Electronic Wavefunction in Empirical Tight-Binding Theory

R. Benchamekh,1,2 F. Raouafi,2 J. Even,3 F. Ben Cheikh Larbi,2 P. Voisin,1 and J-M. Jancu3

1CNRS-Laboratoire de Photonique et de Nanostructures, route de Nozay, F-91460, Marcoussis, France
2Laboratoire de Physico-chimie des Matériaux Polymères, Institut Préparatoire aux Études Scientifiques et Techniques, BP51, 2070 La Marsa, Tunisia
3FOTON, Université Européenne de Bretagne, INSA-Rennes and CNRS, Rennes, France

(Dated: September 24, 2013)

A procedure to obtain single-electron wavefunctions within the tight-binding formalism is proposed. It is based on linear combinations of Slater-type orbitals whose screening coefficients are extracted from the optical matrix elements of the tight-binding Hamiltonian. Bloch functions obtained for zinc-blende semiconductors in the extended-basis spds* tight-binding model demonstrate very good agreement with first-principles wavefunctions. We apply this method to the calculation of electron-hole exchange interaction, and obtain the dispersion of excitonic fine structure of bulk GaAs. Beyond semiconductor nanostructures, this work is a fundamental step toward modeling many-body effects from post-processing single particle wavefunctions within the tight-binding theory.

PACS numbers:

Tight-binding is widely used as a conceptual frame to account for the kinetic energy operator in solid state theory. It served as a basis for major contributions such as Anderson’s strong localization[1] and magnetic impurity[2] theories, Hubbard’s model of interacting electrons[3], and many others[4]. In these theories, interactions are generally introduced using symmetry considerations and ad-hoc parameters, and the goal is to have the simplest model integrating the physically relevant features of hopping and interaction integrals while getting rid of the complexities of underlying atomic physics. On the other hand, empirical-parameter tight-binding (EPTB) is known as a powerful modeling technique for the electronic structure of semiconductors, metals, and all kind of nanoscale structures and devices. The systematic procedure for constructing the EPTB Hamiltonian was discussed in a seminal paper by Slater and Koster (SK) in 1954[5], but it is only many years later that computers have allowed systematic implementations[6,7]. A major step was achieved in the late 1990’s with the development of EPTB models using an extended spds* orbital basis[8,9] and allowing accurate full-band representation of single particle states. However, when it comes to calculating short-range interactions between quasi-particles, EPTB models (like the k.p theory itself) are hampered by the lack of an explicit orbital basis. This is a strong limitation for the use of advanced EPTB schemes to explore highly topical problems of strongly correlated electron systems. In this Letter, we present a method that reconciles the “conceptual frame” and “modeling tool” faces of tight-binding, by self-consistently determining the orbital basis out of the EPTB hamiltonian. We thus obtain the local wavefunctions, which allows parameter-free calculation of short-range interactions. We illustrate this by calculating electron-hole exchange and the fine structure of excitons in GaAs. Beyond semiconductors, this method can be used for many different materials and can handle million-atom supercells that are still out of reach of first-principle methods.

In the SK formalism, the crystal potential is approximated as a sum of spherically symmetrical potentials around each atom. This allows the electronic wavefunctions to be developed on a set of atomic-like orbitals (called the Löwdin orbitals) having well defined angular properties, but unknown radial dependencies. The Hamiltonian matrix elements between Löwdin orbitals are treated as “disposable constants” with which one can fit band structures that have been experimentally determined or calculated using more accurate techniques. The fit is performed in k-space, removing any necessity to further characterize the local wavefunctions in real space. Very interestingly, interaction with the electromagnetic field is built-in without any ambiguity from a point of view of the gauge invariance using a derivation of the Hamiltonian matrix elements in momentum space: $p_{lm,l'm'} = i\hbar \langle \Psi_{lm} | \nabla_k H | \Psi_{l'm'} \rangle$[10]. Optical properties can be consequently calculated in EPTB models without adding parameters, and good accuracy is obtained providing that the orbital basis is rich enough[11]. In this context, the spds* TB model is known to give a description of dielectric properties equivalent to best ab initio calculations within the one-electron approximation[12]. Altogether, the model major qualities are the transferability of parameters from bulk materials to nanostructures, the unique ability to describe with the same accuracy electronic properties in any region of the Brillouin zone, and the capacity to handle large supercells. However, interactions (in particular, short range interactions) between quasi-particles involve the local wavefunctions, and can not be calculated in the present frame. Existing calculations of Coulomb matrix elements use severe approximations on the radial dependence of the basis orbitals[13]. In order to solve this theoretical issue that has remained open ever since
the seminal work of Slater and Koster, we start with a set $\mathcal{B}$ of $spds^*$ basis functions in the form of normalized Slater-type orbitals (STO) $\Phi_{\alpha nlm}(r) = \sqrt{(2\alpha)^{2n+1}/(2\pi)^n}Y_{nlm}(\theta, \phi)r^{n-1}e^{-\alpha r}$, where $n$ is the first quantum number and $\alpha$ is a screening parameter [14]. STOs are largely employed in quantum chemistry but do not fulfill the orthogonality condition, since finite overlap exists between two STOs localized at different sites of the crystal. Firstly, the orbital overlap matrix $S$ is calculated including all orbitals up to a cut-off distance $R_0$ that must be taken large enough so that overlap with remote atoms be negligible. This step is done analytically, and in practice, we found that for $\alpha = 0.5$, overlap with neighbors located farther than 3 lattice parameters (17Å) can safely be neglected. Note that, unlike real atomic orbitals, $s$ and $s^*$ on-site STOs are not orthogonal. The most physical solution to this small difficulty consists in substituting $s^*$ with a Gramm-Schmidt combination $\tilde{s}^* = (s^* - s \cdot s^*)s/\sqrt{1 - (s \cdot s^*)^2}$. Then, an orthogonal basis $\mathcal{B}_{\text{orth}}$ can be obtained by using the Löwdin orthogonalization procedure $\mathcal{B}_{\text{orth}} = S^{-1}B$ [13]. The orthogonalized STOs will serve as trial functions for the unknown Löwdin orbitals. The expansion of electronic eigenfunctions (Bloch functions) in the basis $\mathcal{B}$ is obtained by multiplying the eigenvectors of the $spds^*$ Hamiltonian matrix by the matrix $S^2$, which provides definitely their representation in real space. In practice however, the 9150 x 9150 $S$-matrix obtained for $R_0 = 17\AA$ contains considerably redundant information and its large size makes its inversion computationally difficult. Since we are in-fine interested in the crystal eigenstates (Bloch functions), instead of the procedure sketched above, we equivalently construct, orthogonalize and invert a 40 x 40 $\tilde{S}$ overlap matrix between Bloch sums of STOs, truncated to the cut-off distance $R_0$. Although this may be not obvious at first glance, $S$ and $\tilde{S}$ overlap matrices contain the same information and produce identical results in the large $R_0$ limit. Then the momentum matrix elements are calculated in real space from the Bloch functions by the relation $p_{ilm,l'm'} = \hbar\langle \Psi_{ilm} | \nabla r | \Psi_{l'm'} \rangle$. This derivation involves a sum of matrix elements between two STOs that are calculated analytically [16]. Finally the screening parameters are fitted into a genetic algorithm until the optical matrix elements calculated in real space compare satisfactorily with those derived in k-space from the electronic Hamiltonian. The consideration of different bands and different high symmetry points in the Brillouin zone provides more than necessary information for the fit convergence. However, we observe that somewhat different sets of screening parameters can give similar “fitness” parameters. This problem is linked with the fact that upper-band dispersions are necessarily incorrect, due to the non-completeness of finite basis. Hence, optical matrix elements involving $s^*$ and d orbitals are poorly represented in k-space, and using them to fit screening parameters can lead to unphysical optimum in the algorithm. This is a fundamental difficulty of the present method. It is actually cured by constraining the parameter space in such a way that the $d(\Gamma_{12})$ orbital, which is close to a free-electron state, agrees with independent empirical pseudo-potential or ab initio calculations.

We applied the procedure explained above to the prototype systems of Ge and GaAs [17]. The electron configuration of Ge is [Ar]$3d^{10}4s^24p^2$. In the $spds^*$ model, the deep 3d states are discarded and the basis is formed by the orbitals 4s, the three orbitals 4p, the five empty orbitals 4d and the empty orbitals 5s. When building the STO basis $\mathcal{B}$, we keep fixed the first quantum number $n$ of these orbitals and introduce one adjustable screening parameter $\alpha$ for each symmetry type. Alternatively, as often done in quantum chemistry, we can improve parametrical flexibility by considering that each element of the starting basis is a linear combination of $q$ STOs instead of one. This does not change much the model, but increases to 4$q$ the number of fitted parameters. For GaAs, since there are two different atoms in the unit cell, the number of parameters is twice that for Ge. The fitted screening parameter for Ga and As are given in Table I, and contrasted with the Slater atomic screening constants. Those obtained independently for Ge are close to averaged values for Ga and As. At the end of the fitting procedure, the global discrepancy on the sum of all interband matrix elements, calculated at the $\Gamma$, X and L points of the Brillouin zone, is less than 15% with one Slater orbital per atomic state and less than 7% with a linear combination of two Slater orbitals for each atomic state. By changing the relative weights of different spectral or Brillouin zone regions in the genetic algorithm fitness function, the discrepancy at e.g. $\Gamma$ can be minimized down to the percent range. The residual discrepancy has three distinct physical origins: i) the difference between orthogonalized STOs and the actual Löwdin orbitals, ii) the lack of completeness of the spds$^*$ basis, and iii) the missing intra-atomic contribution in the k-space derivation method. Table II shows the main momentum elements $P_0 = -i\langle s_z | p_x | x_0 \rangle$, $P_1 = -i\langle s_z | p_x | x_1 \rangle$.
and $Q_0 \equiv -i\langle x_e|p_y|z_v \rangle$ obtained for Ge and GaAs at the $\Gamma$ point. Agreement is very good, but a discrepancy observed for the weak matrix element $P_3$, for which real space calculations agree, but differ from the hamiltonian derivation value $^{15}$. This might be a trace of the methodological limitation relative to intra-atomic contributions. While further work is required to explore the method limitations and improve the results, the present achievement is already sufficient for most practical purposes.

Once screening parameters best reproducing the interband matrix elements are obtained, the different Bloch functions can be plotted and compared with ab initio calculations. The latter were performed using the ABINIT code $^{19, 20}$ in the local density approximation (LDA), completed by self-consistent GW correction. Figures $^{12}$ show (110) plane isodensity contours of wavefunctions in bulk GaAs at the Brillouin zone center. Fig. 1 shows valence band states $s_e$ and $y'_e \equiv y_e - x_v$, in both TB and ab initio calculations. The overall quantitative agreement is very good, since the overlap between TB and ab initio densities is always better than 95%. Yet, TB wavefunctions appear somewhat less localized in the sense that they have larger density in regions where the $s_c$ density is almost zero, and the Ga / As asymmetry is more pronounced in the ABINIT result. The most significant difference is for the deep $s_c$ state near the atomic sites, for which TB density is significantly smaller. This probably reflects the difference of projection basis between the two models: TB wavefunctions are expanded in the basis of Slater orbitals which have a node at the atomic sites while the ABINIT wavefunctions are expanded in a basis of plane-wave functions that may be maximum on the atomic sites. The wavevector cut-off used in the ABINIT calculations is important, because this approach cannot describe the region located less than $1/k_{cut-off}$ from atomic sites. Yet, cut-off does not suffice to explain the observed difference. In our TB approach, finite on-site value for the $s_c$ state results from the contribution of neighboring atoms and is quite sensitive to the STO screening parameters.

In Fig. 2 we show the conduction Bloch functions also calculated with the same two models. Again TB wavefunctions are very similar to those calculated in the LDA+GW approximation, except for a significant difference for the $s_c$ state density in the vicinity of atomic sites. In order to clarify this issue, we used the SIESTA code, which is based on DFT expanded in a strictly localized orbitals set. SIESTA results for $s_c$ actually agree very well with our TB results$^{17}$. We note that electron hyperfine interaction constants, that are well documented, scale as $s_c^2(r = 0)$ and could serve as a quantitative test. A most striking result is the TB ability to reproduce the wavefunctions of the nearly free electron states $s^*$ and $d$.

The major interest of having a real space representation of wavefunctions is the ability to study many-body problems. In the following, we illustrate the potential of our wavefunction derivation method by calculating the exciton binding energy and its fine structure due to electron-hole exchange interaction. Coulomb interaction can be specified by the matrix elements $\langle m', k'; n', k'_h|U^{eh}|m, k; n, k_h \rangle$, where $U^{eh} = e^2/\kappa|r_e - r_h|$. Here, $|m, k; n, k_h \rangle$ is the two-particle excited state, $\kappa$ is the permittivity, and $\Psi_{m,k} \equiv \langle r|m, k_c \rangle$ and $\Psi_{n,k_h} \equiv \langle r|n, k_h \rangle$ are the Bloch wave functions in electron and hole representations, respectively $^{21, 22}$. The present expansion of Bloch functions as linear combinations of Slater orbitals allows to expand the electron-hole interaction in terms of Coulomb matrix elements between STOs $V_{l_1m_1l_2m_2l_3m_3l_4m_4} = \langle \Phi_{l_1m_1}(r - R_1), \Phi_{l_2m_2}(r - R_2)|U^{eh}|\Phi_{l_3m_3}(r - R_3), \Phi_{l_4m_4}(r - R_4) \rangle$. Restricting the expansion to two-center contributions ($R_1 = R_3$ and $R_2 = R_4$), the evaluation of the integrals can be done quasi analytically using the expansion of the Coulomb potential in terms of spherical harmonics centered on the same site when $R_1 = R_2$ $^{23}$, and a

TABLE II: Main interband matrix elements (in eV Å) for Ge and GaAs at the $\Gamma$ point, calculated from different models

|         | WF1$^a$ | WF2$^b$ | LDA$^c$ | Hamiltonian$^d$ |
|---------|---------|---------|---------|-----------------|
| Ge      |         |         |         |                 |
| $P_0$   | 7.69    | 10.18   | 8.49    | 10.14           |
| $Q_0$   | 8.29    | 8.42    | 7.32    | 8.70            |
| GaAs    |         |         |         |                 |
| $P_3$   | 0.80    | 0.93    | 1.38    | 0.11            |
| $Q_0$   | 8.16    | 8.31    | 7.37    | 8.72            |

$^{a,b}$real space calculation from TB wavefunctions with respectively one and two Slater orbital for each basis element

$^c$real space calculation from LDA wavefunctions, ABINIT code

$^d$calculated from Hamiltonian derivation

FIG. 1: Isodensity contours of the S and Y' = X+Y valence Bloch function at the zone center in bulk GaAs in the plane (110). TB calculation (left) is compared with ABINIT calculations (right).
bipolar expansion when \( \mathbf{R}_1 \neq \mathbf{R}_2 \). Following [24], we introduce a \( r \)-dependent dielectric constant such that short-range (on-site) interaction is unscreened, while long range interaction is subject to standard dielectric screening. Then, we solve the Bethe-Salpeter equation (BSE), expressed in terms of calculated electron-hole interactions and energies that are obtained from TB calculation. The BSE is an eigenvalue problem of infinite dimensionality:

\[
(E_{e,k} + Q/2 - E_{e,k} - Q/2) A_{veck} + \int_{V_{BZ}} d^3k \sum_{\nu,\nu'} \langle \nu' e k | U | \nu e' k' \rangle A_{\nu' e' k'} = \Omega_S A_{veck}
\]

Where \( E_{e,k} + Q/2 \) and \( E_{e,k} - Q/2 \) are the electron and hole energies respectively. Resolution of BSE gives the exciton wavefunction components \( A_{veck} \) and the excitation energies \( \Omega_S \). To make the problem tractable continuous integration with respect to \( k' \) was replaced by a discrete scheme. Following [27], [28], to calculate the exciton spectra and binding energy at the \( \Gamma \) point, the integration was performed over a small region near the position of the band extrema \( |k| < 0.015 \text{ a.u.} \). This region was divided into a \( 11 \times 11 \times 11 \) uniform grid. For an exciton wave vector \( Q = 0 \), we find an excitonic binding energy \( E_b = 4.75 \text{meV} \). In addition, the eightfold degenerate \( \Gamma_{8e} \times \Gamma_{6c} \) fundamental exciton transition is split by short range exchange interaction into one twofold, and two threefold degenerate excitons. The twofold and threefold \( J = 2 \) “dark excitons” are split by \( \delta_{\text{anis}} = 0.02 \mu\text{eV} \). This anisotropy splitting is due to the zinc-blend structure which does not allow more than threefold degeneracy. We expectedly find a very small value for \( \delta_{\text{anis}} \). The \( J = 1 \) “bright exciton” threefold state is separated from the \( J = 2 \) states by the short range exchange splitting \( \Delta_{\text{exc}} \). We get \( \Delta_{\text{exc}} = 20.6 \mu\text{eV} \), in agreement with recent experimental determination [29]. When one moves away slightly from \( Q = 0 \), the \( J = 1 \) excitons are further split by the long range exchange interaction into twofold degenerate, optically active transverse excitons, and a longitudinal exciton. The energy difference corresponds to the longitudinal-transverse splitting \( \Delta_{\text{LT}} \), for which we find a value \( \Delta_{\text{LT}} = 105.3 \mu\text{eV} \) in very good agreement with the well documented experimental value.

Then, we examine the evolution of the exciton fine structure in function of the exciton wavevector \( Q \). Fig. 3 shows the calculated dispersion curves. For large \( Q \), when the heavy hole light hole splitting becomes larger than exciton binding energy, the exciton splits into a “heavy” exciton formed of two twofold degenerate states and a “light” exciton formed of one threefold degenerate state and one singlet state. Our calculation shows how energy levels interpolate between the small and large \( Q \) regimes. Finally, we note that when \( Q \) is along the [110] direction, our results show the full details of exciton state spin splittings, including both contributions of electron and hole spin splittings.
In conclusion, we have devised a method that allows self-consistent definition of local wavefunctions within the EPTB theory, and successfully used bulk exciton fine-structure as a parameter-free testbed. Extension to nanostructures is straightforward as long as bulk screening parameters are, like other tight binding parameters, transferable to nanostructures. While some fundamental aspects of the method, like the problem of on-site optical matrix elements, still require further clarification, this approach opens a route towards reconciling tight-binding and predictive evaluation of interactions between quasi-particles.

[1] P. W. Anderson, Phys. Rev. 109, 1492 (1958).
[2] P. W. Anderson, Phys. Rev. 124, 41 (1961).
[3] J. Hubbard, Royal Society of London Proceedings Series A 276, 238 (1963).
[4] Not less than 13 theoretical papers that appeared in PRL during the first half of 2013 contain tight binding in their abstract.
[5] J.-C. Slater and G. F. Koster, Phys. Rev. 94, 1498 (1954).
[6] P. Vogl, H. P. Hjalmarson, and J. D. Dow, Journal of Physics and Chemistry of Solids 44, 365 (1983).
[7] J. N. Schulman and T. C. McGill, Phys. Rev. Lett. 39, 1680 (1977).
[8] J.-M. Jancu, R. Scholz, F. Beltram, and F. Bassani, Phys. Rev. B 57, 6493 (1998).
[9] C. Barreteau, D. Spanjaard, and M. C. Desjonquères, Phys. Rev. B 58, 9721 (1998).
[10] T. B. Boykin and P. Vogl, Phys. Rev. B 65, 035202 (2001).
[11] B. A. Foreman, Phys. Rev. B 66, 165212 (2002).
[12] One should however mention that the question of on-site transition matrix elements, not included in this approach, is still debated within the TB community.
[13] S. Schulz, S. Schumacher, and G. Czycholl, Phys. Rev. B 73, 245327 (2006).
[14] J.-C. Slater, Phys. Rev. 36, 57 (1930).
[15] P.-O. Löwdin, J. Chem. Phys. 18, 365 (1950).
[16] Note that the orthogonalization procedure induces intratomic matrix elements of the momentum operator, compatible with the gauge invariance as discussed by Sandu, Phys. Rev. B 72, 125105 (2005). This, however, does not solve completely the issue of on-site optical matrix elements.
[17] See supplementary information.
[18] Note that the “Hamiltonian derivation” value of $P_1$ also differs from that ($\tilde{P}_1 = 0.41eV$) obtained when fitting a 14-band k.p model to tight binding dispersion. $\tilde{P}_1$ corresponds to a renormalized value acting in a restricted basis see J.-M. Jancu et al., Phys. Rev. B 72, 193201 (2005).
[19] X. Gonze et al., Computer Phys. Commun. 180, 2582 (2009).
[20] X. Gonze et al., Zeit. Kristallogr. 220, 552 (2005).
[21] G. L. Bir and G. E. Pikus, Symmetry and Strain-Induced Effects in Semiconductors (Wiley, New York, 1975).
[22] S. Goupalov and E. Ivchenko, Physics of the Solid State 43, 1867 (2001).
[23] J. D. Jackson, Classical Electrodynamics (John Wiley, New York, 1975).
[24] R. Nozawa, J. Math. Phys. 7, 1841 (1966).
[25] R. Resta, Phys. Rev. B 16, 2717 (1977).
[26] H. Fu, L.-W. Wang, and A. Zunger, Phys. Rev. B 59, 5568 (1999).
[27] M. Rohlfing and S. G. Louie, Phys. Rev. Lett. 81, 2312 (1998).
[28] M. Rohlfing and S. G. Louie, Phys. Rev. B 62, 4927 (2000).
[29] W. Ekardt, K. Losch, and D. Bimberg, Phys. Rev. B 20, 3303 (1979).