Bulk-boundary correspondence from the inter-cellular Zak phase

Jun-Won Rhim, Jan Behrends, and Jens H. Bardarson
Max-Planck-Institut für Physik komplexer Systeme, 01187 Dresden, Germany

The Zak phase $\gamma$, the generalization of the Berry phase to Bloch wave functions in solids, is often used to characterize inversion-symmetric 1D topological insulators; however, since its value can depend on the choice of real-space origin and unit cell, only the difference between the Zak phase of two regions is believed to be relevant. Here, we show that one can extract an origin-independent part of $\gamma$, the so-called inter-cellular Zak phase $\gamma_{\text{inter}}$, which can be used as a bulk quantity to predict the number of surface modes as follows: a neutral finite 1D tight-binding system has $n_s = \gamma_{\text{inter}}/\pi \pmod{2}$ number of in-gap surface modes below the Fermi level if there exists a commensurate bulk unit cell that respects inversion symmetry. We demonstrate this by first verifying that $\pm e^{i\gamma_{\text{inter}}/2\pi} \pmod{e}$ is equal to the extra charge accumulation in the surface region for a general translationally invariant 1D insulator, while the remnant part of $\gamma$, the intra-cellular Zak phase $\gamma_{\text{intra}}$, corresponds to the electronic part of the dipole moment of the bulk’s unit cell. Second, we show that the extra charge accumulation can be related to the number of surface modes when the unit cell is inversion symmetric. This bulk-boundary correspondence using $\gamma_{\text{intra}}$ reduces to the conventional one using $\gamma$ when the real-space origin is at the inversion center. Our work thereby clarifies the usage of $\gamma$ in the bulk-boundary correspondence. We study several tight binding models to quantitatively check the relation between the extra charge accumulation and the inter-cellular Zak phase as well as the bulk-boundary correspondence using the inter-cellular Zak phase.

I. INTRODUCTION

The Berry phase is a geometric phase of eigenstates obtained when cyclically varying external parameters. For a Hamiltonian $\mathcal{H}(\mathbf{R})$ that depends on the external parameters $\mathbf{R}$, the Berry phase is given by

$$\gamma_n = i \int_C \langle n(\mathbf{R})|\nabla_{\mathbf{R}}|n(\mathbf{R})\rangle \cdot d\mathbf{R},$$

where $|n(\mathbf{R})\rangle$ is the $n$-th eigenvector and $C$ is a closed loop in the parameter space of $\mathbf{R}$. In a Brillouin zone, as pointed out by Zak, a natural choice for the cyclic parameter is the crystal momentum $\mathbf{k}$. The explicit form of the resulting Zak phase of the $n$-th band is

$$\gamma_n = i \int_{\text{BZ}} dk \langle u_{n,k}|\partial_k|u_{n,k}\rangle,$$

where BZ represents the 1D Brillouin zone, and $u_{n,k} = \sqrt{N} e^{-ikx} \psi_{n,k}$ is the periodic part of the Bloch function $\psi_{n,k}$. The total Zak phase $\gamma$ is obtained by summing $\gamma_n$ over filled bands.

The Zak phase was endowed with a physical meaning with the modern definition of the polarization $\mathbf{P}$, introduced by Vanderbilt and King-Smith. The ambiguity of the classical polarization, which depends on the shape of the unit cell boundary, was resolved by redefining the electronic part of the polarization $\mathbf{P}^{\text{el}}$ as an integral of Wannier functions over the whole space, rather than over a unit cell. With this definition, one can accurately predict the bound surface charge, $\sigma = \mathbf{P} \cdot \hat{n}$ with $\hat{n}$ the surface orientation, measured in capacitance experiments.

Furthermore, the electronic polarization $P_\perp$ along $\hat{n}$ is related to the Zak phase (evaluated along $\hat{n}$ with $\mathbf{k}_\parallel$ kept fixed) via

$$P_\perp = P^{\text{cl}} \cdot \hat{n} = -\frac{e}{(2\pi)^3} \int_A d\mathbf{k}_\parallel \sum_{n=1}^Z \gamma_n,$$

where $A$ is the 2D Brillouin zone projected to the 2D plane perpendicular to $\hat{n}$, $\mathbf{k}_\parallel$ is the momentum component in this plane, and $Z$ the number of filled bands.

In the currently active study of topological aspects of materials, the Zak phase has been utilized as a topological number to classify various genuine 1D topological insulators, as well as effective ones, such as those obtained by fixing one or two momenta of 2D or 3D Hamiltonians. However, it was naturally extended to the concept of non-abelian Wilson loops in the multi-band case and used for classifications of topological insulators with inversion or non-symmorphic symmetries and topological crystalline insulators. Furthermore, the Zak phase has been widely used for the $\mathbb{Z}_2$ classification of inversion symmetric 1D systems where it is quantized to 0 or $\pi$ (mod 2$\pi$). In this case, the conventional bulk-boundary correspondence states that there are boundary modes if the Zak phase is non-trivial, $\gamma = \pi$, while $\gamma = 0$ is considered a trivial insulator without surface modes.

Recent work, however, pointed out that the Zak phase depends on the gauge choice of choosing the origin of the real space, and how one defines boundaries of the unit cell, although it is invariant under gauge transformation of the form $u_{n,k} \to e^{i\phi} u_{n,k}$. This means that the Zak phase itself is not a well-defined topological number since it cannot characterize the bulk uniquely. In an attempt to resolve this ambiguity, Atala et al. suggested that the difference of the Zak phase between different states could be a proper topological number, and Juan et. al. revised the Zak phase by adding a unit cell dependent term such that the resultant $\mathbb{Z}_2$ number
plays the role of a gauge invariant topological number. In spite of these issues, the conventional bulk-boundary
 correspondence using the Zak phase has been successfully applied in many cases\cite{Wang2017, Wang2018}. Furthermore, additional conditions for the applicability of the correspondence to finite systems have been given, such as that terminated edges should not break the inversion symmetry of the bulk\cite{Wang2017, Wang2018} or that the finite system should be commensurate with the bulk’s unit cell\cite{Wang2019}. However, the necessity of those assumptions has never been demonstrated in general 1D systems, and we even find a counterexample for the former one in Sec. IV B where the conventional bulk-boundary correspondence using $\gamma$ does not work even in the presence of inversion symmetry both in the bulk and the terminated system.

In this paper, we clarify these issues by providing a more detailed analysis of the Zak phase, concentrating on general 1D tight-binding models. To this end, we split the Zak phase into two terms, the \textit{intra-cellular} and \textit{inter-cellular} Zak phase (this splitting was earlier introduced by Kudin et al\cite{Kudin2002}, and provide them with their proper physical interpretations. The intra-cellular Zak phase $\gamma_{\text{intra}}$ describes the electronic part of the classical polarization of the bulk’s unit cell, and the inter-cellular Zak phase $\gamma_{\text{inter}}$ represents the difference between the net weight of the Wannier functions in the left and right sides of the 1D system with respect to a unit cell boundary, with their centers belonging to opposite sides as illustrated in Fig. II. We further show that, if the system is terminated, this interpretation of the inter-cellular Zak phase leads to an accurate prediction of the extra charge accumulations in a region around the surface. We assume that these surface regions are commensurate with the bulk’s unit cell, and large enough that effects of the termination are negligible deep inside the bulk. Here, the extra charge accumulation is just the total charge in the surface region including ionic contributions. Note that it is a different quantity from the \textit{bound surface charge} introduced by Vanderbilt et al\cite{Vanderbilt1990}, which can be measured in capacitance experiments. For instance, the bound surface charge can be nonzero even when the surface regions remain neutral while the extra charge accumulation vanishes in this case. Although we have obtained a general relation between $\gamma_{\text{inter}}$ and the extra charge accumulation around surfaces for any translationally invariant systems, it cannot generally be used to count the number of surface modes in the gap below the Fermi level, since it includes charge densities both from bulk bands and surface states; however, we show that it is possible if inversion symmetry is respected in the unit cell. In this case, there are $n_s = \frac{\gamma_{\text{inter}}}{\pi}$ (mod 2) surface modes below the Fermi level in a finite system, if it is commensurate with an inversion symmetric unit cell of the bulk and $\gamma_{\text{inter}}$ is evaluated from this commensurate unit cell. We can replace $\gamma_{\text{inter}}$ with $\gamma$ if the real-space origin is chosen to be located at the inversion center of the unit cell because $\gamma_{\text{intra}}$, an origin-dependent quantity, vanishes in this case. In other words, the conventional bulk-boundary correspondence using the total Zak phase is valid if there exist an inversion symmetric bulk unit cell commensurate with the finite system and the real-space origin is at the inversion center.

The rest of the paper is organized as follows: in Sec. II and III we provide the physical meaning of the intra- and inter-cellular Zak phase and show how they are related, respectively, to the electronic part of the classical polarization of the bulk and the extra charge accumulation in the surface region. In Sec. IV based on this physical interpretation of the parts of the Zak phase, we develop a reformulated bulk-boundary correspondence of one-dimensional systems employing the inter-cellular Zak phase. In Sec. V we apply our interpretation of the inter-cellular Zak phase and the bulk-boundary correspondence to the Rice-Mele 1D-chain model and two 2D toy models which can be treated as effective 1D systems by fixing one of the momenta. Finally, discussions and concluding remarks are given in Sec. VI.

II. INTRA- AND INTER-CELLULAR ZAK PHASE

Let us consider a general translationally-invariant one-dimensional system described within the tight-binding approximation. The system consists of $N_b$ atomic sites per unit cell, each with $N_{\text{orb}}$ orbitals, including the spin degree of freedom for spinful systems. The ionic charge at site $i$ is $Z_i e$, and the total ionic charge in the unit cell is $Z e = \sum_{i=1}^{N_b} Z_i e$. In order for the insulating system to be neutral, there are then necessarily $Z$ fully filled bands below the Fermi level. Within the Born-von Karman boundary condition, $\psi(x + Na) = \psi(x)$, the Bloch eigenfunctions of this system take the form

$$\psi_{n,k}(x) = \frac{1}{\sqrt{N}} \sum_{m=1}^{N_b} \sum_{\xi=1}^{N_{\text{orb}}} \alpha_{n,i,\xi}^{m,i,\xi} e^{i k ma}$$ (4)

where $N$ is the number of unit cells, and $m$ and $n$ are the unit cell and band index. Here,

$$\phi_{m,i,\xi}^{i,\xi}(x) = \phi_{i,\xi}(x - ma - b_i a)$$ (5)

is the $\xi$-th atomic orbital centered at $ma + b_i a$ which is the position of the $i$-th atomic site in the $m$-th unit cell. Atomic orbitals at the same site are orthonormal and the overlaps between orbitals on different sites are assumed to be exponentially vanishing in accordance with the tight-binding condition. The coefficients $\alpha_{n,i,\xi}^{m,i,\xi}$ are obtained from solving the eigenvalue problem with the tight-binding Hamiltonian. We choose the gauge in which every atomic orbital in the $m$-th unit cell has the same phase factor $e^{ikma}$ in (4); the coefficients $\alpha$ then satisfy the periodic boundary condition $\alpha_{n,i,\xi}^{m,i,\xi} = \alpha_{k+G,i,\xi}$, where $G$ is a reciprocal lattice vector. In another widely used gauge, the phase factor for the orbitals at $ma + b_i a$ is instead given by $e^{ik(ma+b_i a)}$\cite{Wang2017, Wang2018, Wang2019}. In this case,
the same Bloch eigenfunction \( \psi_{n,k}(x) \) is obtained by replacing the coefficient \( \alpha_{k}^{n,i,\zeta} \) with \( \tilde{\alpha}_{k}^{n,i,\zeta} = \alpha_{k}^{n,i,\zeta} e^{-iGb_k} \), which satisfies the twisted boundary condition \( \tilde{\alpha}_{k+G}^{n,i,\zeta} = \tilde{\alpha}_{k}^{n,i,\zeta} e^{-iGb_k} \). While both gauges yield the same Bloch eigenfunction \( \psi_{n,k}(x) \), the former allows for a natural separation of the Zak phase, since they describe the same Bloch eigenfunction \( \psi \), the former allows for a natural separation of the Zak phase from the lattice periodic part of the Bloch function.

The Zak phase, following Kudin et al. \(^{43}\), can be split into two terms. The \( n \)-th band Zak phase \( \gamma_{n} \) is obtained from the lattice periodic part of the Bloch function

\[
\psi_{n,k}(x) = \sqrt{N} e^{-ikx} \psi_{n,k}(x) \tag{6}
\]

which satisfies the orthonormality condition

\[
\langle u_{n,k} | u_{n,k}' \rangle = \int_{\Omega} dx u_{n,k}(x)^* u_{n,k}'(x) = \delta_{k,k}' \tag{7}
\]

The Zak phase is defined as

\[
\langle u_{n,k} | \partial_{k} u_{n,k} \rangle = \int dx u_{n,k}(x) \frac{\partial}{\partial k} u_{n,k}(x), \tag{8}
\]

with \( \Omega \) a unit cell. Then, the Zak phase is split into the intra- and inter-cellular Zak phase as

\[
\gamma_{n} = \gamma_{n}^{\text{intra}} + \gamma_{n}^{\text{inter}} \tag{9}
\]

where

\[
\gamma_{n}^{\text{intra}} = N \int_{BZ} dk \int_{\Omega} dx |\psi_{n,k}(x)|^2 - 2\pi m_{\Omega}, \tag{10}
\]

with \( m_{\Omega} \) the index of the unit cell \( \Omega \), and

\[
\gamma_{n}^{\text{inter}} = \sum_{i=1}^{N_{\text{orb}}} \sum_{\zeta=1}^{N_{\text{el}}} \int_{BZ} dk a_{k}^{n,i,\zeta} \frac{\partial}{\partial k} \alpha_{k}^{n,i,\zeta}. \tag{11}
\]

(See Appendix A for further details). Note that \( \gamma_{n}^{\text{intra}} \) depends on the real-space origin while \( \gamma_{n}^{\text{inter}} \) does not, because \( \alpha_{k}^{n,i,\zeta} \), being an element of the eigenvector of the momentum space representation of the Hamiltonian, is origin-independent.

To give a physical interpretation of the Zak phases just introduced, we start by relating the electronic part of the classical polarization of the unit cell to the intra-cellular Zak phase. We consider contributions from each filled band separately. For the \( n \)-the band, we have

\[
P_{n,\text{el}} = \frac{1}{a} \int_{\Omega} dx x \rho_{\text{bulk},n}(x) \tag{12}
\]

where \( \rho_{\text{bulk},n}(x) \) is the electronic density corresponding to the \( n \)-th band. That is, \( P_{n,\text{el}} \) can be evaluated from the intra-cellular Zak phase up to mod \( e \).

Secondly, we give a physical interpretation of the inter-cellular Zak phase. To this end, we employ the Wannier functions given by

\[
W_{n,m}(x) = \frac{1}{\sqrt{N}} \sum_{k \in \text{BZ}} \psi_{n,k}(x) e^{-ikma} \tag{13}
\]

where

\[
A_{n,m}^{i,\zeta} = \sum_{m'} \sum_{i=1}^{N_{\text{orb}}} \sum_{\zeta} \int_{BZ} dk a_{k}^{n,i,\zeta} \frac{\partial}{\partial k} a_{k}^{n,i,\zeta}. \tag{14}
\]

They satisfy the orthonormality condition

\[
\langle W_{n,m} | W_{n',m'} \rangle = \delta_{n,n'} \delta_{m,m'}. \tag{15}
\]

FIG. 1. (Color online) A schematic figure describing a general 1D system and its Wannier functions. Gray dashed boxes are the unit cells, which contain atomic sites marked by filled and empty circles as an example. Each unit cell can in general have any number of atomic sites. The 1D bulk is divided into left and right sides with respect to a given unit cell boundary \( x_b \), which is marked by the red dashed line. We call these regions the left and right side according to their centers. \( Q_{n}^{R \rightarrow L} \) represents how much weight of right side’s Wannier functions are in the left side, and vice versa for \( Q_{n}^{L \rightarrow R} \).
Wannier function $W_{n,m}(x)$ is in general guaranteed to be exponentially localized around a position in the $m$-th unit cell\cite{45–47} while, in 2D and 3D, this is true if and only if the Chern number of the band vanishes\cite{28,49,50}. Then, we express the $n$-th band inter-cellular Zak phase, via the inverse transformation of (16), as

$$\gamma_{n}^{\text{inter}} = 2\pi \sum_{m} N_{0} N_{n}^{m} \sum_{i=1}^{N_{m}} m \left| A_{n,m,i}^{i} \right|^{2}. \quad (17)$$

By fixing a point $x_{b}$ at the boundary between neighbouring unit cells, which we assume for concreteness to be between the $m = -1$ and $m = 0$ unit cells, the system is split into the left ($L$) and right ($R$) regions as depicted in Fig. 1. With these definition, the inter-cellular Zak phase can be split into two parts as

$$\gamma_{n}^{\text{inter}} = -\gamma_{n}^{R \rightarrow L} + \gamma_{n}^{L \rightarrow R}, \quad (18)$$

where

$$\gamma_{n}^{R \rightarrow L} = -2\pi \sum_{m=\infty}^{1} N_{m} \sum_{i=1}^{N_{m}} m \left| A_{m,n,i}^{i} \right|^{2} \quad (19)$$

and

$$\gamma_{n}^{L \rightarrow R} = 2\pi \sum_{m=0}^{\infty} \sum_{i=1}^{N_{m}} m \left| A_{m,n,i}^{i} \right|^{2} \quad (20)$$

Further details are given in Appendix A. From (20) and (21) we note that $Q_{n}^{R \rightarrow L} = -e/(2\pi) \sum_{m} \gamma_{n}^{R \rightarrow L}$ represents the charge weight in the $L$ region of all Wannier functions with their centers belonging to the $R$ region, and vice versa for $Q_{n}^{L \rightarrow R} = -e/(2\pi) \sum_{m} \gamma_{n}^{L \rightarrow R}$. Finally, their difference in the $L$ region is simply expressed by

$$\Delta Q_{L} = Q_{n}^{R \rightarrow L} - Q_{n}^{L \rightarrow R} = \frac{e}{2\pi} \sum_{n=1}^{Z} \gamma_{n}^{\text{inter}} \quad (23)$$

where the sum is over the filled bands. In the $R$ region, this difference is just given by $\Delta Q_{R} = -(e/2\pi) \sum_{n=1}^{Z} \gamma_{n}^{\text{inter}}$. Due to the translational symmetry of the system, this result is valid for any unit-cell boundary. Also, this quantity does not depend on the position of the origin since the inter-cellular Zak phase does not as mentioned previously. Note that, from the point of view of the electronic density, every unit cell is neutral and there is no extra charge in the $L$ and $R$ regions of the bulk. However, we want to eventually obtain the extra charge accumulation around the surfaces when the system is terminated; we demonstrate how it is related to $Q_{n}^{R \rightarrow L}$ and $Q_{n}^{L \rightarrow R}$ in the next section.

### III. EXTRA CHARGE ACCUMULATIONS AROUND SURFACES

In this section we consider a terminated system, divided it into three regions as illustrated in Fig. 2: the left surface region ($LS$) from $m = 1$ to $m = \ell_{L}$ and the right surface ($RS$) region from $m = N - \ell_{R} + 1$ to $m = N$, separated by the bulk region ($B$) from $m = \ell_{L} + 1$ to $m = N - \ell_{R}$. We choose $\ell_{L}$, $\ell_{R}$ and $N$ all large enough that there are no boundary effects in the bulk, and all regions are commensurate with the bulk’s unit cell, namely they do not contain any partial unit cells. While the extra charge accumulation in the surface region is defined as the additional charge over the bulk’s charge distribution, it is equivalent to the total charge including ions in this region, because of the commensurability condition and the neutrality of the bulk’s unit cell. Finally, we assume that the finite system is also insulating because otherwise partially filled degenerate states at the Fermi level would yield ambiguity in evaluating physical quantities depending on which states we choose to be occupied. Then, we show that, if the surface region is commensurate with the bulk’s unit cell, the extra charge accumulation $Q_{n}^{LS(RS)}$ in the (left) surface region of a neutral 1D insulator is given by

$$Q_{n}^{LS(RS)} = +\left(-\frac{e}{2\pi}\right) \gamma_{n}^{\text{inter}} \quad (24)$$

where $\gamma_{n}^{\text{inter}} = \sum_{n=1}^{Z} \gamma_{n}^{\text{inter}}$.

The overall strategy for the demonstration of the relation (24) is as follows. We prepare a set of orthonormal wave functions that are related to the set of occupied eigenfunctions of the finite system by a unitary transformation, following Vanderbilt et al. All of these states are localized, i.e., Wannier-like, such that their characteristic widths are much smaller than sizes of the three regions. First, there are exchanges of weights of the Wannier functions between the surface and bulk region across their boundary as expressed by $\Delta Q_{L}$ and $\Delta Q_{R}$ in the previous section. Then, from the orthonormality and localized feature of the prepared basis set, we show that the total charge in the left and right surface region should be $\Delta Q_{L}$ and $\Delta Q_{R}$ modulo $e$.

The orthonormal localized wave functions that we employ for the description of the terminated system are given by

$$\varphi = \{ \varphi_{1}^{LS}(x), \ldots, \varphi_{N_{L}}^{LS}(x), W_{1,m_{L}}(x), \ldots, W_{N_{R}}^{m_{R}}(x), \varphi_{1}^{RS}(x), \ldots, \varphi_{N_{R}}^{RS}(x) \}^T. \quad (25)$$

While the Wannier functions are the proper basis set for the infinite system, after the termination we need to construct this new set of basis wave functions that satisfy the open boundary condition. $\varphi$ span the same Hilbert space as the $M$ occupied eigenstates of the terminated system denoted by $\psi_{1}^{n_{L}}(x) = \{ \psi_{1}^{n_{L}}(x), \ldots, \psi_{M}^{n_{L}}(x) \}^T$. First, we pick all occupied Wannier functions from the unit cell.
\[ m_L \text{ (in the left surface region)} \] to \[ m_R \text{ (in the right surface region)} \]. The choice of \[ m_L \] and \[ m_R \] is arbitrary but chosen to be far enough from the surfaces so that corresponding Wannier functions satisfy the boundary conditions with exponentially vanishing error. We then replace the Wannier functions around the surfaces with another set of wave functions that satisfy the open boundary condition; these we denote by \( \psi_{x,s}^{L,S}(x) \) and \( \psi_{x,s}^{R,S}(x) \) around the left and right boundary each. The set of wave functions in (25) can always be made orthonormal through the Gram-Schmidt orthonormalization process.45-52

The charge density of the occupied electrons in the terminated system is given by \( \rho_{x,s}^{el}(x) = -e \psi_{x,s}^{L,S}(x) \) \( \psi_{x,s}^{L,S}(x) \). Since \( \psi_{x,s}^{L,S}(x) \) and \( \varphi(x) \), as orthonormal bases, span the same Hilbert space of occupied states, they are related by a unitary matrix, and the electronic charge is rewritten as

\[
\rho_{x,s}^{el}(x) = -e \psi_{x,s}^{L,S}(x) \tag{27}
\]

The total charge in the left surface region, which is the extra charge accumulation, is then evaluated as

\[
Q_{acc}^{LS} = \int_{0}^{x_L} dx \rho_{x,s}^{el}(x) + \ell_L Z e \tag{28}
\]

where \( \ell_L Z e \) is the total ionic charge in the left surface region—we assume that the left-end of the system is at \( x = 0 \) and \( x_m = am \) is the position of the right boundary of the \( m \)-th unit cell. In this region, the integrals of \( \varphi_{x,s}^{RS}(x) \)'s in (27) vanish, while the sum of integrals of \( \varphi_{x,s}^{LS}(x) \)'s, which are well localized to the right and left surface, is an integer number \( s_L \). Then, \( Q_{acc}^{LS} \) becomes

\[
Q_{acc}^{LS} = -e \int_{0}^{x_L} dx \sum_{n=1}^{Z} \sum_{m=m_L}^{m_R} |W_{n,m}(x)|^2 + q_L e \tag{29}
\]

\[
= -e \sum_{n=1}^{Z} \sum_{m=m_L}^{m_R} \int_{0}^{x_L} dx |W_{n,m}(x)|^2 \tag{30}
\]

\[
= \sum_{n=1}^{Z} \sum_{m=m_L}^{m_R} \int_{0}^{x_L} dx |W_{n,m}(x)|^2 + q_L e \tag{31}
\]

This leads to the conclusion

\[
Q_{acc}^{LS} = -e \sum_{n=1}^{Z} (\gamma_{n-L} - \gamma_{n-R}) + q_L e \tag{32}
\]

where \( q_L = -s_L + \ell_L Z \). Since all the Wannier functions are normalized, we have the identity

\[
\sum_{n=1}^{Z} \sum_{m=m_L}^{m_R} \int_{0}^{x_L} dx |W_{n,m}(x)|^2 = (\ell_L - m_L + 1) - \sum_{n=1}^{Z} \sum_{m=m_L}^{m_R} \int_{0}^{x_L} dx |W_{n,m}(x)|^2 . \tag{31}
\]

FIG. 2. (Color online) A schematic of the three regions (bulk, left and right surface region) of a finite 1D system with \( N \) unit cells. Classical bound surface charges (\( Q_{acc}^{LS(RS)} = P_{el} \cdot \hat{n} \)) and extra charge accumulations (\( Q_{acc}^{LS(RS)} \)) in two surface regions are also illustrated schematically. The system is terminated at \( x = 0 \) and \( x_N \). Boundaries between those three regions are represented by red dashed lines at \( x_L \) and \( x_{N-L} \). Bulk unit cells are drawn by gray dashed boxes and their indices are given below them. In each bulk unit cell, the dipole moment (\( P_{el} \)) of the bulk unit cell is drawn symbolically.
role of \( x_b \) in (20) and (22) is taken by \( x_{L_i} \) in (3) and (4). In the same way, the extra charge accumulation in the right surface region is given by

\[
Q_{acc}^{RS} = -\frac{e}{2\pi} \sum_{n=1}^{Z} \gamma_{n}^{inter} + q'_R e
\]

where \( q'_R = -s_R + Z(N - m_R) \). As in Ref. 4, we are not interested in the integer numbers \( q'_L \) and \( q'_R \) which cannot be determined from the bulk. Instead, we conclude that the inter-cellular Zak phase can predict the extra charge accumulation in the surface region modulo \( e \) as described in (24).

Finally, we remark on the difference between the bound surface charge and the extra charge accumulation. The bound surface charge \( \sigma \) is a quantity measured in a capacitance measurement and is related to the modern definition of the polarization by \( \sigma = \mathbf{P} \cdot \hat{n} \), where \( \hat{n} \) is the surface orientation (\( \hat{n} = -(+\hat{x}) \) for the left (right) edge in the 1D case). The bound surface charge of the total charge density \( \rho_{t.s.}(x) \) of the finite system is evaluated explicitly as:

\[
\sigma^{LS} = \frac{1}{a} \int_{-\infty}^{x_e} dx \int_{x-a}^{x+\frac{a}{2}} dx' \rho_{t.s.}(x') \tag{35}
\]

for the bound surface charge at the left edge, and

\[
\sigma^{RS} = \frac{1}{a} \int_{x_d}^{\infty} dx \int_{x-a}^{x+\frac{a}{2}} dx' \rho_{t.s.}(x') \tag{36}
\]

at the right edge, where \( x_e \) and \( x_d \) are arbitrary positions in the middle of the finite system far away from the surfaces. As shown in Appendix B \( \sigma^{LS} \) and \( \sigma^{RS} \) are independent of \( x_e \) and \( x_d \). To clearly compare the bound surface charge and the extra charge accumulation, we choose them as \( x_e = x_{L_1} + a/2 \) and \( x_d = x_{N - L_R} - a/2 \). Then the left and right bound surface charge becomes

\[
\sigma^{LS} = -\frac{1}{a} \int_{x_{L_1}}^{x_{L_1} + a} dx \ x \rho_{t.s.}(x) + \int_{0}^{x_{L_1}} dx \ x \rho_{t.s.}(x), \tag{37}
\]

and

\[
\sigma^{RS} = \frac{1}{a} \int_{x_{N - L_R} - a}^{x_{N - L_R}} dx \ x \rho_{t.s.}(x) + \int_{x_{N - L_R}}^{x_{N}} dx \ x \rho_{t.s.}(x). \tag{38}
\]

Derivations of the above are given in Appendix B. The first terms of (37) and (38) are exactly \( \mathbf{P}_{cl} \cdot \hat{n} \) because \( x_{L_1} \) and \( x_{N - L_R} \) are far enough from the edges so that \( \rho_{t.s.}(x) \) can be considered a bulk charge density. We call them the \textit{classical bound surface charges} \( Q_{cl}^{LS(RS)} \) because in classical electrodynamics they are the bound surface charges in a dielectric material with a uniform dipole distribution through the whole finite system. On the other hand, the second terms of (37) and (38) are just the total charge in the left and right surface regions, which are equal to the extra charge accumulations in those regions in the neutral systems. Consequently, we have shown that the bound surface charge consists of two contributions, one from the bulk’s dipole moment, and the other from the extra charge accumulation. This is consistent with the splitting of the Zak phase into the intra- and inter-cellular Zak phase. Also, it indicates that the modern polarization is actually composed of two kinds of polarizations, one is the classical one from the bulk’s dipole moment, and the other is polarization from the extra charge accumulations at opposite edges. This conclusion is illustrated schematically in Fig. 2.

\[\text{IV. BULK-BOUNDARY CORRESPONDENCE}\]

Our general statement about the number of surface modes in the gap of a finite 1D insulator with charge neutrality is as follows: there are \( n_s = \gamma^{inter}/\pi \) (mod 2) surface modes below the Fermi level if there is inversion symmetry both before and after termination, and the finite system is commensurate with the bulk unit cell used for the calculation of \( \gamma^{inter} \). We justify our general statement in the light of the physical interpretation of the inter-cellular Zak phase \( \gamma^{inter} \). As discussed in the previous section, the inter-cellular Zak phase explains the amount of the extra charge accumulation around the edge which is closely related to the surface modes, while the intra-cellular Zak phase, as a bulk dipole moment, has nothing to do with the surface modes.

First, we note that if the bulk respects inversion symmetry, \( \gamma^{inter} \) is quantized to \( \pi \). When the real-space origin is at one of the inversion centers, the total Zak phase \( \gamma \) is quantized to \( \pi \) while its intra-cellular part vanishes because the dipole moment of the bulk unit cell is zero in this case. Therefore, \( \gamma^{inter} = \gamma \) is also quantized to \( \pi \), and it is independent of the choice of the origin as mentioned in Sec. III.

The surface modes are the eigenstates generated in the bulk gaps as a result of the edge termination. As described in the previous section, they are exponentially localized at one of the edges. Due to inversion symmetry, if we have a surface state localized at the left edge, we always find its counterpart localized at the right edge with the same energy. The degeneracy might increase if the system preserves additional symmetries such as the time reversal symmetry.

We proceed to the justification of our general statement by examining two different cases. One is when the Fermi level of the finite system is not at any of the surface modes, so that the finite system is also an insulator. And, the other is the case when the Fermi level is at the partially filled degenerate surface modes.

First, let us consider the case when the finite system is already insulating. Then, the only allowed value of the total charge accumulations is \( Q_{acc}^{LS} = Q_{acc}^{RS} = 0 \) due to charge neutrality and inversion symmetry. According to (24), the inter-cellular Zak phase should be \( 0 \) (mod \( 2\pi \)) in this case. Since every surface mode has an inversion
partner with the same energy, there should be an even number of surface states below the Fermi level. This means the general statement holds for insulating phases with inversion symmetry.

Second, we consider the case when the Fermi level is located at the surface modes. In this case, we cannot use the results of the previous section since there we assume the finite system is insulating. However, we can resolve this obstacle by opening tiny gaps between degenerate surface modes without changing the total number of surface modes in the bulk gap. We assume that this can be done by applying local perturbative potentials on both edges that break symmetries corresponding to those degeneracies. For instance, different on-site potentials at opposite edges would break the degeneracy responsible for inversion symmetry. Similarly, a local Zeeman field could be used to break spin degeneracy.

Let us investigate the dependence of the surface modes on the tiny inversion symmetry breaking, which is relevant for the quantization of the extra charge accumulations on the edges as explained below. This is achieved by applying an on-site potential \( \delta \) to the first unit cell and \(-\delta\) to the \(N\)-th unit cell of the finite system where \( \delta \) is nonzero but small compared with the bulk gap. Then, the eigenstates at opposite edges have different energies, and it is impossible to construct surface modes localized to both edges simultaneously from them. As a result, the eigenstate of each lifted surface mode is localized to only one of the left and right edges. This implies that each surface mode yields and integer charge \( Q_{LS}^{RS} \) is a half-integer multiple of \( e \) implies that the inter-cellular Zak phase should be a half-integer multiple of \( 2\pi \) according to [21]. In conclusion, the number of surface modes is equal to \( \gamma_{\text{inter}}/\pi \) mod 2 in this case.

In similar fashion, when \( Q_{B,\text{acc}}^{LS(RS)} \) = \( pe \), the surface modes' extra charge accumulations on both edges should be \( \sum_{E_{\gamma}<E_F} Q_{LS}^{RS} + Q_{RS}^{LS} \) = \(-2pe\) to maintain charge neutrality, implying that there is an even number of surface modes in the main gap and that the inter-cellular Zak phase is even. So, the number of surface modes is again equal to \( \gamma_{\text{inter}}/\pi \) mod 2 in this case.

Finally, we note that this bulk-boundary correspondence can be restated by using the Zak phase when the real-space origin is at an inversion center. This is because the intra-cellular Zak phase becomes zero (mod 2\( \pi \)) according to [9], so that the Zak phase is identical to the inter-cellular Zak phase in this case.

V. EXAMPLES

In this section we provide a few examples demonstrating the general results of the earlier sections.

A. Rice-Mele 1D chain

In this subsection, we calculate the extra charge accumulation in the Rice-Mele model[45,55] and show that it is accurately predicted by the inter-cellular Zak phase and is different from the bound surface charge. Also, for inversion symmetric cases of this model, we show that the bulk-boundary correspondence works well.

The Rice-Mele model is given by

\[
\mathcal{H}_{\text{RM}} = \sum_{\sigma,j} \epsilon_j c_{\sigma j}^\dagger c_{\sigma j} + \sum_{\sigma,j} [V_{\sigma,j+1} c_{\sigma j}^\dagger c_{\sigma j+1} + \text{h.c.}] \tag{39}
\]

where \( \sigma \) and \( j \) are indices for the spin and the lattice sites. For neutrality of the system at half-filling, we introduce spin degeneracy. The tight binding parameters are given by \( \epsilon_{2p} = -\Delta \), \( \epsilon_{2p+1} = \Delta \), \( V_{2p-1,2p} = -t - \delta \) and \( V_{2p,2p+1} = -t + \delta \) where \( p \) is an integer. Note that there are two sites per unit cell. When \( \Delta = 0 \), the model reduces to the Su-Schrieffer-Heeger model[22].
The Bloch wave function for $\mathcal{H}_{\text{RM}}$ is described by

$$\psi_{n,k}(x) = \frac{1}{\sqrt{N}} \sum_{m} e^{i k m} [\alpha_{n,k} \phi_{m}^{A}(x) + \beta_{n,k} \phi_{m}^{B}(x)]$$

(40)

where

$$\phi_{m}^{A}(x) = \phi(x - m)$$

and

$$\phi_{m}^{B}(x) = \phi(x - m - \frac{1}{2})$$

(41)

and $\phi(x)$ is the atomic wave function at the lattice sites. In this section we set $a = 1$ for convenience.

The intra- and inter-cellular Zak phases are given by

$$\gamma_{n,\sigma}^{\text{intra}} = \int_{0}^{2\pi} dk \frac{1}{2} |\beta_{n,k}|^2$$

(42)

and

$$\gamma_{n,\sigma}^{\text{inter}} = \int_{0}^{2\pi} dk \left( \alpha_{n,k}^{*} \frac{\partial \alpha_{n,k}}{\partial k} + \beta_{n,k}^{*} \frac{\partial \beta_{n,k}}{\partial k} \right).$$

(43)

Since there is no couplings between the two spin species, we have $\gamma_{n,\uparrow} = \gamma_{n,\downarrow}$. According to the results in Sec. IIII the classical polarization for a given unit cell is

$$P_{\text{cl}} = -\frac{e}{2\pi} \sum_{\sigma} \sum_{n=1}^{Z} \gamma_{n,\sigma}^{\text{intra}} + \frac{e}{2}$$

(44)

and the extra charge accumulations at the edges are

$$Q_{\text{acc}}^{\text{LS(RS)}} = +(-) \frac{e}{2\pi} \sum_{\sigma} \sum_{n=1}^{Z} \gamma_{n,\sigma}^{\text{inter}} \mod (2e).$$

(45)

Here, $Z = 1$ since we consider two spins independently.

We check this relation between $Q_{\text{acc}}^{\text{LS(RS)}}$ and the inter-cellular Zak phase numerically in Fig. 3(b) where the system is parameterized by $t = 1$, $\Delta = 0.6 \cos \theta$ and $\delta = 0.6 \sin \theta$. Also, for the bound surface charge ($\sigma = \mathbf{P} \cdot \hat{n}$), we applied the formulae $\delta^{53,54}$ and $\delta^{55}$ and reproduced exactly the results obtained by Vanderbilt et al. as shown in Fig. 3(a). The difference between those two quantities corresponds to the classical bound surface charge $Z_{\text{cl}} = \mathbf{P}_{\text{cl}} \cdot \hat{n}$ originating from the dipole moment of the unit cell as we discussed in Sec. IIII Since the dipole moment vanishes when the unit cell respects inversion symmetry, those two quantities become identical to each other at $\theta = \pi/2$ and $\theta = 3\pi/2$ where $\Delta = 0$ as shown in Fig. 3.

As in Fig. 3(b), the inter-cellular Zak phase becomes $2\pi$ for $\theta = \pi/2$ and zero for $\theta = 3\pi/2$, which predicts an even number of surface bands below the Fermi level according to the bulk-boundary correspondence in the previous section. In the finite system with an even number of sites, we find two surface bands for $\theta = \pi/2$ and no surface bands for $\theta = 3\pi/2$ below the Fermi energy (data not shown), consistent with the bulk-boundary correspondence. However, if we consider two spins separately, we have more information. For each spin species, the inter-cellular Zak phase is just the half of those in Fig. 3(b), namely $\pi$ for $\theta = \pi/2$ and zero for $\theta = 3\pi/2$ which implies there is an odd and even number of surface states in each case. This means we must have surface modes when $\theta = \pi/2$ while one cannot determine the existence of the surface bands when $\theta = 3\pi/2$.

B. 2D toy model-I

Here, we suggest a 2D toy model as a counterexample for the conventional bulk-boundary correspondence of the Zak phase. From this, we clarify its right usage by emphasizing the assumption of the existence of the commensurate bulk’s unit cell. We also test whether the inter-cellular Zak phase predicts the extra charge accumulation accurately or not.

Our toy model-I is illustrated in Fig. 4(a) and described...
FIG. 4. (Color online) The lattice structure of the 2D toy model-I. The unit cell used for the calculations of the intercellular Zak phases is the dashed box. (b) The Brillouin zone of the given system. Four Dirac nodes at \( k_y = \pi/2 \) are marked by red dots. The Brillouin zone is divided into several regions depending on the value of the Zak phase: \( \gamma = \pi \) is gray and \( \gamma = 2\pi \) is white. (c) Band spectrum at \( k_y = \pi/2 \). The parameters used are \( t = \epsilon_0 = 0.5 \).

by the Hamiltonian

\[
\mathcal{H}^{\text{I}}_{\text{2D}} = \sum_{\mathbf{R},i} \left( \epsilon_A a_{i,\mathbf{R}}^\dagger a_{i,\mathbf{R}} + \epsilon_B b_{i,\mathbf{R}}^\dagger b_{i,\mathbf{R}} \right) + t \sum_{\mathbf{R},\delta_x} \left( a_{2,\mathbf{R}+\delta_x,1,\mathbf{R}}^\dagger a_{1,\mathbf{R}} + b_{2,\mathbf{R}+\delta_x,2,\mathbf{R}}^\dagger b_{1,\mathbf{R}} \right) + \text{h.c.} \tag{46}
\]

where \( \mathbf{R} \) is the position vector of the unit cell, and \( \delta_x = \pm a\hat{x} \) and \( \delta_y = \pm a\hat{y} \) are the nearest neighbor vectors.

There are two orbitals at each site denoted by \( i = 1, 2 \). The electron hops between different orbitals along \( x \) and between the same orbitals along \( y \). Along the \( y \) direction, the sign of the hopping parameter of the \( i = 1 \) and \( i = 2 \) orbitals are opposite.

For \( \epsilon_A = -\epsilon_B = \epsilon_0 \), the energy spectrum is given by

\[
E^{\text{I}}_{\text{2D}}(k_x, k_y) = \pm \left( \epsilon_0 + 2t \cos k_x + 4t^2 \cos^2 k_y \right)^{1/2} \tag{47}
\]

where \( a \) is set to be unity. The spectrum can become gapless only at \( k_y = \pi/2 \), where it becomes...
$E_{2D}^{I}(k_x,\pi/2) = \pm (\epsilon_0 \pm 2t \cos k_x)$ which has four Dirac
nodes at $k_x = \pm \cos^{-1}(\pm \epsilon_0/2t)$ for $0 < \epsilon_0 < 2t$ as shown
in Fig. 4(b) and (c) where $t = \epsilon_0 = 0.5$.

For the Zak phase analysis, we consider an effective
1D Hamiltonian obtained by fixing the momentum
$k_x$. From this, we can calculate the Zak phase by integrating
the Berry connection along $k_y$. The results are presented
in Fig. 4(b) by the gray ($\gamma = \pi$) and white ($\gamma = 2\pi$)
regions. Due to reflection symmetry, which corresponds
to inversion symmetry in the effective Hamiltonian, we
have quantized Zak phases. Whenever we cross one of
the Dirac nodes, the Zak phase changes discontinuously
by $\pm \pi$.

Now, let us consider ribbon-shaped finite systems that
are finite in the $y$-direction but infinite in the $x$-direction.
We consider two kinds of ribbon geometries, which we call
Ribbon-I(A) and Ribbon-I(S) as illustrated in the insets of
Fig. 4(b) and (c). Ribbon-I(A) does not respect reflection
symmetry along the $x$ axis while Ribbon-I(S) does.

As a result, only Ribbon-I(S) remains gapless without
breaking the Dirac nodes of the bulk.

According to the conventional bulk-boundary corre-
spondence, we expect Ribbon-I(S) to have surface modes
at $k_x$’s in the gray or white region in Fig. 4(b) because
both the bulk and terminated system have reflection sym-
metry. However, it turns out that there are no boundary
modes in the Brillouin zone for both ribbon geometries as
shown in Fig. 5. This is a good example of the wrong us-
age of the bulk-boundary correspondence using the Zak
phase. In the revised correspondence given in the pre-
vious section, it requires the commensurability between
the finite system and the bulk’s unit cell in addition to
inversion symmetry. Ribbon-I(A) is commensurate with
the unit cell described in Fig. 5(a), but does not respect
mirror symmetry while Ribbon-I(S), which is reflection
symmetric, cannot be commensurate with any bulk’s unit
cell as it consists of an odd number of dimer lines. There-
fore, those two systems are actually beyond the scope of
the revised bulk-boundary correspondence.

We check that the inter-cellular Zak phase accurately
predicts the extra charge accumulation in the surface re-
regions in Fig. 6. From (24), the extra charge accumulation for a given $k_x$ is given by

$$Q_{\text{acc}}^{B(T)}(k_x) = \frac{e}{2\pi} \sum \gamma_{\text{inter}}(k_x) \quad \text{(mod } e)$$

where $B(T)$ and the plus(minus) sign correspond to the bottom(top) surface, and the sum is over occupied bands. In the case of Ribbon-I(A), both bottom and top surfaces are commensurate with a single unit cell and share the same inter-cellular Zak phase. As a result, the extra charge accumulations in the bottom and top surfaces show opposite signs to each other as shown in Fig. 6(a). On the other hand, for Ribbon-I(S), the commensurate unit cell for the top surface is the one obtained by shifting the commensurate unit cell for the bottom by $ay$. We find that the inter-cellular Zak phases for those unit cells have the same magnitude but opposite signs. While the extra charge accumulations in the bottom and top surfaces should be identical due to the mirror symmetry, it is consistent with the above property of inter-cellular Zak phases. This is described in Fig. 6(b).

C. 2D toy model-II

Finally, we study another 2D toy model which manifests the vitalness of the conditions for the bulk-boundary correspondence, such as inversion symmetry and insulating terminated system. We again confirm that the extra charge accumulation in the surface regions is captured by the inter-cellular Zak phase.

The toy model-II, depicted in Fig. 7(a), has a tight-binding model given by

$$\mathcal{H}_{\text{2D}}^{II} = \sum_{\mathbf{R}} \mathbf{c}_{\mathbf{R}}^\dagger \mathbf{H}_0 \mathbf{c}_{\mathbf{R}} + \sum_{\mathbf{R}, \delta_x} \mathbf{c}_{\mathbf{R} + \delta_x}^\dagger H_x \mathbf{c}_{\mathbf{R}} + \sum_{\mathbf{R}, \delta_y} \mathbf{c}_{\mathbf{R} + \delta_y}^\dagger H_y \mathbf{c}_{\mathbf{R}}$$

(49)

where $\mathbf{c}_{\mathbf{R}} = (a_{1,\mathbf{R}}^\dagger, a_{2,\mathbf{R}}^\dagger, \mathbf{b}_{1,\mathbf{R}}^\dagger, \mathbf{b}_{2,\mathbf{R}}^\dagger, \mathbf{c}_{1,\mathbf{R}}^\dagger, \mathbf{c}_{2,\mathbf{R}}^\dagger)$ consists of the creation operators of the orbitals at $A$, $B$ and $C$ sites. There are two orbitals per site. The matrices are defined as $H_0 = \text{diag}(\epsilon_{A,1}, \epsilon_{A,2}, \epsilon_{B,1}, \epsilon_{B,2}, \epsilon_{C,1}, \epsilon_{C,2})$, $H_x = 1 \otimes (\sigma^x + \sigma^z)$ and $H_y = (\lambda_1 + \lambda_6) \otimes \sigma_z$, where $I$ is the $3 \times 3$ identity matrix, $\sigma^x$ is the Pauli matrix and $\lambda_i$ is the Gellman matrix. The band-structure at $k_y = 0$ is depicted in Fig. 7(b). While there are four Dirac points at $k_y = 0$, as marked by red dots, it is insulating otherwise. We consider this model because we can investigate a finite system with both reflection symmetry and commensurate bulk’s unit cell as shown below.

As in the previous 2D toy model-I, we consider the effective 1D Hamiltonian for a fixed $k_x$ and calculate the inter-cellular Zak phases by integrating the Berry connection along the $k_y$ axis. While $\gamma_{\text{inter}}$ depends on the choice of the unit cell, we consider the two unit cells, called UC-II(A) and UC-II(S), illustrated in Fig. 7(a). Here, we assume that $\epsilon_{A,i} = \epsilon_{C,i}$ so that UC-II(S) preserves reflection symmetry while UC-II(A) does not. Calculations of $\gamma_{\text{inter}}$ are shown in Fig. 8 by red solid curves along $k_x$. The inter-cellular Zak phase of UC-II(S) is quantized to a multiple of $\pi$ due to the reflection symmetry while it has continuous values for UC-II(A). For both cases, the inter-cellular Zak phase shows the discrete jump at every Dirac node as in the 2D toy model-I.

We consider two terminated systems called Ribbon-II(A) and Ribbon-II(S), which are commensurate with UC-II(A) and UC-II(S) each. As in the toy model-I, they are terminated along the $x$ axis. We plot the extra charge accumulations in the bottom surface of those systems in Fig. 8. For the case of Ribbon-II(S), we break the reflection symmetry slightly so that one of the equivalent inter-cellular Zak phases $\pi$ and $-\pi$ is selected to describe the extra charge accumulation in the bottom surface. One can note that the extra charge accumulation is well described by the inter-cellular Zak phase for all values of $k_x$ except those in the shaded interval. This failure in this region is because the condition that the finite system is...
also insulating is violated. As shown in Fig. 9(b), in the shaded region, the topmost band at half-filling joins the conduction band continuum. Here, ‘half-filling’ means the effective 1D system at each \( k_x \) is half-filled. Also, in this region, the states on the topmost band are no longer exponentially localized. On the other hand, in Ribbon-II(S), one can see the extra charge accumulation perfectly at the surface, we represent it by a solid curve. On the other hand, if the eigenstate is bulk-like, it is drawn with a dashed line.

VI. CONCLUSIONS

In this work, we have demonstrated that the inter-cellular Zak phase \( \gamma_{\text{inter}} \) can predict whether the number of surface modes below the Fermi level in 1D insulators is even or odd, when the commensurate bulk unit cell respects inversion symmetry. While the Zak phase itself cannot do this due to its arbitrariness depending on the choice of the real-space origin and the unit cell, we have shown that \( \gamma_{\text{inter}} \), as an origin-independent quantity, can be exploited for this bulk-boundary correspondence. Although \( \gamma_{\text{inter}} \) also depends on the unit cell choice it is not arbitrary once we select a unit cell that is commensurate with the finite system. Our bulk-boundary correspondence using \( \gamma_{\text{inter}} \) was justified with a microscopic interpretation of \( \gamma_{\text{intra}} \) and \( \gamma_{\text{inter}} \). We explicitly showed that \( \gamma_{\text{intra}} \) is the electronic part of the bulk dipole moment of the unit cell, and \( \gamma_{\text{inter}} \) represents how much weight of the Wannier functions are exchanged with respect to a unit cell boundary. When the system is terminated, \( \gamma_{\text{intra}} \) is interpreted as the classical bound surface charge, while \( \gamma_{\text{inter}} \) is understood as the extra charge accumulation around surfaces. Since the number of surface modes is closely related to the extra charge accumulation, we argue how it is related to \( \gamma_{\text{inter}} \) when the commensurate unit cell preserves inversion symmetry. If the origin is at the inversion center, \( \gamma_{\text{inter}} \) becomes identical to the Zak phase, and our bulk-boundary correspondence reduces to the conventional one. Thereby, our work also clarifies the conditions under which the conventional bulk-boundary correspondence using the Zak phase works.

We expect that the extra charge accumulation can be measured by scanning quantum dot microscopy (SQDM). SQDM offers three-dimensional images of electrostatic potentials down to the subnanometer level from which one could infer the total amount of its source charge. Since the electric field caused by extra charge accumulations at opposite edges of a long enough 1D chain can be considered independent, the extra charge accumulation at one edge can be obtained from the local electrostatic potential profile. Therefore, SQDM could be the characterizing experiment for the inter-cellular Zak phase like the capacitance measurement for the Zak phase.
ACKNOWLEDGMENTS

We thank D. Vanderbilt, L. Balents, J. E. Moore, F. de Juan, A. G. Grushin and S. Park for useful discussions. This work was supported by the ERC Starting Grant No. 679722.

Appendix A: Intra- and inter-cellular Zak phases

When we calculate the Zak phase, the momentum derivative operates on the exponential factor $e^{ik(ma - x)}$ and the coefficient $\alpha_{n,i,\zeta}^m$ of the lattice periodic part of the Bloch function (6). That is,

$$\gamma_n = i \int_{BZ} dk \langle u_{n,k} | \partial_k u_{n,k} \rangle$$

First, the intra-cellular Zak phase is evaluated as follows.

$$\gamma_{\text{intra}}^n = i \int_{BZ} dk \int_\Omega dx \sum_{m,m'} \sum_{i,i'=1}^{N_{\text{orb}}} \sum_{\zeta,\zeta'=1}^{\delta_{m,m'}} \alpha_{n,i,\zeta}^m \phi_m^i(x) \phi_m'^{i'}(x) e^{-ik(ma - x)} \frac{\partial e^{ik(m'a - x)}}{\partial k}$$

$$= \int_{BZ} dk \int_\Omega dx \sum_{m,m'} \sum_{i,i'=1}^{N_{\text{orb}}} \sum_{\zeta,\zeta'=1}^{\delta_{m,m'}} (x - m'a) \alpha_{n,i,\zeta}^m \phi_m^i(x) \phi_m'^{i'}(x) e^{ik(m'a - x)}$$

$$= \int_{BZ} dk \int_\Omega dx \sum_{m,m'} \sum_{i,i'=1}^{N_{\text{orb}}} \sum_{\zeta,\zeta'=1}^{\delta_{m,m'}} \alpha_{n,i,\zeta}^m \phi_m^i(x) \phi_m'^{i'}(x) e^{ik(m'a - x)}$$

First, the intra-cellular Zak phase is evaluated as follows.

$$\gamma_{\text{intra}}^n = \int_{BZ} dk \int_\Omega dx \sum_{m,m'} \sum_{i,i'=1}^{N_{\text{orb}}} \sum_{\zeta,\zeta'=1}^{\delta_{m,m'}} \alpha_{n,i,\zeta}^m \phi_m^i(x) \phi_m'^{i'}(x) e^{ik(m'a - x)}$$

$$= \int_{BZ} dk \int_\Omega dx \sum_{m,m'} \sum_{i,i'=1}^{N_{\text{orb}}} \sum_{\zeta,\zeta'=1}^{\delta_{m,m'}} \alpha_{n,i,\zeta}^m \phi_m^i(x) \phi_m'^{i'}(x) e^{ik(m'a - x)}$$

The first term is defined as the intra-cellular Zak phase and the second as the inter-cellular Zak phase.

where $m_\Omega$ is the index of the unit cell $\Omega$. From (A4) to (A5), we used the orthonormality condition of the Löwdin functions, $\int_\Omega dx \phi_m^i(x)^* \phi_m'^{i'}(x) = \delta_{m,m'} \delta_{i,i'} \delta_{\zeta,\zeta'}$. The second term of (A6) is obtained because the coefficients of the eigenstate for every $n$ and $k$ are normalized.

Second, the inter-cellular Zak phase is calculated as
\[ \gamma_{\text{inter}} = i \int_{BZ} dk \int_{BZ} dx m A_{m}^{n,i,\zeta} e^{-ikma} \]

Substituting into Eq. (A10), we have

\[ \gamma_{n}^{\text{inter}} = \sum_{i,\zeta} \sum_{m,m'} \int_{BZ} dk m A_{m}^{n,i,\zeta} A_{m'}^{n,i,\zeta} e^{ik(m-m')a} \]

where the summation ranges of \( i \) and \( \zeta \) are the same as those of (A10).

We further split \( \gamma_{n}^{\text{inter}} \) into \( \gamma_{n}^{R-L} \) and \( \gamma_{n}^{R-L} \), which satisfies \( \gamma_{n}^{\text{inter}} = -\gamma_{n}^{R-L} + \gamma_{n}^{R-L} \), as follows.

\[ \gamma_{n}^{R-L} = -2\pi \sum_{m=-\infty}^{\infty} \sum_{i=1}^{N_b} \sum_{\zeta=1}^{N_{\sigma_{hb}}} m |A_{m}^{n,i,\zeta}|^2 \]

\[ = 2\pi \sum_{m=-\infty}^{\infty} \sum_{i=1}^{N_b} \sum_{\zeta=1}^{N_{\sigma_{hb}}} |A_{m}^{n,i,\zeta}|^2 \]

\[ = 2\pi \sum_{m'}=0 \int_{-x_b}^{x_b} dx |W_{n,m'}(x)|^2 \]

and

\[ \gamma_{n}^{L-R} = 2\pi \sum_{m=0}^{\infty} \sum_{i=1}^{N_b} \sum_{\zeta=1}^{N_{\sigma_{hb}}} m |A_{m}^{n,i,\zeta}|^2 \]

\[ = 2\pi \sum_{m'=0}^{\infty} \sum_{i=1}^{N_b} \sum_{\zeta=1}^{N_{\sigma_{hb}}} |A_{m'}^{n,i,\zeta}|^2 \]

\[ = 2\pi \sum_{m'=0}^{\infty} \int_{x_0}^{\infty} dx |W_{n,m'}(x)|^2 \]

where \( x_b \) is the boundary between the \( m = -1 \) and the \( m = 0 \) unit cells. Although we specify the boundary \( x_b \), the result is independent of it due to translational invariance, as it is clear from the dependence on \( m - m' \) in Eqs. (A15) and (A18).

**Appendix B: Bound surface charge**

In this section, we show that the bound surface charge can be split into the classical bound surface charge and the extra charge accumulation. The bound surface charge of the left edge is given by

\[ \sigma^{LS} = \frac{1}{a} \int_{-\infty}^{x_0} dx \int_{x-x_c}^{x-x_c} dx' \rho_{L,S}(x') = \int_{-\infty}^{x_0} dx \tilde{\rho}_{L,S}(x) \]

where \( \rho_{L,S}(x) \) is the total charge density including both the electronic and ionic contributions, \( \tilde{\rho}_{L,S}(x) = a^{-1} \int_{x-x_c}^{x-x_c} dx \rho_{L,S}(x') \), and \( x_c \) is an arbitrary position in the finite system, far away from the surfaces compared with the widths of the \( \varphi^L_{\delta} \)s.

First, note that \( \sigma^{LS} \) is independent of \( x_c \) since

\[ \frac{\partial \sigma^{LS}}{\partial x_c} = \frac{1}{a} \int_{x_c-x_c/2}^{x_c-x_c+2} dx' \rho_{L,S}(x') = \tilde{\rho}_{L,S}(x_c) \]

vanishes when \( x_c \) is far from the surfaces. The integral of the charge density over one unit cell length \( a \) is zero far from the surfaces because our system is assumed to be neutral.

Second, the expression for the bound surface charge can transformed as follows:

\[ \sigma^{LS} = -\int_{-\infty}^{x_c} dx \frac{d\tilde{\rho}_{L,S}(x)}{dx} x + \tilde{\rho}_{L,S}(x) [x]_{-\infty}^{x_c} \]

\[ = \frac{1}{a} \int_{-\infty}^{x_c} dx \left[ \rho_{L,S}(x+a/2) - \rho_{L,S}(x-a/2) \right] x \]

\[ = \frac{1}{a} \int_{x_c-x_c/2}^{x_c-x_c+2} dx \rho_{L,S}(x) + \int_{x_c-x_c/2}^{x_c-x_c+2} dx \rho_{L,S}(x) \]

where \( x_{\pm} = x \pm a/2 \). We have used the fact that \( \rho_{L,S}(x_c) = \tilde{\rho}_{L,S}(-\infty) = 0 \) in going from (B3) to (B4), and that \( \rho_{L,S}(x) = 0 \) for \( x < 0 \) from (B4) to (B5). If \( x_c = x_L + a/2 \), we have

\[ \sigma^{LS} = -\frac{1}{a} \int_{x_L}^{x_L+a} dx x \rho_{L,S}(x) + \int_{0}^{x_L} dx \rho_{L,S}(x) \]
In the same way, the bound surface charge at the right edge becomes

$$\sigma^{RS} = \frac{1}{a} \int_{xN-tR-a}^{xN-tR} dx \, x \rho_{s.s.}(x) + \int_{xN-tR}^{xN} dx \, x \rho_{s.s.}(x).$$  \hspace{1cm} (B7)