martins [12] pseudopotentials [14], and a theoretically optimized lattice constant of $a = 10.62$ a.u. As can be seen in Fig. 3, there is a drastic difference between the E versus P curves computed using the SRV method (ionic response only) and the one found with our new method (ionic and electronic response). In order to quantify this difference, the dielectric constant in both cases was computed from the curvature at the minimum of each function as $\epsilon = 1 + 4\pi/(d^2E/dP^2)$. When computed with the SRV method, $\epsilon = 3.0$, while, when we use our new method, $\epsilon = 10.3$. The experimental result is $\epsilon = 10.1$ [2].

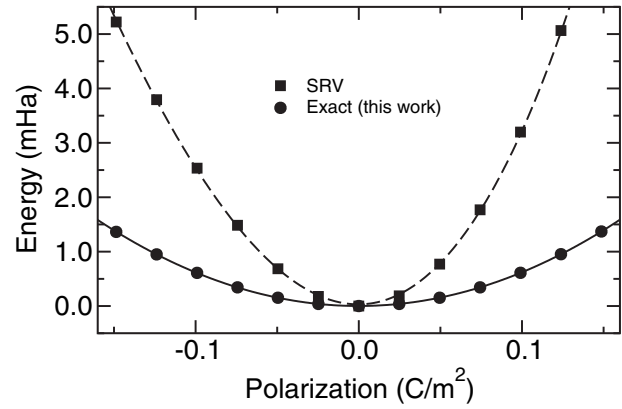


FIG. 3. Energy versus magnitude of polarization in the line joining two nearest Al and As neighbors. The solid (dashed) line is a polynomial fit of the values obtained using our (the SRV) method.

Our last example involves potassium nitrate, which has an interesting phase diagram that includes a reentrant ferroelectric phase (phase III, $R3m$) and that has been proposed as a promising material to be used in random-access memory devices [15]. The unit cell of phase III KNO_3 is a five-atom rhombohedral cell that results in alternating planes of K atoms and NO_3 groups arranged in such a way that the K atom is not equidistant from the NO_3 groups above and below it, but instead is slightly displaced in the vertical direction, giving rise to a polar structure. Figures 4(a) and 4(b) show the conventional 15-atom hexagonal cell that is most convenient for visualization. The calculations are performed using ABINIT [7], the LDA [9], a plane-wave cutoff of 30 Ha, a reciprocal space grid with 6 inequivalent points, and Troullier-Martins [12] pseudopotentials [16]. The theoretical structural optimization of bulk ferroelectric KNO_3 gives hexagonal lattice parameters of $a = 9.68$ a.u. and

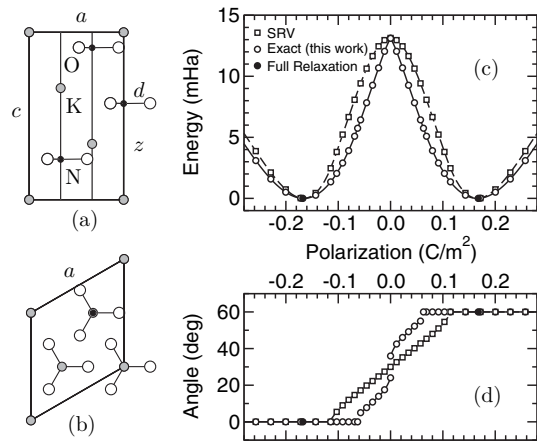


FIG. 4. (a) Side view of the ferroelectric phase of KNO_3 (phase III), with the c hexagonal lattice axis running vertically. (b) Top view of the same structure. (c) Energy versus polarization $\mathbf{P} = (0, 0, P)$. (d) Angle of rotation of the NO_3 group as a function of polarization P .

$c = 14.86$ a.u., to be compared with the ones found experimentally at 91°C , $a = 10.37$ a.u. and $c = 17.30$ a.u. [17]. As for the internal degrees of freedom, the distance between a K atom and the N atom above it is found to be $z = 0.57c$ (experimentally, $z = 0.59c$ [17]), while the N-O distance is $d = 2.35$ a.u. (experimentally, $d = 2.42$ a.u. [17]). The computation of the polarization gives 0.16 C/m^2 , higher than the experimental values of $0.08\text{--}0.11\text{ C/m}^2$ given in Ref. [18]. (This disagreement is related to the reduction in volume found theoretically; when we performed an optimization of the atomic positions using the experimental lattice constants, we found a polarization value of 0.075 C/m^2 .)

Figure 4(c) shows the energy of the ferroelectric phase of KNO_3 as a function of the magnitude of the polarization along the c axis. As in the case of perovskites, we can see a double-well potential, but now the change of polarization given by the movement of the NO_3 atoms parallel to the c axis is accompanied by a rotation of these groups in their plane, as shown in Fig. 4(d). When doing SRV calculations, both the derivative of the energy and the angle of rotation are continuous functions of the polarization, and the NO_3 group rotates from 0° to 60° as the polarization goes from negative to positive, with the paraelectric configuration that corresponds to the maximum of $E(P)$ being the highly symmetrical one in which $z = 0.5$ and the rotation angle is 30° . However, when the new exact method is used, this behavior changes noticeably: the $E(P)$ curve has a discontinuous derivative at $P = 0$, and the angle-rotation curve is no longer continuous.

Although it may seem puzzling that the response of the system changes so significantly just by including the electronic response, this behavior can be understood on the basis of a simple model in which the energy is a function not only of polarization P , but also of the rotation angle θ . To understand the qualitative features, it is sufficient to consider a low-order expansion $E(P, \theta) = \cos 6\theta + \alpha \cos 12\theta + \beta P \cos 3\theta + P^2$. We assume that $\beta > 2\sqrt{2}$, in which case $E(P)$ has two minima. Then, depending on the value of α , the behavior can be continuous ($\alpha < 1/4$) or discontinuous ($\alpha > 1/4$) in P . Here, it appears that the system was already close to this critical value of α , and inclusion of electronic effects happened to shift the system from the continuous to the discontinuous regime. A more detailed discussion of this material, and its modeling along the lines sketched above, will be given elsewhere. In any case, the ability of our approach to describe the complexity of the structural behavior of KNO_3 under polarization reversal provides an excellent example of the power of the method.

To summarize, we have presented a method for finding the most stable structural configuration of an insulating crystal when its electric polarization is constrained to take on a given value. Our method builds upon an earlier

approach [4] that makes the approximation of including only the lattice response to applied electric fields. By using the recently developed theory of finite electric fields [2] to include also the electronic response of the system, we have developed a new approach that is, instead, exact. We have illustrated the method by obtaining $E(P)$ curves for three rather different kinds of insulating materials, and have illustrated how the method is capable of describing the complexity of the nonlinear structural and dielectric response.

O.D. acknowledges useful discussions with José-Luis Mozos, and the instruction and encouragement received from him during visits to the ICMAB in Barcelona. The authors thank J. Scott for suggesting the application to KNO_3 . This work was supported by ONR Grants No. N0014-05-1-0054, No. N00014-00-1-0261, and No. N00014-01-1-0365.

-
- [1] R.D. King-Smith and D. Vanderbilt, Phys. Rev. B **47**, R1651 (1993).
 - [2] I. Souza, J. Íñiguez, and D. Vanderbilt, Phys. Rev. Lett. **89**, 117602 (2002).
 - [3] S. de Gironcoli, S. Baroni, and R. Resta, Phys. Rev. Lett. **62**, 2853 (1989); P. Giannozzi, S. de Gironcoli, P. Pavone, and S. Baroni, Phys. Rev. B **43**, 7231 (1991); X. Gonze, Phys. Rev. B **55**, 10337 (1997).
 - [4] N. Sai, K.M. Rabe, and D. Vanderbilt, Phys. Rev. B **66**, 104108 (2002).
 - [5] H. Fu and L. Bellaiche, Phys. Rev. Lett. **91**, 057601 (2003).
 - [6] X. Wu, O. Diéguez, K.M. Rabe, and D. Vanderbilt (to be published).
 - [7] The ABINIT code is a common project of the Université Catholique de Louvain, Corning Incorporated, and other contributors (www.abinit.org); see X. Gonze *et al.*, Comput. Mater. Sci. **25**, 478 (2002).
 - [8] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
 - [9] D.M. Ceperley and B.J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
 - [10] W. Kohn and L.J. Sham, Phys. Rev. **140**, A1133 (1965).
 - [11] H.J. Monkhorst and J.D. Pack, Phys. Rev. B **13**, 5188 (1976).
 - [12] N. Troullier and J.L. Martins, Phys. Rev. B **43**, 1993 (1991).
 - [13] More details about the pseudopotentials can be found in Ref. [4].
 - [14] The actual pseudopotentials files used are available at www.abinit.org/Psps/?text=../Psps/LDA_TM/lda.
 - [15] J.F. Scott, M.S. Zhang, R.B. Godfrey, C. Araujo, and L. McMillan, Phys. Rev. B **35**, 4044 (1987).
 - [16] The actual pseudopotentials files used are available at www.abinit.org/Psps/?text=../Psps/LDA_FHI/fhi.
 - [17] J.K. Nimmo and B.W. Lucas, Acta Crystallogr. B **B32**, 1968 (1976).
 - [18] S. Sawada, S. Nomura, and Y. Asao, J. Phys. Soc. Jpn. **16**, 2486 (1961).