Spectral and Fermi surface properties from Wannier interpolation

Jonathan R. Yates,[1,2] Xinjie Wang,[3] David Vanderbilt,[3] and Ivo Souza[1,2]

[1]*Department of Physics, University of California, Berkeley, CA 94720*
[2]*Materials Science Division, Lawrence Berkeley Laboratory, Berkeley, CA 94720*
[3]*Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey 08854-8019*

(Dated: February 23, 2007)

We present an efficient first-principles approach for calculating Fermi surface averages and spectral properties of solids, and use it to compute the low-field Hall coefficient of several cubic metals and the magnetic circular dichroism of iron. The first step is to perform a conventional first-principles calculation and store the low-lying Bloch functions evaluated on a uniform grid of k-points in the Brillouin zone. We then map those states onto a set of maximally-localized Wannier functions, and evaluate the matrix elements of the Hamiltonian and the other needed operators between the Wannier orbitals, thus setting up an "exact tight-binding model." In this compact representation the k-space quantities are evaluated inexpensively using a generalized Slater-Koster interpolation. Because of the strong localization of the Wannier orbitals in real space, the smoothness and accuracy of the k-space interpolation increases rapidly with the number of grid points originally used to construct the Wannier functions. This allows k-space integrals to be performed with *ab-initio* accuracy at low cost. In the Wannier representation, band gradients, effective masses, and other k-derivatives needed for transport and optical coefficients can be evaluated analytically, producing numerically stable results even at band crossings and near weak avoided crossings.

I. INTRODUCTION

In electronic structure calculations for solids, the evaluation of an observable requires integrating a periodic function in reciprocal space.[1] We will distinguish between three kinds of properties: those where the integral is over the Brillouin zone (type-I), over the Fermi surface (type-II), and over an energy-difference isosurface (type-III). In many cases those integrals take, at $T = 0$, the form

$$I^{(\mathrm{I})} = \sum_n \int_{\mathrm{BZ}} \frac{d\mathbf{k}}{(2\pi)^3}\, F_{nn}(\mathbf{k})\theta(\mathcal{E}_{n\mathbf{k}} - E_{\mathrm{f}}), \qquad (1)$$

$$I^{(\mathrm{II})} = \sum_n \int_{\mathrm{BZ}} \frac{d\mathbf{k}}{(2\pi)^3}\, F_{nn}(\mathbf{k})\delta(\mathcal{E}_{n\mathbf{k}} - E_{\mathrm{f}}), \qquad (2)$$

and

$$I^{(\mathrm{III})}(\omega) = \sum_n^{\mathrm{occ}} \sum_m^{\mathrm{empty}} \int_{\mathrm{BZ}} \frac{d\mathbf{k}}{(2\pi)^3}\, F_{nm}(\mathbf{k})\delta\big[\hbar\omega - (\mathcal{E}_{m\mathbf{k}} - \mathcal{E}_{n\mathbf{k}})\big]. \qquad (3)$$

Here E_{f} is the Fermi level, $\mathcal{E}_{n\mathbf{k}}$ are the eigenenergies of the one-electron states, and $F_{nm}(\mathbf{k})$ involves matrix elements of periodic operators which commute with the crystal translations. Ground-state properties such as the total energy, and dc response functions such as the Hall conductivity, are examples of the first and second type of property, respectively. The third type includes optical absorption in the dipole approximation; other response and spectral functions can be expressed in a similar form.[2] Eqs. (1)–(3) are by no means exhaustive. While properties such as the electron-phonon interaction [described by matrix elements of the form $F_{nm}(\mathbf{k}, \mathbf{k}+\mathbf{q})$ associated with phonon wavevector $\mathbf{q}$] are not explicitly covered above, the methods discussed in this paper can be extended to handle such cases.[3]

In a practical calculation the continuous integral is replaced by a summation over a finite number N of points in the Brillouin zone (BZ),

$$\frac{V_{\mathrm{cell}}}{(2\pi)^3} \int_{\mathrm{BZ}} d\mathbf{k} \to \frac{1}{N} \sum_{\mathbf{k}} w(\mathbf{k}), \qquad (4)$$

where V_{cell} is the cell volume and $w(\mathbf{k})$ are the k-point weights that arise upon restricting the summation to the irreducible wedge of the BZ. For type-I properties of insulators, the integrand varies smoothly across the BZ and this summation converges rapidly with the number of sampled points.[1] In metals the BZ integral must be treated carefully, as the integrand is now discontinuous due to the partial filling of the energy bands. Properties of type II and III pose the additional challenge of sampling isosurfaces in k-space accurately and efficiently. In all these other cases the rate of convergence is much slower than for ground state properties of insulators, and a very large number of k-points may be needed. This sampling problem severely limits the efficiency and accuracy of *ab-initio* calculations for many properties. Examples of such difficulties abound even in the recent literature, and include the magnetocrystalline anisotropy of ferromagnets[4] and optical absorption in hot liquid metals,[5] to name just a few.

In this paper we describe a widely-applicable WF-based scheme for interpolating both the energy bands $\mathcal{E}_{n\mathbf{k}}$ and the matrix elements $F_{nm}(\mathbf{k})$. The method was used in Ref. 6 to compute the anomalous Hall conductivity of iron. Where possible, we have adopted a notation consistent with that work. Ref. 6 dealt with a type-I problem (the quantity being integrated over the BZ was the Berry curvature of the occupied states), and here we extend the method to problems of type II and III.

As an example of a type-II problem, we study the low-field classical Hall coefficient of several cubic metals. This

and other transport coefficients pose an additional challenge to existing *ab-initio* methodologies: how to evaluate the first and possibly also the second k-derivatives of the energy bands at the Fermi level. Early work[7] employed tight-binding (TB) parameterizations of *ab-initio* bands and the derivatives were calculated by numerical differentiation using the linear tetrahedron method. In other work, an analytic evaluation of the TB band gradients has been used to achieve improved numerical stability,[8] but the second derivatives were still computed by finite differences. Other interpolation strategies, such as the SKW scheme[9,10,11] and spectral differentiation,[12] have also been used.

All previous interpolation schemes have one feature in common: the only information retained from the original *ab-initio* calculation is the set of energy eigenvalues on a grid of k-points. Hence the information about the connectivity of the bands is lost, and the interpolation becomes unreliable or even unstable in the vicinity of band crossings, avoided crossings, and near-degeneracies. Moreover, retaining only the eigenenergies strongly restricts the type of matrix elements $F_{nm}(\mathbf{k})$, and hence observables, that can be computed.

A more powerful interpolation scheme can be obtained by keeping one more piece of information, namely, the overlap matrices between the Bloch states at neighboring grid points as in Eq. (5) below. It is perhaps not widely appreciated that the information about band connectivity is encoded in those overlap matrices. Indeed, they are the key input for the WF-construction method,[13,14] and the connectivity can be recovered from the Wannier representation of the bandstructure. Thus, not only do the Wannier-interpolated bands reproduce the *ab-initio* bands with essentially no loss of accuracy, but their k-derivatives can also be evaluated analytically. Like the SKW scheme, the present method is based on Fourier interpolation. Unlike SKW,[10] however, it produces stable and reliable results even in the presence of band crossings and avoided crossings.

Remarkably, a knowledge of the overlap matrices of Eq. (5) allows for the interpolation of properties that are not determined by the energy bands alone, but also depend on the position (or velocity) matrix elements. (More generally, any one-electron operator can be interpolated if, in addition, its matrix elements between the WFs are tabulated.) As an example, we compute the magnetic circular dichroism of iron, a type-III property.

The paper is organized as follows. Section II contains the methodological aspects of the work. We start by reviewing the WF construction methods. We then describe the Wannier-interpolation strategy for a generic periodic operator. The interpolation of the velocity operator, as well as of band gradients and inverse effective masses, is discussed separately. We conclude Sec. II by presenting an improved broadening scheme for performing the k-space integrals. In Sec. III we apply the technique to the low-field Hall effect of several cubic metals, and in Sec. IV to the magnetic circular dichroism of bcc Fe. In Sec. V we provide a brief discussion and conclusion. The Appendix contains some convergence studies.

II. WANNIER INTERPOLATION

A. Construction of the Wannier functions

Ab-initio calculations provide a certain number of low-lying Bloch eigenstates $\psi_{n\mathbf{q}}(\mathbf{r}) = e^{i\mathbf{q}\cdot\mathbf{r}}u_{n\mathbf{q}}(\mathbf{r})$ on a mesh of k-points in the BZ, which we take to be uniform. We will denote points on this "*ab-initio* mesh" by $\mathbf{q}$, to distinguish them from arbitrary or interpolation points, which will be denoted by $\mathbf{k}$.

Consider a type-II (Fermi-surface) problem; two situations may occur. The first one, which is seen in Pb for example, occurs when the Fermi level lies within an isolated group of M bands, where by "isolated" we mean separated from all higher and lower bands by a gap throughout the BZ. In this case it is possible to construct a set of M WFs per unit cell spanning the Hilbert space of the isolated Bloch manifold. This can be done using the method of Marzari and Vanderbilt[13] to obtain so-called maximally localized Wannier functions for that isolated group of bands.

The second scenario occurs when the bands of interest are "entangled" with other bands. Then it is still possible, using the approach of Souza, Marzari, and Vanderbilt,[14] to construct a small number M of maximally localized WFs which describe those bands exactly. The number M of WFs per cell is now to some extent an adjustable parameter. The first step is to identify the subspace of states of interest. Usually this is done by selecting the bands inside an energy window spanning from $E_{\rm min}$ to $E_{\rm max}$. (For type-I and type-III problems $E_{\rm min}$ is normally in the gap below the lowest valence bands and the position of $E_{\rm max}$ depends on the problem, but is always above $E_{\rm f}$. For a type-II problem the only requirement is that $E_{\rm min} < E_{\rm f} < E_{\rm max}$, and in practice the range is adjusted so that WFs with good localization and symmetry properties result.) The number $N_{\mathbf{q}}$ of states within this window can vary from one $\mathbf{q}$-point to another, and we require that $M \geq N_{\mathbf{q}}$ for all $\mathbf{q}$, so that the space spanned by the WFs (the "projected space") can be chosen to contain as a subspace all the window states. In the method of Ref. 14 a second (outer) energy window is used which encloses the previously defined (inner) window. At each $\mathbf{q}$, the M-dimensional projected space is a subspace of the $\mathcal{N}_{\mathbf{q}}$-dimensional space of states contained in the outer window. For the special case of an isolated group of M bands it is natural to choose $M = N_{\mathbf{q}} = \mathcal{N}_{\mathbf{q}}$

Only two pieces of information from the *ab-initio* calculation are needed as an input to the WF-generation algorithm: the $\mathcal{N}_{\mathbf{q}}$ band-energy eigenvalues $\mathcal{E}_{n\mathbf{q}}$, and the $\mathcal{N}_{\mathbf{q}} \times \mathcal{N}_{\mathbf{q}+\mathbf{b}}$ overlap matrices between the cell-periodic Bloch eigenstates at neighboring points $\mathbf{q}$ and $\mathbf{q}+\mathbf{b}$,

$$S_{nm}(\mathbf{q},\mathbf{b}) = \langle u_{n\mathbf{q}} | u_{m,\mathbf{q}+\mathbf{b}} \rangle. \tag{5}$$

The output consists of an $\mathcal{N}_\mathbf{q} \times M$ matrix $\mathcal{U}(\mathbf{q})$ for each $\mathbf{q}$. These matrices relate the original set of $\mathcal{N}_\mathbf{q}$ *ab-initio* Bloch eigenstates selected by the outer window to a new set of M orthonormal Bloch-like states

$$|u_{n\mathbf{q}}^{(\mathrm{W})}\rangle = \sum_{m=1}^{\mathcal{N}_\mathbf{q}} |u_{m\mathbf{q}}\rangle \mathcal{U}_{mn}(\mathbf{q}) \tag{6}$$

that vary smoothly with $\mathbf{q}$. These states are labeled with a superscript (W) to indicate that the WFs are obtained from them by a direct Fourier sum

$$|n\mathbf{R}\rangle = \frac{1}{N_0} \sum_\mathbf{q} e^{-i\mathbf{q}\cdot\mathbf{R}} |u_{n\mathbf{q}}^{(\mathrm{W})}\rangle, \tag{7}$$

where the sum runs over a grid of N_0 $\mathbf{q}$-points and $|n\mathbf{R}\rangle$ is the n-th Wannier function in the unit cell at $\mathbf{R}$.

Although the explicit construction of the WFs obviously requires a knowledge of the $|u_{n\mathbf{q}}\rangle$'s, only the eigenvalues $\mathcal{E}_{n\mathbf{q}}$ and the overlaps $S(\mathbf{q},\mathbf{b})$ are needed to obtain the $\mathcal{U}(\mathbf{q})$ matrices. Retaining this minimal information from the *ab-initio* calculation is thus sufficient for many applications, including the ones presented in this work.

An important object in what follows is the $M \times M$ Hamiltonian matrix in the projected subspace,

$$\begin{aligned} H_{nm}^{(\mathrm{W})}(\mathbf{q}) &= \langle u_{n\mathbf{q}}^{(\mathrm{W})} | \hat{H}(\mathbf{q}) | u_{m\mathbf{q}}^{(\mathrm{W})} \rangle \\ &= \left[\mathcal{U}^\dagger(\mathbf{q}) H(\mathbf{q}) \mathcal{U}(\mathbf{q}) \right]_{nm}, \end{aligned} \tag{8}$$

where $H_{nm}(\mathbf{q}) = \mathcal{E}_{n\mathbf{q}}\delta_{nm}$ is a diagonal $\mathcal{N}_\mathbf{q} \times \mathcal{N}_\mathbf{q}$ matrix and $\hat{H}(\mathbf{q}) = e^{-i\mathbf{q}\cdot\hat{\mathbf{r}}} \hat{H} e^{i\mathbf{q}\cdot\hat{\mathbf{r}}}$. We diagonalize $H^{(\mathrm{W})}(\mathbf{q})$ by finding an $M \times M$ unitary matrix $U(\mathbf{q})$ such that

$$U^\dagger(\mathbf{q}) H^{(\mathrm{W})}(\mathbf{q}) U(\mathbf{q}) = H^{(\mathrm{H})}(\mathbf{q}), \tag{9}$$

where $H_{nm}^{(\mathrm{H})}(\mathbf{q}) = \mathcal{E}_{n\mathbf{q}}^{(\mathrm{H})}\delta_{nm}$. Then $\mathcal{E}_{n\mathbf{q}}^{(\mathrm{H})}$ will be identical to the original *ab-initio* $\mathcal{E}_{n\mathbf{q}}$ for all bands inside the inner window. The corresponding Bloch states

$$|u_{n\mathbf{q}}^{(\mathrm{H})}\rangle = \sum_m |u_{m\mathbf{q}}^{(\mathrm{W})}\rangle U_{mn}(\mathbf{q}) \tag{10}$$

will also coincide with the *ab-initio* eigenstates $|u_{n\mathbf{q}}\rangle$ inside the inner window. We shall refer to a quantity with a (W) or (H) superscript as belonging to the Wannier or Hamiltonian gauge respectively.

B. Wannier interpolation of a periodic operator

The problem we pose for ourselves is the following one. Suppose we are given a periodic operator operator $\hat{\mathcal{O}}$, and we have computed at every $\mathbf{q}$

$$\mathcal{O}_{nm}(\mathbf{q}) = \langle u_{n\mathbf{q}} | \hat{\mathcal{O}}(\mathbf{q}) | u_{m\mathbf{q}} \rangle, \tag{11}$$

its matrix elements between the $\mathcal{N}_\mathbf{q}$ *ab-initio* eigenstates in the outer energy window. How can we interpolate this matrix onto an arbitrary point $\mathbf{k}$? We now show that this can be achieved once the matrices $\mathcal{U}(\mathbf{q})$ and the eigenvalues $\mathcal{E}_{n\mathbf{q}}$ ($n = 1, \ldots, \mathcal{N}_\mathbf{q}$) are known. Naturally, we can only expect the interpolation onto a given $\mathbf{k}$ to be meaningful for those matrix elements (n,m) for which both $\mathcal{E}_{n\mathbf{k}}$ and $\mathcal{E}_{m\mathbf{k}}$ fall within the inner window.

We start by describing in Sec. II B 1 the interpolation strategy as it applies to most ("conventional") properties. Transport and optical properties merit a separate discussion, given in Sec. II B 2.

1. Conventional properties

By analogy with Eq. (8), we define the $M \times M$ matrix

$$\begin{aligned} \mathcal{O}_{nm}^{(\mathrm{W})}(\mathbf{q}) &= \langle u_{n\mathbf{q}}^{(\mathrm{W})} | \hat{\mathcal{O}}(\mathbf{q}) | u_{m\mathbf{q}}^{(\mathrm{W})} \rangle \\ &= \left[\mathcal{U}^\dagger(\mathbf{q}) \mathcal{O}(\mathbf{q}) \mathcal{U}(\mathbf{q}) \right]_{nm}. \end{aligned} \tag{12}$$

Next we find its Fourier sum

$$\mathcal{O}_{nm}^{(\mathrm{W})}(\mathbf{R}) = \frac{1}{N_0} \sum_\mathbf{q} e^{-i\mathbf{q}\cdot\mathbf{R}} \mathcal{O}_{nm}^{(\mathrm{W})}(\mathbf{q}). \tag{13}$$

This operation is done once and for all for each of the N_0 lattice vectors $\mathbf{R}$ lying in a supercell conjugate to the $\mathbf{q}$-mesh. (If the sum is performed using a fast Fourier transform (FFT), the vectors $\mathbf{R}$ will be disposed in a parallelepipedal supercell.) Using Eq. (7) we recognize in $\mathcal{O}_{nm}^{(\mathrm{W})}(\mathbf{R})$ the matrix element of $\hat{\mathcal{O}}$ between WFs,

$$\mathcal{O}_{nm}^{(\mathrm{W})}(\mathbf{R}) = \langle n\mathbf{0} | \hat{\mathcal{O}} | m\mathbf{R} \rangle. \tag{14}$$

In the above equations, the specification of the lattice vectors $\mathbf{R}$ can be left ambiguous with respect to supercell translations ($\mathbf{R} \to \mathbf{R} + \mathbf{R}_\mathrm{sup}$) since $\exp(i\mathbf{q}\cdot\mathbf{R}_\mathrm{sup}) = 1$ for all mesh points $\mathbf{q}$, and thus $\mathcal{O}^{(\mathrm{W})}(\mathbf{R}+\mathbf{R}_\mathrm{sup}) = \mathcal{O}^{(\mathrm{W})}(\mathbf{R})$. However, we now wish to perform the inverse (slow) Fourier transform

$$\mathcal{O}_{nm}^{(\mathrm{W})}(\mathbf{k}) = \sum_\mathbf{R} e^{i\mathbf{k}\cdot\mathbf{R}} \mathcal{O}_{nm}^{(\mathrm{W})}(\mathbf{R}), \tag{15}$$

which yields the interpolation of Eq. (12) onto an arbitrary point $\mathbf{k}$. At this point the set of lattice vectors must be defined more precisely, since for points $\mathbf{k}$ not on the $\mathbf{q}$-mesh $\exp(i\mathbf{k}\cdot\mathbf{R}_\mathrm{sup}) \neq 1$, and the smoothness of interpolation will depend on the choice of set. Using the FFT parallelepipedal supercell, for example, is generally not optimal. Instead, one wants to choose lattice vectors lying inside the Wigner-Seitz (WS) supercell centered on the origin,[6,14] but the details may vary (e.g., sharing weights of $\mathbf{R}$-vectors lying on the boundary of the WS supercell, or truncation to a sphere lying within the WS supercell). In practice the $|O_{nm}^{(\mathrm{W})}(\mathbf{R})|$ decay exponentially with $|\mathbf{R}|$, as expected if the WFs are exponentially localized, so the results should not be very sensitive to this choice, when using a sufficiently dense $\mathbf{q}$-mesh. These remarks are illustrated in the Appendix.

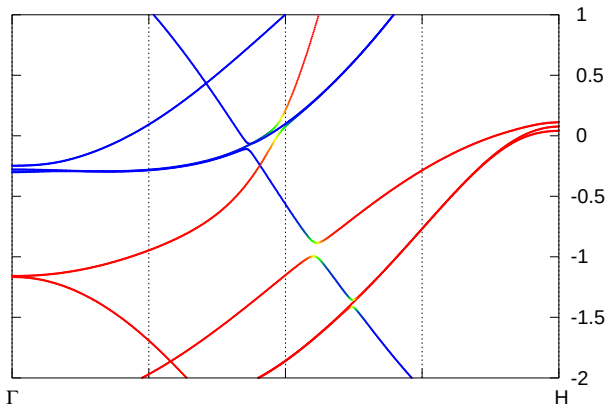

FIG. 1: (Color online.) Wannier-interpolated bands of bcc Fe along Γ–H. The bands are color-coded according to the value of $\langle S_z \rangle$: red for spin up and blue for spin down. The energies are given in eV and the Fermi level is at 0 eV. The vertical dashed lines indicate k-points on the *ab-initio* mesh used for constructing the WFs.

The final step is to transform the matrix of Eq. (15) from the Wannier to the Hamiltonian gauge. To find the required unitary matrix $U(\mathbf{k})$ we repeat the above steps for $\hat{\mathcal{O}} = \hat{H}$ to obtain $H^{(\mathrm{W})}(\mathbf{k})$. The matrix $U(\mathbf{k})$ is then given by Eq. (9), with the replacement $\mathbf{q} \to \mathbf{k}$. Finally,

$$\mathcal{O}^{(\mathrm{H})}(\mathbf{k}) = U^{\dagger}(\mathbf{k}) \mathcal{O}^{(\mathrm{W})}(\mathbf{k}) U(\mathbf{k}), \tag{16}$$

where $M \times M$ matrix products are implied on the right-hand side. This solves the problem posed above at the beginning of Sec. II B.

Once the WF matrix elements of both the operator of interest and the Hamiltonian are tabulated, the interpolation onto an arbitrary k-point requires only inexpensive operations on small $M \times M$ matrices. When $\hat{\mathcal{O}} = \hat{H}$, the present scheme reduces to Slater-Koster interpolation, with the maximally localized WFs playing the role of the TB basis orbitals.[14]

Fig. 1 shows the interpolated band structure of bcc Fe along Γ–H, using the same WFs and computational details as in Ref. 6. If one were to superimpose on this plot the energy bands obtained by performing an *ab-initio* calculation for a large number of points along the same line in k-space, they would be essentially indistinguishable from the interpolated ones. Following Ref. 15, we indicate with vertical dashed lines k-points on the $\mathbf{q}$-mesh used for constructing the WFs (an $8 \times 8 \times 8$ grid in the full BZ). It is apparent that the Wannier interpolation procedure *succeeds in resolving details on a scale much smaller than the spacing between those points.* In particular, *the correct band connectivity is obtained.* This means that spin-orbit-induced avoided crossings, for example, are never mistaken for actual crossings, no matter how weak the spin-orbit interaction.

We note in passing that one could have formulated the problem at the beginning of Sec. II B somewhat differently: rather than viewing the matrix elements of $\hat{\mathcal{O}}$ and $\hat{H}$ between the $\mathcal{N}_{\mathbf{q}}$ *ab-initio* Bloch states at each $\mathbf{q}$ as the basic ingredients of the method, we could have assigned that role to the matrix elements of the two operators between the WFs. Even if the latter viewpoint is in some ways the more fundamental one, in practical implementations one often obtains the Wannier matrix elements (14) via Eqs. (11)–(13). When doing so, the Wannier orbitals are never explicitly constructed.

2. *Transport and optical properties*

The treatment of transport and optical properties in crystals is more subtle. We will restrict our discussion to the electric-dipole approximation, where those properties depend on matrix elements of the velocity operator. The formulation of the previous subsection could in principle be used to interpolate the full velocity operator $\hat{v}_\alpha = -(i/\hbar)[\hat{r}_\alpha, \hat{H}]$ ($\alpha = 1, 2, 3$). Its matrix elements, as those of any other periodic operator, transform between the Wannier and Hamiltonian gauges according to Eq. (16) (such a matrix will be called "gauge-covariant"[6]). They are given by[16]

$$v_{nm,\alpha}(\mathbf{k}) = \langle \psi_{n\mathbf{k}} | \hat{v}_\alpha | \psi_{m\mathbf{k}} \rangle = \frac{1}{\hbar} \left\langle u_{n\mathbf{k}} \middle| \frac{\partial \hat{H}(\mathbf{k})}{\partial k_\alpha} \middle| u_{m\mathbf{k}} \right\rangle. \tag{17}$$

However, when describing the dynamics of electrons in crystals it is convenient, from the points of view of both numerics and physics, to decompose the velocity operator into two terms.[16] This is achieved by taking $\partial_\alpha = \partial/\partial k_\alpha$ outside the brackets on the right-hand side of Eq. (17) and compensating for the extra terms that appear. After a few manipulations one obtains

$$v_{nm,\alpha}(\mathbf{k}) = \frac{1}{\hbar} \frac{\partial \mathcal{E}_{n\mathbf{k}}}{\partial k_\alpha} \delta_{nm} - \frac{i}{\hbar} (\mathcal{E}_{m\mathbf{k}} - \mathcal{E}_{n\mathbf{k}}) A_{nm,\alpha}(\mathbf{k}) \tag{18}$$

where

$$A_{nm,\alpha}(\mathbf{k}) = i \langle u_{n\mathbf{k}} | \partial_\alpha u_{m\mathbf{k}} \rangle. \tag{19}$$

The first (second) term on the right-hand side of Eq. (18) is diagonal (off-diagonal) in the band index. Note that neither is separately gauge-covariant. For example, the second one contains k-derivatives acting on the eigenstates. According to Eq. (10), additional terms involving $\partial U(\mathbf{k})/\partial k_\alpha$ will therefore appear when transforming between the Wannier and Hamiltonian gauges. Let us define, for every matrix object $\mathcal{O}$,

$$\overline{\mathcal{O}}^{(\mathrm{H})} = U^{\dagger} \mathcal{O}^{(\mathrm{W})} U \tag{20}$$

so that, by definition, $\overline{\mathcal{O}}^{(\mathrm{H})} = \mathcal{O}^{(\mathrm{H})}$ only for gauge-covariant objects. This notation will be used in the next section for expressing the intraband (diagonal) velocity matrix elements and the effective mass tensor, and in Sec. IV for the interband (off-diagonal) velocity.

C. Band gradient and Hessian

1. Notation

We make use of the first and second k-derivatives of the Hamiltonian matrix,

$$H_{nm,\alpha} = \frac{\partial H_{nm}}{\partial k_\alpha}, \tag{21}$$

$$H_{nm,\alpha\beta} = \frac{\partial^2 H_{nm}}{\partial k_\alpha \partial k_\beta}, \tag{22}$$

and define $\overline{H}_{nm,\alpha}$ and $\overline{H}_{nm,\alpha\beta}$ via Eq. (20) as usual. We also define the first and second k-derivatives of the band energy,

$$\mathrm{v}_{n\mathbf{k},\alpha} = \frac{1}{\hbar}\frac{\partial \mathcal{E}_{n\mathbf{k}}}{\partial k_\alpha}, \tag{23}$$

$$\mu_{n\mathbf{k},\alpha\beta} = \frac{1}{\hbar^2}\frac{\partial^2 \mathcal{E}_{n\mathbf{k}}}{\partial k_\alpha \partial k_\beta}, \tag{24}$$

which have the interpretation of group velocity (ignoring Berry-curvature contributions) and inverse effective mass tensor, respectively. The strategy for interpolating these quantities is similar to the one developed in Ref. 6 for the Berry curvature. We will again make extensive use of the antihermitian matrix

$$D^{(\mathrm{H})}_{nm,\alpha} \equiv (U^\dagger \partial_\alpha U)_{nm} = \begin{cases} \dfrac{\overline{H}^{(\mathrm{H})}_{nm,\alpha}}{\mathcal{E}^{(\mathrm{H})}_m - \mathcal{E}^{(\mathrm{H})}_n} & \text{if } n \neq m \\ 0 & \text{if } n = m \end{cases} \tag{25}$$

defined in that work.

2. Non-degenerate bands

First we consider the band-gradient velocity, Eq. (23). In the Hamiltonian gauge $H^{(\mathrm{H})}_{nm} = \mathcal{E}^{(\mathrm{H})}_n \delta_{nm}$, and hence $H^{(\mathrm{H})}_{nm,\alpha} = \hbar \mathrm{v}^{(\mathrm{H})}_{n\alpha}\delta_{nm}$. Differentiating Eq. (9) with respect to k_α,

$$\begin{aligned} H^{(\mathrm{H})}_\alpha &= U^\dagger H^{(\mathrm{W})}_\alpha U + \left\{U^\dagger H^{(\mathrm{W})} \partial_\alpha U + \mathrm{h.c.}\right\} \\ &= \overline{H}^{(\mathrm{H})}_\alpha + \left\{H^{(\mathrm{H})} D^{(\mathrm{H})}_\alpha + \mathrm{h.c.}\right\}, \end{aligned} \tag{26}$$

where each object is an $M \times M$ matrix and h.c. denotes the Hermitian conjugate. Because of the extra terms in curly brackets we have $H^{(\mathrm{H})}_\alpha \neq \overline{H}^{(\mathrm{H})}_\alpha$ and thus H_α, the first term on the right-hand side of Eq. (18), is not gauge-covariant. However, those extra terms do not contribute to the velocity; being the product of a diagonal matrix ($H^{(\mathrm{H})}$) with an antihermitian matrix ($D^{(\mathrm{H})}_\alpha$), they only contain off-diagonal elements which cancel those in $\overline{H}^{(\mathrm{H})}_\alpha$. Thus

$$\hbar \mathrm{v}^{(\mathrm{H})}_{n\alpha} = H^{(\mathrm{H})}_{nn,\alpha} = \overline{H}^{(\mathrm{H})}_{nn,\alpha} = \left[U^\dagger H^{(\mathrm{W})}_\alpha U\right]_{nn}. \tag{27}$$

Differentiating this equation yields the inverse effective mass tensor (24)

$$\begin{aligned} \hbar^2 \mu^{(\mathrm{H})}_{n,\alpha\beta} &= \left[\partial_\beta \overline{H}^{(\mathrm{H})}_\alpha\right]_{nn} \\ &= \left[U^\dagger H^{(\mathrm{W})}_{\alpha\beta} U\right]_{nn} + \left\{U^\dagger H^{(\mathrm{W})}_\alpha \partial_\beta U + \mathrm{h.c.}\right\}_{nn} \\ &= \overline{H}^{(\mathrm{H})}_{nn,\alpha\beta} + \left\{\overline{H}^{(\mathrm{H})}_\alpha D^{(\mathrm{H})}_\beta + \mathrm{h.c.}\right\}_{nn}. \end{aligned} \tag{28}$$

Unlike Eq. (26), here the matrix in curly brackets has nonzero diagonal elements which contribute to $\mu^{(\mathrm{H})}_{n,\alpha\beta}$.

Eqs. (27)–(28) are the desired expressions for the band derivatives, valid away from degeneracies and inside the inner energy window. They involve the $M \times M$ matrices $U(\mathbf{k})$ calculated in Sec. II B 1, $D^{(\mathrm{H})}_\alpha$ given by Eq. (25), $H^{(\mathrm{W})}_\alpha$, and $H^{(\mathrm{W})}_{\alpha\beta}$. The last two involve k-derivatives of Eq. (15) that can be taken analytically, i.e.,

$$H^{(\mathrm{W})}_{nm,\alpha}(\mathbf{k}) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} i R_\alpha \langle n\mathbf{0}|\hat{H}|m\mathbf{R}\rangle \tag{29}$$

and

$$H^{(\mathrm{W})}_{nm,\alpha\beta}(\mathbf{k}) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} (-R_\alpha R_\beta) \langle n\mathbf{0}|\hat{H}|m\mathbf{R}\rangle. \tag{30}$$

3. Discussion

In order to interpret the above expressions it is illuminating to introduce $\|\phi_n\rangle\!\rangle$, the n-th M-component column vector of U.[6] $\|\phi_n\rangle\!\rangle$ is an eigenvector of $H^{(\mathrm{W})}$, the Hamiltonian operator projected onto the WF space. We then recognize in Eq. (27) the Hellmann-Feynman result $\hbar \mathrm{v}^{(\mathrm{H})}_{n\alpha} = \langle\!\langle \phi_n \| H^{(\mathrm{W})}_\alpha \| \phi_n \rangle\!\rangle$, and in Eq. (28) the expression for the effective mass tensor in empirical TB theory.[17] Eq. (25) is the standard result from $\mathbf{k}\cdot\mathbf{p}$ perturbation theory, in terms of the TB states. It involves the operator $(1/\hbar)H^{(\mathrm{W})}_\alpha$, which differs from the full velocity operator in that the position-operator-dependent terms are absent.[6,17] We note that all the formulas given so far and in the rest of the paper remain valid when the *ab-initio* Hamiltonian contains non-local and spin-orbit terms.

The advantage of this reformulation of $\mathbf{k}\cdot\mathbf{p}$ perturbation theory is that it is done strictly in terms of the small number M of M-dimensional states $\|\phi_n\rangle\!\rangle$ at an arbitrary $\mathbf{k}$, and yet it is exact within the inner energy window. In contrast, the formulation in terms of the original *ab-initio* states on the $\mathbf{q}$-grid is considerably more expensive and usually entails a truncation error.[18]

4. Degeneracies

While meaningful band derivatives can be defined via degenerate $\mathbf{k}\cdot\mathbf{p}$ perturbation theory even at degeneracy points in the BZ,[19] this is not possible when the only information available about the bandstructure is a list of eigenenergies on a predetermined coarse k-point grid. In that case, the information about the band connectivity is lost, and finite-difference estimates of the derivatives become ill-defined at points of degeneracy, which have to be carefully avoided.[2]

In the present formulation the k-gradients of the degenerate states are the eigenvalues of the submatrix

$$\left[\overline{H}_\alpha^{(\mathrm{H})}\right]_{\mu\nu} = \langle\!\langle \phi_\mu \| H_\alpha^{(\mathrm{W})} \| \phi_\nu \rangle\!\rangle, \tag{31}$$

where the indices μ and ν run over the degenerate states only. We update the $M\times M$ matrix U by replacing the columns corresponding to those states with the rotated states that diagonalize $H_\alpha^{(\mathrm{W})}$. The Hessian matrix can then be obtained from Eq. (28) using the updated U and a modified form of Eq. (25),

$$D_{nm,\alpha}^{(\mathrm{H})} = \begin{cases} \dfrac{\overline{H}_{nm,\alpha}^{(\mathrm{H})}}{\mathcal{E}_m^{(\mathrm{H})} - \mathcal{E}_n^{(\mathrm{H})}} & \text{if } \mathcal{E}_n^{(\mathrm{H})} \neq \mathcal{E}_m^{(\mathrm{H})} \\ 0 & \text{if } \mathcal{E}_n^{(\mathrm{H})} = \mathcal{E}_m^{(\mathrm{H})}. \end{cases} \tag{32}$$

In cases of degeneracies (at band edges, for example) where some of the eigenvalues of the matrix (31) are equal, a first-order treatment is inadequate and the correct rotation between eigenstates needed to compute the Hessian is found by going to second order in degenerate perturbation theory. This amounts to diagonalizing the submatrix obtained from the right-hand side of Eq. (28) by replacing the subscripts nn therein by $\mu\nu$, and letting μ and ν run over the the first-order-degenerate bands. The desired Hessian matrix elements are the eigenvalues of that submatrix.

In our calculations we employ the first-order form of degenerate perturbation theory when two or more energy eigenvalues lie within 10^{-4} eV of each other and we subsequently use the second-order form if in addition two or more band gradients differ by less than 0.1 eV-Å.

5. Validation

As an illustration, we used the above formulas to calculate the inverse effective mass of the three p-like valence bands of Pb along the Γ–K direction in k-space. (We have chosen this example because it displays a threefold band-edge degeneracy at Γ and a band crossing between Γ and K.)

In all the calculations in this work, lattice constants are taken from Ref. 20. The Bloch states are obtained with the `PWSCF` code[21] using density-functional theory in the local-density approximation, together with the planewave-pseudopotential formalism.[1] Norm-conserving pseudopotentials are employed, and spin-orbit effects are included in Sec. IV and Fig. 1 only. The WFs are generated using the `WANNIER90` code.[22]

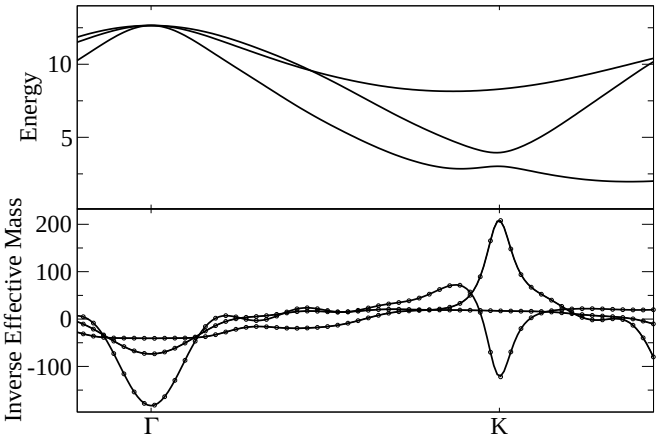

FIG. 2: Upper panel: Dispersion of the three p-like energy bands of Pb along the Γ–K direction, obtained by interpolating a non-relativistic *ab-initio* calculation. Lower panel: Inverse effective masses of those bands along the same direction, calculated in two ways: from the interpolated eigenenergies on a regular mesh of points using a spline fit (circles), and from perturbation theory in the Wannier representation (solid lines).

In Pb the lowest four valence bands crossing the Fermi level are separated everywhere in the Brillouin zone from higher bands, so that the original prescription of Marzari and Vanderbilt[13] can be used to generate the corresponding maximally localized WFs. They are atom-centered and have sp^3 character, with the axis of each orbital pointing towards a nearest neighbor.

The inverse effective masses are shown as solid lines in the lower panel of Fig. 2. For comparison, we also plot as circles in the lower panel the values obtained by fitting a spline function to the energy eigenvalues on a dense grid of k-points along the path. We remark that whereas in the analytic method band crossings are handled automatically, in order to obtain a smooth spline fit it was necessary to manually reorder the eigenvalues close to the band crossing. Once that is done, the agreement between the two is essentially perfect.

D. Adaptive broadening scheme for k-space integration

We conclude this section by discussing the evaluation of type-I, II, and III integrals, Eqs. (1–3). In order to accelerate the convergence of type-I integrals with respect to the number of sampling points in Eq. (4), a broadening scheme can be used.[23,24] This amounts to replacing the step function in Eq. (1) with a Fermi-Dirac-like smear-

ing function. In the case of type-I integrals, smearing is most important when relatively few sampling points are used, as tends to be the case whenever the evaluation of the integrand is expensive.[24] If, however, the integrand is inexpensive, as is the case when using Wannier interpolation, then it is possible to converge the BZ integral without resorting to smearing. For example, no smearing was used in Ref. 6 for integrating the Berry curvature over the occupied states of bcc Fe.

Smearing plays a more fundamental role in integrals of types II and III: when replacing the BZ integral in Eqs. (2-3) by a grid summation, the δ-functions must be replaced by normalized functions with non-zero width, such as Gaussians. For example, in Eq. (2) one would replace $\delta(E_{\rm f} - \mathcal{E}_{n\mathbf{k}})$ by

$$g_{n\mathbf{k}}(E_{\rm f}) = \frac{1}{\sqrt{2\pi}W} \exp\left(\frac{-(E_{\rm f} - \mathcal{E}_{n\mathbf{k}})^2}{2W^2}\right). \qquad (33)$$

Ideally the Gaussian width should be, for a given grid spacing Δk, comparable with the level spacing $\Delta\mathcal{E}_{n\mathbf{k}}$. The level spacing is however difficult to estimate, and the common practice is to set W to a constant for all bands and k-points. As a result, FS sheets arising from steep and flat bands are not described consistently. This is a serious disadvantage of broadening schemes with respect to the linear tetrahedron method, as discussed in Ref. 25.

This drawback is easy to remedy within the Wannier-interpolation method, since the band derivatives are readily available (Sec. II C), and can be used to estimate the level spacing. The simple estimate $\Delta\mathcal{E}_{n\mathbf{k}} \sim |\partial\mathcal{E}_{n\mathbf{k}}/\partial\mathbf{k}|\Delta k$ suggests using a state-depending broadening width

$$W_{n\mathbf{k}} = a\left|\frac{\partial\mathcal{E}_{n\mathbf{k}}}{\partial\mathbf{k}}\right|\Delta k \qquad (34)$$

for type-I and type-II integrals (a is a dimensionless constant of the order of unity), and

$$W_{nm,\mathbf{k}} = a\left|\frac{\partial\mathcal{E}_{m\mathbf{k}}}{\partial\mathbf{k}} - \frac{\partial\mathcal{E}_{n\mathbf{k}}}{\partial\mathbf{k}}\right|\Delta k \qquad (35)$$

for type-III integrals. With this prescription W is no longer an independent adjustable parameter from Δk, guaranteeing that the $\Delta k \to 0$ and $W \to 0$ limits are approached consistently. Several smearing functions beyond a simple Gaussian have been proposed[24,26] and can be used straight forwardly with the adaptive smearing scheme. For all of the calculations presented in this work we use the first-order Hermite polynomial scheme introduced by Methfessel and Paxton.[24]

The above first-order adaptive smearing should be reliable whenever the level-spacing is gradient-dominated. In practice we find that it works rather well even near critical points.This is illustrated in Fig. 3, where we show the density of states of diamond calculated using both the adaptive and conventional (fixed width) smearing, with a $50 \times 50 \times 50$ interpolation mesh. When using a fixed

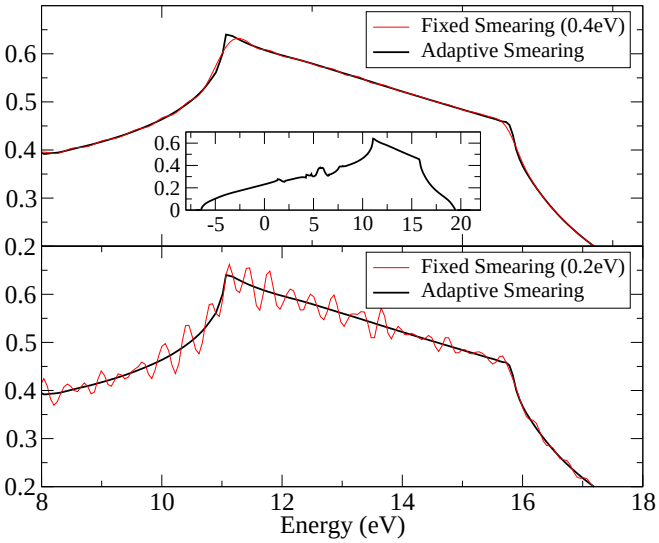

FIG. 3: (Color online.) Density of states of bulk diamond calculated in the range 8-18eV using the conventional Gaussian broadening approach (light thin lines) with a fixed width of 0.4 eV in the upper panel and 0.2 eV in the lower panel, versus the adaptive broadening approach (dark thick lines). The inset shows the density of states in the full valence band range computed with the adaptive broadening approach.

width of $W = 0.4$ eV, the sharp van-Hove features are not well described. Reducing it to 0.2 eV improves the situation for some of them, but introduces spurious oscillations whenever the level-spacing becomes larger than W. With the adaptive scheme such oscillations do not occur for a sensible choice of a, and the sharp features are well-described. We have used $a = 1.0$, but find the results to be quite robust for $0.8 < a < 1.3$.

III. LOW-FIELD HALL COEFFICIENT OF CUBIC METALS

As a first benchmark application we compute the "classical" low-field Hall coefficient of cubic metals, which is given by

$$R_{\rm H} = \frac{\sigma_{xy,z}}{\sigma_{xx}^2}, \qquad (36)$$

$$\sigma_{xx} = q_e^2 \sum_n \int \frac{d\mathbf{k}}{(2\pi)^3}\, \tau_{n\mathbf{k}} \mathrm{v}_{n\mathbf{k},x}^2 \left(-\frac{\partial f}{\partial\mathcal{E}}\right), \qquad (37)$$

and

$$\sigma_{xy,z} = q_e^3 \sum_n \int \frac{d\mathbf{k}}{(2\pi)^3}\, \tau_{n\mathbf{k}}^2 \left(-\frac{\partial f}{\partial\mathcal{E}}\right) \times \\ \times \left(\mathrm{v}_{n\mathbf{k},x}^2 \mu_{n\mathbf{k},yy} - \mathrm{v}_{n\mathbf{k},x}\mathrm{v}_{n\mathbf{k},y}\mu_{n\mathbf{k},xy}\right). \qquad (38)$$

(For the systems studied in this Section, which are non-ferromagnetic and do not include the spin-orbit interaction, the sum over spin-degenerate bands will give

TABLE I: Hall coefficient $R_{\rm H}$, in units of $10^{-11}{\rm m}^3{\rm C}^{-1}$. References to the experimental data can be found in Ref. 27.

	This work	Ref. 7	Ref. 10	Experiment
Li	−12.7	−12.8	−12.4	−15.5
Al	−2.5	−1.7	−3.4	−3.43
Cu	−4.9	−5.2	−4.9	−5.17
Pd	−11.9	−6.0	−17	−7.6

rise to factors of two, which are not written explicitly.) σ_{xx} is the longitudinal conductivity, and and $\sigma_{xy,z} = \partial\sigma_{xy}/\partial B_z$, where σ_{xy} is the antisymmetric (Hall) conductivity. The above expressions are obtained from a Bloch-Boltzmann description of transport; for a detailed derivation, see Ref. 27. We note that Eq. (38) is written in a form which is specific to cubic metals. The quantities $\mathrm{v}_{n\mathbf{k},\alpha}$ and $\mu_{n\mathbf{k},\alpha\beta}$ are given by Eqs. (23) and (24), $f(\mathcal{E})$ is the Fermi-Dirac distribution function, and $q_e < 0$ is the electron charge. At low temperatures $(-\partial f/\partial\mathcal{E})$ tends to $\delta(\mathcal{E} - E_{\rm f})$, and both σ_{xx} and $\sigma_{xy,z}$ become FS integrals of the form of Eq. (2). In the constant relaxation-time approximation $\tau_{n\mathbf{k}}$ drops out from Eq. (36) so that $R_{\rm H}$ is completely specified by the first and second band derivatives at $E_{\rm f}$.

Calculations were done for Li, Al, Cu and Pd. Unlike Pb, in these metals the set of bands crossing the Fermi level is not isolated. Therefore, in order to generate maximally localized WFs we first used the disentanglement procedure summarized in Sec. II A to obtain an optimal projected space. The number of bands contained therein must be at least equal to the number of bands crossing the Fermi level. However, it is often desirable to extract a somewhat larger projected space if this produces a more symmetric set of Wannier functions.

For lithium we obtained four atom-centered WFs per primitive cell with sp^3 character. For aluminum we extracted a nine-dimensional projected subspace. The resulting WFs are atom-centered, but have no distinct symmetry characteristics. For Cu and Pd we used seven WFs: five with d character on atom centers, and two with s character located at the tetrahedral interstitial sites. These have been previously described for Cu in Ref. 14.

The computational details are the same as in Section II C 5. We obtain the self-consistent ground state using a 16×16×16 Monkhorst-Pack mesh of k-points and a fictitious Fermi smearing[24] of 0.02 Ry for the Brillouin-zone integration. We use the local density approximation; for the materials studied we find the results are not significantly altered by using a GGA such as PBE.[28] To compute the Hall coefficient we use an *ab-initio* grid of 12×12×12 **q**-points and obtain the required quantities on a uniform mesh of 60×60×60 **k**-points. We implement an adaptive mesh refinement scheme in which we identify those points of the k-space mesh at which at least one band lies with 0.5eV of the Fermi energy and obtain the required quantities on a $7 \times 7 \times 7$ submesh spanning the original cell associated with this mesh point. We find these parameters give converged values of the Hall coefficient for the four metals studied. This is particularly reassuring in the case of Pd, where previous techniques encountered difficulties because of the occurrence of bands crossings near the Fermi surface.[10]

The results for the Hall coefficient $R_{\rm H}$ are compiled in Table I. For Li, Al, and Cu we find excellent agreement with previous calculations based on empirical TB fitting to *ab-initio* bands,[7] as well as *ab-initio* calculation combined with SKW interpolation.[10] The case of Pd is more delicate as $R_{\rm H}$ depends critically on the details of the *ab-initio* calculation. For example a shift upwards (downwards) in the Fermi level of 0.002Ry causes a change of -3 $(+2)$ $\times 10^{-11}{\rm m}^3{\rm C}^{-1}$. In view of this we refine the position of the Fermi level in a final non-self-consistent step by interpolating the band energies and gradients onto a $60 \times 60 \times 60$ **k**-mesh and using the adaptive broadening scheme. Our converged value of $R_{\rm H}$ is intermediate between the two previously computed values, and shows a relatively large discrepancy with experiment. Previous authors have suggested[7,29] that it maybe necessary to go beyond the constant relaxation time approximation to give a good description of the transport properties of Pd.

IV. MAGNETIC CIRCULAR DICHROISM

Magneto-optical effects in ferromagnets result from a combination of exchange splitting and spin-orbit coupling (SOC).[30,31] The former breaks time-reversal (TR) in the spin channel, and the latter transmits the TR-breaking to the orbital motion of the electrons, endowing the optical conductivity tensor with an antisymmetric component. The simplest such effect to evaluate is magnetic circular dichroism (MCD), the difference in absorption between left- and right-circularly-polarized light, and we have chosen it for illustrative purposes. It is given by the imaginary part of the antisymmetric conductivity, $\sigma^{(2)}_{{\rm A},\alpha\beta}(\omega) = -\sigma^{(2)}_{{\rm A},\beta\alpha}(\omega)$.

A. Evaluation of the Kubo formula

Ab-initio calculations of magneto-optical effects demand high accuracy and dense k-space sampling. The spin-orbit interaction is typically a small perturbation on top of the much larger exchange splitting, and the modifications that it produces on the electronic structure (both in the energy bands and in the matrix elements) are subtly and strongly dependent on k-point and band index.

The conductivity $\sigma^{(2)}_{{\rm A},\alpha\beta}(\omega)$ is evaluated from the Kubo formula of linear-response theory in the electric-dipole approximation.[31] The needed ingredients are the energy eigenvalues of the states involved in the optical transitions and the transition matrix elements. We will evaluate the interband contribution to the magneto-optical

absorption using Eq. (18) for the electric-dipole transition matrix elements, where it is now understood that all Bloch functions $|u_{n\mathbf{k}}\rangle$ are spinors determined from a Hamiltonian that includes the spin-orbit interaction. One finds

$$\sigma^{(2)}_{\mathrm{A},\alpha\beta}(\omega) = -\frac{\pi e^2\omega}{\hbar}\sum_{n}^{\mathrm{occ}}\sum_{m}^{\mathrm{empty}}\int\frac{d\mathbf{k}}{(2\pi)^3}\,\mathrm{Im}\big(A_{nm,\alpha}A_{mn,\beta}\big) \times\ [\delta(\omega-\omega_{mn})-\delta(\omega+\omega_{mn})]\,, \tag{39}$$

where $\hbar\omega_{mn}=\mathcal{E}_m-\mathcal{E}_n$.

Eq. (39) is a type-III integral of the form of Eq. (3). When evaluating it by Wannier interpolation it must be kept in mind that the Wannier-derived bands reproduce the *ab-initio* ones only inside the inner energy window, and therefore its range must be adjusted according to the maximum desired absorption frequency. The matrix elements $A_{nm,\alpha}$ are to be evaluated in the Hamiltonian gauge, and the interpolation of $A^{(\mathrm{H})}_{nm,\alpha}$ is based on the two relations[6]

$$A^{(\mathrm{H})}_{\alpha} = \overline{A}^{(\mathrm{H})}_{\alpha} + iD^{(\mathrm{H})}_{\alpha} \tag{40}$$

and

$$A^{(\mathrm{W})}_{nm,\alpha}(\mathbf{k}) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}}\,\langle\mathbf{0}n|\hat{r}_\alpha|\mathbf{R}m\rangle\,, \tag{41}$$

where $D^{(\mathrm{H})}_{\alpha}$ is given by Eq. (25) and $\overline{A}^{(\mathrm{H})}_{\alpha}$ and $A^{(\mathrm{W})}_{\alpha}$ are related by Eq. (20). Inserting Eq. (40) into Im(...) in Eq. (39), we find

$$\begin{aligned}\mathrm{Im}\big(A^{(\mathrm{H})}_{nm,\alpha}A^{(\mathrm{H})}_{mn,\beta}\big) &= \mathrm{Im}\left(\overline{A}^{(\mathrm{H})}_{nm,\alpha}\overline{A}^{(\mathrm{H})}_{mn,\beta}\right)\\ &+ \mathrm{Re}\left(\overline{A}^{(\mathrm{H})}_{nm,\alpha}D^{(\mathrm{H})}_{mn,\beta}+D^{(\mathrm{H})}_{nm,\alpha}\overline{A}^{(\mathrm{H})}_{mn,\beta}\right)\\ &- \mathrm{Im}\left(D^{(\mathrm{H})}_{nm,\alpha}D^{(\mathrm{H})}_{mn,\beta}\right).\end{aligned} \tag{42}$$

The contributions to $\sigma^{(2)}_{\mathrm{A},\alpha\beta}(\omega)$ from the three terms on the right-hand side will be denoted as $\overline{A}$–$\overline{A}$, D–$\overline{A}$, and D–D, respectively.

B. Results for bcc Fe

Unlike the calculations presented earlier in the paper, to calculate the MCD spectrum we have used relativistic pseudopotentials which explicitly include spin-orbit effects.[34] The computational details for the *ab-initio* calculation, WF-generation, and treatment of the spin-orbit interaction are the same as in Ref. 6. The integral in Eq. (39) was evaluated on a uniform **k**-mesh containing $125\times125\times125$ points using the adaptive broadening scheme, which we find to be essential for resolving the fine details in the MCD spectrum. The spin magnetization is along the z-axis, so that $\sigma^{(2)}_{\mathrm{A},xy}(\omega)$ is the only independent non-zero component.

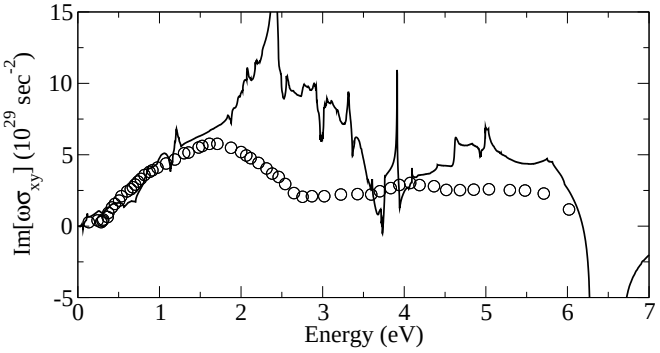

FIG. 4: Magnetic circular dichroism spectrum of bcc iron. The calculated spectrum (solid lines) is compared with the experimental spectrum from Ref. 32 as reproduced in Ref. 33 (open circles).

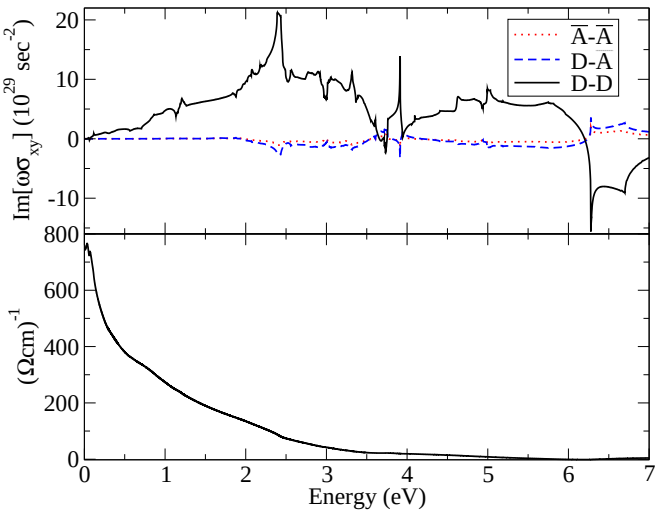

FIG. 5: (Color online.) Upper panel: decomposition of the magnetic circular dichroism spectrum into the three terms defined by the Wannier-interpolation procedure. Lower panel: cumulative anomalous Hall conductivity (AHC) versus ω, defined as the contribution to the AHC from frequencies higher than ω in Eq. (43).

It is conventional to plot the MCD spectrum as $\omega\sigma^{(2)}_{\mathrm{A},xy}(\omega)$, and adopt this convention in Fig. 4. Our calculated spectrum for bcc Fe is in excellent agreement with the one computed in Ref. 33 using a completely different electronic structure method. (Previous calculations of magneto-optical effects have most commonly used all-electron methods.) As remarked in Ref. 30, this level of agreement between two different calculations is non-trivial when it comes to the spin-orbit-induced $\sigma^{(2)}_{\mathrm{A},xy}(\omega)$. It provides a strong validation of the Wannier-interpolation scheme combined with the pseudopotential-planewave method as a viable way of computing magneto-optical effects.

The upper panel of Fig. 5 shows the decomposition of the calculated MCD spectrum into the three contribu-

tions ($\overline{A}$–$\overline{A}$, D–$\overline{A}$, and D–D) defined by the Wannier-interpolation procedure, as discussed following Eq. (42). It is clear that the D–D contribution tends to dominate the spectrum in the frequency range from 0 to 7 eV, especially at the lowest frequencies. For frequencies above 7 eV (not shown), the $\overline{A}$–$\overline{A}$ and D–$\overline{A}$ terms become significant.

The interband MCD spectrum $\sigma^{(2)}_{\mathrm{A},xy}(\omega)$ is related to the Karplus-Luttinger anomalous Hall conductivity[6] (AHC) $\sigma^{(1)}_{\mathrm{A},xy}(0)$ by the Kramers-Krönig relation

$$\sigma^{(1)}_{\mathrm{A},xy}(0) = \frac{2}{\pi}\int_0^\infty \frac{1}{\omega}\sigma^{(2)}_{\mathrm{A},xy}(\omega)\,d\omega. \tag{43}$$

In the lower panel of Fig. 5 we show the cumulative AHC versus ω, defined as the contribution to the AHC from frequencies higher than ω in Eq. (43). In practice we use as the upper frequency limit in Eq. (43) the difference from the Fermi energy to the top of the inner energy window (18 eV). It is clear that the AHC is completely dominated by the low-frequency contributions below $\sim$ 5.5 eV.

It can be shown that applying the transformation (43) separately to the D–D term of the MCD spectrum yields the D–D term of the AHC, as defined in Ref. 6. This explains the intriguing result that more than 99% of the anomalous Hall conductivity can be recovered from the D–D term alone.[6] This is a consequence of (i) the low-frequency part of the spectrum being weighted more in the integral as a result of the $1/\omega$ factor in the integrand, and (ii) the D–D term overwhelming the other two at very low frequencies.

V. CONCLUSIONS

We have presented a Wannier-interpolation scheme to compute efficiently and accurately Fermi-surface and spectral properties from first principles. As an example of the former we computed the low-field Hall conductivity for several cubic metals. As an example of the latter we calculated the magnetic circular dichroism spectrum of bcc Fe.

The scheme naturally resolves a number of difficulties which have plagued existing interpolation schemes. Firstly, by preserving the information about band connectivity, band crossings and avoided crossings are treated correctly. In addition, the evaluation of the velocity matrix elements needed to compute both the Hall coefficient and the MCD spectrum can be done analytically in the Wannier representation. Furthermore, the scheme does not become any more complex upon inclusion of the spin-orbit interaction in the Hamiltonian. In particular, there are no additional contributions to the velocity matrix elements; all the spin-orbit-related corrections are contained in the spinor WFs. Also, the Wannier-interpolation scheme is decoupled from the particular choice of basis set used for performing the original *ab-initio* calculation, nor does it depend on the specific level of single-particle theory. As such, the calculation of a given property can be implemented in a universal way inside the Wannier module, which can then be interfaced with any desired electronic structure code.

The appeal of the present approach is that it combines the simplicity of a tight-binding-like scheme with the power and accuracy of *ab-initio* methods. Most importantly, *it allows operators other than the Hamiltonian to be interpolated in the same manner as the Slater-Koster interpolation of energy bands.* As such, it can be applied to a wide variety of problems in condensed matter physics. It should be particularly useful for studying metallic systems. A number of properties of metals remain extremely challenging to compute from first-principles, as a result of difficulties in sampling the Fermi surface with sufficient accuracy. Wannier interpolation provides an elegant and powerful framework for investigating such problems with *ab-initio* techniques.

VI. ACKNOWLEDGMENTS

This work was supported by the Laboratory Directed Research and Development Program of Lawrence Berkeley National Laboratory under the Department of Energy Contract No. DE-AC02-05CH11231, and by NSF Grant 0549198.

APPENDIX: CONVERGENCE PROPERTIES OF THE INTERPOLATION SCHEME

For a given operator $\hat{\mathcal{O}}$ the agreement inside the inner energy window between $\mathcal{O}^{(\mathrm{H})}_{nm}(\mathbf{k})$ obtained by Wannier interpolation and $\mathcal{O}_{nm}(\mathbf{k})$ calculated using a full first-principles calculation is determined by N_0, the number of points in the **q**-grid. The resulting WFs are periodic over the the conjugate real-space supercell spanning N_0 unit cells. For any finite N_0 there is some overlap between a WF and its neighboring periodic images, which affects the matrix $\mathcal{O}(\mathbf{R})$. It is generally accepted that WFs decay exponentially; numerical studies have confirmed this for several materials,[35] and recently there has been a claim of a formal proof for multiband time-reversal-invariant insulators.[36] The error in $\mathcal{O}(\mathbf{R})$, and therefore in the interpolation, should accordingly also decrease exponentially beyond some supercell size.

We report numerical tests for two cases: the isolated set of four valence bands in Pb, and the low-lying bands of Li, using the same WFs as in Secs. II C 5 and III, respectively. The band energies are computed via both Wannier interpolation and non-self-consistent diagonalization of the planewave Hamiltonian on a $200\times200\times200$ BZ grid. For Li we collect data from the bottom of the inner energy window to 0.5eV below the top of the inner energy window; points close to the top of the inner window may show larger discrepancies, as they result from

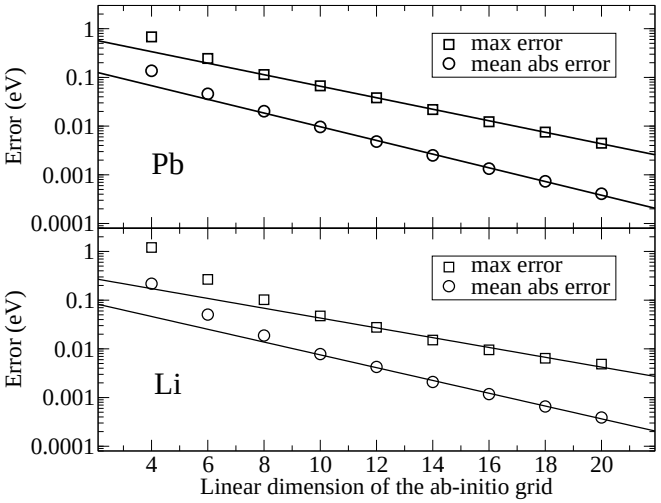

FIG. 6: Convergence of the Wannier interpolated band energies as a function of the linear dimensions $N_0^{1/3}$ of the *ab-initio* **q**-point grid. We plot the maximum error (squares) and mean absolute error (circles), where the error is the difference between the Wannier interpolated band energy and the value obtained from a full non-self consistent diagonalization of the planewave Hamiltonian. The lines are linear fits to the points with $N_0^{1/3} > 8$.

an interpolation between **q**-points inside and outside the inner window. Fig. 6 shows several measures of the difference in the energies as a function of $N_0^{1/3}$. In both cases we find that the error decreases exponentially for $N_0 \gtrsim 10$. It is particularly reassuring that this occurs in Li, since the decay properties of disentangled WFs has yet to be investigated thoroughly, and they probably fall outside the scope of existing formal proofs of exponential decay.

Finally, we examine the optimal choice of supercell in which to define the set of lattice vectors **R** for the Fourier transform in Eq. (15). To illustrate the discussion introduced earlier in the vicinity of Eq. (15), we compare the results for parallelepipedal and Wigner-Seitz supercells. In the upper part of Fig. 7 we compare the interpolated energy bands for a $4 \times 4 \times 4$ grid of **q**-points. For such a sparse **q**-grid the interpolated bands do not agree precisely with the exact *ab-initio* bands from a non-self-consistent diagonalization of the planewave Hamiltonian; this is most noticeable in the deviation of the curvature of the three upper bands between K and Γ. However, it is clear that the Wigner-Seitz supercell yields significantly better results than the parallelepipedal cell. This is most clear for the upper band from L to Γ, which displays large oscillations for the parallelepipedal cell. In the lower portion of Fig. 7, we show several measures of the error in the interpolated bands as a function of $N_0^{1/3}$; for any given **q**-grid the Wigner-Seitz cell gives the more accurate results. The superiority of the Wigner-Seitz choice can be easily understood, as it ensures the largest minimum distance between a WF and its periodic images.

[1] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge, 2004).
[2] G. Gilat, J. Comp. Phys. **10**, 432 (1972).
[3] F. Giustino, J. R. Yates, I. Souza, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. **98**, 047005 (2006).
[4] M. Colarieti-Tosti, S. I. Simak, R. Ahjuja, L. Nordstrom, O. Eriksson, and D. Aberg, Phys. Rev. Lett. **91**, 157201 (2003).
[5] Y. Ping, D. Hanson, I. Koslow, T. Ogitsu, D. Prendergast, E. Schwegler, G. Collins, and A. Ng, Phys. Rev. Lett. **96**, 255003 (2006).
[6] X. Wang, J. R. Yates, I. Souza, and D. Vanderbilt, Phys. Rev. B **74**, 195118 (2006).
[7] W. Schulz, P. Allen, and N. Trivedi, Phys. Rev. B **45**, 10886 (1992).
[8] I. I. Mazin, D. A. Papaconstantopoulos, and D. J. Singh, Phys. Rev. B **61**, 5223 (2000).
[9] W. E. Pickett, H. Krakauer, and P. B. Allen, Phys. Rev. B **38**, 2721 (1988).
[10] K. Uehara and J. Tse, Phys. Rev. B **61**, 1639 (2000).
[11] G. K. H. Madsen and D. J. Singh, Comput. Phys. Commun. **175**, 67 (2006).
[12] L. Chaput, P. Pecheur, J. Tobola, and H. Scherrer, Phys. Rev. B **72**, 085126 (2005).
[13] N. Marzari and D. Vanderbilt, Phys. Rev. B **65**, 12847 (1997).
[14] I. Souza, N. Marzari, and D. Vanderbilt, Phys. Rev. B **65**, 035109 (2002).
[15] Y. S. Lee, M. B. Nardelli, and N. Marzari, Phys. Rev. Lett. **95**, 076804 (2005).
[16] E. N. Blount, Solid State Phys. **12**, 305 (1962).
[17] M. Graf and P. Vogl, Phys. Rev. B **51**, 4940 (1995).
[18] C. J. Pickard and M. C. Payne, Phys. Rev. B **62**, 4383 (2000).
[19] T. Boykin, Phys. Rev. B **52**, 16317 (1995).
[20] D. A. Papaconstantopoulos, *Handbook of the Band Structure of Elemental Solids* (Plenum, New York, 1986).
[21] S. Baroni, A. D. Corso, S. de Gironcoli, P. Giannozzi, C. Cavazzoni, G. Ballabio, S. Scandolo, G. Chiarotti, P. Focher, A. Pasquarello, et al., `http://www.pwscf.org/`.
[22] A. A. Mostofi, J. R. Yates, N. Marzari, I. Souza, and D. Vanderbilt, `http://www.wannier.org/`.
[23] C.-L. Fu and K.-M. Ho, Phys. Rev. B **28**, 5480 (1983).
[24] M. Methfessel and A. T. Paxton, Phys. Rev. B **40**, 3616 (1989).
[25] P. E. Blöchl, O. Jepsen, and O. K. Andersen, Phys. Rev. B **49**, 16223 (1994).
[26] N. Marzari, D. Vanderbilt, and M. Payne, Phys. Rev. Lett. **79**, 1337 (1997).
[27] C. M. Hurd, *The Hall Effect in Metals and Alloys* (Plenum, New York, 1972).
[28] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[29] T. Beaulac and P. B. Allen, J. Phys. F **13**, 383 (1982).

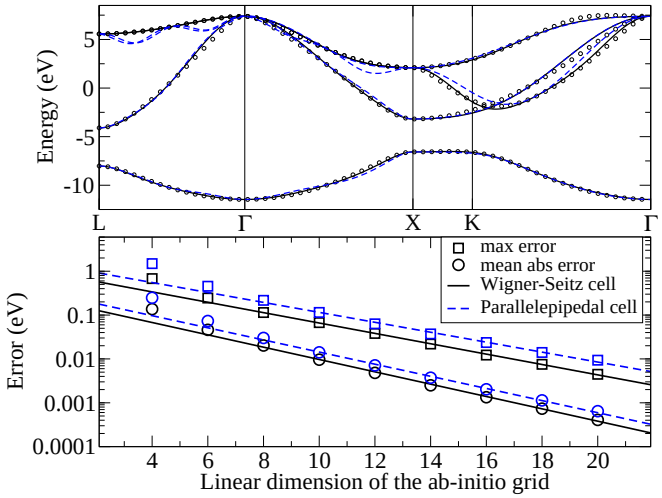

FIG. 7: (Color online.) Comparison of interpolated band-energies for Pb obtained using sets of lattice vectors defined within Wigner-Seitz (WS) and parallelepipedal (P) supercells. Upper figure: Energy bands interpolated using a $4 \times 4 \times 4$ **q**-point grid (WS cell – solid lines, P cell – dashed lines). Full *ab-initio* results are shown as open circles. Lower figure: Convergence of the Wannier-interpolated band energies for the two supercells as a function of the linear dimensions $N_0^{1/3}$ of the **q**-point grid. Details as in Fig. 6.

[30] H. Ebert, Rep. Prog. Phys. **59**, 1665 (1996).

[31] V. Antonov, B. Harmon, and A. Yaresko, *Electronic Structure and Magneto-Optical Properties of Solids* (Kluwer Academic, 2004).

[32] G. S. Krinchik and V. A. Artem'ev, Zh. Eksp. Teor. Fiz. **53**, 1901 (1967).

[33] Y. Yao, L. Kleinman, A. H. MacDonald, J. Sinova, T. Jungwirth, D. S. Wang, E. Wang, and Q. Niu, Phys. Rev. Lett. **92** (2004).

[34] A. D. Corso and A. Conte, Phys. Rev. B **71** (2005).

[35] L. He and D. Vanderbilt, Phys. Rev. Lett. **86**, 5341 (2001).

[36] C. Brouder, G. Panati, M. Calandra, C. Mourougane, and N. Marzari, Phys. Rev. Lett. **98** (2007).