Spin-orbit splitting of surface states is analyzed within and beyond the Rashba model using as examples the (111) surfaces of noble metals, Ag$_2$Bi surface alloy, and topological insulator Bi$_2$Se$_3$. The ab initio analysis of relativistic velocity proves the Rashba model to be fundamentally inapplicable to real crystals. The splitting is found to be primarily due to a spin-orbit induced in-plane modification of the wave function, namely to its effect on the nonrelativistic Hamiltonian. The usual Rashba splitting – given by charge distribution asymmetry – is an order of magnitude smaller.

I. INTRODUCTION

Spin structure of nonmagnetic surfaces has attracted much interest in the last decade due to the promising properties of spin-split surface states for spintronics applications. The idea of using the Rashba effect in the Datta-Das spin transistor as well as for spin filtering has stimulated great activity in the search of enhanced spin-orbit splitting at surfaces. Upon the discovery in 1996 of the splitting of the surfaces state on Au(111) much attention has turned to metal surfaces: spin-split surface states were promptly found on (110) surfaces of W and Mo covered with thin overlayers and on clean surfaces of Bi and Sb. The giant spin splitting found in Bi and Pb/Ag(111) surface alloys has inspired a search for a strong Rashba effect in heavy-element adsorbates on technologically more attractive (111) surfaces of Si and Ge. Recently, a giant Rashba splitting was observed in ternary semiconductors BiTeI, BiTeCl, BiTeBr, and BiTeBr$_3$. The way to practical spintronics depends on a deep understanding of fundamental mechanisms by which material properties determine the magnitude of the spin-orbit splitting. The early theoretical analyses of the splitting effect at the high-Z crystal surfaces emphasized its similarity to the Rashba effect in semiconductor heterojunctions, which is well understood within the two-dimensional electron gas model. The model is described by the Rashba-Bychkov (RB) Hamiltonian:

$$\hat{H}_{RB} = \hat{p}^2 / 2m^* + \alpha_R \sigma \cdot (\mathbf{n} \times \hat{p}_\parallel).$$ (1)

In application to semi-infinite crystals, the microscopic model behind Eq. (1) is obtained from the true two-component Hamiltonian (in Rydberg atomic units)

$$\hat{H} = \hat{p}^2 + V(r) + \beta \sigma \cdot [\nabla V(r) \times \hat{p}]$$ (2)

by neglecting the variation of the crystal potential $V(r)$ parallel to the surface, see Fig. 1(a). Thus, equation (1) assumes a free motion along the surface, and the material dependence of the spin-orbit splitting comes through the Rashba parameter $\alpha_R$. In the RB model, $\alpha_R$ has a clear physical meaning of the expectation value of the potential gradient in the surface normal direction $\mathbf{n}$ (along $y$ axis):

$$\alpha_R = \beta \langle \psi | \partial V / \partial y | \psi \rangle = \beta \int_{-\infty}^{+\infty} \frac{\partial V}{\partial y} \rho(y) \, dy,$$ (3)

with $\beta = \hbar / 4m^2c^2$. Owing to the perfectly parabolic dispersion $E(k \parallel)$ of the Au(111) surface state and to the nearly total spin polarization of the split branches, the Au(111) case is often considered a textbook example of Rashba effect in metals and a demonstration of the applicability of Eq. (1). This view is detailed in Refs. 39 and 40 since the spin-orbit interaction is significant only in a close vicinity of the nucleus, where the potential is spherically symmetric, the value of $\alpha_R$ is determined by the surface-perpendicular asymmetry of the probability density distribution $\rho(r) = |\psi(r)|^2$ at the nuclei. This idea, expressed by Eqs. (1) and (3), has been the basis for the analysis of individual atomic contributions to the spin splitting and polarization at real surfaces. In addition, in Refs. 29, 22, 41, 42 an important role is attributed to the in-plane inversion asymmetry and to the related surface-parallel potential gradient. In spite of their different view on the relative importance of surface-normal and in-plane asymmetry, all the studies have focused on the shape of $\rho(r)$ and how it may affect the $\nabla V$ expectation value, but so far no attempts have been made to express the asymmetry underlying the effect in energy units and to quantify different contributions to the spin-orbit splitting. The microscopic origin of the Rashba effect, i.e., its relation to the shape of the wave function in real crystals, thus remains an open question.

The aim of this work is to understand the physics behind the Rashba parameter in real crystals. An ab initio analysis of the Hamiltonian is performed to reveal the actual value of the RB term. A surprising result is that the pure Rashba contribution to the splitting—the one given by Eqs. (1) and (3)—typically yields only a few percent of the whole effect, while the major part arises from a relativistic modification of the wave function through its influence on the nonrelativistic energy operator.
II. MICROSCOPIC RASHBA MODEL

Before we turn to real crystals let us take a closer look at how the microscopic Rashba model works. To be specific about terminology, here by Rashba model is understood a system in which electrons move freely along the surface. In the geometry of Fig. 1(a), with the spin quantization axis along $z$ and with $k_\parallel$ along $x$, the spatial and spin variables in the Schrödinger equation separate, and for a given $k_\parallel$ the problem reduces to a pair of independent scalar equations for the two eigenfunctions $\psi_\uparrow(y)$ and $\psi_\downarrow(y)$:

$$-\psi''_\uparrow(y) + \hat{V}_\uparrow(y)\psi_\uparrow(y) = (E - k_\parallel^2)\psi_\uparrow(y).$$  \hspace{1cm} (4)

The effective potential $\hat{V}_\uparrow(y) = V(y) + \beta k_\parallel V'(y)$ depends upon spin ("\uparrow" for $\hat{V}_\uparrow$ and "\downarrow" for $\hat{V}_\downarrow$) and upon $k_\parallel$. For a bulk crystal with inversion symmetry, $V(y) = V(-y)$, the effective potentials are related as $\hat{V}_\uparrow(y) = \hat{V}_\downarrow(-y)$, as illustrated by Figs. 1(b) and 1(c) (bulk is to the left from the surface plane $y_S$). Equation (4) then yields the same band structure $E(k_\parallel)$ for both spins, for real and for complex $k_\parallel$ [in the spectral gap, Fig. 1(e)], which is known as Kramers degeneracy. The existence of a surface state in the gap depends on whether the regular solution of Eq. (4) for $y > y_S$ matches the bulk evanescent wave at the same energy in both value and slope,\textsuperscript{47,48} see Fig. 1(d). Being equivalent in an infinite crystal, the potentials $\hat{V}_\uparrow$ and $\hat{V}_\downarrow$ are different when looked at from the surface, see Fig. 1(c): $\hat{V}_\uparrow(y)$ is slightly shifted to the left and $\hat{V}_\downarrow(y)$ to the right relative to $V(y)$. Therefore, the evanescent waves of spin $\uparrow$ and spin $\downarrow$ are different at $y_S$, and the matching occurs at different energies for spin $\uparrow$ and spin $\downarrow$, which is known as Rashba splitting.

The dispersion lines $E(k_\parallel)$ calculated in the nonperturbative complex-band-structure approach\textsuperscript{49} are presented in Fig. 1(f) with full lines for the potential $V$ [shown in Fig. 1(b)] and with dashed lines for the same bulk potential but for twice larger surface barrier, $V_0 \to 2V_0$. Figure 1(g) shows the results for the potential $2V$ (i.e., both $V_0 \to 2V_0$ and $V_2 \to 2V_2$ for a larger "atomic number"). The spin-orbit effect on the wave function leads to a dependence of $E(k_\parallel)$ on $k_\parallel$, see Fig. 1(h), i.e., to a nonparabolicity of $E(k_\parallel)$. In the RB model the slope at $k_\parallel = 0$ (point $\Gamma$) equals zero\textsuperscript{25} so at finite $k_\parallel$ we define $2\alpha_R = d(E_\Gamma - E_j)/dk_\parallel$. More important and counterintuitive result is that the potential scaling does not always increase the splitting – it may even become smaller because at sufficiently small $k_\parallel$ $\alpha_R$ decreases with increasing the barrier height, see Figs. 1(f) and 1(h). Microscopically, this means that the effect of a larger potential gradient is completely cancelled by a modification of the wave function. From the complex band structure point of view this stems from the fact that in the $2V$ case [Fig. 1(g)] the surface state is pushed toward the gap edge, where Im $k_\parallel$ and, consequently, the wave function changes faster with energy, see Fig. 1(e).

The role of the surface is demonstrated by Fig. 1(i), which compares the splitting as a function of the surface barrier $V_0$ for a "clean" surface and for a surface with a "lighter adatom", see Fig. 1(k). For small $k_\parallel$ the splitting grows as $\alpha_R^2k_\parallel$, so Figs. 1(i) and 1(j) show the coefficient $\alpha_R$ determined by regression. [The bulk potential is the one of Fig. 1(b), so the two circles in the dashed curve in Fig. 1(i) correspond to the data in Fig. 1(f).] In both cases $\alpha_R$ rapidly grows at small $V_0$, but at larger barriers it keeps growing in the adatom case and decreases for the smooth barrier. This seemingly different behavior, however, looks rather similar when $\alpha_R^2$ is plotted as a function of the eigenenergy, see Fig. 1(j): in both cases $\alpha_R^2$ is seen to diminish when the energy is pushed to a gap edge, i.e., it shows the same trend as with scaling the potential, Figs. 1(f) and 1(g). According to Eq. (6), the peculiar dependence of $\alpha_R$ on the system parameters results from a complicated redistribution of the charge density, whereby different contributions to the potential
The fundamental difference between the spin-orbit splitting in real crystals and in the RB model is best seen at the $\Gamma$ point, where the relativistic effect manifests itself as a nonzero group velocity $dE/dk$, with the opposite sign for the two branches. The expression for $dE/dk$ follows from Eq. (2), cf. Eq. (3) in Ref. [5]

$$dE/dk = 2 \langle \psi | \hat{p}_\parallel | \psi \rangle + \beta \langle \psi | \sigma \cdot [\nabla \times \tau] | \psi \rangle,$$

where $\tau = k_j/k$ and $\hat{p}_\parallel = \hat{p} \cdot \tau$. In the RB model, the wave function is $\psi(r) = u(y) \exp(ik_j \cdot r_j)$, so the first (nonrelativistic) term vanishes at $k_j = 0$, and the splitting is solely due to the relativistic term of Eq. (5). This is not the case in real crystals: here it is the nonrelativistic term that gives the major contribution to the slope at $\Gamma$. This is illustrated in Fig. 2 by the examples of three well-studied hexagonal surfaces: the classical case of Au(111), the surface alloy Bi/Ag(111) exhibiting giant splitting, and the topological insulator Bi$_2$Se$_3$. In the upper row of Fig. 2 the numerical derivative of the eigenenergy $dE(k_j)/dk_j$ (total velocity) is compared with the nonrelativistic (classical) velocity $v_c = 2 \langle \psi | \hat{p}_\parallel | \psi \rangle$.

The wave functions $\psi$ were calculated with the Hamiltonian of Eq. (2) with linear augmented plane wave method in a repeated slab geometry. The density-functional-theory calculations for symmetric slabs (19 layers for noble metals 9 for Bi/Ag(111) 25 and 25 for Bi$_2$Se$_3$ 26), and a small perturbation at one of the surfaces was introduced afterwards to disentangle the surface states at the opposite surfaces. Relativistic effects are treated within the second-variational two-component Koelling-Harmon approximation. The potential gradient is taken into account only in the muffin-tin spheres, where the spin-orbit term takes the form $\beta[V(r)]/r |\sigma \cdot \mathbf{L}|$. Its contribution to the group velocity is then given by the expectation value $v_r = \langle \psi | \beta \sigma \cdot [r \times \tau] V(r)/r | \psi \rangle$.

The velocities calculated by this formula are shown in the middle row of Fig. 2. The main message is that at $\Gamma$ the spin-orbit part is an order of magnitude smaller than the full velocity (circles in the upper row). This result is corroborated by the calculation of the classical velocity $v_r$ (solid lines in the upper row), which is seen to be very close to the full velocity (found by numerical differentiation). Clearly, the splitting here follows a strikingly different scenario from the one of the standard Rashba model: crucial is the influence of the relativistic term on the wave function, which then gives rise to the strong splitting through the purely nonrelativistic velocity operator. In other words, a perturbation theory based on non- or scalar-relativistic wave functions, which in the Rashba model is exact at $\Gamma$, is here manifestly inapplicable. The parameter $\alpha_R$ that can be derived by fitting experimental or ab initio energies to Eq. (4), is, thus, not the measure of either normal or in-plane gradient (or the respective charge density asymmetry). Moreover, the true Rashba parameters in the three crystals (see the middle row of Fig. 2) differ much less than the full slope at $\Gamma$ (upper row).

The dependence of the velocity ingredients on the parameters in the three crystals (see the middle row of Fig. 2) differ much less than the full slope at $\Gamma$ (upper row).

The dependence of the velocity ingredients on the parameters in the three crystals (see the middle row of Fig. 2) differ much less than the full slope at $\Gamma$ (upper row).

In order to understand how the spin-orbit term modifies the wave function, let us consider the classical current density distribution in relativistic and scalar...
FIG. 3. (Color online) In-plane distribution of the $k_\parallel$ projection of the classical current density $j(r_\parallel)$ (in the units of Ry-a.u.) for the surface state at Bi/Ag(111) with $k_\parallel$ along the horizontal axis. (a) relativistic $j(r_\parallel)$ map for $k_\parallel = 10^{-4}$ Å$^{-1}$ for the inner-circle surface state, i.e., for the one with the smaller $k_\parallel$ at a given $E$, hence the negative group velocity. (b) scalar relativistic map for $k_\parallel = 10^{-3}$ Å$^{-1}$.

The above current-density considerations offer a rather transparent picture of how the spin-orbit interaction affects the surface state: its principal role is to modify the wave function in such a way that the function $j(r_\parallel)$ acquire a large amplitude at $\bar{\Gamma}$ (which in the scalar relativistic case would happen at large $k_\parallel$). To put it most concisely, the spin-orbit term shifts the distribution $j(k_\parallel; r_\parallel)$ along $k_\parallel$. This relativistic effect is omitted in the Rashba model because there it is $j(r_\parallel) = 0$.

TABLE I. Group velocity at $\bar{\Gamma}$ for the surface states at some (111) surfaces. Values obtained by numerical differentiation (upper row) are compared to nonrelativistic and spin-orbit parts obtained from the wave functions. Identity Eq. (3) is satisfied only with a certain accuracy due to the variational character of the wave functions.

| Material   | Cu  | Ag  | Au  | Ag$_2$Bi | Bi$_2$Se$_3$ |
|------------|-----|-----|-----|----------|--------------|
| $dE(k_\parallel)/dk_\parallel$ | 0.02 | 0.01 | 0.10 | 0.32 | 0.26 |
| nonrelativistic $v_c$ | 0.02 | 0.01 | 0.12 | 0.32 | 0.29 |
| relativistic $v_r$ | 0.002 | 0.006 | 0.014 | 0.022 | 0.010 |

IV. CONCLUSIONS

We can conclude that the relativistic splitting of surface states has the same microscopic nature in all the materials, regardless of its size or topology (trivial surface states in the metals or topologically protected Dirac cone in Bi$_2$Se$_3$). In particular, Au(111) turns out to be not a paradigm case of an RB system, which explains the recently encountered “departure from the Rashba model” in the spin-flip excitations at Au(111). Thus, neither the parabolicity of $E(k_\parallel)$ nor the full in-plane spin polarization of the states are an indication of the RB scenario. The relativistic wave function modification revealed by Figs. 2 and 3 has important implications in a wide range of scattering and excitation processes: this property should be taken into account in modeling the inelastic scattering and lifetime effects and especially electron and spin transport by the surface states, which is directly connected to the current density distribution. Furthermore, the spatial structure of the wave function reflects itself in angle-resolved photoemission: the component of the electric field of the incident light along $k_\parallel$ emphasizes the relevant feature of the initial state via the matrix element of the classical velocity operator $\langle \Phi_{\text{initial}} | p_\parallel | \psi \rangle$. According to the present theory the $k_\parallel$ dependence of the matrix element would be different for the splitting of the RB type and of the present type.

To summarize, in order to manipulate the spin-orbit splitting it is necessary to know how it arises. The present study has established that in crystals built of high-Z atoms the splitting comes primarily from a relativistic modification of the wave function that makes the classical current carried by the surface state finite at $\bar{\Gamma}$. This spin-orbit induced transformation can be described as a shift of the current density distribution along $k_\parallel$. This is by contrast to the Rashba model, in which the spin-orbit interaction does not modify the in-plane structure of the wave function. The discovered scenario is followed equally closely by noble metals, surface alloys, and topological insulators irrespective of the strength of the effect and topological nature.

ACKNOWLEDGMENTS

The author is grateful to R. Kuzian, I. Nechaev, and I. Tokatly for the critical reading of this manuscript and many useful discussions. Partial support is acknowledged from the Spanish Ministerio de Ciencia e Innovació (Grant No. FIS2010-19609-C02-02).
This follows from the Ehrenfest theorem, as discussed in Chapter 6.2.2 in R. Winkler, *Spinorbit coupling effects in two-dimensional electron and hole systems* (Springer, 2003).

T. Oguchi and T. Shishidou, J. Phys.: Condens. Matter 21, 092001 (2009)

The self-consistent potential is obtained in local density approximation with the full-potential linear augmented plane waves method, see E. E. Krasovskii, F. Starrost, and W. Schattke, Phys. Rev. B 59, 10504 (1999).

Properties of the surface states in noble metals are addressed in detail in J. Lobo-Checa, J. E. Ortega, A. Mascaraque, E. G. Michel, and E. E. Krasovskii, Phys. Rev. B 84, 245419 (2011); P. Borghetti, J. Lobo-Checa, E. Goiri, A. Mugarza, F. Schiller, J. E. Ortega, and E. E. Krasovskii, J. Phys.: Condens. Matter 24, 395006 (2012).

For the crystal structure of the Ag₂Bi surface alloy and the properties of the surface state see, e.g., Refs. 21–24.

For the electronic structure of Bi₂Se₃ see, e.g., S. Kim, M. Ye, K. Kuroda, Y. Yamada, E. E. Krasovskii, E. V. Chulkov, K. Miyamoto, M. Nakatake, T. Okuda, Y. Ueda, K. Shimada, H. Namatame, M. Taniguchi, and A. Kimura, Phys. Rev. Lett. 107, 056803 (2011).

D. D. Koelling and B. N. Harmon, J. Phys. C 10, 3107 (1977).

A. H. MacDonald, W. E. Picket, and D. D. Koelling, J. Phys. C 13, 2675 (1980).

The uncertainty in $v_c$ as large as 0.1 Ry·a.u. has been reported, e.g., for bulk states in Be in E. E. Krasovskii, V. V. Nemoshkalenko, and V. N. Antonov, Z. Phys. B 91 (1993).

J. Ibañez Azpiroz, A. Bergara, E. Y. Sherman, and A. Eiguren, Phys. Rev. B 88, 125404 (2013).

I. A. Nechaev, P. M. Echenique, and E. V. Chulkov, Phys. Rev. B 81, 195112 (2010).