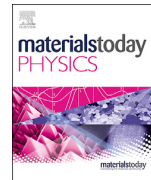




Contents lists available at ScienceDirect

Materials Today Physics

journal homepage: <https://www.journals.elsevier.com/materials-today-physics>

Interband tunneling effects on materials transport properties using the first principles Wigner distribution

Andrea Cepellotti*, Boris Kozinsky

Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, 02138, USA

ARTICLE INFO

Article history:

Received 16 December 2020

Received in revised form

7 April 2021

Accepted 8 April 2021

Available online 16 April 2021

Keywords:

Thermoelectrics

Electron transport

Wigner distribution

Narrow gap semiconductors

ab-initio simulations

ABSTRACT

Electronic transport in narrow gap semiconductors is characterized by spontaneous vertical transitions between carriers in the valence and conduction bands, a phenomenon also known as Zener tunneling. However, this effect is not captured by existing models based on the Boltzmann transport equation. In this work, we propose a new fully first principles model for electronic transport using the Wigner distribution function and implement it to solve the equations of motion for electrons. The formalism generalizes the Boltzmann equation to materials with strong interband coupling and include transport contributions from off-diagonal components of the charge current operator. We illustrate the method with a study of Bi_2Se_3 , showing that interband tunneling dominates the electron transport dynamics at experimentally relevant small doping concentrations, a behavior that is likely shared with other semiconductors, including topological insulators. Surprisingly, Zener tunneling occurs also between band subvalleys separated by energy much larger than the band gap.

© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Small band gap semiconductors are expected to possess transport characteristics that differ significantly from conventional wide-gap materials due to Zener (or Klein) tunneling, a phenomenon in which carriers undergo vertical transitions across the band gap, which can substantially increase electrical conductivity. This phenomenon is relevant for a number of novel systems, such as small band gap nanotubes or multi-layer graphene-based systems [1,2]. Topological insulators are another important example, thanks to their variety of interesting physical properties and promising applications, such as low-power electronics and quantum computing [3–7]. Topological insulators are characterized by conductive surfaces, while their bulk phases have a small band gap, and they may therefore display Zener tunneling and enhanced bulk conduction, which warrants investigation.

The de-facto tool of choice for first-principles simulation studies of electronic transport properties in crystals is the ab initio Boltzmann Transport Equation (aiBTE), which provides estimates of transport properties in good agreement with experimental measurements (e.g. [8–13]). However, this semiclassical approach is

not always sufficient to model electronic transport properties. Zener tunneling, for example, is not captured by semiclassical models, as they omit contributions from off-diagonal components of the flux operators. Sophisticated models based on the non-equilibrium Green's function methods [14] are capable of overcoming the limitations of semiclassical models. However, lack of efficient numerical and theoretical formulations prevents the combination of these approaches with first-principles computations that can treat realistic materials, and neglect the critical influence of scattering on the charge flux. Here we explore a new approach based on the Wigner function: this formalism is particularly appealing as it is exact in principle and, as discussed below, reduces to the aiBTE when off-diagonal flux components are neglected. The use of the Wigner function to study materials' electronic properties has been limited so far to studies of model Hamiltonians (see e.g. Ref. [15] for a recent review). In contrast, first-principles modeling can provide not only a more accurate description of real materials' Hamiltonian, but also an accurate description of electronic scattering that is responsible for the relaxation of the out-of-equilibrium system. This work formulates the generalized transport approach building on recent progress in first-principles computations of electron scattering rates and fills this methodological gap, opening possibility for the study of first-principles electronic properties of realistic materials.

In this work, we derive and implement from first-principles an

* Corresponding author.

E-mail address: acepellotti@g.harvard.edu (A. Cepellotti).

equation of motion for electrons, termed the ab initio Wigner transport equation (aiWTE) based on the single-particle Wigner distribution function. The formalism is capable of describing the space-time evolution of electrons, including effects due to electronic scattering, and captures the off-diagonal contributions of flux operators to transport properties. This equation is explicitly solved for the set of electronic transport coefficients, in particular electrical conductivity and Seebeck coefficient, showing how they can be significantly different from their semiclassical counterpart due to the presence of off-diagonal components of the Wigner distribution function. We apply the formalism to a first-principles study of the bulk transport properties of the topological insulator Bi₂Se₃. We show that at small doping concentrations, the aiWTE estimates of bulk electronic transport properties deviate significantly from semiclassical estimates, due to the presence of Zener tunneling, i.e. vertical transitions that couple carriers of valence and conduction bands.

2. Theory

$$\begin{aligned} \frac{\partial f_{bb'ss'}(\mathbf{x}, \mathbf{k}, t)}{\partial t} + \frac{i}{\hbar} [\mathcal{E}(\mathbf{k}) + \mathbf{d}(\mathbf{k}) \cdot \mathbf{E}, f_{bb'ss'}(\mathbf{x}, \mathbf{k}, t)]_{bb'ss'} + \frac{1}{2} \left\{ \mathbf{v}(\mathbf{k}), \frac{\partial f_{bb'ss'}(\mathbf{x}, \mathbf{k}, t)}{\partial \mathbf{x}} \right\}_{bb'ss'} - \\ - e\mathbf{E} \cdot \frac{\partial f_{bb'ss'}(\mathbf{x}, \mathbf{k}, t)}{\partial \mathbf{k}} = - \left. \frac{\partial f_{bb'ss'}(\mathbf{x}, \mathbf{k}, t)}{\partial t} \right|_{coll} \end{aligned} \quad (5)$$

We start by considering the ground state Hamiltonian H_0 of a crystal, which we assume to be an independent-particle Hamiltonian with eigenvalues $\epsilon_{bs}(\mathbf{k})$ and Bloch states $\psi_{kbs}(\mathbf{x})$, where \mathbf{x} is the position, \mathbf{k} the wavevector, b the band index and $s = \pm 1$ the spin index. The ground-state is perturbed by a constant electric-field \mathbf{E} which couples with the carriers' charge e and by the electron-phonon interaction H_{el-ph} , so that the total Hamiltonian is $H = H_0 + e\mathbf{x} \cdot \mathbf{E} + H_{el-ph}$. To derive an equation of motion for such a system, we use the single-particle Wigner function f of the system [16], defined as the Wigner transform of the density matrix ρ as

$$f_{bb'ss'}(\mathbf{x}, \mathbf{k}, t) = \sum_{\Delta \mathbf{k}} e^{i\Delta \mathbf{k} \cdot \mathbf{x}} \rho_{bb'ss'} \left(\mathbf{k} + \frac{\Delta \mathbf{k}}{2}, \mathbf{k} - \frac{\Delta \mathbf{k}}{2}; t \right), \quad (1)$$

where t is the time. We build the Wigner function through a transformation of the density matrix $\rho_{bb'ss'}(\mathbf{k}, \mathbf{k}')$ in the reciprocal space representation. Such Wigner transform consists in a rotation of variables $\mathbf{k}, \mathbf{k}' \rightarrow \frac{\mathbf{k} + \mathbf{k}'}{2}, \mathbf{k}' - \mathbf{k}$ combined with a Fourier transform on one variable. The Wigner function operates in a phase-space representation, which is especially useful to draw connections between quantum and classical mechanics.

The evolution of the Wigner function [17,18] is found through a Wigner transform of the equation of motion of the density matrix, and has been shown to be

$$\frac{\partial f_{bb'ss'}(\mathbf{x}, \mathbf{k}, t)}{\partial t} = - \{ \{ f(\mathbf{x}, \mathbf{k}, t), H(\mathbf{x}, \mathbf{k}) \} \}_{bb'ss'} \quad (2)$$

$$\begin{aligned} &:= \frac{i}{\hbar} \left(f(\mathbf{x}, \mathbf{k}, t) \star H(\mathbf{x}, \mathbf{k}) \right. \\ &\quad \left. - H(\mathbf{x}, \mathbf{k}) \star f(\mathbf{x}, \mathbf{k}, t) \right) \end{aligned} \quad (3)$$

where $\{ \{ f, H \} \}$ is the Moyal bracket (the quantum mechanical

extension of the Poisson bracket) and the Moyal product \star is defined as

$$f \star H = f(\mathbf{x}, \mathbf{k}) \exp \left(\frac{i}{2} \left(\overleftarrow{\frac{\partial}{\partial \mathbf{x}}} \cdot \overrightarrow{\frac{\partial}{\partial \mathbf{k}}} - \overleftarrow{\frac{\partial}{\partial \mathbf{k}}} \cdot \overrightarrow{\frac{\partial}{\partial \mathbf{x}}} \right) \right) H(\mathbf{x}, \mathbf{k}), \quad (4)$$

where $H(\mathbf{x}, \mathbf{k})$ is the Wigner transform of the Hamiltonian, and the arrows indicate that the derivative operator is acting to the left/right operators.

We now simplify the Hamiltonian supposing that the electron-phonon interaction is weak and evaluate the Moyal bracket for the single-particle part of the Hamiltonian; the electron-phonon interaction is added later as a perturbation. Since we are interested in macroscopic properties, we can make the assumption that only slow spatial variations of the Wigner function are relevant. Therefore, we expand the Moyal product in Taylor series to the lowest orders of \hbar and find an equation of motion which we term the ab initio Wigner Transport Equation (aiWTE), that is

where $\{, \}$ is an anticommutator, $\mathcal{E}_{bb'ss'}(\mathbf{k}) = \delta_{bb'} \delta_{ss'} \epsilon_{bs}(\mathbf{k})$ is a tensor of electronic energies, $\mathbf{d}_{bb'ss'}(\mathbf{k}) = (1 - \delta_{bb'}) \langle \mathbf{k}bs | e\mathbf{x} | \mathbf{k}b's' \rangle$ is a tensor of electric dipoles between two Bloch states (typically used to describe optical excitations), and $\mathbf{v}_{bb'ss'}(\mathbf{k})$ is the velocity operator, defined from the commutator of the Hamiltonian and the position operator as $\mathbf{v} = \frac{i}{\hbar} [H, \mathbf{r}]$. The electron-phonon scattering operator $\left. \frac{\partial f_{bb'ss'}(\mathbf{x}, \mathbf{k}, t)}{\partial t} \right|_{coll}$ is added as a perturbation to the aiWTE and is built using scattering rates from the Fermi Golden rule [19–23]. We refer to the Supplementary Information for a more detailed derivation of the aiWTE.

The aiWTE needs to be solved to obtain an estimate of the single-particle Wigner distribution function. As a first comment, we note that the aiBTE is recovered as a limiting case of the aiWTE, when the off-diagonal terms $b \neq b'$ and $s \neq s'$ are set to zero. This corresponds physically to a situation when different bands do not couple via Zener tunneling. This can happen, for example, when neither thermal excitation nor dipole interaction provide sufficient energy to allow for the vertical transition of one particle from one band to another. Therefore, the most interesting terms to discuss in the aiWTE are the off-diagonal terms, which introduce the possibility of additional electronic transitions, or couplings, between different electronic states at a given wavevector \mathbf{k} . We further note that some of the off-diagonal terms of the aiWTE shown here are absent in other works based on either the density matrix or the Wigner function [24–29]. Additionally, we note that the electronic aiWTE is conceptually similar to a formalism developed for phonon transport in Ref. [19], although here we use a simplified derivation and include the effect of external forces (the electric field).

The aiWTE can be solved to estimate transport coefficients with a technique similar to the one used for the aiBTE. For steady-state transport in a bulk system, the Wigner distribution is stationary in time and independent from the particular position inside the bulk; therefore, we look for a solution in the form $f_{bb'ss'}(\mathbf{k})$. Next, we look for the linear response to an applied external electric field \mathbf{E} or

a temperature gradient ∇T and write $f_{bb's's'}(\mathbf{k}) = \bar{f}_{bs}(\mathbf{k}) + f_{bb's's'}^{E_i}(\mathbf{k})E_i$ or $f_{bb's's'}(\mathbf{k}) = \bar{f}_{bs}(\mathbf{k}) + f_{bb's's'}^{T_i}(\mathbf{k})\nabla_i T$, where i is the direction of the applied perturbations and $\bar{f}_{bs}(\mathbf{k})$ is the Fermi–Dirac distribution function. We further adopt the relaxation time approximation, which simplifies the treatment of the scattering operator. After inserting these two proposed solutions in the aiWTE, and retaining only terms up to linear order in \mathbf{E} or ∇T , we find two equations for the diagonal part of the aiWTE:

$$ev_{bb's's'}^i(\mathbf{k}) \frac{\partial f_{bb's's'}(\mathbf{k})}{\partial \epsilon} = \Gamma_{bs}(\mathbf{k}) f_{bb's's'}^{E_i}(\mathbf{k}). \quad (6)$$

and

$$v_{bb's's'}^i(\mathbf{k}) \cdot \frac{\partial f_{bb's's'}(\mathbf{k})}{\partial T} = \Gamma_{bs}(\mathbf{k}) f_{bb's's'}^{T_i}(\mathbf{k}), \quad (7)$$

where $\Gamma_{bs}(\mathbf{k})$ is the carrier's linewidth. These equations are equivalent to the standard aiBTE for describing electronic transport within the relaxation time approximation. However, the aiWTE also gives rise to two equations for the off-diagonal terms of the Wigner distribution, which are

$$i(\epsilon_{bs}(\mathbf{k}) - \epsilon_{b's'}(\mathbf{k})) f_{bb's's'}^{E_i}(\mathbf{k}) + i(\bar{f}_{b's'}(\mathbf{k}) - \bar{f}_{bs}(\mathbf{k})) d_{bb's's'}^i(\mathbf{k}) \\ = -\frac{\Gamma_{bs}(\mathbf{k}) + \Gamma_{b's'}(\mathbf{k})}{2} f_{bb's's'}^{E_i}(\mathbf{k}), \quad b \neq b', s \neq s', \quad (8)$$

and

$$i(\epsilon_{bs}(\mathbf{k}) - \epsilon_{b's'}(\mathbf{k})) f_{bb's's'}^{T_i}(\mathbf{k}) + \frac{1}{2} \left(\frac{\partial \bar{f}_{b's'}(\mathbf{k})}{\partial T} + \frac{\partial \bar{f}_{bs}(\mathbf{k})}{\partial T} \right) v_{bb's's'}^i(\mathbf{k}) \\ = -\frac{\Gamma_{bs}(\mathbf{k}) + \Gamma_{b's'}(\mathbf{k})}{2} f_{bb's's'}^{T_i}(\mathbf{k}), \quad b \neq b', s \neq s', \quad (9)$$

describing the response of the system to electrical and thermal perturbations, respectively. Now the aiWTE is in a form that can be readily solved for f^{E_i} and f^{T_i} with some algebra, allowing us to reconstruct the Wigner distribution of a crystal in presence of an electric field or a thermal gradient.

Having found the Wigner distribution, transport properties are readily obtained. For example, the charge current density is

$$\mathbf{j} = \frac{e}{2V N_k} \sum_{\mathbf{k}bs} \{ \mathbf{v}(\mathbf{k}), f(\mathbf{k}) \}_{bb's's'}, \quad (10)$$

where V is the crystal unit cell volume and N_k the number of wavevectors used to integrate the Brillouin zone. Additionally, the heat flux is

$$\mathbf{q} = \frac{1}{2V N_k} \sum_{\mathbf{k}bs} (\epsilon_{bs}(\mathbf{k}) - \mu) \{ \mathbf{v}(\mathbf{k}), f(\mathbf{k}) \}_{bb's's'}. \quad (11)$$

These definitions readily allow us to compute transport coefficients using an approach typical of transport theory. After inserting the solution to the aiWTE in the definition of \mathbf{j} and \mathbf{q} , it is readily seen that charge current and heat flux can be written in the form $\mathbf{j} = L_{EE} \mathbf{E} + L_{ET} \nabla T$ and $\mathbf{q} = L_{TE} \mathbf{E} + L_{TT} \nabla T$, where L denotes the Onsager coefficients. Transport coefficients can be expressed in terms of these Onsager coefficients: the electrical conductivity $\sigma = L_{EE}$, the Seebeck coefficient $S = -L_{ET}^{-1} L_{TE}$, and the thermal

conductivity $k = -(L_{TT} - L_{TE} L_{EE}^{-1} L_{ET})$.

We now inspect the electrical conductivity in more details. Since \mathbf{j} is linear in $f_{bb's's'}(\mathbf{k})$, the total electrical conductivity σ is a sum of diagonal and off-diagonal contributions $\sigma_{ij} = \sigma_{ij}^{aiBTE} + \Delta\sigma_{ij}$, where i, j are cartesian labels. Here, σ_{ij}^{aiBTE} comes from the diagonal terms of the Wigner distribution ($b = b'$ and $s = s'$) and is equal to the estimate from the aiBTE:

$$\sigma_{ij}^{aiBTE} = \frac{e^2}{V N_k} \sum_{\mathbf{k}bs} \frac{\partial \bar{f}_{bs}(\mathbf{k})}{\partial \epsilon} v_{bb's's'}^i(\mathbf{k}) v_{bb's's'}^j(\mathbf{k}) \tau_{bs}(\mathbf{k}), \quad (12)$$

where $\tau_{bs}(\mathbf{k}) = \hbar/\Gamma_{bs}(\mathbf{k})$ is the carriers' lifetime. This semiclassical conductivity is corrected by an additional term $\Delta\sigma_{ij}$ that can be found inserting $f_{bb's's'}^{E_i}(\mathbf{k})$ in the definition of \mathbf{j} and is

$$\Delta\sigma_{ij} = \frac{2e^2}{V N_k} \sum_{\substack{\mathbf{k}bsb's's' \\ bs \neq b's'}} \frac{\bar{f}_{b's'}(\mathbf{k}) - \bar{f}_{bs}(\mathbf{k})}{\epsilon_{b's'}(\mathbf{k}) - \epsilon_{bs}(\mathbf{k})} \times \\ \times \frac{v_{bb's's'}^i(\mathbf{k}) v_{b'b's's'}^j(\mathbf{k}) (\Gamma_{bs}(\mathbf{k}) + \Gamma_{b's'}(\mathbf{k}))}{4(\epsilon_{b's'}(\mathbf{k}) - \epsilon_{bs}(\mathbf{k}))^2 + (\Gamma_{bs}(\mathbf{k}) + \Gamma_{b's'}(\mathbf{k}))^2}. \quad (13)$$

We can now understand the effects of the off-diagonal corrections. First, the correction $\Delta\sigma$ is positive (note that $\bar{f}_{bs}(\mathbf{k})$ is a monotonic function of $\epsilon_{bs}(\mathbf{k})$), and therefore the aiWTE will always adjust the semiclassical prediction of conductivity to higher values. Second, the correction depends on a few quantities: the energy difference between electrons and holes, their linewidths, and velocity. One scenario where this correction is relevant, for example, is whenever the energy difference between an electron and a hole is comparable to their linewidth, so that the two carriers interact. The strength of such interaction is determined by the velocity matrix element $v_{bb's's'}(\mathbf{k})$, i.e. the matrix element for the optical transition. In addition, a strong dipole coupling can also occur between states that are far from the conduction or valence band edge, as discussed below. Whenever this off-diagonal correction is large, the system allows for an additional transport mechanism, known as Zener tunneling, in which electrons propagate by tunneling across the band gap. In contrast, we stress that Zener tunneling, or similar spontaneous vertical electronic transitions, are entirely omitted in the aiBTE formalism, since the aiBTE doesn't include off-diagonal contributions from the velocity operator.

3. Computational methods

All quantities appearing in the aiWTE are available from first-principles codes and we can therefore apply this formalism using fully ab-initio parameters.

We use density functional theory as implemented in the plane-wave software suite Quantum-ESPRESSO [30,31]. To compute the ground state, we use ultrasoft pseudopotentials from the GBRV library [32], with the PBEsol functional. We use an energy cutoff of 80 Ry, and integrate the Brillouin zone with a $8 \times 8 \times 8$ mesh of \mathbf{k} -points. We build the trigonal unit cell using experimental estimates of the crystal structure [33], i.e. with a lattice parameter of 9.839 Å, and an angle α such that $\cos\alpha = 0.91068$. The Wannier functions are computed using p orbitals on both Bi and Se atoms as initial guesses.

Phonon properties, and the electron-phonon matrix elements are computed with density-functional perturbation theory [34] on a coarse grid of $4 \times 4 \times 4$ \mathbf{q} -points. Electron-phonon matrix elements are subsequently interpolated on a finer grid of $39 \times 39 \times 39$ Brillouin zone wavevectors using a mixed Wannier and linear

interpolation [12], while electronic energies and velocities are interpolated using Wannier90 [35]. Transport properties have been implemented in a custom-made software. The Dirac-delta ensuring energy conservation during an electron-phonon scattering event has been approximated using an adaptive-smearing scheme [36]. Transport properties have been converged with respect to the k -points mesh used to integrate the Brillouin zone. The scattering operator, detailed in the Supplementary Information, is built considering only intrinsic electron-phonon scattering. The electron-phonon coupling includes the effect of long-range polar interaction, which contributes significantly to transport coefficients, and is evaluated according to the methodology of Ref. [37]. While the scattering operator may be treated beyond the relaxation time approximation with techniques similar to those of Refs. [8,9,12], all results of this study are obtained within the relaxation time approximation, as it allows for significantly faster simulations. We note that results beyond the relaxation time approximation would only affect the diagonal components of the Wigner distribution, but not the off-diagonal components since the action of the scattering operator on the off-diagonal components of the Wigner distribution involve only the carriers' lifetimes/linewidths (see Eq. 18 of Supplementary Information).

4. Results

We now apply the formalism to study the intrinsic phonon-limited electronic transport of bulk Bi_2Se_3 . In Fig. 1 we report the band structure [38,39] and the density of states (DOS) for this narrow-gap semiconductor. We estimate a quasiparticle gap of 0.2 eV, in agreement with experimental estimates [40]. We also mention that, while DFT may not always be accurate in estimating the band gap, the qualitative aspects of the results discussed below do not strongly depend on the precise band gap value. It is worth noting that the DOS increases away from the Fermi level (set at 0 eV at the middle of the band gap) and flattens at energies of approximately -0.8 eV and 1.0 eV for the valence and conduction bands, respectively, indicating that the subvalleys are separated by an energy of approximately 1.8 eV.

In Fig. 2a (b), we estimate the electron-phonon limited electrical conductivity σ of Bi_2Se_3 in the in-plane direction as a function of temperature for different values of n-type (p-type) doping concentrations. Dashed lines represent the semiclassical estimates σ^{aiBTE} , while aiWTE estimates are shown in solid lines. For the

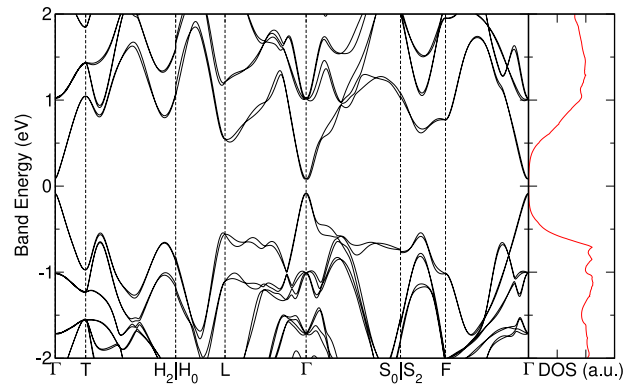


Figure 1. Electronic band structure and density of states of Bi_2Se_3 . Electronic band structure and density of states (DOS) of Bi_2Se_3 near the Fermi energy, set at the center of the band gap. The bulk crystal is characterized by a small gap, opened by the spin-orbit coupling. We also note from the DOS that subvalleys of valence and conduction bands are approximately 1.8 eV apart in energy.

highest doping values, when the chemical potential shifts into the conduction (valence) band, the conductivity has the typical metallic-like behavior of decreasing with temperature. Under these conditions, aiBTE and aiWTE do not differ significantly, except at higher temperatures. In the Supplementary Information, we briefly discuss a comparison with available experimental results which, for the purpose of the present study, shows good qualitative agreement especially at large doping concentrations.

For lower doping concentrations the chemical potential lies in the band gap and we thus observe a semiconducting behavior of σ increasing with temperature. The semiclassical model predicts a smaller conductivity than the aiWTE estimate. In fact, when only a few carriers from the bottom of the conduction (top of the valence) band are excited, the average carriers' group velocity is small, due to the quadratic nature of the band minimum. Therefore, the semiclassical contribution to electrical conductivity tends to be rather small. The aiWTE corrects this picture, including the Zener tunneling effect [2]. As carriers from valence and conduction band are close in energy, they can interact and contribute to the electrical transport through the tunneling effect, as discussed above. For small dopings, the aiWTE correction is significant, and can be much larger than the aiBTE conductivity value. For the smallest value of doping reported (10^{16} cm^{-3}), this correction is largest at lower temperatures. The doping of 10^{18} cm^{-3} is an intermediate case, with metallic behavior at low temperatures (and thus smaller aiWTE correction) and semiconducting (with larger aiWTE correction) at higher temperatures as the chemical potential moves from the conduction band into the band gap. We can thus conclude that a substantial portion of electrical current at low doping is carried through the Zener tunneling included in the aiWTE formalism: the current is not only caused by the carriers traveling at a finite group velocity, but also by carriers' tunneling between single-particle Bloch states.

In Fig. 2, panels c and d, we report the Seebeck coefficient S for n-type (p-type) doping concentrations, with aiBTE results in dashed lines and aiWTE in solid lines. The Seebeck coefficient is usually expected to have negative values for n-type doping and positive for p-type. However, there are deviations from this behavior, and the Seebeck coefficient can even change sign when temperature is varied at fixed doping concentrations. This complex behavior is not caused by the off-diagonal terms of the Wigner distributions, and the Seebeck coefficient sign changes are observed when using the aiBTE as well, e.g. due to the bipolar effect where minority carriers are thermally excited across the band gap. Such behavior has been found for Bi_2Te_3 [41] and CoSb_3 [42], where the Seebeck coefficient exhibits marked temperature dependence.

The Wigner correction tends to make the Seebeck coefficient smaller in absolute value, which may be attributed to the electrical conductivity appearing at the denominator of the definition of the Seebeck coefficient. As for the case of electrical conductivity, the off-diagonal correction is more prominent at small values of doping concentrations and low temperatures. This may have a significant consequence when estimating thermoelectric properties, such as the power factor that depends on the square of the Seebeck coefficient. In this case, the inclusion of off-diagonal components of the Wigner distribution lead to a revision of power factor estimates to smaller values.

We now examine the relationship of electrical conductivity σ and open-circuit electronic thermal conductivity κ_{el} . The Wiedemann-Franz (WF) law defines the Lorenz number $L = \frac{\kappa_{\text{el}}}{\sigma T}$, which in the ideal metallic limit is a constant $L_0 = 2.44 \cdot 10^{-8} \text{ W}\Omega\text{K}^{-2}$. Knowledge of L is necessary to decouple the electronic contribution κ_{el} and the lattice contribution κ_{ph} from measurements of the total κ . In Fig. 2 e and f we plot the computed ratio $\frac{L}{L_0}$

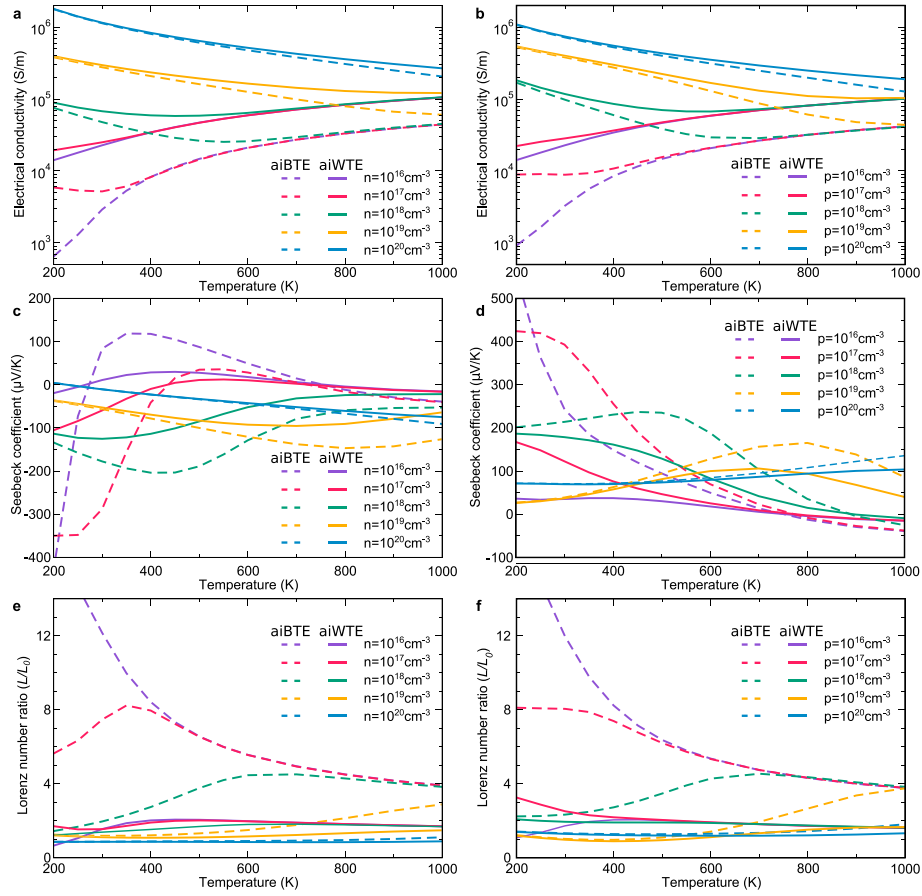


Figure 2. Transport properties of Bi₂Se₃ from semiclassical and Wigner formalism. Computational estimates of the in-plane electrical conductivity (panels a and b), Seebeck coefficient (panels c and d), and Lorenz number ratio (panels e and f, see text for description) of Bi₂Se₃ as a function of temperature, for different values of n-type or p-type doping concentration. Solid lines are estimated using the Wigner transport equation, while dotted lines are semiclassical estimates obtained solving the Boltzmann transport equation. For small doping concentrations, the interaction between electrons and holes significantly affect the estimates of transport coefficients.

for several temperatures and n-type or p-type kind of doping concentrations, using both aiBTE and aiWTE in dashed and solid lines respectively. At high doping the system has metallic character, and both predictions closely follow the WF law. In the case of small doping, in the bipolar regime, the semiclassical aiBTE predicts large deviations from the WF law, as has been discussed previously [13, 43–45]. Remarkably, in the aiWTE solution the Lorenz numbers are much closer to the expected range for semiconductors, indicating that quantum transport effects included in the aiWTE strongly suppress deviations and work towards restoring the validity of the WF law.

In Fig. 3a, we analyze the contributions to the aiBTE electrical conductivity as a function of the carriers' energy at doping concentration of 10¹⁸ cm⁻³, and temperature of 700 K. This histogram is built such that the area under the curve integrates to the total electrical conductivity. Within the semiclassical relaxation time approximation, the quantity plotted is an energy-resolved histogram of $\frac{e^2}{V N_k} \frac{\partial \bar{f}_{bs}(\mathbf{k})}{\partial \epsilon} t_{bss}^2(\mathbf{k}) \tau_{bs}(\mathbf{k})$, i.e. the contribution of a single mode to the aiBTE electrical conductivity. As expected, the dominant contributions to electrical conductivity come from carriers whose energy is close to the chemical potential (set at 0 eV). The contributions of other carriers decay exponentially as their energy gets further from the chemical potential.

The aiWTE correction $\Delta\sigma$ cannot be resolved in terms of a single

carrier's energy, since it involves the tunneling between two states at different energies. Therefore, in Fig. 3b, we plot the contributions to the electrical conductivity as a function of two interacting carriers energies. On the diagonal, we find again the aiBTE-like terms shown in Fig. 3a. In addition, we can see important off-diagonal contributions to the electrical conductivity that are not present in the aiBTE and are introduced with the aiWTE. In particular, there are two peaks of contributions to electrical conductivity, that couple electrons of energy 1.0 eV with holes at -0.8 eV. These two values correspond to the average energies of the valence and conduction bands, when the DOS reaches the corresponding maximum values. Therefore, in contrast to the typical intuition of the Zener tunneling, we find that the most significant coupling between carriers takes place far from the chemical potential, with carriers of energy much larger than the thermal energy. For this material, the dipole interaction between carriers in subvalleys of the valence and conduction is thus particularly strong, allowing for high-energy carriers to contribute to transport. As a result, we speculate that Zener tunneling may take place also in semiconductors with a wide gap and contribute significantly to electrical conductivity, provided that the inter-band dipole interaction is sufficiently strong, for instance in materials with high optical absorption character.

We can gain further insight by looking at how different points in

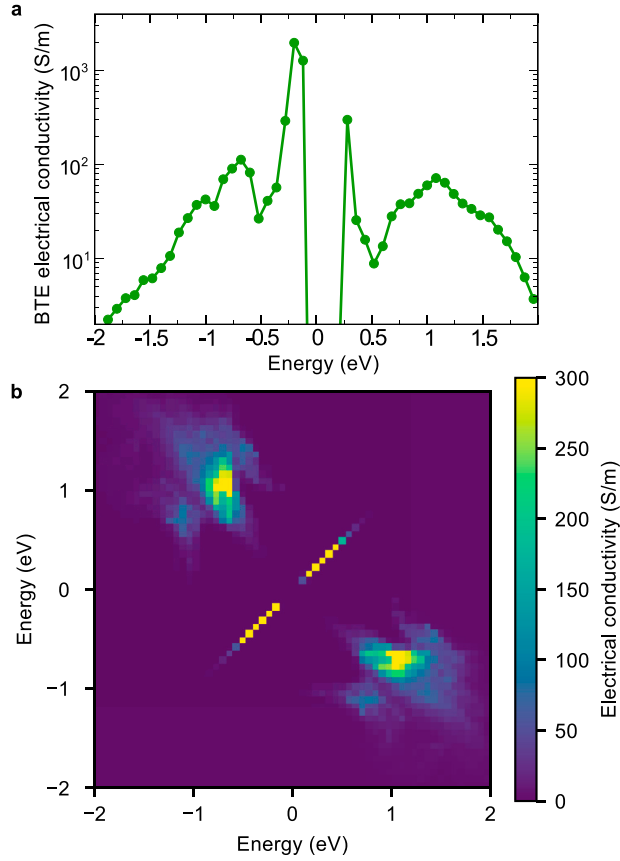


Figure 3. Microscopic contributions to electrical conductivity from Boltzmann and Wigner formalisms. Panel a: histogram of contributions to semiclassical electrical conductivity as a function of the carrier energy, for Bi_2Se_3 at 700K and n-doping at 10^{18} cm^{-3} . Panel b, 2D histogram of contributions to the electrical conductivity for the same system as estimates with the Wigner transport equation against the energy of two coupled carriers. Off-diagonal contributions represent electrical conductivity arising from the coupling between electrons and holes.

the Brillouin zone contribute to the conductivity. In Fig. 4, we plot the contributions to the in-plane electrical conductivity from carriers with wavevector lying on high-symmetry lines of the Brillouin zone, the y-coordinate of the dot in figure is positioned according to the carrier's energy. The green dots' radius is proportional to the contribution of such carrier to the aiBTE electrical conductivity, i.e. the contributions from the diagonal part of the Wigner distribution, whereas the red dots radius is proportional to the electrical conductivity contribution from the off-diagonal components of the Wigner distribution. Quantities are computed for the electrical conductivity at temperature of 700K and doping concentration $n = 10^{18} \text{ cm}^{-3}$. As one would expect from semiclassical arguments, the diagonal contributions are mostly originating from carriers of energy close to the band gap. The contributions from the off-diagonal components instead extend to carriers that are further from the band gap edge. States close to the band gap at the Γ point have a sizeable contribution to the off-diagonal conductivity. However, subvalleys overall contribute more significantly to the conductivity, especially due to the larger availability of states for the vertical transitions to take place. We further note that there is not a single

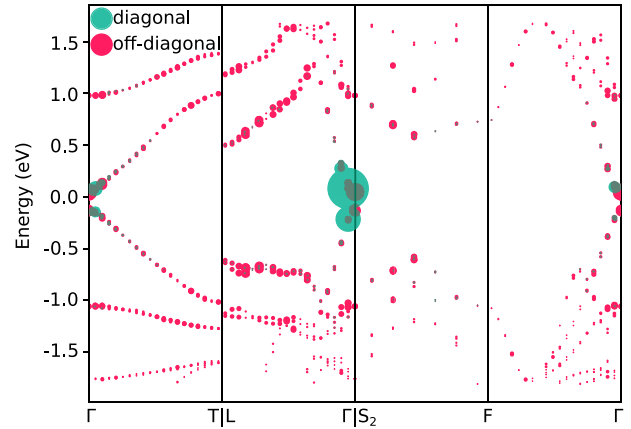


Figure 4. Microscopic contributions to electrical conductivity from Boltzmann and Wigner formalisms. The dispersion relation of carriers on some high-symmetry lines are plotted together with the contributions to the in-plane electrical conductivity of Bi_2Se_3 at 700K and doping concentration of $n = 10^{18} \text{ cm}^{-3}$. In particular, the radius of green dots (mostly visible only near the band gap at the Γ point) is proportional to the contribution to electrical conductivity from the diagonal components of the Wigner distribution, while red dots radius is proportional to the contributions of the off-diagonal Wigner distribution. While semiclassical transport mostly originates close to the band gap at Γ , the corrections from the aiWTE are coming from several states especially in the conduction and valence subvalleys.

transition that dominates the off-diagonal effects. Lastly, we stress that diagonal and off-diagonal contributions have not been drawn to scale with respect to one another. In fact, diagonal contributions are mostly determined by carriers close to the chemical potential and sum up to a conductivity of $\sigma^{\text{aiBTE}} = 2.9 \cdot 10^4 \text{ S/m}$. The off-diagonal conductivity $\Delta\sigma = 4.4 \cdot 10^4 \text{ S/m}$ receives contributions from carriers over a much larger energy range. As a result, if the diagonal contributions were drawn to the same scale of the off-diagonal ones, green circles would appear much larger than what is shown; therefore, green radii have been shrunk by a factor about 40 in order to fit in the figure.

So far, we have seen that the off-diagonal contributions to electrical conductivity at energies far from the chemical potential arise from a large presence of available states. However, Eq. (13) depends on several other factors. In Fig. 5a, we plot the average lifetime as a function of carrier's energy, at the same doping density and temperature of Fig. 3. Here we can see that lifetimes tend to be large for states close to the band edges and decrease away from them: to a first approximation, the average lifetime as a function of energy scales inversely proportional to the density of states. Therefore, the decrease of lifetimes is offset by the larger availability of states, which still allows for a sizeable off-diagonal contributions from states away of the band gap. In Fig. 5b we also plot the energy-resolved histogram of the velocity operator $|v_{bb's'}(\mathbf{k})|$ in the in-plane direction, where the two axis represent the energies of the carriers $\varepsilon_{bs}(\mathbf{k})$ and $\varepsilon_{b's'}(\mathbf{k})$ respectively, and the color intensity represents the absolute value of the velocity operator matrix element. Clearly, in order to have a sizeable Wigner correction, it is essential for the velocity operator elements to be sufficiently large. However, one can see that the energy values where the velocity elements are largest do not always coincide with the energies of states that contribute most to the conductivity. Therefore, the off-diagonal effects of the Wigner distribution are not easily interpreted in terms of a single factor appearing in Eq. (13): the overall

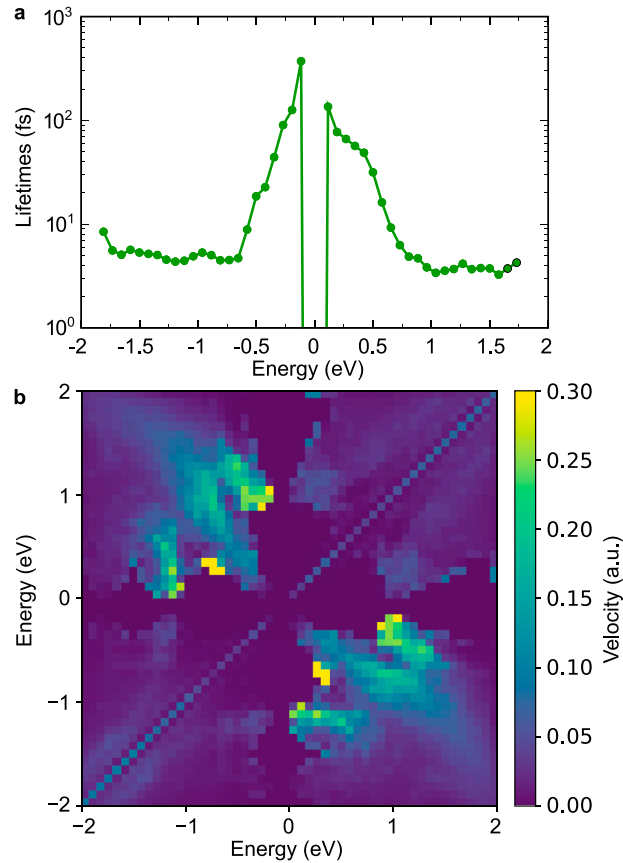


Figure 5. Analysis of electronic lifetime and velocity. Panel a: average of electronic lifetimes $\tau_{bs}(\mathbf{k})$ as a function of energy, for Bi_2Se_3 at 700K and n-type doping concentration of 10^{18} cm^{-3} . The energy dependence of lifetimes is approximately proportional to the inverse of the density of states. Panel b: 2D histogram of the velocity operator $v_{bb/ss}(\mathbf{k})$ as a function of the energies of the two states $\epsilon_{bs}(\mathbf{k})$ and $\epsilon_{b's'}(\mathbf{k})$. Diagonal matrix elements are the electronic group velocity, while off-diagonal elements are related to the dipole operator. The presence of states with large off-diagonal velocity allows the coupling needed for vertical transitions.

correction arises due to a combination of large linewidth and velocity (dipoles), as well as high availability of states for the transitions.

5. Conclusions

First-principles simulations of electronic transport are often limited to the semiclassical approximation, that excludes contributions to transport coming from non-semiclassical particle dynamics. In this study, we have shown that the Moyal equation of motion leads to the Wigner transport equation, an equation of motion for the Wigner function that can be solved using first-principles techniques. The formalism has a broader range of applicability than the semiclassical Boltzmann transport equation, and includes off-diagonal contributions to charge and heat currents. In particular, the formalism captures contribution to transport coming from spontaneous vertical transitions between carriers' states, that are especially relevant for narrow-gap semiconductors. Moreover, the semiclassical Boltzmann transport equation is found as a simplified limit to the Wigner transport equation, provided that off-diagonal components are neglected. The formalism is well-suited to first-principles simulations and has

been tested with a study of the transport properties of Bi_2Se_3 . We have shown that, while at high doping concentrations the Boltzmann equation provides a fairly accurate description of transport, it fails at low doping concentrations. At low dopings, spontaneous vertical transitions across the band gap, also known as the Zener tunneling effect, contribute significantly to electronic transport, modifying both electrical conductivity and Seebeck coefficient. Surprisingly, Zener tunneling does not just take place across the states closest to the band gap, but involves states that are significantly further apart in energy, provided that the dipole interaction is sufficiently strong. As a result, we have extended the range of applicability of ab-initio transport simulations to materials where band tunneling couples carriers and a semiclassical description is no longer adequate.

Author contributions

A. C. derived results and performed numerical simulations. A.C. and B. K. conceived the project and wrote the paper.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We acknowledge funding support from the Star-Friedman Fund for Promising Scientific Research, the Harvard Quantum Initiative, the STC Center for Integrated Quantum Materials, NSF Grant No. DMR-1231319, and computational resources at Harvard FAS Research Computing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.mtphys.2021.100412>.

References

- [1] Niels Vandecasteele, Amelia Barreiro, Michele Lazzeri, Adrian Bachtold, Francesco Mauri, Current-voltage characteristics of graphene devices: interplay between zener-klein tunneling and defects, *Phys. Rev. B* 82 (2010): 045416.
- [2] Clarence Zener, A theory of the electrical breakdown of solid dielectrics, *Proc. Roy. Soc. Lond. A* 145 (1934) 523.
- [3] Binghai Yan, Shou-Cheng Zhang, Topological materials, *Rep. Prog. Phys.* 75 (2012): 096501.
- [4] N.P. Armitage, E.J. Mele, Ashvin Vishwanath, Weyl and Dirac semimetals in three-dimensional solids, *Rev. Mod. Phys.* 90 (2018): 015001.
- [5] Lukas Muechler, Frederick Casper, Binghai Yan, Stanislav Chadov, Claudia Felser, Topological insulators and thermoelectric materials, *Phys. Status Solidi RRL* 7 (2013) 91–100.
- [6] Joseph P. Heremans, Robert J. Cava, Nitin Samarth, Tetradymites as thermoelectrics and topological insulators, *Nat. Rev. Mater.* 2 (2017) 17049.
- [7] Pengzi Liu, James R. Williams, Judy J. Cha, Topological nanomaterials, *Nat. Rev. Mater.* 4 (2019) 2058–8437.
- [8] Zhao Wang, Shidong Wang, Sergey Obukhov, Nathalie Vast, Jelena Sjakste, Valery Tyuterev, Natalio Mingo, "Thermoelectric transport properties of silicon: toward an ab initio approach, *Phys. Rev. B* 83 (2011) 205208.
- [9] Li Wu, "Electrical transport limited by electron-phonon coupling from Boltzmann transport equation: an ab initio study of si, al, and mos_2 , *Phys. Rev. B* 92 (2015): 075405.

- [10] S. Ponc , E.R. Margine, C. Verdi, F. Giustino, Epw: electron-phonon coupling, transport and superconducting properties using maximally localized wannier functions, *Comput. Phys. Commun.* 209 (2016) 116–133.
- [11] Samuel Ponc , R. Elena, Margine, and Feliciano Giustino, "Towards predictive many-body calculations of phonon-limited carrier mobilities in semiconductors, *Phys. Rev. B* 97 (2018) 121201.
- [12] Mattia Fiorentini, Nicola Bonini, "Thermoelectric coefficients of *n*-doped silicon from first principles via the solution of the Boltzmann transport equation, *Phys. Rev. B* 94 (2016): 085204.
- [13] Georgy Samsonidze, Boris Kozinsky, Thermoelectric materials: accelerated screening of thermoelectric materials by first-principles computations of electron-phonon scattering, *Advanced Energy Materials* 8 (2018) 1870095.
- [14] Branislav K. Nikoli , Kamal K. Saha, Troels Markussen, Kristian S. Thygesen, First-principles quantum transport modeling of thermoelectricity in single-molecule nanojunctions with graphene nanoribbon electrodes, *J. Comput. Electron.* 11 (78) (2012) 92.
- [15] J. Weinbub, D.K. Ferry, Recent advances in wigner function approaches, *Appl. Phys. Rev.* 5 (2018): 041104, <https://doi.org/10.1063/1.5046663>.
- [16] E. Wigner, On the quantum correction for thermodynamic equilibrium, *Phys. Rev.* 40 (1932) 749–759.
- [17] J.E. Moyal, Quantum mechanics as a statistical theory, *Math. Proc. Camb. Phil. Soc.* 45 (1949) 99–124.
- [18] H.J. Groenewold, On the principles of elementary quantum mechanics, *Physica* 12 (1946) 405–460.
- [19] Michele Simoncelli, Nicola Marzari, Francesco Mauri, Unified theory of thermal transport in crystals and glasses, *Nat. Phys.* 15 (2019) 809–813.
- [20] Rita Claudia Iotti, Fabrizio Dolcini, Fausto Rossi, Wigner-function formalism applied to semiconductor quantum devices: need for nonlocal scattering models, *Phys. Rev. B* 96 (2017) 115420.
- [21] Z. Zhan, E. Colom s, X. Oriols, Unphysical features in the application of the Boltzmann collision operator in the time-dependent modeling of quantum transport, *J. Comput. Electron.* 15 (2016) 1206–1218.
- [22] M. Nedjalkov, S. Selberherr, D.K. Ferry, D. Vasileska, P. Dollfus, D. Querlioz, I. Dimov, P. Schwaha, Physical scales in the wigner-Boltzmann equation, *Ann. Phys.* 328 (2013) 220–237.
- [23] M. Nedjalkov, D. Querlioz, P. Dollfus, H. Kosina, Wigner function approach, in: *Nano-Electronic Devices: Semiclassical and Quantum Transport Modeling*, Springer New York, New York, NY, 2011, pp. 289–358. Dragica Vasileska and Stephen M. Goodnick.
- [24] Fausto Rossi, Tilmann Kuhn, Theory of ultrafast phenomena in photoexcited semiconductors, *Rev. Mod. Phys.* 74 (2002) 895–950.
- [25] R. H bner, R. Graham, Landau-zener transitions and dissipation in a mesoscopic ring, *Phys. Rev. B* 53 (1996) 4870–4885.
- [26] J.B. Krieger, G.J. Iafrate, Quantum transport for bloch electrons in a spatially homogeneous electric field, *Phys. Rev. B* 35 (1987) 9644–9658.
- [27] Gaston Kan , Michele Lazzeri, Francesco Mauri, Zener tunneling in the electrical transport of quasimetallic carbon nanotubes, *Phys. Rev. B* 86 (2012) 155433.
- [28] Gaston Kan , Michele Lazzeri, Francesco Mauri, High-field transport in graphene: the impact of zener tunneling, *J. Phys.: Condens. Matter* 27 (2015) 164205.
- [29] G.J. Iafrate, V.N. Sokolov, J.B. Krieger, Quantum transport and the wigner distribution function for bloch electrons in spatially homogeneous electric and magnetic fields, *Phys. Rev. B* 96 (2017) 144303.
- [30] Paolo Giannozzi, Stefano Baroni, Nicola Bonini, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, L. Guido, Chiarotti, Matteo Cococcioni, Ismaila Dabo, Andrea Dal Corso, Stefano de Gironcoli, Stefano Fabris, Fratesi Guido, Ralph Gebauer, Uwe Gerstmann, Christos Gougousis, Kokalj Anton, Michele Lazzeri, Layla Martin-Samos, Nicola Marzari, Francesco Mauri, Riccardo Mazzarello, Stefano Paolini, Alfredo Pasquarello, Lorenzo Paulatto, Carlo Sbraccia, Sandro Scandolo, Gabriele Sclauzero, Ari P. Seitsonen, Smogunov Alexander, Paolo Umari, Renata M. Wentzcovitch, QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials, *J. Phys. Condens. Matter* 21 (2009) 395502.
- [31] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. Buongiorno Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. Dal Corso, S. de Gironcoli, P. Delugas, R.A. DiStasio Jr., A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. K    benli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N.L. Nguyen, H.V. Nguyen, A. Otero de-la Roza, L. Paulatto, S. Ponc , D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A.P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, S. Baroni, Advanced capabilities for materials modelling with Quantum ESPRESSO, *J. Phys. Condens. Matter* 29 (2017) 465901.
- [32] Kevin F. Garrity, Joseph W. Bennett, Karin M. Rabe, David Vanderbilt, Pseudopotentials for high-throughput dft calculations, *Comput. Mater. Sci.* 81 (2014) 446–452.
- [33] K. Schubert, K. Anderko, M. Kluge, H. Beeskow, M. Ilshner, E. D rre, P. Esslinger, Strukturuntersuchung der legierungsphasen Cu_2Te , Cu_3Sb , InTe , Bi_2Se_3 , Pd_5Sb_3 und Pd_5Bi_3 , *Naturwissenschaften* 40 (1953) 269.
- [34] Stefano Baroni, Stefano de Gironcoli, Andrea Dal Corso, Paolo Giannozzi, Phonons and related crystal properties from density-functional perturbation theory, *Rev. Mod. Phys.* 73 (2001) 515–562.
- [35] Pizzi Giovanni, Valerio Vitale, Ryotaro Arita, Stefan Bl gel, Freimuth Frank, et al., Wannier90 as a Community Code: New Features and Applications, 2019 arXiv:1907.09788 [cond-mat.mtrl-sci].
- [36] Jonathan R. Yates, Xinjie Wang, David Vanderbilt, Ivo Souza, Spectral and fermi surface properties from wannier interpolation, *Phys. Rev. B* 75 (2007) 195121.
- [37] Carla Verdi, Feliciano Giustino, Fr hlich electron-phonon vertex from first principles, *Phys. Rev. Lett.* 115 (2015) 176401.
- [38] Atsushi Togo, Isao Tanaka, Spglib: a Software Library for Crystal Symmetry Search, 2018 arXiv:1808.01590 [cond-mat.mtrl-sci].
- [39] Yoyo Hinuma, Pizzi Giovanni, Yu Kumagai, Fumiyasu Oba, Isao Tanaka, Band structure diagram paths based on crystallography, *Comput. Mater. Sci.* 128 (2017) 140–184.
- [40] G. Martinez, B.A. Piot, M. Haki, M. Potemski, Y.S. Hor, A. Materna, S.G. Strzelecka, A. Hruban, O. Caha, J. Nov k, A. Dubroka, C. Dra ar, M. Orlita, Determination of the energy band gap of Bi_2Se_3 , *Sci. Rep.* 7 (2017) 6891.
- [41] Ian T. Witting, Thomas C. Chasapis, Francesco Ricci, Matthew Peters, Nicholas A. Heinz, Geoffroy Hautier, G. Jeffrey Snyder, The thermoelectric properties of bismuth telluride, *Advanced Electronic Materials* 5 (2019) 1800904. <https://onlinelibrary.wiley.com/doi/pdf/10.1002/aelm.201800904>.
- [42] Daehyun Wee, Boris Kozinsky, Nicola Marzari, Marco Fornari, Effects of filling in CoSb_3 : local structure, band gap, and phonons from first principles, *Phys. Rev. B* 81 (2010): 045204.
- [43] Hyun-Sik Kim, Zachary M. Gibbs, Yinglu Tang, Heng Wang, G. Jeffrey Snyder, Characterization of lorenz number with seebeck coefficient measurement, *Appl. Mater.* 3 (2015): 041506, <https://doi.org/10.1063/1.4908244>.
- [44] K.C. Lukas, W.S. Liu, G. Joshi, M. Zebarjadi, M.S. Dresselhaus, Z.F. Ren, G. Chen, C.P. Opeil, Experimental determination of the lorenz number in $\text{Cu}_{0.01}\text{Bi}_{2-\text{Te}_{2.7}\text{Se}_{0.3}}$ and $\text{Bi}_{0.88}\text{Sb}_{0.12}$, *Phys. Rev. B* 85 (2012) 205410.
- [45] Yani Chen, Jinlong Ma, Li Wu, Understanding the thermal conductivity and lorenz number in tungsten from first principles, *Phys. Rev. B* 99 (2019): 020305.