

REVIEW ARTICLE

Manifestations of Berry's phase in molecules and condensed matter

Raffaele Resta

INFM—Dipartimento di Fisica Teorica, Università di Trieste, Strada Costiera 11, I-34014 Trieste, Italy

Received 28 July 1999

Abstract. Since the appearance of Berry's seminal paper in 1984, geometric phases have been discovered in virtually all fields of physics. Here we address molecules and solids, and we limit our scope to the Berry's phases of the many-electron wavefunction. Many advances have occurred in very recent years relating to the theory of such phases and their observable consequences. After discussing the basic features of Berry's phases in a generic quantum system, we specialize to selected examples taken from molecular physics and condensed matter physics; in each of these cases, a Berry's phase of the electronic wavefunction leads to measurable effects.

Contents

1. Introduction
2. Fundamentals
 - 2.1. The discrete (Pancharatnam's) geometric phase
 - 2.2. Berry's geometric phase
 - 2.3. Connection and curvature
 - 2.4. The paradigm: the Aharonov–Bohm effect
3. Features of the geometric phase
 - 3.1. Parallel transport
 - 3.2. Computing a Berry's phase
 - 3.3. The open-path geometric phase
 - 3.4. The single-point Berry's phase
4. The electronic Berry's phase
 - 4.1. Wavefunctions and density matrices
 - 4.2. Independent electrons
 - 4.3. Bloch orbitals
 - 4.4. Zak's phase
5. Manifestations of Berry's phase
 - 5.1. The molecular Aharonov–Bohm effect
 - 5.2. Adiabatic approximation in a magnetic field
 - 5.3. Semiclassical electron dynamics in crystals
 - 5.4. Bloch oscillations and the Wannier–Stark ladder
 - 5.5. Spin-wave dynamics in crystals

6. Macroscopic polarization
 - 6.1. The problem
 - 6.2. Polarization as a Berry's phase
 - 6.3. A lone electron in a periodic box
 - 6.4. The position operator in extended systems
 - 6.5. A crystalline system of independent electrons
 - 6.6. King-Smith and Vanderbilt's formula
 - 6.7. Non-crystalline systems
 - 6.8. Correlated electrons and topological phase transitions
7. Conclusions

1. Introduction

The Berry's phase takes its name from a very influential paper by Michael Berry [1], circulated in 1983 although published in 1984. In this review I will use the name 'Berry's phase' to mean a geometric phase in quantum mechanics, although analogous phases occur in other domains (e.g. in classical optics or classical mechanics [2]). Berry realized that there is a very general phenomenon, leading to observable effects in several quite different physical cases. A few manifestations of what we now call a Berry's phase were known of (and understood) before: the Aharonov-Bohm (AB) effect [3-5] dating from 1959, and the so-called molecular AB effect [6, 7] discovered even earlier [8].

Since the first one appeared in 1983, a plethora of publications have dealt with the theory of geometric quantum phases and its applications in various fields. A very useful book [2] containing reprints and some original articles was published in 1989. Other useful publications of various kinds exist: general reviews [9, 10]; pedagogical and/or historical presentations for non-specialists [11-13]; and a guide to the published literature [14]. Furthermore, the concept of a Berry's phase is even presented in some modern quantum mechanics textbooks [15-17].

The Berry's phase is almost ubiquitous in present-day physics: because of this, unavoidably the Berry's phase means somewhat different things to different people. Here we narrow the focus to the geometric quantum phases of many-electron systems and their observable effects. Even after thus narrowing the focus, there remain several rather different fields and phenomena where the Berry's phase plays a relevant role; therefore further selectivity is in order. The most notable omissions here concern the quantum Hall effect [18], and the physics of superfluids in general. In these areas, which are not covered by the present review, a large amount of work is related to geometric phases; some key references are [19] and [20].

I delivered a series of lectures on the role of Berry's phases in electronic wavefunctions in 1995. The present review differs from the lecture notes [21] in two main respects: it is less pedagogical, and, more importantly, it covers several outstanding advances which occurred after that date: in the semiclassical theory of electron dynamics in crystals [22] (sections 5.3 and 5.4), in the theory of spin-wave dynamics in crystals [23, 24] (section 5.5), and the novel formulation of polarization theory [25-27] (section 6). In order to keep both the size and the scope of the present review within reasonable limits, I am going to exclude discussion of any non-Abelian geometric phase [2, 21], or any 'geometric distance'. The latter concept is in a sense complementary to that of geometric phase [28-30]: here let me just mention that some very recent work has shown that the *localization* of the electronic wavefunction in insulating condensed systems can be measured as a geometric quantum distance [31, 32].

The plan of this work is as follows. In section 2 I introduce the basic concept of a geometric phase, starting from a discrete approach, and I show how the geometric phase is observable, using the venerable example of the AB effect [3, 4]. In section 3 I present some more advanced

or less known features, which are relevant to the case studies which I am going to discuss in the following sections. In particular, the very unconventional ‘single-point Berry’s phase’ is emphasized. In section 4 I narrow the focus to a many-electron system, and I discuss the relationship between the Berry’s phase of the wavefunction and those of the single-particle orbitals. I then specialize to a crystalline system in the thermodynamic limit, where the orbitals have the Bloch form: the occurrence of a Berry’s phase in this area was first recognized by Zak in 1989 [33]. In section 5 I discuss various observable manifestations of an electronic Berry’s phase in molecular physics and in solid-state physics: the so-called molecular AB effect [6, 7, 34]; a surprising (and little known) feature of the adiabatic approximation in magnetic fields [35, 36]; the semiclassical electron dynamics in crystals [22] and the related problem of Wannier–Stark ladders; the theory of spin-wave dynamics in crystals [23, 24]. Section 6 is entirely devoted to the modern theory of dielectric polarization. The occurrence of a Berry’s phase in the phenomenon of polarization was discovered by King-Smith and Vanderbilt in 1993 [37]: this phase plays a very important role in electronic structure theory both as a matter of principle [38–40] and as a computational tool [41]. Outstanding advances have occurred since: the work of reference [37] can nowadays be regarded as a special case of a more general—and conceptually simpler—theory, first outlined in reference [25]. The account given here of polarization theory is based on this novel viewpoint, where a single-point Berry’s phase plays a major role; the implications for correlated and/or non-crystalline condensed systems are discussed.

2. Fundamentals

2.1. The discrete (Pancharatnam’s) geometric phase

The concept of ‘geometric phase’ made its first appearance in 1956 in a paper written by the young Indian physicist S Pancharatnam [42, 43], which at the time went largely unnoticed in the Western world. This paper concerns the phases of light beams, but the ideas behind it can be applied with basically no change to the phases of the quantum states of any physical system. The essential feature of Pancharatnam’s work is that of considering *discrete* phase changes, at variance with the more recent work [1, 2], where *continuous* phase changes are usually addressed.

The present review concerns mostly the geometric properties of the many-electron wavefunction of a condensed system. When dealing with this problem, several good reasons (which will be made clear in the rest of this work) lead us to regard the discrete approach to the geometric phase as the most fundamental one. We use therefore Pancharatnam’s discrete approach, translated into quantum mechanical language [44], as the starting point for introducing our main subject.

Let us assume that a generic quantum Hamiltonian has a parametric dependence:

$$H(\xi)|\psi(\xi)\rangle = E(\xi)|\psi(\xi)\rangle \quad (1)$$

where the parameter ξ is defined in a suitable domain: a two-dimensional real ξ has been chosen for display in figure 1. I start by discussing the most general case, and therefore for the time being I do not specify which quantum system is described by this Hamiltonian, nor what the physical meaning of the parameter ξ is. To give an idea with one particular example, $H(\xi)$ could be the electronic Hamiltonian of a molecule in the Born–Oppenheimer approximation, and ξ a suitable nuclear coordinate.

The state vectors $|\psi(\xi)\rangle$ are all supposed to reside in the same Hilbert space: this amounts to saying that the wavefunctions are supposed to obey ξ -independent boundary conditions. We

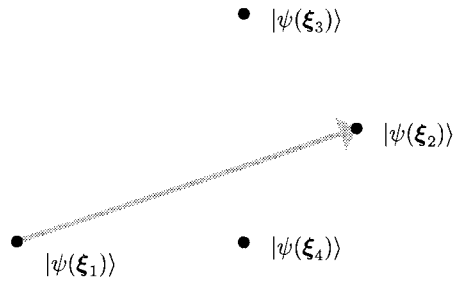


Figure 1. Four state vectors, at different parameter values. $\Delta\varphi_{12}$, as defined in the text, is the phase difference between the state vectors at the two points connected by the arrow.

will further assume that the ground state $|\psi_0(\xi)\rangle$ is non-degenerate for any value of ξ in the parameter domain; the 0 subscript is omitted in most of this work, in order to avoid an undue burden of notation.

We start by defining the *phase difference* between the ground eigenstates at two different ξ -points in the most natural way:

$$e^{-i\Delta\varphi_{12}} = \frac{\langle\psi(\xi_1)|\psi(\xi_2)\rangle}{|\langle\psi(\xi_1)|\psi(\xi_2)\rangle|} \quad (2)$$

$$\Delta\varphi_{12} = -\text{Im} \log \langle\psi(\xi_1)|\psi(\xi_2)\rangle. \quad (3)$$

For any given choice of the two states, equations (2) and (3) provide a $\Delta\varphi_{12}$ which is unique modulo 2π , except in the very special case where the states are orthogonal. However, it is also clear that such $\Delta\varphi_{12}$, despite being well defined, *cannot* have any physical meaning. In fact any quantum mechanical state vector is arbitrary as regards a constant phase factor: we refer to the choosing of this phase as the choice of the gauge. A simple change of gauge at ξ_1 and/or at ξ_2 will give to $\Delta\varphi_{12}$ the most arbitrary value.

We now proceed by choosing a finite number of ξ -points, and consider the *total* phase difference along the closed path which joins them *in a given order*, as shown in figure 2:

$$\begin{aligned} \gamma &= \Delta\varphi_{12} + \Delta\varphi_{23} + \Delta\varphi_{34} + \Delta\varphi_{41} \\ &= -\text{Im} \log \langle\psi(\xi_1)|\psi(\xi_2)\rangle \langle\psi(\xi_2)|\psi(\xi_3)\rangle \langle\psi(\xi_3)|\psi(\xi_4)\rangle \langle\psi(\xi_4)|\psi(\xi_1)\rangle. \end{aligned} \quad (4)$$

It is now clear that all the gauge-arbitrary phases cancel in pairs, making the overall phase γ a gauge-invariant quantity. Of course, such a phase *does* depend on the chosen path.

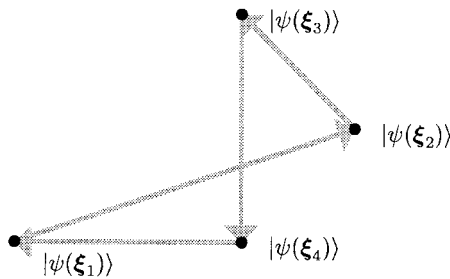


Figure 2. A closed path in parameter space.

The above very simple algebra leads to a result of overwhelming physical importance: in fact, a gauge-invariant quantity is potentially a physical observable. We are accustomed to regarding any observable effect in quantum mechanics as the eigenvalue (or expectation

value) of some Hermitian operator; here we find a kind of 'exotic' observable, which *cannot* be expressed in terms of any Hermitian operator, being instead a gauge-invariant phase of the state vectors.

2.2. Berry's geometric phase

We now proceed to define the most usual [1, 2] continuum Berry's phase, as a limiting process starting from the discrete case considered so far. Suppose we have a smooth closed curve C in the parameter domain, such as that in figure 3, and we discretize it with a set of points on it. Using equation (2), we write the phase difference between any two contiguous points as

$$e^{-i\Delta\varphi} = \frac{\langle\psi(\xi)|\psi(\xi + \Delta\xi)\rangle}{|\langle\psi(\xi)|\psi(\xi + \Delta\xi)\rangle|}. \quad (5)$$

If we further assume that the gauge is so chosen that the phase varies in a *differentiable* way along the path, then from equation (5) we get to leading order in $\Delta\xi$

$$-i\Delta\varphi \simeq \langle\psi(\xi)|\nabla_{\xi}\psi(\xi)\rangle \cdot \Delta\xi. \quad (6)$$

In the limiting case of a set of points which becomes dense on the continuous path, the total phase difference γ defined as in the previous section converges to a circuit integral:

$$\gamma = \sum_{s=1}^M \Delta\varphi_{s,s+1} \longrightarrow \oint_C d\varphi \quad (7)$$

where the linear differential form $d\varphi$ is defined as

$$d\varphi = i\langle\psi(\xi)|\nabla_{\xi}\psi(\xi)\rangle \cdot d\xi. \quad (8)$$

Since the state vectors are assumed to be normalized at any ξ , the linear differential form is *real*.

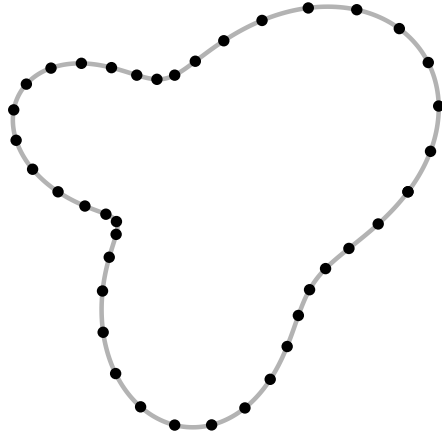


Figure 3. A closed curve in parameter space, and its discretization.

The Berry's phase γ is thus defined as the loop integral over C of the linear differential form $d\varphi$, defined in some parameter domain by equation (8), and which will be referred to as the Berry's connection in the following. Whenever a non-trivial γ occurs, the differential form is of course non-exact. Since there is no risk of ambiguity here, we will alternatively give the same name of Berry's connection to the real vector field $i\langle\psi(\xi)|\nabla_{\xi}\psi(\xi)\rangle$ itself, defined in the parameter domain, and which (except in trivial cases) is non-conservative.

As in the discrete case outlined above, γ is a gauge-invariant quantity and therefore is potentially observable. The main discovery of Berry's milestone paper can in fact be spelled out as follows: there exist a whole class of observables which cannot be cast as the expectation values of *any* operator. The work which has followed Berry's seminal paper has shown that such observables occur in the most disparate physical phenomena.

In order to understand why these observables originate, we must re-examine our initial assumption of a parametric Hamiltonian, equation (1), and discuss its most fundamental meaning. In general, a quantum system having a parametric dependence in its Hamiltonian cannot be isolated: the parameter schematizes a kind of coupling with other variables not included in the Hilbert space, or more generally with 'the rest of the Universe', to use Berry's words [1]. In a truly isolated system, there can be no manifestation of a Berry's phase, and therefore observable effects may only occur in the familiar way, as eigenvalues of some Hermitian operator. In this sense, the Berry's phase is an unnecessary semiclassical concept. Its value stems from the fact that a parametric Hamiltonian allows dealing in several cases with a part of a larger system *as if* it were isolated: as a trade-off, some observable effects may occur as gauge-invariant phases. As long as a parametric Hamiltonian is useful for describing the physics of a phenomenon, the corresponding Berry's phase is also very useful and unavoidable. This viewpoint helps in understanding why the Berry's phase is almost ubiquitous, and manifests itself in many unrelated physical problems.

2.3. Connection and curvature

We write again the main definition of the Berry's phase for clarity:

$$\gamma = \oint_C d\varphi = i \oint_C \langle \psi(\xi) | \nabla_\xi \psi(\xi) \rangle \cdot d\xi. \quad (9)$$

The outstanding feature of this definition is that the integrated value γ is gauge invariant and potentially observable, while the integrand itself is largely arbitrary and lacking in any physical meaning.

From a purely mathematical viewpoint, the above concepts are reminiscent of a classical case study in electromagnetism. We can think of our Berry's connection as of an abstract 'vector potential' $\mathbf{A}(\xi)$ in the parameter space. The vector potential has a large gauge arbitrariness, and has no physical meaning, while its circuit integral is indeed gauge invariant and observable. Suppose for the sake of simplicity that ξ is a three-dimensional real parameter, and that our closed curve is contained in a simply connected domain, where the Berry's connection is regular enough (say of class C^2). Then by Stokes's theorem

$$\gamma = \oint_C \mathbf{A}(\xi) \cdot d\xi = - \int \int \mathbf{B}(\xi) \cdot \mathbf{n} d\sigma \quad (10)$$

where $d\sigma$ denotes the area element in ξ -space, and the integral is performed over any surface enclosed by the contour C . The \mathbf{B} -field is defined as

$$\mathbf{B}(\xi) = \nabla_\xi \times \mathbf{A}(\xi) \quad (11)$$

and in the electromagnetic analogue is obviously gauge invariant and observable.

As has been emphasized, Berry's phases occur in many different physical problems. In the general case, there is *no* genuine magnetic field, nor is ξ a space coordinate in three dimensions. Only in some special examples (as e.g. for the AB phase considered below) may it happen that *there is* a magnetic field lurking around, and that ξ is a space coordinate. However, this is by no means general. Only the *mathematical structure* has this close electromagnetic analogue.

To emphasize this fact, I will not use the symbols A and B in the general case. Instead, I will indicate the vector field (or Berry's connection) as \mathcal{X} :

$$\mathcal{X}(\xi) = i\langle\psi(\xi)|\nabla_{\xi}\psi(\xi)\rangle. \quad (12)$$

The analogue of the magnetic field in the most general case is an antisymmetric tensor field of second rank (or equivalently a 2-form [45]) called the 'curvature'. It will be indicated as \mathcal{Y} , and its components are

$$\mathcal{Y}_{\alpha\beta} = \frac{\partial}{\partial\xi_{\alpha}}\mathcal{X}_{\beta}(\xi) - \frac{\partial}{\partial\xi_{\beta}}\mathcal{X}_{\alpha}(\xi) = -2\operatorname{Im}\left\langle\frac{\partial}{\partial\xi_{\alpha}}\psi(\xi)\left|\frac{\partial}{\partial\xi_{\beta}}\psi(\xi)\right.\right\rangle. \quad (13)$$

The curvature is gauge invariant; hence in principle it is physically observable at any given ξ -point.

Whenever the physics of the problem is such that the wavefunction can be taken as real, the curvature \mathcal{Y} vanishes and the Berry's connection is locally exact. Therefore a non-trivial Berry's phase (i.e. $\gamma = \pi$) may only occur if the ξ -domain is *not* simply connected.

If, instead, the nature of the problem makes the wavefunctions unavoidably complex, then the curvature in general does not vanish, and one gets non-trivial Berry's phases even with simply connected ξ -domains. In this case the Berry's phase γ can take *any* value, not just 0 or π .

2.4. The paradigm: the Aharonov–Bohm effect

We illustrate the physical implications of the above mathematics for the simplest and best known example of a geometric phase in quantum mechanics: the AB effect, whose manifestation has been experimentally verified in several occurrences [5].

Suppose we have an electron in a box (infinite potential well) centred at the origin. We take the ground wavefunction as real, and we write it as $\chi(\mathbf{r})$. The time-independent Schrödinger equation is

$$\left[\frac{p^2}{2m} + V(\mathbf{r})\right]\chi(\mathbf{r}) = E\chi(\mathbf{r}). \quad (14)$$

Displacing the centre of the box to position \mathbf{R} changes the Hamiltonian to

$$H(\mathbf{R}) = \frac{p^2}{2m} + V(\mathbf{r} - \mathbf{R}). \quad (15)$$

We will identify the ξ -parameter with the box position \mathbf{R} . Because of translational invariance, the \mathbf{R} -dependence of the state vectors is as follows:

$$\langle\mathbf{r}|\psi(\mathbf{R})\rangle = \chi(\mathbf{r} - \mathbf{R}) \quad (16)$$

while the eigenvalue is \mathbf{R} -independent.

Suppose now that a magnetic field is switched on somewhere in space. Then the Hamiltonian becomes

$$H(\mathbf{R}) = \frac{1}{2m}\left[\mathbf{p} - \frac{e}{c}\mathbf{A}(\mathbf{r})\right]^2 + V(\mathbf{r} - \mathbf{R}) \quad (17)$$

where \mathbf{A} is the vector potential and e is the electron charge. It can be easily verified that a solution of the Schrödinger equation can be formally written in the form

$$\langle\mathbf{r}|\psi(\mathbf{R})\rangle = \exp\left(\frac{ie}{\hbar c}\int_{\mathbf{R}}^{\mathbf{r}}\mathbf{A}(\mathbf{r}')\cdot d\mathbf{r}'\right)\chi(\mathbf{r} - \mathbf{R}). \quad (18)$$

But such a solution is in general *not* a single-valued function of \mathbf{r} , since the phase factor depends on the path. Therefore we restrict ourselves to a less general case, where the magnetic

field is generated by a solenoid: the \mathbf{B} -field is non-zero only within a given cylinder, and we prevent our box from overlapping this cylinder by suitably restricting the domain of \mathbf{R} . With such a choice, the wavefunction, equation (18), is a single-valued function of \mathbf{r} for any fixed \mathbf{R} , and is therefore a legitimate ground wavefunction. As for the dependence on \mathbf{R} , equation (18) only guarantees *local* single valuedness, since the domain is not simply connected: when the system is transported on a closed path winding once round the solenoid, the electron wavefunction picks up a Berry's phase. This phase difference can be actually detected in interference experiments.

The Berry's connection of the problem is

$$\mathcal{X}(\mathbf{R}) = i\langle\psi(\mathbf{R})|\nabla_{\mathbf{R}}\psi(\mathbf{R})\rangle = \frac{e}{\hbar c}\mathbf{A}(\mathbf{R}) + i\int d\mathbf{r}\chi(\mathbf{r}-\mathbf{R})\nabla_{\mathbf{R}}\chi(\mathbf{r}-\mathbf{R}) \quad (19)$$

where the last term vanishes. Therefore the Berry's connection is proportional—in the specific case of the AB effect—to the ordinary vector potential. A gauge transformation in the quantum mechanical sense also coincides with an electromagnetic gauge transformation, which changes \mathbf{A} while leaving \mathbf{B} invariant. In fact, in this example \mathbf{B} is essentially the Berry's curvature. The Berry's phase is

$$\gamma = \frac{e}{\hbar c}\oint_C \mathbf{A}(\mathbf{R}) \cdot d\mathbf{R} \quad (20)$$

and is therefore proportional to the flux of the magnetic field across the *interior* of the solenoid, a space region *not accessed* by the quantum system.

This intriguing result was first published in 1959 [3] and was initially perceived by a part of the physicists' community as a paradox. Instead, the AB effect is an outstanding and inescapable feature of quantum mechanics. As early as 1962–1963, R P Feynman included it in the basic physics topics he chose to teach sophomore classes [4]. Nowadays we know that the AB effect is just the archetype of a general class of phenomena: for these phenomena, the work of Berry gave an elegant and universal description.

The reason for the surprise engendered by the AB paper is that the quantum system resides in a region of space where there is *no* magnetic field \mathbf{B} . In classical electromagnetism, the 'real' field is \mathbf{B} , and the vector potential \mathbf{A} is just a mathematical construct, affected by gauge arbitrariness and non-measurable. A classical particle can only be affected by the field \mathbf{B} at its position: no measurable effect is then possible for a charged particle which orbits outside the solenoid. The quantum result is qualitatively different, and somewhat counterintuitive: the system 'feels'—through measurable effects—the presence of \mathbf{B} in regions not accessed, or in other words it feels—in a gauge-invariant way—the 'reality' of the vector potential \mathbf{A} . We may look at this effect from a very general viewpoint, already pointed out in the original AB paper, and emphasized in Feynman's textbook [4]. The concepts of forces and fields are central in classical physics, but they fade away in quantum physics. The Schrödinger equation does not involve forces and fields, and is necessarily expressed in terms of scalar and vector potentials: the 'real' physical quantity is \mathbf{A} , not \mathbf{B} . Experimental verifications of the effect have been sought for (and found) since the early 1960s [5].

3. Features of the geometric phase

3.1. Parallel transport

In writing equation (7) we have implicitly assumed $|\psi(\xi_1)\rangle \equiv |\psi(\xi_{M+1})\rangle$, i.e. a single-valued state vector along the continuous path. A different—and in a sense more 'natural'—choice is possible. Suppose for instance that the physical problem allows a real wavefunction at any ξ .

If we choose a differentiable gauge, and we require the wavefunction to be indeed real at any ξ , then this gauge is unique, and necessarily yields a vanishing connection everywhere on the path. This implies that when closing the path the state possibly returns on itself with a sign change, i.e. the state vector is multiply valued. This state of affairs is quite general: even in the case where the wavefunctions are complex, there is a special gauge where the phase of the state vector is kept constant as ξ is varied by an infinitesimal amount. This gauge goes under the name of 'parallel transport': it is expedient to illustrate it by means of perturbation theory.

We need to restore—for a while only—the 0 subscript for the ground state, in order to be able to deal with the excited states as well. We therefore rewrite the Berry's connection as

$$\mathcal{X}(\xi) = i\langle\psi_0(\xi)|\nabla_{\xi}\psi_0(\xi)\rangle. \quad (21)$$

At first sight, the most natural tool for evaluating the gradient in equation (21) seems to be perturbation theory:

$$|\psi_0(\xi + \Delta\xi)\rangle \simeq |\psi_0(\xi)\rangle + \sum_{n \neq 0} |\psi_n(\xi)\rangle \frac{\langle\psi_n(\xi)|H(\xi + \Delta\xi) - H(\xi)|\psi_0(\xi)\rangle}{E_0(\xi) - E_n(\xi)} \quad (22)$$

$$|\nabla_{\xi}\psi_0(\xi)\rangle = \sum_{n \neq 0} |\psi_n(\xi)\rangle \frac{\langle\psi_n(\xi)|\nabla_{\xi}H(\xi)|\psi_0(\xi)\rangle}{E_0(\xi) - E_n(\xi)}. \quad (23)$$

Since the excited states are orthogonal to the ground one, we get in this way a vanishing Berry's connection at any ξ .

The key point is that in the standard expression of perturbation theory, equation (22), a specific choice of the gauge is implicitly made: namely, the parallel transport, which forces the infinitesimal change in $|\psi_0(\xi)\rangle$ to be orthogonal to $|\psi_0(\xi)\rangle$ itself. This is a perfectly legitimate choice, but when we go continuously round the closed path the state vector in general returns on itself only modulo a phase factor, which is indeed the Berry's phase of the path. In other words parallel transport is incompatible with a state vector which is globally single valued—as a function of ξ —on the closed path.

In the previous sections, we have used a single-valued state vector throughout. This choice can be made compatible with perturbation theory only if we rewrite equation (22) in the more general form

$$|\psi_0(\xi + \Delta\xi)\rangle \simeq e^{-i\Delta\varphi} \left[|\psi_0(\xi)\rangle + \sum_{n \neq 0} |\psi_n(\xi)\rangle \frac{\langle\psi_n(\xi)|H(\xi + \Delta\xi) - H(\xi)|\psi_0(\xi)\rangle}{E_0(\xi) - E_n(\xi)} \right] \quad (24)$$

where $\Delta\varphi$ is an arbitrary gauge phase. Since we want it to vanish when $\Delta\xi = 0$, the gauge phase must be linear in $\Delta\xi$; we thus write equation (24) to leading order as

$$|\psi_0(\xi + \Delta\xi)\rangle \simeq (1 - i\Delta\varphi)|\psi_0(\xi)\rangle + \sum_{n \neq 0} |\psi_n(\xi)\rangle \frac{\langle\psi_n(\xi)|H(\xi + \Delta\xi) - H(\xi)|\psi_0(\xi)\rangle}{E_0(\xi) - E_n(\xi)}. \quad (25)$$

In this way we easily get an arbitrary phase difference—and in general a non-vanishing Berry's connection—since

$$\langle\psi_0(\xi)|\psi_0(\xi + \Delta\xi)\rangle \simeq 1 - i\Delta\varphi. \quad (26)$$

A suitable choice of the gauge phase in equation (24) restores a single-valued behaviour of the state vector round the path.

3.2. Computing a Berry's phase

It has been shown in the previous section that a straightforward implementation of perturbation theory cannot provide the Berry's phase as a loop integral of the connection. The discrete

approach of section 2.1, instead, is the practical tool which is implemented in actual calculations. Berry's phases in condensed matter are routinely evaluated as [39]

$$\gamma = \sum_{s=1}^M \Delta\varphi_{s,s+1} = -\text{Im} \log \prod_{s=1}^M \langle \psi(\xi_s) | \psi(\xi_{s+1}) \rangle \quad (27)$$

in the limit of large M . This expression, besides its usefulness, is even important as a matter of principle.

The discrete formulation is more general than the (more traditional [2]) continuum one, since it does *not* assume any regularity of the phase along the path. In fact, following the present discrete approach, one has a well defined limiting γ even if the phase of $|\psi(\xi + \Delta\xi)\rangle$ is allowed to vary discontinuously for $\Delta\xi \rightarrow 0$. The differentiability of the local phase becomes unnecessary for defining a Berry's phase over a continuum path.

A pathological behaviour of the local phase has some relevance for numerical work, and actually manifests itself in practical calculations. Suppose the state $|\psi(\xi)\rangle$ is obtained by numerical diagonalization over a finite basis. Then the phase at each ξ -point is chosen—essentially at random—by the diagonalization routine, and shows no regularity at all when the set becomes denser and denser. Notwithstanding this, the phase γ defined as in equation (27) does converge to a meaningful value.

The role of perturbation theory could be rescued if, instead of writing γ as the loop integral of the connection, we write it as the surface integral of the curvature. In fact the latter quantity, being gauge invariant, can be safely expressed by means of perturbation theory within the parallel-transport gauge. Using equation (23) we recast equation (13) as

$$\mathcal{Y}_{\alpha\beta} = -2 \text{Im} \sum_{n \neq 0} \frac{\langle \psi_0(\xi) | \partial H(\xi) / \partial \xi_\alpha | \psi_n(\xi) \rangle \langle \psi_n(\xi) | \partial H(\xi) / \partial \xi_\beta | \psi_0(\xi) \rangle}{[E_0(\xi) - E_n(\xi)]^2}. \quad (28)$$

This expression could be actually implemented, although its computational appeal is very limited: it would require in fact calculating all the excited states at any ξ .

The expression of equation (28) shows that the curvature is singular at the values of ξ where the ground state is degenerate with the first excited state. This fact has important consequences, particularly in those cases where the wavefunction can be taken as real. In such cases, the Berry's phase on a curve C can assume the value $\gamma = \pi$ only if C encircles such degeneracy points. From this follows that if ξ is a d -dimensional parameter, the manifold of the singular points of the curvature needs to be at least $(d - 2)$ -dimensional to ensure multiple connectiveness, and hence to produce a non-trivial Berry's phase. In the cases where the wavefunction must be taken as complex, instead, singularities are no longer required to produce a non-trivial Berry's phase.

3.3. The open-path geometric phase

We start again from the discrete four-point introductory example of section 2.1, illustrated in figure 2, but now we consider an open path, as in figure 4. It is then obvious that the total phase difference along the path *cannot* be gauge invariant in general: but we only consider the very special case where—because of some specific symmetry of the physical problem—the Hamiltonian at ξ_4 is unitarily equivalent to the one at ξ_1 ; i.e.,

$$H(\xi_4) = W^{-1} H(\xi_1) W \quad (29)$$

where W is a *fixed* unitary operator, given by the symmetry of the physical problem. Then it is clear that we might like to choose the eigenstate at ξ_4 as $|\psi(\xi_4)\rangle = W^{-1} |\psi(\xi_1)\rangle$, thus imposing a preferred phase relationship amongst the eigenstates at the initial and final points

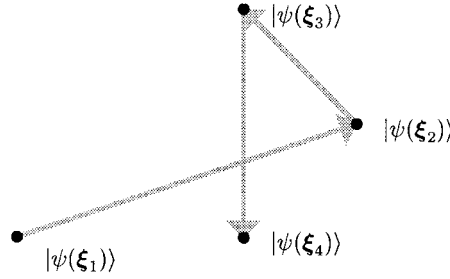


Figure 4. An open path in parameter space

of the path. In this way we are able to define a gauge-invariant phase difference along the open path as

$$\begin{aligned} \gamma &= \Delta\varphi_{12} + \Delta\varphi_{23} + \Delta\varphi_{34} \\ &= -\text{Im} \log \langle \psi(\xi_1) | \psi(\xi_2) \rangle \langle \psi(\xi_2) | \psi(\xi_3) \rangle \langle \psi(\xi_3) | W^{-1} | \psi(\xi_1) \rangle. \end{aligned} \quad (30)$$

We will call such a phase the open-path geometric phase.

I have used four points for pedagogical purposes. It is worth noticing that, in the closed-path case, one needs at least *three* points (and three Hamiltonian diagonalizations) in order to get a non-trivial γ -value. At variance with this, in the open-path case any number of points provides in principle a non-trivial γ , i.e. a γ -value different from zero (modulo 2π).

The open-path phase is easily generalized to the continuum limit like the ordinary (closed-path) Berry's phase:

$$\gamma = i \int_C \langle \psi(\xi) | \nabla_\xi \psi(\xi) \rangle \cdot d\xi \quad (31)$$

where now C is an open curve connecting the distinct points ξ_{initial} and ξ_{final} . This phase depends on the path C and is gauge invariant, provided the gauges are so chosen that

$$|\psi(\xi_{\text{final}})\rangle = W^{-1} |\psi(\xi_{\text{initial}})\rangle \quad (32)$$

thus enforcing relative phase coherence between the final and initial points. This peculiar kind of geometric phase was first discussed by Zak [46].

The alternate viewpoint would be to choose the parallel-transport gauge, in which case the line integral is formally zero, but the state vector at the final point no longer fulfils equation (32). Instead, parallel transport yields

$$|\psi(\xi_{\text{final}})\rangle = e^{i\gamma} W^{-1} |\psi(\xi_{\text{initial}})\rangle \quad (33)$$

where again γ is the geometric phase of the problem.

3.4. The single-point Berry's phase

The extreme case for the open-path geometric phase is a two-point path, requiring only *one* Hamiltonian diagonalization:

$$\gamma = -\text{Im} \log \langle \psi(\xi_1) | W^{-1} | \psi(\xi_1) \rangle. \quad (34)$$

Whether even in this extreme case the phase γ still deserves the name of a geometric phase is a matter of pure semantics: indeed, the form of equation (34) is often referred to as the 'single-point Berry's phase' [26]. It is convenient to perform a double change of sign in equation (34) and to express the single-point Berry's phase as

$$\gamma = \text{Im} \log \langle \psi | W | \psi \rangle \quad (35)$$

where the parameter ξ is no longer needed.

The name ‘single-point Berry’s phase’ looks almost like an oxymoron. In fact we have stressed above that a geometric phase is an observable which cannot be cast as the expectation value of any operator, while instead the main ingredient of equation (35) *is* indeed the expectation value of the operator W (unitary and *not* Hermitian): but the observable is the phase of this expectation value, and *not* the expectation value itself. We could thus say that equation (35) is a Berry’s phase ‘in disguise’.

The single-point Berry’s phase manifests itself as a very important observable in modern electronic structure theory. In fact, the macroscopic polarization of an extended system can be expressed precisely in the form of equation (35), where ψ must be identified with the many-body wavefunction (obeying Born–von Kármán boundary conditions), and the unitary operator W with a suitable many-body operator, first introduced in reference [25], and sometimes referred to as the ‘many-body phase operator’.

A large part of this review is about the modern theory of polarization, and we base our presentation very much on the concept of the single-point Berry’s phase. This very recent development is important both as a matter of principle [25–27, 32] and as a tool used in actual computations [47, 48]. We will show in section 6.6 how the disguised geometric phase of equation (35) is very naturally linked to other, more traditional, geometric phases which occurred in previous formulations of the polarization theory [37–39].

4. The electronic Berry’s phase

4.1. Wavefunctions and density matrices

We wrote down a Hamiltonian at the very beginning, equation (1), but then the Hamiltonian quickly disappeared from our consideration. In fact, the only role of the Hamiltonian is to single out the relevant state in the Hilbert space at a given ξ . We have assumed that the relevant state is the ground state of a given Hamiltonian, but strictly speaking this is unnecessary. In order to define a Berry’s phase, it is enough to specify the projector over the one-dimensional manifold spanned by the relevant state, which coincides with the (pure-state) density matrix:

$$\rho(\xi) = |\psi(\xi)\rangle\langle\psi(\xi)| \quad (36)$$

as a function of ξ . In equation (36), we assume $|\psi(\xi)\rangle$ normalized, and having an arbitrary phase. Of course the density matrix is gauge invariant.

The Berry’s phase of the problem can then be cast in a transparent way as an explicit function of $\rho(\xi)$. Consider the simple discrete example discussed at the beginning. Then equation (4) can be recast as the quantum mechanical trace of a product of density matrices:

$$\gamma = -\text{Im} \log \text{tr} \rho(\xi_1)\rho(\xi_2)\rho(\xi_3)\rho(\xi_4) \quad (37)$$

where the gauge invariance is now transparent. Analogously, the single-point Berry’s phase of equation (35) is cast as

$$\gamma = \text{Im} \log \text{tr} W \rho. \quad (38)$$

In the present work we are mostly concerned with cases where the state vector $|\psi(\xi)\rangle$ is a many-electron wavefunction. Since from now on it is essential to distinguish the N -electron wavefunction from the one-particle orbitals, we will use a capital letter for the former. The wavefunction is therefore written in Schrödinger representation as

$$|\Psi(\xi)\rangle \rightarrow \langle \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N | \Psi(\xi) \rangle \quad (39)$$

where $\mathbf{x}_i \equiv (\mathbf{r}_i, \sigma_i)$ are the space and spin coordinates of the i th electron.

As discussed above, a Hamiltonian is not strictly needed. However, it is convenient to assume that $|\Psi(\xi)\rangle$ is the ground state of a many-electron Hamiltonian $H(\xi)$, whose dependence on the parameter ξ can be due to the one-body term, to the two-body term, or to both. We also assume that the ground state is non-degenerate at any point of the path. A relevant issue in the framework of Berry's phases is the kind of boundary conditions we assume for the wavefunction. First of all—as observed above—the boundary conditions must be ξ -independent. Then we must distinguish between two rather different cases.

- (i) *Case one.* If the electronic system is *finite*, such as for a molecule, we typically assume \mathcal{L}^2 (square-integrable) wavefunctions. Supposing there is no magnetic field (and neglecting spin-orbit effects), the wavefunction can then be safely taken as real. In this case the only allowed value for a non-trivial Berry's phase is π (modulo 2π): this amounts to saying that the wavefunction changes sign when transported parallel round the closed path. This occurrence is the so-called molecular AB effect, whose essential features are discussed below, in section 5.1; for a comprehensive review, refer to reference [7], and for a textbook presentation to reference [34].
- (ii) *Case two.* Whenever the wavefunctions must be complex, the Berry's phase can take any value. This happens in the presence of magnetic fields, such as in the cases discussed in section 2.4 above and in section 5.2 below, or as in the quantum Hall effect, not covered by this review. But wavefunctions are unavoidably complex—even in the absence of a magnetic field—in a non-centrosymmetric system whenever we assume *periodic Born-von Kármán boundary conditions* for the wavefunction, as is almost mandatory in dealing with extended systems. This is precisely the case dealt with in section 6; indeed, a Berry's phase γ measures the macroscopic polarization of a non-centrosymmetric many-electron system in suitable units.

The density matrix in equation (37) must be understood as the full many-body density matrix $|\Psi(\xi)\rangle\langle\Psi(\xi)|$, acting on the N -particle Hilbert space. However, the physical properties which will be related to an electronic Berry's phase are typically of the one-body kind: currents, charge transport, polarization, and so on. So, it is quite natural to ask the question of whether the Berry's phase can be expressed only in terms of the *reduced one-body* electronic density matrix

$$\rho_{\xi}^{(1)}(\mathbf{x}, \mathbf{x}') = N \int d\mathbf{x}_2 \cdots d\mathbf{x}_N \langle \mathbf{x}, \mathbf{x}_2, \dots, \mathbf{x}_N | \Psi(\xi) \rangle \langle \Psi(\xi) | \mathbf{x}', \mathbf{x}_2, \dots, \mathbf{x}_N \rangle \quad (40)$$

as a function of the parameter ξ . This would be a major step forward, since the N -electron wavefunction contains messy and redundant information [49]. The answer to this major question is unfortunately 'no' in general: Berry's phase is essentially a many-body observable. In fact, the reduced density matrix is the partial trace over $N - 1$ variables of the many-body one; but in equation (37) we are in general not allowed to take partial traces *before* evaluating the operator product. There is however a very important exception, discussed next: the case of independent electrons. This concerns mean-field theories such as the Hartree-Fock or Kohn-Sham theory. It also concerns some *approximate* forms of many-body Green's function theory [50, 51].

4.2. Independent electrons

We study here the case where the N -electron wavefunction is uncorrelated, and has therefore the form of a single Slater determinant:

$$|\Psi(\xi)\rangle = \frac{1}{\sqrt{N!}} |\psi_1(\xi)\psi_2(\xi) \cdots \psi_N(\xi)| \quad (41)$$

where $\psi_j(\xi)$ are the occupied spin orbitals, taken as orthonormal. As stated above, we refer to either Hartree–Fock orbitals or to Kohn–Sham orbitals.

An arbitrary unitary transformation of the occupied spin orbitals amongst themselves provides *the same* many-body wavefunction, times a constant gauge phase factor [49]. What determines the relevant state at a given ξ is only the *projector* over the N -dimensional manifold of the occupied one-particle spin orbitals, which coincides with the one-body reduced density matrix of the uncorrelated wavefunction:

$$\rho_\xi^{(1)}(x, x') = \sum_{i=1}^N \langle x | \psi_i(\xi) \rangle \langle \psi_i(\xi) | x' \rangle. \quad (42)$$

We will identify a gauge transformation with an $N \times N$ unitary matrix $U(\xi)$, which leaves the reduced density matrix invariant.

Next we wish to express the Berry's phase of the many-particle state, as defined in section 2, in terms of the spin orbitals. We start from equation (3):

$$\Delta\varphi_{12} = -\text{Im} \log \langle \Psi(\xi_1) | \Psi(\xi_2) \rangle \quad (43)$$

and we recognize that the overlap amongst two determinants is equal to the determinant of the overlap matrix amongst the occupied spin orbitals:

$$\langle \Psi(\xi_1) | \Psi(\xi_2) \rangle = \det S(\xi_1, \xi_2). \quad (44)$$

Explicitly, the elements of the $N \times N$ matrix S are

$$S_{ij}(\xi_1, \xi_2) = \langle \psi_i(\xi_1) | \psi_j(\xi_2) \rangle. \quad (45)$$

Therefore the Berry's phase of the N -electron state over an M -point discrete closed path can be written as

$$\gamma = \sum_{s=1}^M \Delta\varphi_{s,s+1} = -\text{Im} \log \prod_{s=1}^M \det S(\xi_s, \xi_{s+1}). \quad (46)$$

It can be straightforwardly verified that γ defined in equation (46) is gauge invariant, provided the spin orbitals at $s = 1$ and $s = M + 1$ are *the same* (same phases, same N -ordering).

Taking the continuum limit of equation (46) is straightforward. The Berry's connection is

$$\mathcal{X}(\xi) = i \langle \Psi(\xi) | \nabla_\xi \Psi(\xi) \rangle = i \nabla_{\xi'} \log \det S(\xi, \xi') \Big|_{\xi'=\xi}. \quad (47)$$

We may also exploit the well known matrix identity [52]:

$$\det \exp A = \exp \text{tr} A \quad (48)$$

which applied to $A = \log S$ yields

$$\nabla_{\xi'} \log \det S(\xi, \xi') = \text{tr} \nabla_{\xi'} \log S(\xi, \xi') = \text{tr} \{ S^{-1}(\xi, \xi') \nabla_{\xi'} S(\xi, \xi') \}. \quad (49)$$

Since the overlap at $\xi = \xi'$ coincides with the identity, we recast the Berry's connection as

$$\mathcal{X}(\xi) = i \text{tr} \nabla_{\xi'} S(\xi, \xi') \Big|_{\xi'=\xi} = i \sum_{i=1}^N \langle \psi_i(\xi) | \nabla_\xi \psi_i(\xi) \rangle \quad (50)$$

and the continuum Berry's phase as

$$\gamma = \oint_C \mathcal{X}(\xi) \cdot d\xi = i \sum_{i=1}^N \oint_C \langle \psi_i(\xi) | \nabla_\xi \psi_i(\xi) \rangle \cdot d\xi. \quad (51)$$

We recognize in equation (51) precisely the sum of the Berry's phases of each individual spin orbital. It may happen that the occupied orbitals are energy degenerate at some ξ -points, in which case there is some ambiguity in defining the ξ -dependence of each orbital separately.

This fact does not cause any problem for the Berry's phase, which is uniquely defined once the projector on the occupied spin orbitals is provided as a function of ξ . A sufficient condition for having a well defined Berry's phase is that the highest occupied orbital is separated by a finite gap from the lowest unoccupied one at any point of the path.

Finally, it is important to point out that the separation of the many-body Berry's phase into the sum of single-body Berry's phases of the orbitals is valid, in general, only in the continuum limit. In numerical work the integral is evaluated as a finite sum, actually using equation (46), and it is essential to retain the determinantal form in order to preserve gauge invariance.

4.3. Bloch orbitals

So far, we have considered the most general many-electron system, and the most general parameter ξ occurring in the electronic Hamiltonian. We now consider a very special case: independent electrons in a crystalline system.

The orbitals have the Bloch form:

$$\psi_{nq}(\mathbf{r} + \boldsymbol{\tau}) = e^{i\mathbf{q}\cdot\boldsymbol{\tau}} \psi_{nq}(\mathbf{r}) \quad (52)$$

where $\boldsymbol{\tau}$ is a lattice translation and \mathbf{q} is the Bloch quasimomentum: a continuous variable in the thermodynamic limit. The Bloch orbitals are solutions of mean-field (Hartree–Fock or Kohn–Sham) Schrödinger equation:

$$\left[\frac{1}{2m} p^2 + V(\mathbf{r}) \right] \psi_{nq}(\mathbf{r}) = \varepsilon_n(\mathbf{q}) \psi_{nq}(\mathbf{r}) \quad (53)$$

where for the sake of simplicity we assume a *local* potential. No major complication arises for non-local potentials, such as are needed for state-of-the-art pseudopotential calculations, or for dealing with Fock exchange. Notice that we are thus using a \mathbf{q} -independent Hamiltonian, equation (53), and \mathbf{q} -dependent (quasiperiodic over the elementary cell) boundary conditions, equation (52).

The Bloch functions can be rewritten as

$$\psi_{nq}(\mathbf{r}) = e^{i\mathbf{q}\cdot\mathbf{r}} u_{nq}(\mathbf{r}) \quad (54)$$

where the u s obey *periodic* boundary conditions over the elementary cell:

$$u_{nq}(\mathbf{r} + \boldsymbol{\tau}) = u_{nq}(\mathbf{r}). \quad (55)$$

The Schrödinger equation takes then the form

$$\left[\frac{1}{2m} (\mathbf{p} + \hbar\mathbf{q})^2 + V(\mathbf{r}) \right] u_{nq}(\mathbf{r}) = \varepsilon_n(\mathbf{q}) u_{nq}(\mathbf{r}). \quad (56)$$

All this is very trivial indeed, but I wish to point out the dual role of the boundary conditions under such transformation. Switching from equations (52) and (53) to equations (55) and (56) one maps the eigenvalue problem into a \mathbf{q} -dependent Hamiltonian, with \mathbf{q} -independent boundary conditions. This is precisely the typical case where it makes sense investigating the possible occurrence of a geometric phase: the Hamiltonian depends on a parameter, while the eigenstates all reside in the same Hilbert space (that is, obey parameter-independent boundary conditions). We therefore identify the parameter ξ with the Bloch vector \mathbf{q} : notice also that \mathbf{q} appears in the Hamiltonian, equation (56), as a kind of 'vector potential', although no magnetic field is present.

According to our general formulation, the Berry's connection of the n th band is

$$\mathcal{X}(\mathbf{q}) = i \langle u_{nq} | \nabla_{\mathbf{q}} u_{nq} \rangle \quad (57)$$

while the corresponding curvature is

$$\mathcal{Y}_{\alpha\beta}(\mathbf{q}) = \frac{\partial}{\partial q_\alpha} \mathcal{X}_\beta(\mathbf{q}) - \frac{\partial}{\partial q_\beta} \mathcal{X}_\alpha(\mathbf{q}) = -2 \operatorname{Im} \left\langle \frac{\partial}{\partial q_\alpha} u_{n\mathbf{q}} \left| \frac{\partial}{\partial q_\beta} u_{n\mathbf{q}} \right. \right\rangle. \quad (58)$$

This antisymmetric Cartesian tensor can be written in vector form as

$$\Omega_\alpha(\mathbf{q}) = \frac{1}{2} \epsilon_{\alpha\beta\gamma} \mathcal{Y}_{\beta\gamma}(\mathbf{q}) = -\epsilon_{\alpha\beta\gamma} \operatorname{Im} \left\langle \frac{\partial}{\partial q_\beta} u_{n\mathbf{q}} \left| \frac{\partial}{\partial q_\gamma} u_{n\mathbf{q}} \right. \right\rangle \quad (59)$$

or in more compact notation:

$$\Omega(\mathbf{q}) = -\operatorname{Im} \langle \nabla_{\mathbf{q}} u_{n\mathbf{q}} | \times | \nabla_{\mathbf{q}} u_{n\mathbf{q}} \rangle = i \langle \nabla_{\mathbf{q}} u_{n\mathbf{q}} | \times | \nabla_{\mathbf{q}} u_{n\mathbf{q}} \rangle. \quad (60)$$

The curvature is gauge invariant and, therefore, corresponds in principle to a physical observable. In fact, Ω occurs in the semiclassical dynamics of crystalline electrons, discussed below in sections 5.3 and 5.4.

4.4. Zak's phase

We are interested in the integral of the Berry's connection for Bloch electrons, equation (57), over an open curve C connecting the points $\mathbf{q}_{\text{initial}}$ and $\mathbf{q}_{\text{final}}$:

$$\gamma = i \int_C \langle u_{n\mathbf{q}} | \nabla_{\mathbf{q}} u_{n\mathbf{q}} \rangle \cdot d\mathbf{q}. \quad (61)$$

We only consider the case where the difference $\mathbf{q}_{\text{final}} - \mathbf{q}_{\text{initial}}$ is a reciprocal-lattice vector \mathbf{G} ; at these two points, the Bloch eigenfunctions obey the same Schrödinger equation, equation (53), and the same boundary conditions, equation (52). It is therefore very natural to impose $|\psi_{n\mathbf{q}_{\text{final}}}\rangle \equiv |\psi_{n\mathbf{q}_{\text{initial}}}\rangle$, where *the same phase* is chosen: this goes under the name of 'periodic gauge'. Owing to equation (54), we have for u_n the relationship

$$u_{n\mathbf{q}_{\text{final}}}(\mathbf{r}) = e^{-i\mathbf{G}\cdot\mathbf{r}} u_{n\mathbf{q}_{\text{initial}}}(\mathbf{r}). \quad (62)$$

Comparing now equations (61) and (62) to equations (31) and (32), we recognize in γ an open-path geometric phase, where the unitary operator W of equation (32) has to be identified with the multiplicative operator $e^{i\mathbf{G}\cdot\mathbf{r}}$ in Schrödinger representation.

If there are several occupied bands, we may consider the sum of their geometric phases as a whole:

$$\gamma = i \sum_{n=1}^{n_b} \int_C \langle u_{n\mathbf{q}} | \nabla_{\mathbf{q}} u_{n\mathbf{q}} \rangle \cdot d\mathbf{q} \quad (63)$$

where n_b is the number of occupied bands. In the case of band crossings, there is some arbitrariness in identifying the n th band as a function of \mathbf{q} . According to the discussion which follows equation (51), the total phase is gauge invariant and unaffected by such arbitrariness, provided that the n_b occupied bands do not cross the *empty* ones at any point of the path.

The occurrence of the geometric phase γ , equation (61), in the band theory of solids was first discovered by Zak [33], who identified its observable effects in Wannier–Stark ladders (see below, section 5.4). Only much later, owing to King-Smith and Vanderbilt [37], was it realized that Zak's phase has another outstanding manifestation: namely, macroscopic polarization in crystalline dielectrics. More will be said below about this (section 6).

A final comment is in order about how to compute a multi-band Zak's phase, equation (63); more details can be found in references [21, 39]. We define the overlap matrix

$$S_{nn'}(\mathbf{q}, \mathbf{q}') = \langle u_{n\mathbf{q}} | u_{n'\mathbf{q}'} \rangle. \quad (64)$$

Before discretizing the line integral, it is essential to convert the trace of S appearing in equation (63) into a determinant, using the same algebra as in section 4.2. Discretizing the open path with $M + 1$ points \mathbf{q}_s we have

$$\gamma = i \sum_{n=1}^{n_b} \int_C \langle u_{nq} | \nabla_q u_{nq} \rangle \cdot d\mathbf{q} \rightarrow -\text{Im} \log \prod_{s=0}^{M-1} \det S(\mathbf{q}_s, \mathbf{q}_{s+1}) \quad (65)$$

where, according to equation (62), we define

$$S_{nn'}(\mathbf{q}_{M-1}, \mathbf{q}_M) = \langle u_{nq_{M-1}} | e^{-iG \cdot r} | u_{n'q_0} \rangle. \quad (66)$$

5. Manifestations of Berry's phase

5.1. The molecular Aharonov–Bohm effect

The effect which nowadays is called the molecular AB effect [7] was first predicted as early as 1958 [8], and historically is the first occurrence of a geometric phase in quantum mechanics [6, 15]. It concerns the ionic motion in the adiabatic approximation, and is observable through some peculiar features of the rotovibrational spectra.

The real parameter ξ coincides with a set of ionic coordinates, and is in general d -dimensional; the state vector $|\Psi(\xi)\rangle$ is the electronic wavefunction in the Born–Oppenheimer approximation, and Berry's connection, equation (12), acts formally as a vector potential in the effective Hamiltonian which governs the ionic motion. Notice that there is no *genuine* magnetic field in this problem, hence the name 'geometric vector potential' which is often used. Since the wavefunction can be taken as real, the only possible value for a non-trivial Berry's phase is π .

We start from the complete Hamiltonian \mathcal{H} of an isolated molecular system, and we explicitly separate the nuclear kinetic energy:

$$\mathcal{H}(\xi, [\mathbf{x}]) = \frac{1}{2} \sum_{\alpha, \beta=1}^d \mathbf{M}_{\alpha\beta}^{-1} p_\alpha p_\beta + H(\xi, [\mathbf{x}]) \quad (67)$$

where $[\mathbf{x}]$ indicates the electronic degrees of freedom collectively, $p_\alpha = -i\hbar \partial/\partial \xi_\alpha$ is the canonical momentum conjugate to ξ_α , and the inverse mass matrix \mathbf{M}^{-1} in general may be a function of ξ , but not of the momenta.

The Born–Oppenheimer approximation starts by writing the eigenfunctions of equation (67) in the Schrödinger representation as the product $\langle [\mathbf{x}] | \Psi(\xi) \rangle \Xi(\xi)$. Our aim is obtaining an effective Schrödinger equation for the nuclear wavefunction $\Xi(\xi)$, where the electronic degrees of freedom have been integrated out. We start by considering the effect of the canonical nuclear momentum p on the product *ansatz*:

$$p|\Psi(\xi)\rangle \Xi(\xi) = -i\hbar |\Psi(\xi)\rangle \nabla_\xi \Xi(\xi) - i\hbar |\nabla_\xi \Psi(\xi)\rangle \Xi(\xi). \quad (68)$$

We then multiply by the electronic eigenbra $\langle \Psi(\xi) |$ on the left, thus integrating over the electronic degrees of freedom. We get the effective nuclear momentum π acting on Ξ as

$$\pi \Xi(\xi) = -i\hbar [\nabla_\xi + \langle \Psi(\xi) | \nabla_\xi \Psi(\xi) \rangle] \Xi(\xi). \quad (69)$$

We easily recognize the 'geometric vector potential' anticipated above.

Whenever the timescales of nuclear and electronic motions are well separated the coupling between different electronic states can be neglected, and the adiabatic approximation allows us to treat the slow variable ξ in $H(\xi, [\mathbf{x}])$ as a classical parameter. The electronic eigenvalue

$E(\xi)$ of a given state (say the ground one) plays therefore the role of a (scalar) potential for nuclear motion, whose effective Hamiltonian acting on $\Xi(\xi)$ is then

$$H_{\text{eff}} = \frac{1}{2} \sum_{\alpha, \beta=1}^d \mathbf{M}_{\alpha\beta}^{-1} \pi_{\alpha} \pi_{\beta} + E(\xi). \quad (70)$$

Whenever the ionic motion is purely classical, governed by Newton's equation, the vector-potential-like term in equation (69) is irrelevant: the corresponding 'magnetic field' is in fact identically vanishing along the nuclear trajectory on the Born–Oppenheimer surface.

The interesting effects arise when one quantizes the nuclear degrees of freedom, using equation (70) as the effective Hamiltonian in the Schrödinger equation for nuclear motion, whose eigenvalues provide the rotovibrational spectrum of the molecule. Notice that we have explicitly neglected any coupling to the higher electronic states, and thus we apparently are strictly within the adiabatic approximation. But in fact the Berry's connection in equation (69) accounts for effects which are absent within the 'naive' adiabatic approximation, and therefore it amounts to the first step beyond it.

In most of the pre-Berry literature, the vector-potential-like term in equation (69) is neglected, or not even mentioned: π is directly identified with the canonical nuclear momentum $-i\hbar \nabla_{\xi}$. The reasons are the following. In a molecular problem, one typically assumes \mathcal{L}^2 boundary conditions, i.e. vanishing wavefunctions at infinity. Then, in the absence of an external magnetic field and of spin–orbit interaction, the electronic wavefunctions can be taken as real at any given ξ , apparently with no loss of generality. For real electronic wavefunctions, the nuclear kinetic momentum π is identical to the canonical one p . The problem arises when we follow the adiabatic evolution along a closed path in ξ -space.

If we insist in requiring a real electronic wavefunction, having a continuous dependence on ξ , then when we close the loop the electronic wavefunction may undergo a sign change. In such a case, corresponding to an electronic Berry's phase of π , the ionic wavefunction $\Xi(\xi)$ must undergo a sign change as well, since the *total* wavefunction must be single-valued. This sign change is a boundary condition which modifies the spectrum of the effective nuclear Hamiltonian H_{eff} , and which has in fact remarkable observable effects on the rotovibrational spectrum of some molecules, first predicted in 1958 [8]. A sign change in the nuclear wavefunction corresponds in fact to half-odd-integer quantum numbers for some rotovibrational levels.

The alternate description of the same phenomenon—originally due to Mead and Truhlar [53]—requires a single-valued electronic wavefunction, which can be simply realized if we allow it to be complex. In this case the Berry's connection $\mathcal{X}(\xi) = i\langle \Psi(\xi) | \nabla_{\xi} \Psi(\xi) \rangle$ no longer vanishes, and we must write the kinetic momentum as

$$\pi = -i\hbar \nabla_{\xi} - \hbar \mathcal{X}(\xi). \quad (71)$$

The effective Hamiltonian, equation (70), now explicitly contains the geometric vector potential \mathcal{X} . The classical motion of the ions on the Born–Oppenheimer surface is *unaffected* by this extra term, since the corresponding 'magnetic field' \mathcal{Y} , equation (13), vanishes on the classical trajectory of the nuclei. However, as soon as we wish to quantize the nuclear rotovibrations, the Berry's connection shows up in producing non-trivial interference effects in the nuclear wavefunction. These may dramatically affect the spectrum, allowing for half-odd-integer quantum numbers. When the phenomenon is described in this way, one sees very clearly the analogy with the standard AB effect. In fact, although there is no *physical* magnetic field B , we have a curvature \mathcal{Y} which plays the same role. The curvature \mathcal{Y} vanishes on the classical trajectory, but we get a non-trivial Berry's phase π whenever the nuclear trajectory winds once round a locus of δ -like singularities of \mathcal{Y} . What determines the Berry's phase in fact is the

flux of the curvature through a surface whose boundary is the nuclear trajectory, although the nuclear configuration never visits the singular points. As explained in section 2, the effect is possible only if the manifold of the singular points is at least $(d - 2)$ -dimensional: such singular points are those where the given Born–Oppenheimer surface becomes degenerate with the Born–Oppenheimer surface of a different electronic state. Such loci of singularities are called ‘conical intersections’ [6, 7].

The previous considerations show that the effective Hamiltonian for ionic motion, equation (70), picks up the important topological feature of the electronic spectrum, and in particular of the lowest excited state. Surprisingly, this happens despite the fact that equation (70) has been indeed derived with explicit neglect of any coupling to the higher electronic states.

The smallest molecular system where the molecular AB effect is possible is a trimer, having three internal coordinates (e.g. the three internuclear distances). An experiment performed in 1986 indicated the presence of a Berry phase in Na_3 clusters [54]: the interpretation of these data was later challenged [55–58]. A less controversial case appears to be Li_3 , for which more recent measurements are available [59].

Finally, let me mention that a Berry's phase of the same kind as discussed here for the bound states of the ions manifests itself even in processes of scattering between molecular species. The phase affects the scattering amplitudes and the rates for chemical reactions. The simplest and most studied [60] example of this class of phenomena is the reaction $\text{H} + \text{H}_2 \rightarrow \text{H}_2 + \text{H}$.

5.2. Adiabatic approximation in a magnetic field

We study in this section the case where a genuine magnetic field, generated by some external source, acts on the molecular system. The Hamiltonian of equation (67) is then modified by the addition of a vector potential term in the kinetic energies of both the nuclei and the electrons. Proceeding as in the zero-field case, one writes an *ansatz* wavefunction and arrives at the effective Hamiltonian for the ionic motion, equation (70), where an extra term must be added to the kinetic momentum π of equation (71). There are thus *two* vector potentials in the effective nuclear Hamiltonian: a geometric one, and a genuinely magnetic one.

However, with respect to the zero-field case, there is a qualitative difference whose importance is overwhelming. Since the electronic Hamiltonian is no longer invariant under time reversal, the electronic wavefunction is necessarily complex. In this case no degeneracy between different Born–Oppenheimer surfaces is needed in order to produce a non-trivial Berry's phase: on the contrary, the Berry's phase will be in general non-zero along *any* path in ξ -space. Furthermore, the Berry's curvature \mathcal{Y} , equation (13), will be in general non-vanishing everywhere, at variance with the zero-field case where it has only δ -like singularities (on the loci of degeneracy, also known as conical intersections).

Suppose we are interested in the ionic motion at the purely classical level. The Hamiltonian of equation (70)—whose kinetic momentum π includes now the two different vector potentials—yields the Hamilton equations of motion, which can be transformed into the Newton equations of motion; within the latter, the effects of the vector potentials appear in terms of fields, in the form of Lorentz forces. The curl of the magnetic vector potential obviously yields the magnetic field due to the external source; the curl of the geometric vector potential (Berry's curvature) yields an additional ‘magnetic-like’ field which is *non-zero* even on the classical trajectory of the ions. Notice that this is at variance with the zero-field case, where the Berry's phase had no effect on the ionic motion at the classical level, and could only be detected when quantizing the ionic degrees of freedom.

Within a naive Born–Oppenheimer approximation—where Berry's phases are neglected—the magnetic field acts on the nuclei as if they were ‘naked’ charges: a proper treatment must

instead account for electronic screening. This state of affairs was first recognized in 1988 by Schmelcher *et al* [35], who arrived at a ‘screened Born–Oppenheimer approximation’. Only more recently was this effect recognized as a manifestation of the Berry’s phase, in the work of Mead and co-workers [7,36]. Surprisingly, there are very few calculations of the effect: it is pretty clear, however, that the geometric term is no small correction. This is easily illustrated with the simple case of a neutral atom moving at constant speed in a uniform magnetic field. Obviously, a neutral object is not deflected by a Lorentz force, while if we analyse the motion of the ion within a naive Born–Oppenheimer approximation we find that it is deflected according to its naked nuclear charge. The geometric vector potential solves the paradox: remarkably, the ‘magnetic-like’ field due to the Berry’s phase is—in this simple example—exactly opposite to the external magnetic field, thus providing the complete screening which is physically expected.

It is worth showing how this works in detail for the hydrogen atom. In the complete Hamiltonian of equation (67) we identify $\xi \rightarrow \mathbf{R}$ and $[\mathbf{x}] \rightarrow \mathbf{r}$:

$$\mathcal{H}(\mathbf{R}, \mathbf{r}) = \frac{1}{2M} \left[\mathbf{p} + \frac{e}{c} \mathbf{A}(\mathbf{R}) \right]^2 + H(\mathbf{R}, \mathbf{r}) \quad (72)$$

$$H(\mathbf{R}, \mathbf{r}) = \frac{1}{2m} \left[-i\hbar \nabla_{\mathbf{r}} - \frac{e}{c} \mathbf{A}(\mathbf{r}) \right]^2 - \frac{e^2}{|\mathbf{r} - \mathbf{R}|}. \quad (73)$$

Notice that e is negative here. We then consider the electronic Hamiltonian, equation (73), as a function of the parameter \mathbf{R} , in the symmetric gauge where $\mathbf{A}(\mathbf{r}) = \frac{1}{2} \mathbf{B} \times \mathbf{r}$. If $\psi_0(\mathbf{r})$ is the exact ground eigenfunction when the proton sits at $\mathbf{R} = 0$, the eigenfunction at a generic \mathbf{R} is

$$\psi_{\mathbf{R}}(\mathbf{r}) = \exp\left(\frac{ie}{2\hbar c} \mathbf{r} \cdot \mathbf{B} \times \mathbf{R}\right) \psi_0(\mathbf{r} - \mathbf{R}). \quad (74)$$

The Berry’s connection \mathcal{X} is clearly

$$\mathcal{X}(\mathbf{R}) = i \langle \psi_{\mathbf{R}} | \nabla_{\mathbf{R}} \psi_{\mathbf{R}} \rangle = \frac{e}{2\hbar c} \langle \psi_{\mathbf{R}} | \mathbf{B} \times \mathbf{r} | \psi_{\mathbf{R}} \rangle = \frac{e}{2\hbar c} \mathbf{B} \times \mathbf{R} = \frac{e}{\hbar c} \mathbf{A}(\mathbf{R}) \quad (75)$$

since the \mathbf{R} -derivative of $\psi_0(\mathbf{r} - \mathbf{R})$ does not contribute. The geometric vector potential \mathcal{X} adds up to the kinetic momentum π which appears in the effective equation of motion for the proton. Clearly this term, equation (75), multiplied by $-\hbar$ as in equation (71), exactly cancels the magnetic vector potential acting on the naked proton, which appears in equation (72), thus providing the complete screening discussed above.

5.3. Semiclassical electron dynamics in crystals

The semiclassical approach has proven very useful for describing the electronic dynamics in slowly perturbed crystals: it has played, over several decades, a fundamental role in the physics of metals and semiconductors [61,62], and in technology as well. Within the semiclassical approach, the dynamics of an electron in a given band is described by a wave packet centred at \mathbf{r} , and whose crystal momentum is \mathbf{q} : both \mathbf{r} and \mathbf{q} evolve slowly in time. In the absence of collisions, the equations of motion are usually written as

$$\begin{aligned} \dot{\mathbf{r}} &= \frac{1}{\hbar} \frac{\partial \varepsilon_n(\mathbf{q})}{\partial \mathbf{q}} \\ \hbar \dot{\mathbf{q}} &= e \left(\mathbf{E} + \frac{1}{c} \dot{\mathbf{r}} \times \mathbf{B} \right) \end{aligned} \quad (76)$$

where $\varepsilon_n(\mathbf{q})$ is the band structure of the relevant band, and \mathbf{E} and \mathbf{B} are the perturbing fields.

As emphasized in reference [62], the derivation of equation (76), despite the formal simplicity of the result, is 'a formidable task'. The early derivations date from the 1930s [63]; the problem was reconsidered several times in the literature, by Slater [64], Luttinger [65], and Zak [66] among others. The most thorough analysis, due to Sundaram and Niu [22], is very recent: the main finding is that some important Berry's phase terms are missing in equation (76). With hindsight, and having in mind other adiabatic phenomena, this finding is perhaps not too surprising.

According to Sundaram and Niu, three major modifications are needed in the semiclassical equations of motion, equation (76). First: the unperturbed crystal momentum \mathbf{q} has to be replaced with the gauge-invariant crystal momentum $\hbar\mathbf{k} = \hbar\mathbf{q} - (e/c)\mathbf{A}(\mathbf{r}, t)$; this is indeed the analogue of the kinetic momentum $\boldsymbol{\pi}$ usually appearing in Hamiltonians, such as in the simple case of equation (73). Second: the band-structure energy $\varepsilon_n(\mathbf{q})$ needs a correction term which accounts for the orbital magnetization of the wave packet. Third: a Berry's curvature appears explicitly in the equation of motion. In this review, we focus on this last point only, and we refer the reader to the original paper [22] for the most general case. We therefore consider only the special case where the magnetic field vanishes, in which case the Sundaram–Niu equations of motion are

$$\begin{aligned}\dot{\mathbf{r}} &= \frac{1}{\hbar} \frac{\partial \varepsilon_n(\mathbf{q})}{\partial \mathbf{q}} - \dot{\mathbf{q}} \times \boldsymbol{\Omega}(\mathbf{q}) \\ \hbar\dot{\mathbf{q}} &= e\mathbf{E}.\end{aligned}\quad (77)$$

The extra contribution with respect to equation (76) comes from the vector field $\boldsymbol{\Omega}(\mathbf{q})$, having the dimensions of a squared length: this is precisely the standard Berry's curvature of Bloch electrons, equation (60), for the relevant band.

5.4. Bloch oscillations and the Wannier–Stark ladder

In the presence of a constant field the semiclassical equations of motion, equation (77), are easily integrated: $\mathbf{q}(t) = e\mathbf{E}t/\hbar + \text{constant}$, and $\dot{\mathbf{r}}$ is then an explicit function of time. The semiclassical approach is, by construction, a single-band approximation: since both $\varepsilon_n(\mathbf{q})$ and $\boldsymbol{\Omega}(\mathbf{q})$ are periodic functions of \mathbf{q} , one has that $\dot{\mathbf{r}}$ is a periodic function of time. Furthermore, $\dot{\mathbf{r}}$ is an odd function of \mathbf{q} : therefore even the motion of \mathbf{r} is periodic in time. The motion of \mathbf{q} can be thought of as periodic as well, once a reduced-zone scheme is adopted. This periodic motion of the slow variables \mathbf{r} and \mathbf{q} in a constant field goes under the name of 'Bloch oscillations'. We will indicate the trajectory of this motion as $\mathbf{r}(\mathbf{q})$: this trajectory is a closed orbit in the torus geometry of the reciprocal cell.

Next, we are going to quantize the motion of the slow variables: we will content ourselves with a semiclassical Bohr–Sommerfeld quantization. The appropriate multidimensional formula, known as Einstein–Brillouin–Keller quantization, is

$$\oint_C \left(\frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot d\mathbf{q} + \frac{\partial L}{\partial \dot{\mathbf{r}}} \cdot d\mathbf{r} \right) = 2\pi\hbar \left(m + \frac{\alpha}{4} \right) \quad (78)$$

where L is the Lagrangian, C is the closed orbit $\mathbf{r}(\mathbf{q})$ at a given energy W_m , and α is an integer, known as the Maslov index [67].

It is easy to prove that the semiclassical equations of motion, equation (77), can be obtained from the Lagrangian [22]:

$$L = L(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{r}, \dot{\mathbf{r}}) = -\varepsilon_n(\mathbf{q}) + e\mathbf{E} \cdot \mathbf{r} + \hbar\dot{\mathbf{r}} \cdot \mathbf{q} + i\hbar \langle u_{n\mathbf{q}} | \nabla_{\mathbf{q}} u_{n\mathbf{q}} \rangle \cdot \dot{\mathbf{q}}. \quad (79)$$

The only non-trivial term occurring in this proof is

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{q}}} = i\hbar \frac{d}{dt} \langle u_{n\mathbf{q}} | \nabla_{\mathbf{q}} u_{n\mathbf{q}} \rangle = \hbar\boldsymbol{\Omega}(\mathbf{q}) \times \dot{\mathbf{q}}. \quad (80)$$

Replacing the canonical momenta conjugate to q and r into equation (78) we get

$$\oint_C (i\hbar \langle u_{nq} | \nabla_q u_{nq} \rangle \cdot d\mathbf{q} + \hbar \mathbf{q} \cdot d\mathbf{r}) = 2\pi\hbar \left(m + \frac{\alpha}{4} \right) \quad (81)$$

$$\oint_C (i \langle u_{nq} | \nabla_q u_{nq} \rangle - \mathbf{r}) \cdot d\mathbf{q} = 2\pi \left(m + \frac{\alpha}{4} \right).$$

This expression takes a particularly simple form in one dimension, since the trajectory in the (q, x) variables is uniquely determined by the conservation of energy:

$$\varepsilon_n(q) - eEx = \text{constant} = W_m. \quad (82)$$

We thus get the discrete spectrum from equation (81):

$$W_m = \bar{\varepsilon}_n + eEa \left(m + \frac{\alpha}{4} - \frac{\gamma}{2\pi} \right) \quad (83)$$

where a is the lattice constant, $\bar{\varepsilon}_n$ is the band-energy average, and γ is Zak's geometric phase:

$$\bar{\varepsilon}_n = \frac{a}{2\pi} \int_{-\pi/a}^{\pi/a} \varepsilon_n(q) dq \quad \gamma = i \int_{-\pi/a}^{\pi/a} \langle u_n(q) | (d/dq) u_n(q) \rangle dq. \quad (84)$$

The integer m can take any value from $-\infty$ to ∞ : this spectrum is the famous Wannier–Stark ladder, first predicted by Wannier in 1960 [68, 69] without the geometric-phase term. The γ -correction was found in 1968 by Zak [70], who much later (1989) recognized it as a manifestation of the geometric phase [33]. The derivation given here follows Sundaram and Niu [22].

The presence of the γ -term in the spectrum, equation (83), is essential as a matter of principle: without it, the spectrum would be disturbingly unphysical. Suppose we change the potential by an arbitrary constant: this can be written as $Ex \rightarrow E(x - x_0)$, corresponding thus to a shift of origin. Clearly the spectrum must be translationally invariant: the position of the discrete levels (with respect to the band energy) must be independent of the choice of the origin x_0 in the crystal cell. The geometric-phase term in equation (83) removes this drawback. In fact, the Zak's phase *does* depend of the choice of the origin in the cell: the changes in the two terms compensate[†], thus ensuring the translational invariance of the spectrum.

For many years, Wannier–Stark ladders have remained a purely academic subject. In order to observe the phenomenon, one needs collisionless coherent motion of the carriers throughout an entire band. Furthermore, the existence of the phenomenon, even in principle, was controversial [70]: the whole theory seems in fact to rely on the semiclassical approximation and on the neglect of interband coupling. The effect was first detected in 1988 by means of optical techniques [71]. The breakthrough was made possible by the use of a very special 'crystalline solid': a semiconductor superlattice, whose lattice constant a is of the order of 100 Å. A moderate field is then sufficient for driving the carriers all the way to the top of the 'miniband'. The existence of Wannier–Stark ladders was later confirmed by a variety of techniques [72], and for different physical systems [73].

5.5. Spin-wave dynamics in crystals

We discuss here the case of a spin wave in a crystal, which evolves in time (slowly compared to the electronic timescales). The phenomenon has been almost invariably studied in terms of a Heisenberg model Hamiltonian [74, 75], where the continuum nature of the spin distribution is disregarded, and magnetization is schematized as localized at lattice sites. Very recently,

[†] The algebra to prove this is very similar to that used in reference [39] around equation (22).

Niu and Kleinman [23] have given a general formulation, which applies to real crystals and where the continuum nature of the spin distribution is fully accounted for. Their major result is that a Berry curvature plays a key role in the theory. First-principles calculations based on this theory have been performed by Gebauer and Baroni: their results for iron are in excellent agreement with experiment [76].

Quite generally, wave propagation in extended systems occurs whenever the static equilibrium configuration is perturbed by a disturbance. The disturbance has an energy cost which provides a restoring force. In order to establish the dynamics of the wave, one has to identify the inertia of the system in responding to this force. This is pretty clear for the paradigmatic example of elastic waves, where the restoring force is measured by the appropriate elastic modulus, and the inertia is measured by the mass density. For a spin wave within the standard textbook treatment [75], the Heisenberg Hamiltonian immediately provides the equilibrium configuration (ferromagnetic or antiferromagnetic, in the simplest cases), and also provides the energy cost of any disturbance. The restoring force in this case is actually a restoring *torque*; what about the inertia? We observe first of all that the equation of motion is *first order* in time, and not second order as in the most usual wave equations: the torque equals the rate of change of the (localized) spin angular momentum. If we replace the latter with a classical angular momentum, we see that it is clearly partitioned into an inertia factor (the moment of inertia) times a kinematic factor (the angular velocity). But for a spin angular momentum such partition is meaningless: inertia and kinematics are entangled. The modulus is fixed and only the orientation may change: the rate of change is therefore off-diagonal in the spin components. Within the equation of motion [75], the factor which multiplies the time derivative of the local spin—and which represents the ‘inertia’—is simply \hbar times a pure-number matrix. We could regard this as a very simple ‘geometric’ quantity.

We now switch from the idealized Heisenberg model to a real crystal, where the magnetization $\vec{\mathcal{M}}(\mathbf{r})$ is a continuum vector field. For a magnetic system the equilibrium magnetization $\vec{\mathcal{M}}_0(\mathbf{r})$ is obviously non-zero, and can be computed using a variety of first-principles approaches. We wish to define the energy cost of a magnetization fluctuation $\delta\vec{\mathcal{M}}(\mathbf{r})$, where we explicitly make the adiabatic hypothesis: the electronic wavefunction is assumed to follow the instantaneous spin-wave configuration. Owing to this, and exploiting the variational theorem, one defines the total energy of the system for an arbitrary magnetization $E[\vec{\mathcal{M}}]$ by means of a constrained-minimum problem:

$$E[\vec{\mathcal{M}}] = \min\langle\Psi|H|\Psi\rangle \quad (85)$$

where the search is performed over all the wavefunctions Ψ yielding the given magnetization $\vec{\mathcal{M}}(\mathbf{r})$. The unconstrained minimum of equation (85) coincides with the absolute minimum of the functional, which occurs at $\vec{\mathcal{M}}_0(\mathbf{r})$; for an arbitrary $\vec{\mathcal{M}}(\mathbf{r})$, equation (85) defines $E[\vec{\mathcal{M}}]$ and $\Psi[\vec{\mathcal{M}}]$.

The Cartesian components of the harmonic restoring force are therefore

$$K_{\alpha\beta}(\mathbf{r}, \mathbf{r}') = \frac{\delta^2 E[\vec{\mathcal{M}}]}{\delta\mathcal{M}_\alpha(\mathbf{r}) \delta\mathcal{M}_\beta(\mathbf{r}')}. \quad (86)$$

We then need the inertia term: as anticipated above, the outstanding finding of Niu and Kleinman [23] is that the ‘inertia’ of a spin wave within the equation of motion is measured by the appropriate Berry’s curvature of the electronic wavefunction (times \hbar). In the notation of Gebauer and Baroni [76] this curvature is

$$\Omega_{\alpha\beta}(\mathbf{r}, \mathbf{r}') = -2 \operatorname{Im} \left\langle \frac{\delta\Psi[\vec{\mathcal{M}}]}{\delta\mathcal{M}_\alpha(\mathbf{r})} \left| \frac{\delta\Psi[\vec{\mathcal{M}}]}{\delta\mathcal{M}_\beta(\mathbf{r}')} \right. \right\rangle \quad (87)$$

and the equation of motion for the spin wave is

$$\hbar \int d\mathbf{r}' \Omega_{\alpha\beta}(\mathbf{r}, \mathbf{r}') \dot{\mathcal{M}}_{\beta}(\mathbf{r}') = \int d\mathbf{r}' K_{\alpha\beta}(\mathbf{r}, \mathbf{r}') \mathcal{M}_{\beta}(\mathbf{r}'). \quad (88)$$

In practical calculations, the magnetization fluctuation $\delta\vec{\mathcal{M}}(\mathbf{r})$ must be expressed in terms of discrete parameters: the functional derivatives become then partial derivatives, equation (87) becomes a very standard Berry's curvature, and the equation of motion becomes a matrix equation.

In the limiting case of the Heisenberg model the magnetization fluctuation is discretized as a simple lattice variable; the equations of motion, equation (88), take then the usual textbook form [75]. In this limit, the Berry curvature $\Omega_{\alpha\beta}(\mathbf{r}, \mathbf{r}')$ becomes a pure-number matrix, diagonal over the sites and off-diagonal in the Cartesian coordinates [24].

6. Macroscopic polarization of dielectrics

6.1. The problem

The dipole moment of any *finite* N -electron system in its ground state is a simple and well defined quantity. Given the many-body wavefunction Ψ and the corresponding single-particle density $n(\mathbf{r})$ the electronic contribution to the x -component of the dipole is

$$e\langle X \rangle = e \int d\mathbf{r} x n(\mathbf{r}) = e \langle \Psi | \hat{X} | \Psi \rangle \quad (89)$$

where

$$\hat{X} = \sum_{i=1}^N x_i.$$

This looks very trivial, but we are exploiting here an essential fact: the ground wavefunction of any finite N -electron system is square integrable and vanishes exponentially at infinity; the density vanishes exponentially as well.

Considering now a macroscopic solid, the related quantity is macroscopic polarization, which is very much an essential concept in any phenomenological description of dielectric media [77]: this quantity is ideally defined as the dipole of a macroscopic sample, divided by its volume. The point is that, when using equation (89), the integral is dominated by what happens at the surface of the sample: knowledge of the electronic distribution in the bulk region is not enough to unambiguously determine the dipole. This looks like a paradox, since in the thermodynamic limit macroscopic polarization must be an intensive quantity, insensitive to surface effects.

Macroscopic polarization in the bulk region of the solid must be determined by what 'happens' in the bulk as well. This is the case if one assumes a model of discrete and well separated dipoles, à la Clausius and Mossotti: but real dielectrics are very much different from such an extreme model. The valence electronic distribution is continuous, and often very delocalized (particularly in covalent dielectrics). Most textbooks attempt to explain the polarization of a periodic crystal via the dipole moment of a unit cell, or something of the kind [61, 74]. These definitions are incorrect [78]: according to the modern viewpoint, bulk macroscopic polarization is a physical observable *completely independent* of the periodic charge distribution of the polarized crystalline dielectric.

The basic clue to the solution of the polarization problem comes from some experimental facts. The absolute polarization of a crystal in a given state has never been measured as a bulk property, independent of sample termination. Instead, well known bulk properties

are *derivatives* of the polarization with respect to suitable perturbations: permittivity, pyroelectricity, piezoelectricity, effective charges (for lattice dynamics). In one important case—namely, ferroelectricity—the relevant bulk property is inferred from the measurement of a finite difference (polarization reversal). In all cases, the derivative or the difference in the polarization is typically accessed via the measurement of a macroscopic current.

We illustrate the main concepts on the case of the piezoelectric effect: two possible realizations of it are shown schematically in figure 5. In (a) the crystal is uniaxially strained (along a piezoelectric axis) while kept in a shorted capacitor; in (b) the sample is strained while kept isolated. Focusing on (a), we notice that the phenomenon manifests itself as a bulk current traversing the sample, while nothing peculiar happens at the surfaces. Indeed the current flowing across the shorting wire is the quantity which is actually measured in experiments. The modern theory of polarization focuses on currents as well: since it is the *phase* of the wavefunction which carries information about the current, it should not be too surprising that a gauge-invariant phase occurs in the theoretical description of electronic polarization.

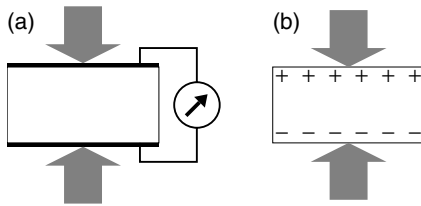


Figure 5. Two possible realizations of the piezoelectric effect: (a) the sample is in a shorted capacitor, and the current is measured; (b) the sample is isolated.

Figure 5(a) is also helpful for the purpose of understanding the root of the Berry's phase in the phenomenon: the quantum system is in interaction with an external apparatus, and 'slaved' to it. We are allowed to study the quantum system *as if* it were isolated, but then—as emphasized in Berry's original paper [1]—the interaction with the 'rest of the Universe' gives rise to a non-trivial phase which is observable: indeed, this Berry's phase measures the piezoelectric polarization in suitable units. This state of affairs has to be contrasted with figure 5(b), where the quantum system is isolated and no Berry's phase may occur: the polarization can be evaluated from the *modulus* of the wavefunction, as in equation (89). The obvious drawback is that the whole effect appears like a surface phenomenon, while piezoelectric polarization must instead be an intensive bulk quantity.

6.2. Polarization as a Berry's phase

In condensed matter physics the standard way of getting rid of undesired surface effects is to adopt periodic Born–von Kármán (BvK) boundary conditions: this makes equation (89) useless, since the integrals are ill defined due to the unbounded nature of the quantum mechanical position operator. Because of this fact, macroscopic polarization remained a major challenge in electronic structure theory for many years [78]. The breakthrough came in 1992, when the problem was approached essentially from the viewpoint of figure 5(a): a definition based on the wavefunctions, and *not* on the charge, was provided [79]. This definition has an unambiguous thermodynamic limit, such that BvK boundary conditions and Bloch states can be used safely. In the following months a modern theory of macroscopic polarization in crystalline dielectrics was completely established [38, 39], thanks to a major advance due to R D King-Smith and D Vanderbilt [37], who expressed polarization as a Berry's phase. A

comprehensive account of the modern theory exists [39]. Other less technical presentations are available as well [80–82]; for a very simplified non-technical outline see reference [40]. First-principles calculations based on this theory have been performed for several crystalline materials [41].

All of the above-quoted work refers to a crystalline system within an independent-electron formulation: the single-particle orbitals have then the Bloch form, and the Berry's phase which provides the polarization value is just the standard Zak's phase of the orbitals, discussed in section 4.4. The related, but substantially different, problem of macroscopic polarization in a correlated many-electron system was first solved by Ortíz and Martin in 1994 [83]. However, according to them, polarization is defined—and computed [84, 85]—by means of a peculiar 'ensemble average', integrating over a set of different electronic ground states; this was much later (1998) shown to be unnecessary. In reference [25], in fact, a simpler viewpoint is taken: the polarization of a correlated solid is expressed as a single-point Berry's phase, of the same kind as in equation (35). This is indeed a 'pure-state' property, although of a rather exotic kind. By the same token, it was also possible to define [21, 25, 26]—and to compute [47, 48]—macroscopic polarization in non-crystalline systems. The presentation given here is based on this very recent advance.

The single-point Berry's phase provides an expression for the expectation value $\langle X \rangle$, derived in the following: equations (98) and (99). This expression must be used in the place of equation (89) when BvK boundary conditions are adopted. The value of our definition resides in the fact that, whenever the electronic Hamiltonian is adiabatically varied, the time derivative of $\langle X \rangle$ coincides with the macroscopic electronic current $\langle J_x \rangle$ traversing the sample. Having in mind the simple example of figure 5(a), this is precisely what is needed to define macroscopic polarization.

All of the different versions of the theory, including the one presented here, concern phenomena where the macroscopic polarization is induced by a source other than an electric field. Even in this case, the polarization may (or may not) be *accompanied* by a field, depending on the boundary conditions chosen for the macroscopic sample. The modern theory, based on a Berry's phase, addresses polarization differences *in zero field*: this concerns therefore lattice dynamics, piezoelectricity—as in the ideal experiment sketched in figure 5(a)—and ferroelectricity. Notably, we do not address the problem of evaluating the dielectric constant: this can be done using alternative approaches, not based on a Berry's phase, such as the well established linear response theory [81, 86], or other more innovative theories [87].

6.3. A lone electron in a periodic box

We start by considering the simple case of a single electron within BvK boundary conditions in one dimension: $\psi(x + L) = \psi(x)$, where L is the imposed periodicity, typically chosen to be large with respect to atomic dimensions. The electron distribution may look like the one of figure 6: intuitively, the distribution appears to have a 'centre' x_0 , which however is defined only modulo the replica periodicity, and furthermore *cannot* be evaluated simply as in equation (89), i.e.

$$\langle x \rangle = \int dx x |\psi(x)|^2$$

precisely because of the BvK boundary conditions. Selloni *et al* [88] solved the problem some years ago by means of a very elegant and far-reaching formula, presented in the following. The work of reference [25] can be regarded as the many-body generalization of it.

According to references [25, 32, 88], the key quantity for dealing with the position operator

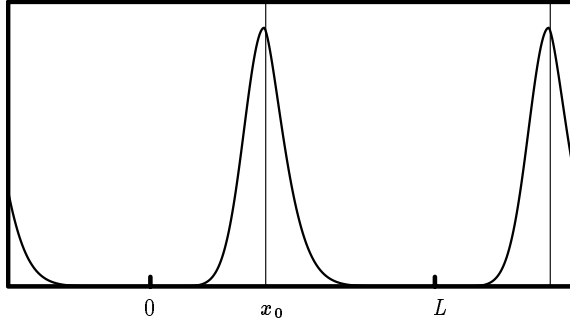


Figure 6. The distribution $|\psi(x)|^2$ of a single-particle orbital within periodic Born-von Kármán boundary conditions.

within BvK boundary conditions is the dimensionless complex number z , defined as

$$z = \langle \psi | e^{i(2\pi/L)x} | \psi \rangle = \int_0^L dx e^{i(2\pi/L)x} |\psi(x)|^2 \quad (90)$$

whose modulus is no larger than 1. The most general electron density, such as the one depicted in figure 1, can always be written as a superposition of a function $n_{\text{loc}}(x)$, normalized over $(-\infty, \infty)$, and of its periodic replicas:

$$|\psi(x)|^2 = \sum_{m=-\infty}^{\infty} n_{\text{loc}}(x - x_0 - mL). \quad (91)$$

Both x_0 and $n_{\text{loc}}(x)$ have a large arbitrariness: we restrict it a little bit by imposing the condition that x_0 is the centre of the distribution, in the sense that

$$\int_{-\infty}^{\infty} dx x n_{\text{loc}}(x) = 0.$$

Using equation (91), z can be expressed in terms of the Fourier transform of n_{loc} as

$$z = e^{i(2\pi/L)x_0} \tilde{n}_{\text{loc}}\left(-\frac{2\pi}{L}\right). \quad (92)$$

If the electron is localized in a region of space much smaller than L , its Fourier transform is smooth over reciprocal distances of the order of L^{-1} and can be expanded as

$$\tilde{n}_{\text{loc}}\left(-\frac{2\pi}{L}\right) = 1 - \frac{1}{2} \left(\frac{2\pi}{L}\right)^2 \int_{-\infty}^{\infty} dx x^2 n_{\text{loc}}(x) + \mathcal{O}(L^{-3}). \quad (93)$$

A very natural definition of the centre of a localized periodic distribution $|\psi(x)|^2$ is therefore provided by the phase of z as

$$\langle x \rangle = \frac{L}{2\pi} \text{Im} \log z \quad (94)$$

which is in fact the formula first proposed by Selloni *et al* [88]. The expectation value $\langle x \rangle$ is defined modulo L , as expected since $|\psi(x)|^2$ is BvK periodic. Comparing with equation (35), we notice that $\langle x \rangle$ has precisely the form of a single-point Berry's phase: the angle is converted into a length by the factor $L/2\pi$, and the unitary operator W of equation (35) has to be identified with the multiplicative operator $e^{i(2\pi/L)x}$ in Schrödinger representation.

The above expressions imply $\langle x \rangle \simeq x_0 \text{ mod } L$; in the special case where $n_{\text{loc}}(x)$ can be taken as an even (centrosymmetric) function, its Fourier transform is real and equation (92)

yields indeed $\langle x \rangle \equiv x_0 \bmod L$. In the case of extreme delocalization we have instead $|\psi(x)|^2 = 1/L$ and $z = 0$: hence the centre of the distribution $\langle x \rangle$, according to equation (94), is ill defined. For a more general delocalized state, we expect z to go to zero at large L [32].

We have therefore arrived at a definition of $\langle x \rangle$ within BvK boundary conditions which has many of the desirable features we were looking for; nonetheless, there is a property that is even more important, and which we are going to discuss next. Suppose the potential which the electron moves in has a slow time dependence—as was the case in reference [88]—and we wish to follow the adiabatic evolution of the electronic state $|\psi\rangle$. If we write $|\varphi_j\rangle$ for the instantaneous eigenstates at time t , the lowest-order adiabatic evolution of the ground-state density matrix is [89]

$$|\psi\rangle\langle\psi| \simeq |\varphi_0\rangle\langle\varphi_0| + i \sum_{j \neq 0} \left(|\varphi_j\rangle \frac{\langle\varphi_j|\dot{\varphi}_0\rangle}{\epsilon_j - \epsilon_0} \langle\varphi_0| - \text{h.c.} \right) \quad (95)$$

where the phases have been chosen in order to make $|\varphi_0\rangle$ orthogonal to its time derivative $|\dot{\varphi}_0\rangle$. The macroscopic electrical current flowing through the system at time t is therefore

$$\langle j \rangle = \frac{e}{L} \langle \psi | p | \psi \rangle \simeq \frac{ie}{L} \sum_{j \neq 0} \frac{\langle \varphi_0 | p | \varphi_j \rangle \langle \varphi_j | \dot{\varphi}_0 \rangle}{\epsilon_j - \epsilon_0} + \text{c.c.} \quad (96)$$

It is then rather straightforward to prove that $\langle j \rangle$ to lowest order in $1/L$ equals $(e/L) d\langle x \rangle/dt$, where $\langle x \rangle$ is evaluated using in equation (94) the instantaneous ground eigenstate:

$$\langle j \rangle \simeq \frac{e}{2\pi} \text{Im} \frac{d}{dt} \log \langle \varphi_0 | e^{i(2\pi/L)x} | \varphi_0 \rangle. \quad (97)$$

This finding proves the value of equations (90) and (94) in studying electron transport [88]: for similar reasons, the many-body generalization of equations (90) and (94) defines the macroscopic polarization of an extended system.

6.4. The position operator in extended systems

We are now going to consider a system of N electrons in a cubic periodic box of side L , where eventually the thermodynamic limit is taken: $L \rightarrow \infty$, $N \rightarrow \infty$, and $N/L^3 = n_0$ is constant. We also assume the ground state non-degenerate, and we deal with insulating systems only: this means that the gap between the ground eigenvalue and the excited ones remains finite for $L \rightarrow \infty$. BvK boundary conditions impose periodicity (of period L) in each Cartesian coordinate and in each electronic variable separately.

We start by defining the many-body analogue of equation (90), namely, the multiplicative operator $\hat{X} = \sum_{i=1}^N x_i$, and the complex number

$$z_N = \langle \Psi | e^{i(2\pi/L)\hat{X}} | \Psi \rangle. \quad (98)$$

It is obvious that the operator \hat{X} is ill defined in the BvK Hilbert space, while its complex exponential appearing in equation (98) is well defined; we will call this exponential the ‘many-body phase operator’ in the following. The main result of reference [25] is that the ground-state expectation value of the position operator is given by the analogue of equation (94), namely

$$\langle X \rangle = \frac{L}{2\pi} \text{Im} \log z_N \quad (99)$$

which is a quantity defined modulo L as above. As in the much simpler case of equation (89), $\langle X \rangle$ is extensive. We stress that nowhere have we assumed crystalline periodicity. Therefore our definition of $\langle X \rangle$ is very general: it applies to any condensed system, either ordered or disordered, either independent electron or correlated.

The right-hand side of equation (99) is not simply the expectation value of an operator: it is in fact the *phase* of it. Indeed, our definition of $\langle X \rangle$ is a single-point Berry's phase, converted into length units by the factor $L/2\pi$, and where the unitary operator W of equation (35) has to be identified with the many-body phase operator $e^{i(2\pi/L)\hat{X}}$. It is important to realize that this operator is a genuine many-body one. In general, one defines an operator to be a one-body one whenever it is the *sum* of N identical operators, acting on each electronic coordinate separately: for instance, the \hat{X} -operator is one such. In order to express the expectation value of a one-body operator the full many-body wavefunction is not needed: knowledge of the one-body reduced density matrix is enough; I stress that, instead, the expectation value of $e^{i(2\pi/L)\hat{X}}$ over a correlated wavefunction *cannot* be expressed in terms of the one-body density matrix.

As in the one-body case, whenever the many-body Hamiltonian is slowly varying in time, the (macroscopically averaged) electrical current density flowing through the system is given by

$$\langle J_x \rangle = \frac{e}{L^3} \frac{d}{dt} \langle X \rangle = \frac{e}{2\pi L^2} \frac{d}{dt} \text{Im} \log z_N \quad (100)$$

where z_N is evaluated using in equation (98) the instantaneous ground eigenstate of the Hamiltonian at time t ; this result is proved in reference [25]. Since $\langle X \rangle$ is an extensive quantity, the current density $\langle J_x \rangle$, equation (100), goes to a well defined thermodynamic limit. Notice that the $L \rightarrow \infty$ limit in equation (100) is a rather unconventional one, since the many-body phase operator in z_N goes formally to the identity, but the size of the system and the number of electrons in the wavefunction Ψ increase with L .

6.5. A crystalline system of independent electrons

We now specialize the above very general formulation to a simple cubic crystal of lattice constant a , where the BvK periodicity is $L = Ma$. The most general crystal structure can be dealt with by means of a simple coordinate transformation, such as the one considered in reference [39]. In the non-interacting case the single-particle orbitals can be chosen in the Bloch form: there are M^3 allowed Bloch vectors \mathbf{q}_s , which lie on a regular mesh in the unit reciprocal cell, where $s \equiv (s_1, s_2, s_3)$ and

$$\mathbf{q}_s = \frac{2\pi}{Ma} (s_1, s_2, s_3) \quad s_\alpha = 0, 1, \dots, M-1. \quad (101)$$

We adopt a plane-wave-like normalization for the Bloch orbitals:

$$\langle \psi_{n\mathbf{q}_s} | \psi_{n'\mathbf{q}_{s'}} \rangle = \frac{1}{a^3} \int d\mathbf{r} \psi_{n\mathbf{q}_s}^*(\mathbf{r}) \psi_{n'\mathbf{q}_{s'}}(\mathbf{r}) = \delta_{nn'} \delta_{ss'} \quad (102)$$

where the integral is performed over the unit cell. If our insulating system has n_b doubly occupied bands, there are $N = 2n_b M^3$ independent spin orbitals: the ground wavefunction is a Slater determinant of size N . This large determinant can be cast as the antisymmetrized product of M^3 small determinants of size $2n_b$; each; i.e.,

$$|\Psi\rangle = \mathbf{A} \prod_s \frac{1}{L^{3n_b}} \frac{1}{\sqrt{(2n_b)!}} |\psi_{1\mathbf{q}_s} \bar{\psi}_{1\mathbf{q}_s} \psi_{2\mathbf{q}_s} \bar{\psi}_{2\mathbf{q}_s} \cdots \psi_{n_b\mathbf{q}_s} \bar{\psi}_{n_b\mathbf{q}_s}| \quad (103)$$

where \mathbf{A} is the antisymmetrizer operator, and the N -body wavefunction is normalized to 1 over the hypercube of side L .

For this wavefunction, the expectation value of the many-body phase operator, equation (98), is conveniently written as

$$z_N = \langle \Psi | \tilde{\Psi} \rangle \quad (104)$$

where $\tilde{\Psi}$ is the Slater determinant of a different set of Bloch spin orbitals:

$$\tilde{\psi}_{nq_s}(\mathbf{r}) = e^{i(2\pi/Ma)x} \psi_{nq_s}(\mathbf{r}) \quad (105)$$

and analogously for the bar (spin-down) ones. According to a well known theorem, the overlap between two single-determinant wavefunctions is equal to the determinant of the $N \times N$ overlap matrix built out of the occupied spin orbitals. We start by noticing that the overlaps between different-spin spin orbitals vanish, while those between equal-spin ones are identical in pairs. We can therefore write

$$z_N = (\det \mathcal{S})^2 \quad (106)$$

where \mathcal{S} is the overlap matrix for spatial orbitals, having size $N/2 = n_b M^3$, and whose elements are

$$\begin{aligned} \mathcal{S}_{nq_s, n'q_{s'}} &= \frac{1}{a^3} \int d\mathbf{r} \psi_{nq_s}^*(\mathbf{r}) \tilde{\psi}_{n'q_{s'}}(\mathbf{r}) \\ &= \frac{1}{a^3} \int d\mathbf{r} u_{nq_s}^*(\mathbf{r}) u_{n'q_{s'}}(\mathbf{r}) \exp\left(i\left[\frac{2\pi}{Ma}x + \mathbf{q}_{s'} \cdot \mathbf{r} - \mathbf{q}_s \cdot \mathbf{r}\right]\right) \end{aligned} \quad (107)$$

where the u s are the periodic factors in the Bloch orbitals.

The matrix \mathcal{S} is very sparse: in fact, given the geometry of the \mathbf{q}_s on the regular reciprocal mesh, the overlap integrals in equation (107) are non-vanishing only if $s_1 = s'_1 + 1$, $s_2 = s'_2$, and $s_3 = s'_3$. We express the non-vanishing elements in terms of a small overlap matrix S , of size $n_b \times n_b$:

$$S_{nn'}(\mathbf{q}, \mathbf{q}') = \langle u_{nq} | u_{n'q'} \rangle = \frac{1}{a^3} \int d\mathbf{r} u_{nq}^*(\mathbf{r}) u_{n'q'}(\mathbf{r}). \quad (108)$$

This same matrix has been introduced previously, in equation (64). Owing to the sparseness of \mathcal{S} , its determinant factors into products of determinants of small matrices S as follows:

$$z_N^{1/2} = \det \mathcal{S} = \prod_{s_1, s_2, s_3} \det S(\mathbf{q}_{s_1+1, s_2, s_3}, \mathbf{q}_{s_1, s_2, s_3}) \quad (109)$$

$$\log z_N = 2 \sum_{s_2, s_3=0}^{M-1} \log \prod_{s_1=0}^{M-1} \det S(\mathbf{q}_{s_1+1, s_2, s_3}, \mathbf{q}_{s_1, s_2, s_3}). \quad (110)$$

We have

$$\langle X \rangle = \frac{2Ma}{2\pi} \sum_{s_2, s_3=0}^{M-1} \text{Im} \log \prod_{s_1=0}^{M-1} \det S(\mathbf{q}_{s_1+1, s_2, s_3}, \mathbf{q}_{s_1, s_2, s_3}) \quad (111)$$

where the thermodynamic limit $M \rightarrow \infty$ (and $N \rightarrow \infty$) is understood. Notice, however, that numerical implementations exactly correspond to taking a finite M in equation (111).

6.6. King-Smith and Vanderbilt's formula

We have stressed in section 6.1 that, when measuring polarization, the quantity actually measured is the current which traverses the sample while a given perturbation is adiabatically switched on. The electronic contribution to this current is, after equation (100):

$$\Delta P_x = \int_0^{\Delta t} dt \langle J_x(t) \rangle = \frac{e}{2\pi L^2} [\text{Im} \log z_N(\Delta t) - \text{Im} \log z_N(0)]. \quad (112)$$

Notice that in the adiabatic limit Δt goes to infinity and $J(t)$ goes to zero, while ΔP_x only depends on the initial and final states. In the special case of a crystalline system of independent

electrons we may replace equation (111) in equation (112). We exchange the arguments in S , which brings up a minus sign. We get

$$P_x = \frac{e\langle X \rangle}{M^3 a^3} = -\frac{2e}{2\pi M^2 a^2} \sum_{s_2, s_3=0}^{M-1} \text{Im} \log \prod_{s_1=0}^{M-1} \det S(\mathbf{q}_{s_1, s_2, s_3}, \mathbf{q}_{s_1+1, s_2, s_3}) \quad (113)$$

where it is understood that equation (113) is to be used twice, with the final and with the initial ground states, in order to evaluate the quantity of interest ΔP_x .

We are interested in the thermodynamic limit of equation (113), where $M \rightarrow \infty$, and the \mathbf{q}_s -vectors, equation (101), become dense in the unit reciprocal cell. We start by considering the limit

$$\gamma_{s_2, s_3} = -\lim_{M \rightarrow \infty} \text{Im} \log \prod_{s_1=0}^{M-1} \det S(\mathbf{q}_{s_1, s_2, s_3}, \mathbf{q}_{s_1+1, s_2, s_3}). \quad (114)$$

Comparing to equation (65), we recognize in equation (114) the discretized form of a Zak's phase, where the path is the segment joining the reciprocal points $\mathbf{q}_{\text{initial}} = (0, 2\pi s_2/L, 2\pi s_3/L)$ and $\mathbf{q}_{\text{final}} = (2\pi/a, 2\pi s_2/L, 2\pi s_3/L)$, differing by the reciprocal-lattice vector $\mathbf{G} = (2\pi/a, 0, 0)$. Equation (114) converges therefore to the integral (Zak's phase)

$$\gamma_{s_2, s_3} = i \sum_{n=1}^{n_b} \int_{C_{s_2, s_3}} \langle u_{nq} | (\partial/\partial q_x) u_{nq} \rangle dq_x \quad (115)$$

where C_{s_2, s_3} is the above-mentioned open path. The quantity γ_{s_2, s_3} can be regarded as the discretization over a mesh of a function $\gamma(q_y, q_z)$ defined in the two-dimensional unit reciprocal cell $[0, 2\pi/a) \times [0, 2\pi/a)$. For a dense mesh one obviously gets

$$\sum_{s_2, s_3} \gamma_{s_2, s_3} \rightarrow \frac{M^2 a^2}{(2\pi)^2} \int dq_y dq_z \gamma(q_y, q_z) \quad (116)$$

and equation (113) becomes in the thermodynamic limit

$$P_x = \frac{2e}{(2\pi)^3} \int dq_y dq_z \gamma(q_y, q_z) = \frac{2ie}{(2\pi)^3} \sum_{n=1}^{n_b} \int d\mathbf{q} \langle u_{nq} | (\partial/\partial q_x) u_{nq} \rangle. \quad (117)$$

This is the—by now famous—Berry's phase formula for the macroscopic polarization of a crystalline system of independent electrons, derived by King-Smith and Vanderbilt in 1993 [37]. The form given here is for a simple cubic lattice; see reference [39] for arbitrary crystal structures.

The present derivation of equation (117) is very different from the original one, and from other derivations in the literature [38, 39] as well. In fact, we have based our derivation on the viewpoint of reference [25], where equation (117) appears as a special case of the most general expressions, equations (98) and (112): these, in fact, apply on the same grounds to any many-electron extended system: either independent electron or correlated, either crystalline or disordered.

The most fundamental expressions used in the present work to define macroscopic polarization, such as equations (98) and (112), are based on a single-point Berry's phase: polarization is essentially the phase of z_N , which is the expectation value of a unitary operator. Since the operator is *local* in Schrödinger representation, the value of z_N only depends on the square modulus of the wavefunction Ψ , and *not* on its phase. For such reasons—as stressed in section 3.4—we have here a very unconventional kind of geometric phase, which could be defined as a Berry's phase 'in disguise'. Instead, the expressions given as equations (114) and (115) are just 'non-exotic' Berry phases, of the kind introduced by Zak [33]: macroscopic

polarization in a three-dimensional system appears in equation (117) simply as the two-dimensional integral of a Zak's phase. We have therefore the remarkable finding that the single-point Berry's phase of the *many-body wavefunction* Ψ coincides with a much more standard (many-point) Berry's phase of the *single-particle orbitals*. Indeed, the square modulus of Ψ embeds some interference features which—in the uncorrelated single-determinant case—can be expressed in terms of the relative phases of the orbitals, as is indeed explicitly shown by equation (114). The many-body phase operator $e^{i(2\pi/L)\hat{X}}$ of reference [25] 'extracts' the phase differences between the orbitals in a gauge-invariant and effective way.

The choice of Bloch orbitals for describing the independent-electron wavefunction Ψ , as in equation (103), is just one of the possible choices and is by no means mandatory. The many-body wavefunction is in fact invariant under unitary transformations of the occupied orbitals among themselves: z_N (and hence macroscopic polarization) is invariant as well. Suppose we make a different choice: for instance, we replace equation (103) with a Slater determinant of Wannier orbitals. In this case the many-body phase operator acts in a completely different way upon the orbitals, and no longer extracts any phase difference out of them. On the contrary, macroscopic polarization obtains from the *modulus* of the Wannier orbitals—more precisely from the centres of their distribution. This fact has been well known since the work of King-Smith and Vanderbilt [37] (see also reference [39]), and can be easily recovered in the present formalism.

6.7. Non-crystalline systems

We consider now a non-crystalline system of independent electrons in a periodic cubic box of side L . The one-particle orbitals $\psi_i(\mathbf{r})$ are periodic functions of period L in each Cartesian coordinate, featuring in general no additional symmetry. We write the many-body wavefunction as

$$|\Psi\rangle = \frac{1}{\sqrt{N!}} |\psi_1 \bar{\psi}_1 \psi_2 \bar{\psi}_2 \cdots \psi_{N/2} \bar{\psi}_{N/2}|. \quad (118)$$

In principle we are interested in the thermodynamic limit $L \rightarrow \infty$; in typical calculations (such as in Car-Parrinello simulations [90–92]) it is enough to take a 'supercell' whose size L is larger than the appropriate correlation length of the disordered system.

The wavefunction in equation (118) can be regarded as a special case of the previous one, equation (103), where only one \mathbf{q} -vector is used (the Γ point): the ψ s coincide with the u s at $\mathbf{q} = 0$. Most of the previous results are correct as they stand, provided one takes $M = 1$, $a = L$, and $n_b = N/2$. The 'small' overlap matrix S coincides now with the 'large' one \mathcal{S} : this matrix is no longer sparse. According to equations (66) and (108), its elements are

$$S_{ij} = \langle \psi_i | e^{-i(2\pi/L)x} | \psi_j \rangle. \quad (119)$$

The single-point Berry's phase formula for macroscopic polarization becomes then, after equations (112) and (113),

$$P_x = -\frac{2e}{2\pi L^2} \text{Im} \log \det \mathcal{S}. \quad (120)$$

This expression was first proposed by the present author in a volume of lecture notes [21], and has been very successfully used since in first-principles calculations of infrared spectra for non-crystalline systems [47, 48]. Within a simulation of the Car-Parrinello type [90–92] one follows the time evolution of the ions at finite temperature along discrete time steps, while the electrons follow adiabatically. The polarization is calculated 'on the fly' using equation (120), and the spectrum is evaluated from its time-correlation function [47].

6.8. Correlated electrons and topological phase transitions

We discuss here the very different case of a crystalline system whose wavefunction is explicitly correlated. Calculations have been performed [84, 85] for a one-dimensional model system, intended to reproduce the main features of an insulator having a mixed ionic/covalent character. The correlation is switched on by means of an interaction parameter U of the Hubbard type.

The model mimics an oxide, with two atoms per cell: a rigid relative displacement of the two sublattices corresponds to a zone-centre optical mode, or equivalently to a ferroelectric distortion. This non-centrosymmetric distortion induces a non-vanishing polarization, which is measured by the Born dynamical charge $Z^*|e|$. In the ideal rigid-ion limit $Z^* = \pm 2$, which is the nominal ionic charge in oxides. Even in the non-interacting case ($U = 0$) the actual dynamical charge is very different from the nominal one, owing to partial covalency [93].

The model has a very interesting behaviour as a function of U : at small U it is a band insulator, while at a critical U_c it undergoes a transition to a Mott-like insulating phase. Considering first the centrosymmetric geometry, z_N is a real number, which changes sign at U_c . The Berry phase γ jumps therefore by π : it turns out that this occurrence is the main feature signalling the transition. In fact, the topological quantum number [20] γ/π can be used as an order parameter to identify the two different phases of the system. This idea was later generalized to models richer than the one considered here, even displaying superconducting transitions [94, 95].

Considering then the distorted geometries of our one-dimensional model oxide and calculating its Z^* -value, it is found that Z^* diverges at the transition and changes sign in the highly correlated regime: we have therefore a swapping of roles between anion and cation as regards the dynamical charges. Notice that the *static* charge, instead, is continuous across U_c . We refer the reader to the original literature [32, 84, 85] for more about the physics of this model, while we give in the following some more details about the Berry's phase aspects of this calculation.

The original work, reference [84], was intended as a numerical implementation of the Ortíz–Martin theory [83], where the polarization of a correlated system is defined as a *continuum* Berry's phase: their main formula was then discretized for purely computational purposes. With hindsight, the calculation of reference [84] can nowadays be reinterpreted as a very interesting application of the single-point Berry phase, and the discretization assumes a most fundamental meaning. This is done in the following, where the same calculation is thus presented on the basis of a completely different logic.

The aim is to evaluate z_N for large N . We compute the correlated wavefunction Ψ by exact Lanczos diagonalization: the affordable sizes are not very large, and most of the results of reference [84] are obtained by means of eight-electron wavefunctions, which are not close enough to the thermodynamic limit. This drawback is overcome by building approximate wavefunctions for much larger sizes: we first exemplify how this works in the simple case of size doubling, and then we will generalize.

Suppose that at given size N we perform two independent calculations, one using periodic (BvK) boundary conditions over the segment of length L , and the other one using *antiperiodic* boundary conditions: let us call these wavefunctions Φ_0 and $\Phi_{\pi/L}$ respectively, which are both BvK periodic with a period $L' = 2L$. We then build an antisymmetric wavefunction for $N' = 2N$ electrons as

$$\Psi(x_1, \dots, x_{2N}) = \mathbf{A}\Phi_0(x_1, \dots, x_N)\Phi_{\pi/L}(x_{N+1}, \dots, x_{2N}) \quad (121)$$

where \mathbf{A} is the antisymmetrizer, and spinless electrons are used for simplicity. Of course, this *ansatz* wavefunction has restricted variational freedom. It is rather straightforward to prove

that $z_{N'}$ factors for even N as

$$z_{N'} = \langle \Psi | e^{i(2\pi/L')\hat{X}} | \Psi \rangle = \langle \Phi_{\pi/L} | e^{i(\pi/L)\hat{X}} | \Phi_0 \rangle \langle \Phi_0 | e^{i(\pi/L)\hat{X}} | \Phi_{\pi/L} \rangle \quad (122)$$

where each \hat{X} acts on the appropriate number of electrons.

The *ansatz* of equation (121) is generalized to larger systems in references [32, 84] using skew (quasiperiodic) boundary conditions with Bloch vector k over each electronic variable separately:

$$\Phi_k(x_1, \dots, x_i + L, \dots, x_N) = e^{ikL} \Phi_k(x_1, \dots, x_i, \dots, x_N) \quad (123)$$

and choosing M equally spaced values of k in the interval $[0, 2\pi/L)$:

$$k_s = \frac{2\pi}{ML} s \quad s = 0, 1, \dots, M-1. \quad (124)$$

The size of the periodically repeated system is $L' = ML$, and we build an *ansatz* wavefunction for $N' = MN$ electrons as the antisymmetrized product of the M different N -particle wavefunctions. In the simple case of $N = 1$ this construction yields the Slater determinant of M orbitals, and is therefore the exact wavefunction for a system of $N' = M$ non-interacting electrons. Upon choosing $N > 1$ one allows the MN electrons to correlate, but only in clusters of N at a time. At any given M , and for even N , we have

$$z_{N'} = z_{MN} = \langle \Psi | e^{i(2\pi/ML)\hat{X}} | \Psi \rangle = \prod_{s=0}^{M-1} \langle \Phi_{k_{s+1}} | e^{i(2\pi/ML)\hat{X}} | \Phi_{k_s} \rangle. \quad (125)$$

We then approximate the thermodynamic limit upon studying the large- M limit at fixed N . In the original references [32, 84], evidence is given that $z_{N'}$ converges fast to the thermodynamic limit.

In the uncorrelated case we may choose $N = 1$, and hence $N' = M$, as observed above: the Φ s become then single-particle orbitals, and equation (125) coincides with the one-dimensional analogue of equations (109) and (110) (for a single band and spinless electrons), i.e. a standard Zak's phase. Finally, it is worth stressing again that both equations (110) and (125) were first used as mere computational algorithms with the purpose of evaluating a *continuum* Berry's phase, while they assume here a much more fundamental meaning.

7. Conclusions

This review starts with a very general presentation of Berry's phase in quantum mechanics (sections 1 and 2). I then specialize to features which are peculiar to many-electron systems, as appropriate for molecules and solids (section 3). In the rest of this review I discuss several selected examples where a Berry's phase of the electronic system leads to a measurable effect in molecular physics and in condensed matter physics (sections 5 and 6).

The general presentation is rather unconventional, in that it is essentially based on a *discrete* viewpoint. There are at least two good reasons for this: (1) Berry's phases are actually computed using the discrete approach; and (2) the most extreme case of a discrete Berry's phase (namely, the single-point one) carries a very important meaning as a matter of principle in polarization theory.

In his milestone paper [1], Berry emphasizes the coupling with 'the rest of the Universe' as the very reason for the occurrence of geometric phases in quantum systems. The fast variables—in our case: the electron coordinates—are 'slaved' to some slowly varying parameter summarizing the coupling. The different cases discussed in sections 5 and 6 are in fact variations on this theme for different slow parameters: ionic coordinates in molecules

(section 5.1 and 5.2); coordinates of the wave-packet centre in crystals (section 5.3 and 5.4); fluctuating magnetization in crystals (section 5.5). The role of time is less evident in the dielectric polarization of a condensed system (section 6), which looks like a static property. We have instead emphasized that the measured bulk quantities are polarization *differences*, and these are induced by a slow parameter controlled by the experimenter (see e.g. figure 5).

An important difference between molecules and condensed systems concerns the boundary conditions adopted. In the latter case, one almost invariably assumes periodic Born–von Kármán boundary conditions: the system has no surface, and all properties are by definition *bulk*. This choice of boundary conditions plays a dominant role when a phenomenon in condensed matter is interpreted in terms of a Berry's phase, as we have shown for all the examples discussed in this review.

Most of the topics covered in this review have witnessed outstanding developments in very recent times. I have tried to present a timely and updated account of such developments. However, the field cannot be considered a mature one: on the contrary, the area of geometric phases in electronic structure theory is a fast moving one and further advances are to be expected in the near future. Let me just mention that, at the time of writing, some very promising developments are under way concerning geometric distances and electronic localization [31, 32, 96, 97].

Acknowledgments

Invaluable discussions with R M Martin, F Mauri, Q Niu, G Ortíz, S Sorella, and D Vanderbilt are gratefully acknowledged. I thank R Gebauer and S Baroni for making reference [76] available prior to publication. This work was partially supported by the Office of Naval Research through grant N00014-96-1-0689.

References

- [1] Berry M V 1984 *Proc. R. Soc. A* **392** 45
- [2] Shapere A and Wilczek F (ed) 1989 *Geometric Phases in Physics* (Singapore: World Scientific)
- [3] Aharonov Y and Bohm D 1959 *Phys. Rev.* **115** 485
Reprinted as Aharonov Y and Bohm D 1989 *Geometric Phases in Physics* ed A Shapere and F Wilczek (Singapore: World Scientific) p 104
- [4] Feynman R P, Leighton R B and Sands M 1964 *The Feynman Lectures in Physics* vol 2 (Reading, MA: Addison-Wesley) section 15-4
- [5] Peshkin M and Tonomura A 1989 *The Aharonov–Bohm Effect* (Berlin: Springer)
- [6] Herzberg G and Longuet-Higgins H C 1963 *Discuss. Faraday Soc.* **35** 77
- [7] Mead C A 1992 *Rev. Mod. Phys.* **64** 51
- [8] Longuet-Higgins H C, Öpik U, Pryce M H L and Sack R A 1958 *Proc. R. Soc. A* **244** 1
- [9] Zwanziger J W, Koenig M and Pines A 1990 *Annu. Rev. Phys. Chem.* **41** 601
- [10] Moore D J 1991 *Phys. Rep.* **210** 1
- [11] Berry M V 1988 *Sci. Am.* **259** (6) 46
- [12] Berry M V 1990 *Phys. Today* **43** (12) 34
- [13] Anandan J 1992 *Nature* **360** 307
- [14] Anandan J, Christian J and Wanelik K 1997 *Am. J. Phys.* **65** 180
- [15] Bohm A 1994 *Quantum Mechanics: Foundations and Applications* 3rd edn (Berlin: Springer)
- [16] Sakurai J J and Tuan S F 1994 *Modern Quantum Mechanics* revised edn (Reading, MA: Addison-Wesley)
- [17] Holstein B 1992 *Topics in Advanced Quantum Mechanics* (Reading, MA: Addison-Wesley)
- [18] Prange R E, Girvin S M, Cage M E and von Klitzing K 1990 *The Quantum Hall Effect* 2nd edn (New York: Springer)
- [19] Morandi G 1992 *The Role of Topology in Classical and Quantum Physics* (Berlin: Springer)
- [20] Thouless D J 1998 *Topological Quantum Numbers in Nonrelativistic Physics* (Singapore: World Scientific)

- [21] Resta R 1996 *Berry Phase in Electronic Wavefunctions (Lecture Notes for the 'Troisième Cycle de la Physique en Suisse Romande' (Lausanne, 1996))* available online (194K) at the URL: <http://www-dft.ts.infn.it/~resta/publ/notes.trois.ps.gz>
- [22] Sundaram G and Niu Q 1999 *Phys. Rev. B* **59** 14 915
- [23] Niu Q and Kleinman L 1998 *Phys. Rev. Lett.* **80** 2205
- [24] Niu Q *et al* 1999 *Phys. Rev. Lett.* **83** 207
- [25] Resta R 1998 *Phys. Rev. Lett.* **80** 1800
- [26] Yaschenko E, Fu L, Resca L and Resta R 1998 *Phys. Rev. B* **58** 1222
- [27] Resta R 1999 *Int. J. Quantum Chem.* **75** 599
- [28] Anandan J and Aharonov Y 1990 *Phys. Rev. Lett.* **65** 1697
- [29] Anandan J 1990 *Phys. Lett. A* **147** 3
- [30] Pati A K 1991 *Phys. Lett. A* **159** 105
- [31] Marzari N and Vanderbilt D 1997 *Phys. Rev. B* **56** 12 847
- [32] Resta R and Sorella S 1999 *Phys. Rev. Lett.* **82** 370
- [33] Zak J 1989 *Phys. Rev. Lett.* **62** 2747
- [34] Bohm A 1994 *Quantum Mechanics: Foundations and Applications* 3rd edn (Berlin: Springer) ch 23
- [35] Schmelter P, Cederbaum L S and Meyer H-D 1988 *Phys. Rev. A* **38** 6066
- [36] Yin L and Mead C A 1994 *J. Chem. Phys.* **100** 8125
- [37] King-Smith R D and Vanderbilt D 1993 *Phys. Rev. B* **47** 1651
- [38] Resta R 1993 *Europhys. Lett.* **22** 133
- [39] Resta R 1994 *Rev. Mod. Phys.* **66** 899
- [40] Resta R 1997 *Europhys. News* **28** 18
- [41] A good account of the first-principles work can be found in the most recent proceedings of the biannual Williamsburg Workshops on 'First-Principles Calculations for Ferroelectrics':
Ferroelectrics 1995 **164**
Ferroelectrics 1997 **194**
 Cohen R E (ed) 1998 *AIP Conf. Proc.* **436**
- [42] Pancharatnam S 1956 *Proc. Indian Acad. Sci. A* **44** 247
 Reprinted as Pancharatnam S 1989 *Geometric Phases in Physics* ed A Shapere and F Wilczek (Singapore: World Scientific) p 51
- [43] Berry M V 1987 *J. Mod. Opt.* **34** 1401
 Reprinted as Berry M V 1989 *Geometric Phases in Physics* ed A Shapere and F Wilczek (Singapore: World Scientific) p 67
- [44] Simon R and Mukunda N 1993 *Phys. Rev. Lett.* **70** 880
- [45] Arnold V I 1989 *Mathematical Methods of Classical Mechanics* (New York: Springer) ch 7
- [46] Zak J 1989 *Europhys. Lett.* **9** 615
- [47] Debernardi A, Bernasconi M, Cardona M and Parrinello M 1997 *Appl. Phys. Lett.* **71** 2692
 Silvestrelli P L, Bernasconi M and Parrinello M 1997 *Chem. Phys. Lett.* **277** 478
 Bernasconi M, Silvestrelli P L and Parrinello M 1998 *Phys. Rev. Lett.* **81** 1235
- [48] Pasquarello A and Car R 1997 *Phys. Rev. Lett.* **79** 1766
- [49] McWeeny R 1992 *Methods of Molecular Quantum Mechanics* 2nd edn (London: Academic)
- [50] Massidda S, Resta R, Posternak M and Baldereschi A 1995 *Phys. Rev. B* **52** 16 977
- [51] Massidda S, Posternak M, Baldereschi A and Resta R 1999 *Phys. Rev. Lett.* **82** 430
- [52] Schiff L I 1968 *Quantum Mechanics* (Singapore: McGraw-Hill)
- [53] Mead C A and Truhlar D G 1979 *J. Chem. Phys.* **70** 2284
- [54] Delacrétaz G, Grant E R, Whetten R L, Wöste L and Zwanziger J W 1986 *Phys. Rev. Lett.* **56** 2598
 Reprinted as Delacrétaz G, Grant E R, Whetten R L, Wöste L and Zwanziger J W 1989 *Geometric Phases in Physics* ed A Shapere and F Wilczek (Singapore: World Scientific) p 240
- [55] Cocchini F, Upton Th H and Andreoni W 1988 *J. Chem. Phys.* **88** 6068
- [56] Ernst W E and Rakowsky S 1995 *Phys. Rev. Lett.* **74** 58
- [57] Kendrick B 1997 *Phys. Rev. Lett.* **79** 2431
- [58] von Busch H *et al* 1999 *Phys. Rev. Lett.* **81** 4584
- [59] Dugourd Ph *et al* 1994 *Chem. Phys. Lett.* **225**
- [60] Goss Levi B 1993 *Phys. Today* **46** (3) 17
- [61] Ashcroft N W and Mermin N D 1976 *Solid State Physics* (Philadelphia, PA: Saunders)
- [62] Ashcroft N W and Mermin N D 1976 *Solid State Physics* (Philadelphia, PA: Saunders) ch 12
- [63] Peierls R 1933 *Z. Phys.* **80** 763
 Jones H and Zener C 1934 *Proc. R. Soc.* **144** 101

- [64] Slater J C 1959 *Phys. Rev.* **76** 1592
- [65] Luttinger J M 1951 *Phys. Rev.* **84** 814
- [66] Zak J 1968 *Phys. Rev.* **168** 686
- [67] Tabor M 1989 *Chaos and Integrability in Nonlinear Dynamical Systems: an Introduction* (New York: Wiley)
- [68] Wannier G H 1960 *Phys. Rev.* **117** 432
- [69] Wannier G H 1962 *Rev. Mod. Phys.* **34** 645
- [70] Zak J 1968 *Phys. Rev. Lett* **20** 1477
- [71] Mendez E E, Agulló-Rueda F and Hong J M 1988 *Phys. Rev. Lett.* **60** 2426
- [72] Mendez E E and Bastard G 1993 *Phys. Today* **46** (6) 34
- [73] Raizen M, Salomon Ch and Niu Q 1997 *Phys. Today* **50** (7) 30
- [74] Kittel C 1996 *Introduction to Solid State Physics* 7th edn (New York: Wiley)
- [75] Kittel C 1996 *Introduction to Solid State Physics* 7th edn (New York: Wiley) ch 15
- [76] Gebauer R and Baroni S 1999 *Preprint cond-mat/9908386*
- [77] Landau L D and Lifshitz E M 1984 *Electrodynamics of Continuous Media* (Oxford: Pergamon)
- [78] Martin R M 1974 *Phys. Rev.* **B 9** 1998
- [79] Resta R 1992 *Ferroelectrics* **136** 51
- [80] Resta R 1994 *Ferroelectrics* **151** 49
- [81] Resta R 1996 *Quantum-Mechanical Ab-Initio Calculation of the Properties of Crystalline Materials* (Springer Lecture Notes in Chemistry vol 67) ed C Pisani (Berlin: Springer) p 273
- [82] Martin R M and Ortíz G 1997 *Solid State Commun.* **102** 121
- [83] Ortíz G and Martin R M 1994 *Phys. Rev.* **B 43** 14 202
- [84] Resta R and Sorella S 1995 *Phys. Rev. Lett.* **74** 4738
- [85] Ortíz G, Ordejón P, Martin R M and Chiappe G 1995 *Phys. Rev.* **B 54** 13 515
- [86] Giannozzi P, de Gironcoli S, Pavone P and Baroni S 1991 *Phys. Rev.* **B 43** 7231
- [87] Nunes R W and Vanderbilt D 1994 *Phys. Rev. Lett.* **73** 712
Dal Corso A and Mauri F 1994 *Phys. Rev.* **B 50** 5756
Fernández P, Dal Corso A and Baldereschi A 1998 *Phys. Rev.* **B 58** R7480
- [88] Selloni A, Carnevali P, Car R and Parrinello M 1987 *Phys. Rev. Lett.* **59** 823
Fois E S, Selloni A, Parrinello M and Car R 1988 *J. Phys. Chem.* **92** 3268
- [89] Thouless D J 1983 *Phys. Rev.* **B 27** 6803
Niu Q and Thouless D J 1984 *J. Phys. A: Math. Gen.* **17** 2453
- [90] Car R and Parrinello M 1985 *Phys. Rev. Lett.* **55** 2471
- [91] Car R and Parrinello M 1988 *Simple Molecular Systems at Very High Density* (NATO ASI Series B, vol 186) ed P Loubeyre, A Polian and N Boccara (New York: Plenum) p 455
- [92] Galli G and Parrinello M 1991 *Computer Simulation in Materials Science* ed M Meyer and V Pontikis (Dordrecht: Kluwer)
- [93] Posternak M, Resta R and Baldereschi A 1994 *Phys. Rev.* **B 50** 8911
- [94] Aligia A A and Gagliano E R 1998 *Physica C* **304** 29
- [95] Aligia A A 1999 *Europhys. Lett.* **45** 411
- [96] Silvestrelli P L and Parrinello M 1999 *Phys. Rev. Lett.* **82** 3308
Silvestrelli P L and Parrinello M 1999 *J. Chem. Phys.* **111** 3572
- [97] Aligia A A and Ortíz G 1999 *Preprint cond-mat/9910490*