THE EUROPEAN
PHYSICAL JOURNAL B

Regular Article

Polarization in Kohn-Sham density-functional theory[*]

Raffaele Resta[1,2,a]

[1] CNR-IOM DEMOCRITOS, Istituto Officina dei Materiali, Consiglio Nazionale delle Ricerche,
Strada Costiera 11, 34151 Trieste, Italy
[2] Donostia International Physics Center, 20018 San Sebastián, Spain

Abstract. The modern theory of polarization, based on a Berry phase, is currently implemented in most first-principle electronic structure codes. Many KS-DFT calculations have addressed various phenomena (ferroelectricity, piezoelectricity, lattice dynamics, infrared spectra of liquid and amorphous systems) in several materials. Notwithstanding, the KS polarization *does not* coincide with the exact one, even when ideally implemented with the exact KS crystalline potential. This is at odds with the fact that the KS electrical dipole of a bounded crystallite coincides by definition with the exact one: we analyze this issue from several viewpoints. According to the modern theory, the polarization of a centrosymmetric crystal does not vanish in general; we show that the polarization of a centrosymmetric quasi-1d systems (stereoregular linear polimer) is a topological invariant: ergo in this case the KS polarization coincides with the exact one.

1 Introduction

The modern theory of polarization, based on a Berry phase, was developed in the early 1990s and is nowadays a mature topic; it is implemented in most of the available first-principle electronic structure codes in the KS-DFT framework. Several reviews and pedagogical presentations appeared over the years [1–5].

The original formulation addresses noninteracting electrons, and therefore provides the polarization of the noninteracting Kohn-Sham (KS) system. In the early papers [1,6–8] it was argued that this coincides with the exact many-body polarization, on the basis that the density $n(\mathbf{r})$ is given exactly. Such reasoning was indeed rather naive: bounded samples (where the KS orbitals are square-integrable) and unbounded crystalline samples (where the KS orbitals are Bloch functions) behave very differently. In the former case $n(\mathbf{r})$ uniquely determines the electric dipole (for neutral samples), while in the latter case $n(\mathbf{r})$ *does not* determine, not even in principle, the macroscopic polarization. It is then pretty clear that the KS polarization and the genuine many-body polarization are different when periodic boundary conditions are adopted, while they coincide by definition when e.g. a large bounded crystallite is addressed. In this work we investigate several aspects of this apparently paradoxical situation.

Because of the three good reasons detailed next – (i) to (iii) – we limit ourselves to quasi-1d systems: stereoregular insulating polymeric chains, either bounded or unbounded.

(i) The modern theory addresses polarization in zero macroscopic electric field: in general, this does not correspond to the situation in the bulk of a 3d macroscopic bounded sample in zero external field. If the material is polarized, a shape-dependent depolarization field is present in its bulk. In general, such field is noniform; it is constant inside the sample only if its shape is ellipsoidal [9]. The problem does not exist for quasi-1d systems: in a polarized polymeric chain of length L, in zero external field, the field in the middle of the chain goes to zero like L^{-2}.

(ii) A 3d bounded sample cut from an insulating material may have metallic surfaces, in which case the dipole may vanish (Faraday-cage effect). Even this problem does not exist for quasi-1d insulating systems.

(iii) In 1d the polarization (dipole per unit length) has the dimensions of a pure charge. In any dimension the Berry-phase polarization is defined – as shown below – modulo a "quantum"; it is only in 1d that the polarization quantum is a fundamental constant.

2 Many-body polarization

2.1 Bounded sample

Consider first the dipole $\mathbf{d}$ of a finite system of electrons with density $n(\mathbf{r})$ and classical nuclei of charge eZ_ℓ at sites $\mathbf{R}_\ell$:

$$\mathbf{d} = -e \int d\mathbf{r}\, \mathbf{r}\, n(\mathbf{r}) + e \sum_\ell Z_\ell \mathbf{R}_\ell. \qquad (1)$$

If $|\Psi_0\rangle$ is the many-body ground state, $\mathbf{d}$ can also be written as

$$\mathbf{d} = e \langle\Psi_0| \left(\sum_\ell Z_\ell \mathbf{R}_\ell - \sum_j \mathbf{r}_j \right) |\Psi_0\rangle. \qquad (2)$$

For charge-neutral systems, the dipole is invariant by translation of the coordinate origin. We also notice that equation (2) applies as it stands even when the quantum nature of the nuclei is considered: it is enough to replace the electronic wavefunction $|\Psi_0\rangle$ with its complete (electronic and nuclear) counterpart.

In the following we focus on the dipole in the x direction, i.e.

$$d_x = e \langle\Psi_0| \left(\sum_\ell Z_\ell X_\ell - \sum_j x_j \right) |\Psi_0\rangle. \qquad (3)$$

2.2 Unbounded sample within periodic boundary conditions

Here we suppose that our quasi-1d system is unbounded in the x direction. We adopt Born-von-Kàrmàn periodic boundary conditions (PBCs) for a system of N electrons over a period L; the thermodynamic limit $N \to \infty$, $L \to \infty$, $N/L = $ const is understood. If $\mathbf{e}_1$ is the unit vector in the x direction, then for any j

$$\langle\mathbf{r}_1, \mathbf{r}_2 \ldots, \mathbf{r}_j, \ldots \mathbf{r}_N|\Psi_0\rangle = \langle\mathbf{r}_1, \mathbf{r}_2 \ldots, \mathbf{r}_j + L\mathbf{e}_1, \ldots \mathbf{r}_N|\Psi_0\rangle, \qquad (4)$$

where irrelevant spin variables have been neglected. The nuclear positions $\mathbf{R}_\ell$ have periodic replicas as well.

The key observation is that within PBCs the x_j coordinates are equivalent to the angles $\theta_j = 2\pi x_j/L$, and therefore the expectation value of the coordinates as in equation (3) does not make any sense: when the x-integration is restricted to a segment of length L, the result depends on the choice of the origin. The expectation value must be expressed in terms of the phase angles themselves. The solution was found in reference [10]:

$$d_x = e \frac{L}{2\pi} \text{Im} \ln \langle\Psi_0|\mathrm{e}^{i\frac{2\pi}{L}(\sum_\ell Z_\ell X_\ell - \sum_j x_j)}|\Psi_0\rangle, \qquad (5)$$

whence macroscopic polarization d_x/L is

$$P = e \frac{\gamma}{2\pi}, \quad \gamma = \text{Im} \ln \langle\Psi_0|\mathrm{e}^{i\frac{2\pi}{L}(\sum_\ell Z_\ell X_\ell - \sum_j x_j)}|\Psi_0\rangle. \qquad (6)$$

The phase γ is a Berry phase in disguise. In the present formulation it comprises the contribution of both electrons and nuclei; alternatively, it can be split into an electronic and a classical phase (see Appendix A). Each of the two phases depends on the choice of the origin on the coordinate axis; the sum of the two, owing to charge neutrality, is translationally invariant.

Notice that we have not used crystalline periodicity so far, and the general expression of equation (6) applies as such to correlated and disordered systems. We even argue

that equation (6) should provide the polarization when the quantum nature of the nuclei is considered, in full analogy to equation (2).

The expectation value in equation (6) is a peculiar one, and its thermodynamic limit deserves a comment. The operator therein *cannot* be expanded in powers of $1/L$, because such expansion is incompatible with PBCs; furthermore, when L is increased, even the number of electrons in $|\Psi_0\rangle$ increases.

Suppose instead for a moment that we implement equation (5) with the ground state $|\Psi_0\rangle$ of a *bounded* sample, at fixed N. Then the $1/L$ expansion is no longer forbidden, and equation (5) coincides to leading order with equation (3).

2.3 Polarization vs. current

We arrived at our main expression via a kind of plausibility argument. In order to prove that equation (5) yields the macroscopic polarization of the sample we resort to a time-dependent viewpoint, in the adiabatic limit. Macroscopic electrostatics requires that $\partial \mathbf{P}/\partial t = \mathbf{j}(t)$, where $\mathbf{j}(t)$ is the macroscopic current density. In our quasi-1d framework we write this as

$$I = \frac{dP}{dt} = \frac{e}{2\pi}\dot{\gamma}(t), \qquad (7)$$

where I is the macroscopic current flowing along the chain while the Hamiltonian is adiabatically varied, and the expectation value in $\gamma(t)$, equation (6), is evaluating using the adiabatic instantaneous eigenstate:

$$\dot{\gamma}(t) = \frac{2\pi}{L} \sum_\ell Z_\ell \dot{X}_\ell(t) + \text{Im} \left(\frac{\langle\dot{\Psi}_0|\mathrm{e}^{-i\frac{2\pi}{L}\sum_j x_j}|\Psi_0\rangle}{\langle\Psi_0|\mathrm{e}^{-i\frac{2\pi}{L}\sum_j x_j}|\Psi_0\rangle} \right.$$
$$\left. + \frac{\langle\Psi_0|\mathrm{e}^{-i\frac{2\pi}{L}\sum_j x_j}|\dot{\Psi}_0\rangle}{\langle\Psi_0|\mathrm{e}^{-i\frac{2\pi}{L}\sum_j x_j}|\Psi_0\rangle} \right). \qquad (8)$$

It is clear that the first line yields the contribution to I from the classical nuclei.

In order to deal with the electronic term, we consider – following a milestone 1964 paper by Kohn [11] – the family of many-body Hamiltonians parametrized by κ as

$$\hat{H}(\kappa) = \frac{1}{2m} \sum_j (\mathbf{p}_j + \hbar\kappa\mathbf{e}_1)^2 + \hat{V}, \qquad (9)$$

where $\hat{V}$ includes one-and two-body terms, and is periodic along x with period L; the ground eigenstate $|\Psi_0(\kappa)\rangle$ obeys PBCs at any κ. The extra "vector potential" in equation (9) is a gauge transformation. As observed by Kohn, PBCs violate gauge-invariance in the conventional sensee, and $|\Psi_0(\kappa)\rangle$ has a nontrivial κ-dependence. The ground-state energy depends on κ and the Drude weight is, according to Kohn (in the quasi-1d case),

$$D = \frac{\pi e^2}{L} \frac{d^2 E_0(\kappa)}{d\kappa^2}\bigg|_{\kappa=0}; \qquad (10)$$

the Drude weight is nonzero in metals and vanishes in insulators.

When we switch on κ continuously, we get to lowest order

$$|\Psi_0(\kappa)\rangle \simeq |\Psi_0\rangle + \frac{\hbar\kappa}{m} \sum_{n\neq 0} |\Psi_n\rangle \frac{\langle \Psi_n| \sum_j p_{xj} |\Psi_0\rangle}{E_0 - E_n}, \qquad (11)$$

where we have simplified the notations setting $|\Psi_n\rangle = |\Psi_n(0)\rangle$, $E_n = E_n(0)$. In the metallic case, owing to equation (10), the state $|\Psi_0(\kappa)\rangle$ has an energy higher than E_0.

Let us consider next the auxiliary state $|\tilde\Psi_0\rangle = e^{-i\frac{2\pi}{L}\sum_j x_j}|\Psi_0\rangle$: this state obeys the Schrödinger equation for $\kappa = 2\pi/L$, and obeys PBCs as well. It is therefore an eigenstate of $\hat H(2\pi/L)$ with energy E_0; in the metallic case it is therefore orthogonal to $|\Psi_0(2\pi/L)\rangle$. In the insulating case we have instead, to leading order in $1/L$,

$$e^{-i\frac{2\pi}{L}\sum_j x_j}|\Psi_0\rangle \simeq |\Psi_0\rangle + \frac{h}{mL} \sum_{n\neq 0} |\Psi_n\rangle \frac{\langle \Psi_n| \sum_j p_{xj} |\Psi_0\rangle}{E_0 - E_n}. \qquad (12)$$

Replacing this into equations (7) and (8) one gets to leading order in $1/L$

$$I^{(\mathrm{el})} \simeq -\frac{ie\hbar}{mL} \sum_{n\neq 0} \langle \dot\Psi_0|\Psi_n\rangle \frac{\langle \Psi_n| \sum_j p_{xj} |\Psi_0\rangle}{E_0 - E_n} + \mathrm{c.c.}, \qquad (13)$$

where c.c. stands for complex conjugate; the $n = 0$ term does not appear because $\langle \dot\Psi_0|\Psi_0\rangle + \mathrm{c.c.} = 0$. This concludes our proof: equation (13) is indeed the expression for the adiabatic electronic current, as shown below in Appendix B. According to the derivation given therein, equation (13) holds even at finite L, to leading order in the adiabaticity parameter; but the time-integrated current cannot be expressed as a difference of finite-L Berry phases – at variance with equation (14) below – owing to the approximate nature of equation (11) at finite κ.

One final point is worth commenting. The two members of equation (12) may differ by a phase factor; given that the current is gauge-invariant, such phase factor is irrelevant. It is also easily verified that an arbitrary phase factor in $|\tilde\Psi_0\rangle$ would cancel in $\dot\gamma(t)$, equation (8).

2.4 Polarization difference vs. polarization itself

The early formulation of the modern theory avoided defining what polarization "itself" was: the focus was on the difference in polarization between two states of the same material which can be connected via an adiabatic switching process, while the system remains insulating [6,7]. This was inspired by the fact that, for bulk 3d materials, one typically measures polarization *differences*. For instance the spontaneous polarization of a ferroelectric is measured via hysteresis cycles.

In the present quasi-1d formulation the polarization difference is given by:

$$\Delta P = \frac{e}{2\pi}[\,\gamma(t_2) - \gamma(t_1)\,] = \int_{t_1}^{t_2} I(t)\,dt, \qquad (14)$$

where it is understood that a continuous branch of the multivalued function "Im ln" must be chosen, and the large-L limit is addressed.

Soon after the major breakthrough by King-Smith and Vanderbilt [7], the same authors [8] realized that one can even define polarization itself (besides polarization differences), at the price of abandoning the idea that bulk polarization is a vector. At variance with the common wisdom in those times (early 1990s), bulk polarization must be regarded as a lattice; such feature is indeed pretty clear in the formulation given here. The macroscopic polarization P of a quasi-1d system is given by equation (5): given that any phase angle is defined modulo 2π, bulk polarization of a quasi-1d system is only defined modulo e. This apparently exotic state of affairs has a profound physical meaning, which we are going to explain.

2.5 Bounded vs. unbounded samples

At this point it is expedient to limit ourselves to a "crystalline" quasi-1d system (a.k.a. stereoregular polymer), with lattice constant a. In the following, we will use the term "crystalline" in this peculiar sense.

The polarization of an unbounded insulating chain is given, as said above, by equation (6) in the large-system limit. Consider now a bounded realization of the same system, made of M cells, and terminated by two arbitrary sets of nuclei attached at its left and right ends; the number of electrons in the bounded sample must ensure charge neutrality. The dipole $d_x(M)$ of the system is given by equation (3), while the definition of its length is somewhat arbitrary. Given that the end groups are nonextensive, for large M the leading term in the length is Ma, hence a possible definition is

$$P = \lim_{M\to\infty} \frac{d_x(M)}{Ma}. \qquad (15)$$

The following alternative definition is asymptotically equivalent but converges faster:

$$P = \frac{1}{a} \lim_{M\to\infty} [\,d_x(M+1) - d_x(M)\,]. \qquad (16)$$

The reasons for faster convergence are given in reference [12].

A chain with nonzero P is neutral in its bulk but is charged at the ends: in the large-M limit the charge Q accumulated at the right end is indeed *equal* to P (and $Q = -P$ at the left end). In fact a classical uncharged rod of length Ma, with charges $\pm Q$ attached at the ends, would have the same dipole, i.e. $d_x(M) = MaQ$. For a semi-infinite insulating chain the value of Q – and equivalently of P – can be determined by accessing the charge

distribution at its end only. A practical way to achieve this is presented in Appendix C.

And now the main question: which is the relationship between the bulk polarization P of an unbounded chain, evaluated as in equation (5), and the end charge Q of a bounded chain in the large-M limit? The answer is that Q depends on the termination, but is quantized: it can only assume one of the modulo e values dictated by the bulk. Which of them depends on the energetics: roughly speaking, by the relative ionicity of the end groups. One can in principle vary such ionicity almost continuously, but – counterintuitively – the end charge Q *does not* vary continuously: only discrete values, dictated by the bulk, are allowed. A perspicuous demonstration of this feature at the Hartree-Fock (HF) level is presented – following reference [13] – in Section 4.3.

What has been presented in this section is just a restatement of the "theorem of quantization of the surface charge". Early occurrences of the theorem date since the 1960s [14–17]; its topological nature was pointed out by Niu in 1986 [18]. The theorem can be regarded as a corollary of the famous Thouless quantization of charge transport [19], extended by Niu and Thouless to the case where disorder and many-body interaction is present [20].

According to references [18–20] quantization manifests itself in the thermodynamic limit.

We have therefore seen that "polarization itself" determines the possible values of the end charges in a long polymeric chain. If we switch to the "polarization difference" viewpoint, the theorem is just a manifestation of continuity equation: the integrated current flowing in the middle of the chain, given by equation (14), equals the charge which accumulates at its ends.

3 Kohn-Sham crystalline polarization

As anticipated in the Introduction, the KS polarization *does not* coincide with the many body polarization. This feature has been first pointed out in 1995 by Gonze, Ghosez, and Godby [21]; some other papers followed [22]. The presentation given here, inspired by the 1998 formulation of reference [10], makes the reasons for the disagreement very perspicuous.

In the KS case the many-body wavefunction $|\Psi_0^{(KS)}\rangle$ is a Slater determinant of one-particle spinorbitals; in the crystalline case (i.e. a stereoregular polymer) the spinorbitals have the Bloch form. If we evaluate the expectation value in equation (6) using such a $|\Psi_0^{(KS)}\rangle$ we get exactly the KS polarization as given by the (by now standard) Berry-phase formula by King-Smith and Vanderbilt [1–5]: we present a proof in Appendix A.

The genuine many-body ground state $|\Psi_0\rangle$ is necessarily different from $|\Psi_0^{(KS)}\rangle$: the two are only guaranteed to have in common the density $n(\mathbf{r})$, which is lattice periodical. But the macroscopic polarization is *not* a function of the periodic density. According to equation (6), P obtains from the expectation value of a unitary operator: it is a many-body one, having the form of the *exponential* of a one-body operator. Knowledge of the density only is

therefore not enough to determine the P value (in both the exact and the KS cases).

The discrepancy between the KS and the exact polarization can also be understood from the viewpoint of equation (14), where the polarization difference between two states of the chain is expressed in terms on the integrated adiabatic transient current. Continuity equation holds for the exact many-body wavefunction, while instead the ordinary DFT neither allows for time-dependence, nor reproduces correctly the current. One would require here a time-dependent current DFT, capable of enforcing continuity equation.

In practical cases the exact KS potential is unknown and first-principle calculations rely on approximate forms of the density functional. Let us consider bounded samples first: the error in the computed dipole is solely due to the functional itself. Incidentally, it turns out that the dipoles of quasi-1d systems are a very severe benchmark for the available functionals [23]. When instead unbounded samples within PBCs are considered, even the exact KS periodic potential would not provide the exact many-body polarization. There are therefore two sources of error: one intrinsic, and one due to the approximate functional. Despite this drawback, a large amount of first-principle work on 3d systems has addressed properties like ferroelectricity, piezoelectricity, lattice dynamics, and infrared spectra in many classes of materials. Such a large database proves that the modern theory, implemented with the most popular functionals, can reproduce and predict the macroscopic polarization of a 3d system quite accurately (within a few percent in many cases).

Finally, when we assume a local (or semilocal) approximation in the KS potential, we may tacitly adopt it even in an adiabatic time-dependent framework, such as e.g. in the ALDA (adiabatic local density approximation). In this case continuity equation holds (like for the exact many-body case), and therefore the polarization of an unbounded chain, as given by the modern theory, coincides (modulo the quantum) with the polarization of a bounded chain in the long-chain limit.

4 Polarization of a centrosymmetric system

The exact polarization and the KS polarization do coincide in the centrosymmetric case; this statement is not as trivial as it might seem because – according to the modern theory – polarization of a centrosymmetric system *is not* necessarily zero. The reason for the coincidence of the two quantities in the quasi-1d centrosymmetric case is topology: in fact P is arguably the simplest $\mathbb{Z}_2$ invariant appearing in electronic structure.

Topological arguments similar to the present ones also prove the quantization of the soliton charge in polyacetilene, first discovered by Su et al. [24]. If we insist that we want a singlet wavefunction – as everywhere in the present work – then the soliton charge can only be $\pm e$. But we may relax this condition, allowing for a nonzero spin density in the soliton region: then the soliton can even be neutral, but in this case it carries spin $\pm 1/2$.
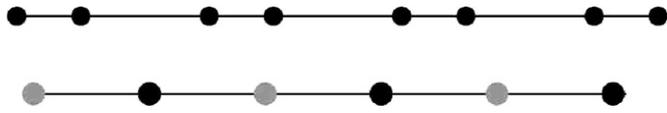
Fig. 1. Two paradigmatic centrosymmetric lattices in 1d.

Fig. 2. A centrosymmetric insulating quasi-1d "crystal" with two different termination: alternant trans-polyacetylene. Here the "bulk" is five-monomer long. After reference [13].

As said above, the polarization of an unbounded system is a lattice, and there are two distinct 1d lattices compatible with centrosymmetry: either $P = 0 \mod e$ ($\mathbb{Z}_2$ even), or $P = e/2 \mod e$ ($\mathbb{Z}_2$-odd). We focus again on a "crystalline" quasi-1d system, invariant by inversion with respect to some origin (Wyckoff position); notice however that P (and γ) are to be computed from equation (6) where instead the origin may be arbitrary, owing to charge neutrality.

4.1 A toy model

The top sketch in Figure 1 refers to an idealized molecular crystal, or to an alternant polymer, like polyacetilene (compare to Fig. 2); the bottom sketch refers to an idealized ionic crystal. Both system are centrosymmetric and can be described at the simplest level by an extreme tight-binding model Hamiltonian (Hückel-like), whose only parameters are the first-neighbor hoppings t and the onsite energies ϵ. The top chain has alternating t's and constant ϵ; the bottom chain has constant t and alternating ϵ's. At half filling the system is a one-band insulator (except in the nonalternating case); if we consider spinful electrons, charge neutrality requires $+1$ classical charges on each site for both chains.

It is simple to show that the top chain in Figure 1 has $P = 0 \mod e$: we define this as $\mathbb{Z}_2$ even. The botton chain has $P = e/2 \mod e$, and is $\mathbb{Z}_2$-odd. It is also a straightforward exercise to show that one cannot "continuously deform" the Hamiltonian (and its ground state) from one case into the other without closing the gap, while conserving centrosymmetry. In modern jargon, the $\mathbb{Z}_2$ invariant is "protected" by centrosymmetry; throughout this work we have also tacitly assumed time-reversal invariance of the many-body Hamiltonian at $\kappa = 0$.

Suppose instead that a centrosymmetric chain (either $\mathbb{Z}_2$ even or $\mathbb{Z}_2$-odd) is closed in form of a loop, and that we break time-reversal symmetry by adiabatically threading the loop by a magnetic flux. This induces a current and varies P, although the potential remains centrosymmetric: the final P may then assume quantized values different from 0 or π.

4.2 Kohn-Sham vs. exact

Given that the polarization of an unbounded centrosymmetric crystalline quasi-1d system is a topological invariant, its value is insensitive to many details of the Hamitonian. Suppose that, for a given interacting chain, we know the periodic KS potential $V_{\mathrm{KS}}(\mathbf{r})$, which by definition yields the exact periodic density. The noninteracting KS system is adiabatically connected to the exact many-body one by adiabatically switching on the interaction, while varying the one-body potential in order to keep the
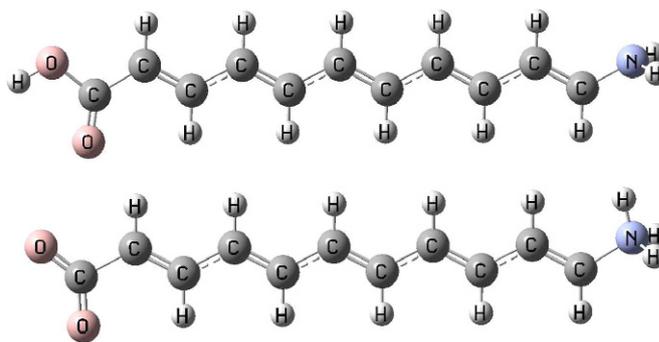
density constant. A basic tenet of topology is that insofar as the chain remains centrosymmetric and insulating, the $\mathbb{Z}_2$ invariant cannot change.

A similar line of reasoning holds for the Hartree-Fock (HF) approximation: even the HF polarization of a centrosymmetric insulating chain is topological ($\mathbb{Z}_2$-even or $\mathbb{Z}_2$-odd), and therefore coincides with the exact one.

4.3 Bounded vs. unbounded samples

As explained above, the dipole per unit length (or equivalently the end charge Q) of a long polymeric chain – with arbitrary terminations – is asymptotically equal to one of the modulo e values of its bulk polarization P. If the bulk of the chain is centrosymmetric, then one asymptotically has either $Q = 0 \mod e$ ($\mathbb{Z}_2$-even) or $Q = e/2 \mod e$ ($\mathbb{Z}_2$-odd).

First-principle simulations demonstrating this key topological feature, at the HF level, have been published in reference [13]. Here we only report the results for the case of alternant trans-polyacetylene, which is topologically equivalent to the top sketch in Figure 1: $\mathbb{Z}_2$-even.

We chose two different terminations, as illustrated in Figure 2. The structure is either "neutral" or "charge-transfer"; the end groups are "donor" (NH_2) and "acceptor" ($COOH$). We computed the dipole divided by the number of monomers M as a function of M up to large M values, shown in Figure 3. In both cases the bulk is centrosymmetric, while the whole molecule is polar: therefore at finite M the two dipoles are nonzero and nonquantized; the topological quantization shows up in the $M \to \infty$ limit. This is in agreement with a basic tenet of references [18–20]: quantization manifests itself in the thermodynamic limit.

For this system the lattice constant is $a = 4.670$ a.u., hence the figure shows that the end charges are either 0 or e for the two cases of Figure 2, respectively. The figure also shows that in the polyacetylene case the asymptotic regime is attained at about 10–20 monomers. This length is essentially the exponential decay length of the one-body density matrix of bulk polyacetylene.
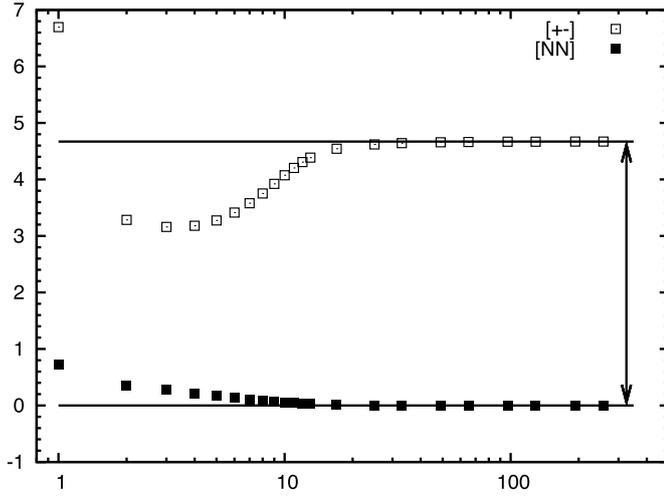
Fig. 3. Dipole per monomer (a.u.) as a function of the number M of monomers in the chain, for the two different terminations of Figure 2, after reference [13]. In the large-M limit the length of the polymer is $L \simeq Ma$, with $a = 4.670$ bohr; hence the asymptotic values of the dipole per unit lenght are 0 and e for the two cases.

5 Conclusions

The dipole of a bounded sample is a trivial function of the density, which in turn depends on the *modulus* of the wavefunction. Instead, the polarization of an unbounded sample within Born-von-Kàrmàn PBCs is a gauge-invariant *phase* of the wavefunction. When a time-dependent viewpoint is assumed, a gauge-invariant phase factor is essential for accessing the electronic current. In this work we have addressed the intimate link between charge and current (in the adiabatic limit), at the root of the modern theory. It is not surprising that DFT breaks this link; in order to restore such link, one would require a time-dependent current DFT, where continuity equation is conserved. In the final part of this work we have also shown that the polarization of a quasi-1d centrosymmetric polymer is a $\mathbb{Z}_2$ topological invariant. The KS-DFT ground state is adiabatically connected to the exact one: the topological invariant cannot vary insofar as the system remains insulating.

Appendix A: King-Smith and Vanderbilt formula

We separate the nuclear and electronic Berry phases in equation (6)

$$\gamma = \text{Im} \ln \, \mathrm{e}^{i\frac{2\pi}{L}\sum_\ell Z_\ell X_\ell} + \text{Im} \ln \, \langle \Psi_0 | \mathrm{e}^{-i\frac{2\pi}{L}\sum_j x_j} | \Psi_0 \rangle, \tag{A.1}$$

and we rewrite the latter as

$$\gamma^{(\text{el})} = \text{Im} \ln \, \mathfrak{z}_N, \quad \mathfrak{z}_N = \langle \Psi_0 | \mathrm{e}^{-i\frac{2\pi}{L}\sum_j x_j} | \Psi_0 \rangle, \tag{A.2}$$

where $|\Psi_0\rangle$ is the ground state of a system on N electrons in a periodic box (along x) of length L. The modulus of the complex number $\mathfrak{z}_N$ is no larger than one; the system is an insulator insofar as $\mathfrak{z}_N$ does not vanish in the large-N limit [27].

Suppose we have a quasi-1d "crystalline" system of lattice constant a, where we impose PBCs over M linear cells: there are then M equally spaced Bloch vectors in the reciprocal cell $[0, 2\pi/a)$:

$$k_s = \frac{2\pi}{Ma}s, \quad s = 0, 1, \ldots, M-1. \tag{A.3}$$

The Born-von-Kàrmàn period is $L = Ma$. $|\Psi_0\rangle$ is a Slater determinant of spinorbitals, which can be chosen to have the Bloch form:

$$\begin{aligned} \psi_{mk_s}^{\uparrow}(\mathbf{r} + \tau\mathbf{e}_1) &= \mathrm{e}^{ik_s\tau}\psi_{mk_s}^{\uparrow}(\mathbf{r}), \\ \psi_{mk_s}^{\downarrow}(\mathbf{r} + \tau\mathbf{e}_1) &= \mathrm{e}^{ik_s\tau}\psi_{mk_s}^{\downarrow}(\mathbf{r}), \end{aligned} \tag{A.4}$$

where $\tau = la$ is a lattice translation, and m is a band index. The Bloch spinorbitals are normalized to one over the unit cell and a singlet ground state is assumed: the up- and down-arrows in equation (A.4) are spin labels.

There are $n_{\text{b}} = N/(2M)$ doubly occupied bands bands in the Slater determinant wavefunction, which we write as

$$|\Psi_0\rangle = \frac{1}{\sqrt{M^N}} \mathscr{A} \prod_{m=1}^{n_{\text{b}}} \prod_{s=0}^{M-1} \psi_{mk_s}^{\uparrow}\psi_{mk_s}^{\downarrow}, \tag{A.5}$$

where $\mathscr{A}$ is the antisymmetrizer; the $1/\sqrt{M^N}$ factor in equation (A.5) owes to the different normalizations: $|\Psi_0\rangle$ is normalized over $L = Ma$, while the Bloch orbitals $|\psi_{m\mathbf{k}}\rangle$ are normalized over a crystal cell of length a.

It is now expedient to define a new set of Bloch orbitals:

$$\tilde{\psi}_{mk_s}^{\uparrow}(\mathbf{r}) = \mathrm{e}^{-i\frac{2\pi}{L}x}\psi_{mk_s}^{\uparrow}(\mathbf{r}), \tag{A.6}$$

and analogously for down spin. We then recast the matrix element $\mathfrak{z}_N$ as

$$\mathfrak{z}_N = \langle \Psi_0 | \tilde{\Psi}_0 \rangle, \tag{A.7}$$

where $|\tilde{\Psi}_0\rangle$ is the Slater determinant of the $\tilde{\psi}$'s (with up and down spin). According to a well known theorem, the overlap amongst two determinants is equal to the determinant of the overlap matrix amongst the spinorbitals. This determinant factorizes into a spin-up and a spin-down sectors, ergo

$$\mathfrak{z}_N = \frac{1}{M^N}(\det \mathcal{S})^2, \tag{A.8}$$

where the (spinless) overlap matrix is

$$\begin{aligned} \mathcal{S}_{sm,s'm'} &= \int_{-\infty}^{\infty}\!\!dy \int_{-\infty}^{\infty}\!\!dz \int_0^L\!\!dx \, \psi_{mk_s}^*(\mathbf{r})\mathrm{e}^{-i\frac{2\pi}{L}x}\psi_{m'k_{s'}}(\mathbf{r}) \\ &= M\langle \psi_{mk_s} | \mathrm{e}^{-i\frac{2\pi}{L}x} | \psi_{m'k_{s'}} \rangle \end{aligned} \tag{A.9}$$

and $\psi_{mk_M}(\mathbf{r}) \equiv \psi_{mk_0}(\mathbf{r})$ is implicitly understood (so-called periodic gauge); the $1/M^N$ factor cancels from now on.

Owing to the orthogonality properties of the Bloch functions, the overlap matrix elements vanish except when $k_{s'} = k_s + 2\pi/L$, that is $s' = s+1$. The $(N/2) \times (N/2)$ determinant can then be factorized into M small determinants:

$$\det \mathcal{S} = \prod_{s=0}^{M-1} \det S(k_s, k_{s+1}), \qquad (A.10)$$

where for the small overlap matrix we use the notation

$$\begin{aligned} S_{m,m'}(k_s, k_{s+1}) &= \langle \psi_{mk_s} | e^{-i\frac{2\pi}{L}x} | \psi_{m'k_{s+1}} \rangle \\ &= \langle u_{mk_s} | u_{m'k_{s+1}} \rangle, \qquad (A.11) \end{aligned}$$

where $|u_{mk}\rangle = e^{-ikx}|\psi_{mk}\rangle$ is the periodic factor in the Bloch orbital. We finally get (for doubly occupied bands)

$$\gamma^{(\mathrm{el})} = 2 \, \mathrm{Im} \, \ln \prod_{s=0}^{M-1} \det S(k_s, k_{s+1}) : \qquad (A.12)$$

this is the well known expression of the modern theory of polarization, as implemented in many electronic-structure codes. The formula was originally obtained by King–Smith and Vanderbilt by defining a (continuum) Berry phase as a line integral, and then discretizing it [7]. Here we are providing a reverse derivation, inspired by reference [10].

For the sake of simplicity, we assume there is only one occupied band; equations (A.11) and (A.12) become then

$$S(k_s, k_{s+1}) = \langle u_{k_s} | u_{k_{s+1}} \rangle \qquad (A.13)$$

$$\begin{aligned} \gamma^{(\mathrm{el})} &= 2 \, \mathrm{Im} \, \ln \prod_{s=0}^{M-1} \langle u_{k_s} | u_{k_{s+1}} \rangle \\ &= 2 \sum_{s=0}^{M-1} \mathrm{Im} \, \ln \langle u_{k_s} | u_{k_{s+1}} \rangle. \qquad (A.14) \end{aligned}$$

If we adopt a gauge such that $|u_k\rangle$ is a differentiable function of k, then

$$\mathrm{Im} \, \ln \langle u_k | u_{k+\Delta k} \rangle \simeq -i \langle u_k | du_k/dk \rangle \, \Delta k. \qquad (A.15)$$

In the large-M limit the sum converges to the integral

$$\gamma^{(\mathrm{el})} = -2 \int_0^{2\pi/a} i \langle u_k | du_k/dk \rangle \, dk. \qquad (A.16)$$

This integral is a Berry phase, of the kind first addressed by Zak [28]; the integrand goes under the name of Berry connection. Since $|u_k\rangle$ is gauge-dependent (i.e. arbitrary by a k-dependent phase factor), the Berry connection is gauge-dependent as well; the Berry phase, instead, is gauge-invariant modulo 2π.

As observed above, the continuum formulation of equation (A.16) requires $|u_k\rangle$ to be a differentiable function of k, while the discrete formulation of equations (A.12) and (A.14) is more general. This is of foremost importance in numerical work: in fact, when the orbitals are obtained by numerical diagonalization, the choice of the gauge is erratic; further ambiguity arises from band ordering in the many-band case. We stress that this ambiguity is harmless in both equations (A.12) and (A.14), which are *numerically* gauge-invariant in-form. To this aim, it is essential that the Hamiltonian is diagonalized M times, not $M+1$, in order to enforce the periodic gauge, i.e.

$$\langle u_{mk_0} | u_{m'k_M} \rangle \equiv \langle u_{mk_0} | e^{-i2\pi x/a} | u_{m'k_0} \rangle. \qquad (A.17)$$

Appendix B: The adiabatic current

Let us consider again the many-body Hamiltonian of equation (9)

$$\hat{H}(t) = \frac{1}{2m} \sum_j \mathbf{p}_j^2 + \hat{V}(t), \qquad (B.1)$$

where $\kappa = 0$, and the t-dependence is slow with respect to the characteristic time $\hbar/(E_1 - E_0)$; let $|\Psi_n(t)\rangle$ be the adiabatic instantaneous eigenstates.

The (electrical) current operator along x is

$$\hat{j} = -\frac{e}{mL} \sum_j p_{xj}. \qquad (B.2)$$

The time evolution of the ground state, in the lowest order adiabatic approximation, is $|\Psi(t)\rangle \simeq |\Psi_0(t)\rangle$. This is not enough to evaluate the electronic macroscopic current I: in fact the expectation value $\langle \Psi_0(t)|\hat{j}|\Psi_0(t)\rangle$ vanishes at all times, owing to time-reversal symmetry (for a nondegenerate ground state). We thus need the next order in the adiabatic approximation.

The derivation can be done in various ways; here we adopt the approach of Niu and Thouless [25], where the density-matrix formulation avoids the fastidious occurrence of dynamical phases. We write the exact density matrix of the time-evolved ground state as

$$\hat{\rho}(t) = |\Psi_0(t)\rangle\langle\Psi_0(t)| + \Delta\hat{\rho}(t). \qquad (B.3)$$

Since the adiabatic term commutes with $\hat{H}(t)$, the equation of motion is

$$[\hat{H}(t), \Delta\hat{\rho}(t)] = i\frac{d}{dt}\hat{\rho}(t) \simeq i\frac{d}{dt}|\Psi_0(t)\rangle\langle\Psi_0(t)|, \qquad (B.4)$$

where we have neglected terms of higher order in the adiabaticity parameter. Taking then the matrix elements between $\langle\Psi_0(t)|$ and $|\Psi_n(t)\rangle$ one gets

$$(E_0 - E_n)\langle\Psi_0(t)|\Delta\hat{\rho}(t)|\Psi_n(t)\rangle = i\hbar\langle\dot{\Psi}_0(t)|\Psi_n(t)\rangle; \qquad (B.5)$$

the $n = 0$ term vanishes because of norm conservation. Notice also that, to this order, $\langle\Psi_m(t)|\Delta\hat{\rho}(t)|\Psi_n(t)\rangle = 0$ when both $m \neq 0$ and $n \neq 0$; therefore

$$\Delta\hat{\rho}(t) = i\hbar \sum_{n\neq 0} |\Psi_0(t)\rangle \frac{\langle\dot{\Psi}_0(t)|\Psi_n(t)\rangle}{E_0 - E_n} \langle\Psi_n(t)| + \text{H.c.,} \quad \text{(B.6)}$$

where H.c. stays for Hermitian conjugate.

The adiabatic electronic current is therefore

$$I^{(\text{el})}(t) = \text{Tr}\left\{\hat{j}\,\Delta\hat{\rho}(t)\right\}$$
$$= i\hbar \sum_{n\neq 0} \langle\dot{\Psi}_0(t)|\Psi_n(t)\rangle \frac{\langle\Psi_n(t)|\hat{j}|\Psi_0(t)\rangle}{E_0 - E_n} + \text{c.c.,}$$

which coincides indeed with equation (13).

Equation (B.7) yields the *macroscopic* current in an unbounded sample within PBCs. It is worth mentioning that the local version of equation (B.7) holds as well for a bounded sample with square-integrable eigenstates, and provides the *microscopic* current density. It is easy to show that such microscopic current density obeys continuity equation (to leading order in adiabaticity).

Appendix C: End charges

The net charges $\pm Q$ at the end of a long polymeric chain have been defined by means of a neutral rod having the same dipole d_x as the real system. There is an alternative pictorial way to visualize how and where there is charge accumulation on a long polymeric chain, and such approach also applies to semi-infinite chains.

The appropriate tool is rather popular within the theory of semiconductor interfaces, where it has been given the name of "macroscopic average" [26]. In the present context, we are going to call the analogous tool a "sliding-window convolution".

The total charge density of our system is

$$\rho(\mathbf{r}) = e\sum_\ell Z_\ell \delta(\mathbf{r} - \mathbf{R}_\ell) - e\,n(\mathbf{r})\,; \qquad \text{(C.1)}$$

we define the 1d function $\bar{\rho}(x)$ as the yz integral of $\rho(\mathbf{r})$:

$$\bar{\rho}(x) = e\sum_\ell Z_\ell \delta(x - X_\ell) - e\int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dz\; n(\mathbf{r})$$
$$= e\sum_\ell Z_\ell \delta(x - X_\ell) + \bar{\rho}_{\text{el}}(x). \qquad \text{(C.2)}$$

Since we are addressing stereoregular quasi-1d chains, in their bulk region the function $\bar{\rho}(x)$ is periodic with period a, and deviates from periodicity in the end regions only. Clearly, for long insulating chains (i.e. for large enough M) the size of the end regions may depend on the chemical nature of the system, but is *independent* of the actual value of M.

Besides being periodic, the function $\bar{\rho}(x)$ averages to zero over one period a. We exploit this fact by defining

the sliding-window convolution $\bar{\bar{\rho}}(x)$ as

$$\bar{\bar{\rho}}(x) = \frac{1}{a}\int_{-a/2}^{a/2} dx'\; \bar{\rho}(x - x')$$
$$= \frac{1}{a}\int_{-\infty}^{\infty} dx'\; \theta(a/2 - |x - x'|)\,\bar{\rho}(x'), \qquad \text{(C.3)}$$

where θ is the Heaviside step function. In essence, the window convolution filters out the irrelevant periodic density oscillations, and blows up net charge accumulation. By construction, $\bar{\bar{\rho}}(x)$ vanishes in the bulk region of the chain, in either the case of finite (large M) or semi-infinite ($M = \infty$) chains.

Therefore $\bar{\bar{\rho}}(x)$ visualizes the spatial extent of the "perturbation" due to the chain termination. By using equation (C.2) into equation (C.3) we have

$$\bar{\bar{\rho}}(x) = \sum_\ell \frac{eZ_\ell}{a}\theta(a/2 - |x - X_\ell|) + \frac{1}{a}\int_{-a/2}^{a/2} dx'\; \bar{\rho}_{\text{el}}(x - x'),$$
$$\text{(C.4)}$$

which also explicitly shows how the filter broadens the δ-like nuclear charge density, resulting in a kind of hystogram.

References

1. R. Resta, Rev. Mod. Phys. **66**, 899 (1994)
2. D. Vanderbilt, R. Resta, in *Conceptual foundations of materials: A standard model for ground- and excited-state properties*, edited by S.G. Louie, M.L. Cohen (Elsevier, 2006), p. 139
3. R. Resta, D. Vanderbilt, in: *Physics of Ferrelectrics: a Modern Perspective, Topics in Applied Physics*, Vol. 105, edited by Ch.H. Ahn, K.M. Rabe, J.-M. Triscone (Springer-Verlag, 2007), p. 31
4. R. Resta, J. Phys.: Condens. Matter **22**, 123201 (2010)
5. N.A. Spaldin, J. Solid State Chem. **195**, 2 (2012)
6. R. Resta, Ferroelectrics **136**, 51 (1992)
7. R.D. King-Smith, D. Vanderbilt, Phys. Rev. B **47**, 1651 (1993)
8. D. Vanderbilt, R.D. King-Smith, Phys. Rev. B **48**, 4442 (1993)
9. L.D. Landau, E.M. Lifshitz, *Electrodynamics of continuous media* (Pergamon Press, Oxford, 1984).
10. R. Resta, Phys. Rev. Lett. **80**, 1800 (1998)
11. W. Kohn, Phys. Rev. **133**, A171 (1964)
12. K.N. Kudin, R. Car, R. Resta, J. Chem. Phys. **122**, 134907 (2005)
13. K.N. Kudin, R. Car, R. Resta, J. Chem. Phys. **127**, 194902 (2007)
14. V. Heine, Phys. Rev. **145,** 593 (1966)
15. J.A. Appelbaum, D.R. Hamann, Phys. Rev. B **10**, 4973 1974
16. L. Kleinman, Phys. Rev. B **11**, 858 (1975)
17. F. Claro, Phys. Rev. B **17**, 699 (1977)
18. Q. Niu, Phys. Rev. **33**, 5368 (1986)
19. D.J. Thouless, Phys. Rev. B **27**, 6083 (1983)
20. Q. Niu, D.J. Thouless, J. Phys A **17**, 2453 (1984)
21. X. Gonze, Ph. Ghosez, R.W. Godby, Phys. Rev. Lett. **74**, 4035 (1995)

22. D. Vanderbilt, Phys. Rev. Lett. **79**, 3966 (1997)
23. M. van Faassen, P.L. de Boeij, R. van Leeuwen, J.A. Berger, J.G. Snijders, J. Chem. Phys. **118**, 1044 (2003)
24. W.P. Su, J.R. Schrieffer, A.J. Heeger, Phys. Rev. Lett. **42**, 1698 (1979)
25. Q. Niu, D.J. Thouless, Y.S. Wu, Phys. Rev. B **31**, 3372 (1985)
26. A. Baldereschi, S. Baroni, R. Resta, Phys. Rev. Lett. **61**, 734 (1988)
27. R. Resta, S. Sorella, Phys. Rev. Lett. **82**, 370 (1999)
28. J. Zak, Phys. Rev. Lett. **62**, 2747 (1989)