Multiband Hamiltonians of the Luttinger-Kohn Theory and Ellipticity Requirements

Dmytro Sytnyk, Sunil Patil and Roderick Melnik

M2NeT Laboratory, Wilfrid Laurier University, 75 University Avenue West, Waterloo, ON, Canada, N2L 3C5.

Abstract: Modern applications require a robust and theoretically strong tool for the realistic modeling of electronic states in low dimensional nanostructures. This theory allows us to get theoretical insight into the electronic properties of dominating bands near the extremum point. Furthermore, the theory establishes a robust computational framework for simulating observable quantum-mechanical states and corresponding energies in the low-dimensional systems including quantum wells, wires, nanodots. In the original Luttinger–Kohn work, authors applied the theory to the Schrödinger equation perturbed by a smooth potential and constructed a representation for valence bands Hamiltonian near high symmetry point Γ of the first Brillouin zone in bulk Zinc-Blende (ZB) crystals with large fundamental band gap. Soon after that, Kane showed how to expand the model to the narrow gap materials such as InSb and Ge for instance, where one can also account for the influence of the conduction band. One of the advantages of the \( k \cdot p \) theory is in its universality. Indeed, the theory had also been extended to cover Wurtzite (WZ) type of crystals, materials with inclusions, heterostructure materials and superlattices. Another advantage of the effective mass theory that has recently been explored in some details is its flexibility, as one can easily adjust the models to include additional effects like strain, piezoelectricity, magnetic field, and respective nonlinear effects. These inbuilt multi-scale effects are crucial for such applications as light-emission diodes, lasers, high precision sensors, photo-galvanic elements, hybrid bio-nanodevices, and many others.

For a wide range of applications these models have provided good, computationally feasible and efficient approximations that agree well with experimental results. However, for some types of crystal materials band structure calculations based on such multiband models lead to the solutions with unphysical properties or so called spurious solutions. As a result, there have been various attempts to explain the origin of the spurious solutions and develop some reliable procedures on how to avoid them. These approaches rely on three main ideas: (a) to modify the original Hamiltonian and remove the terms responsible for the spurious solutions, (b) to change bandstructure parameters, and (c) to identify and exclude physically inadequate observable states. All mentioned approaches suffer from the common weakness – the lack of clear justification of the underlying theoretical procedure and thus from limitations in their applicability.

In this work we show that spurious solutions are not just a reason but rather an imminent consequence of another fundamental problem in applications of the classical theory – the non-ellipticity of the multiband Hamiltonian. The widely adopted effective mass approximations of the original Schrödinger elliptic Hamiltonian turns out to be non-elliptic for a broad class of known material parameters (cf. Table). This fact leads to the following consequence: since any qualitative approximation methods must preserve the topological structure of the spectrum of a general linear operator, and as an implication symmetric properties in case of Schrödinger Hamiltonian, this evident discrepancy signifies the mathematical invalidity of performed approximation procedures for those materials. Among these multiband Hamiltonians there are a number of widely used \( 6 \times 6 \) and \( 8 \times 8 \) Hamiltonians.

The paper is organized as follows. First, we revise basic properties of the original Schrödinger equation, and outline the mathematical model of the electronic bandstructure problem. Next, we derive the exact constraints on the material parameters for typical \( 6 \times 6 \) Hamiltonians in ZB and WZ materials. Direct calculations using conventional Luttinger parameters (e.g. Table) show that such parameters (e.g. Table) for many important semiconductor materials entail the violation of ellipticity requirements. As a result, the corresponding bandstructure model for them potentially susceptible to unphysical solutions, even in the bulk case, and therefore ought to be modified. Moreover, the corresponding time dependent Schrödinger equation loses the fundamental property of state conservation.

The material properties (such as fundamental band-gaps and spin-orbit splitting energies) obtained experimentally, represent real phenomena, whereas models...
based on finite bands Hamiltonians are meant to approximate them. The very last step in such approximation schemes\textsuperscript{10} enables us to calculate interband corrections to the main part of the Hamiltonians with help of perturbation theory.\textsuperscript{11,12,22} However, this last step lacks a rigorous theoretical foundation as it does not guaranty the convergence of the perturbation expansion. The result is that the derived Hamiltonian, although directly based on experimental parameters (column 4,\textsuperscript{1}), represents a totally different mathematical object compared to its origin. The physical evidences, to support this claim have been already known for GaAs\textsuperscript{23} and recently been reported for Si\textsuperscript{24}.

We start with the Schrödinger equation with the potential \( V(x) \)

\[
H_0 \psi(x) \equiv \frac{p^2}{2m^*} \psi(x) + V(x) \psi(x) = E \psi(x),
\]

where \( E_0^r \) is the band edge energy of the system, \( p \) is a momentum operator, \( x \in \Omega \subset \mathbb{C}^n, n \leq 3, m^* \) is a piecewise-constant effective mass parameter. In \( \Omega \) we supplement \( H_0 \) by usual Dirichlet conditions on the boundary \( S \equiv \partial \Omega \)

\[
\psi(x) = 0, \quad x \in S.
\]

Our major focus is on the heterostructure case, where the difference in the effective mass and probability current conservation imposes the discontinuity of \( \nabla \psi(x) \) on the interface between materials (assuming that the basis Bloch functions are the same in all constituents\textsuperscript{23}). In this case an appropriate choice of the functional space for coefficients from \( H_0 \) is the Hilbert space of distributions with the compact support \( \mathcal{H}^1(\Omega) \textsuperscript{20,27} \).

Then, the operator \( H_0 \) with the domain of definition \( D(H_0) \subset \mathcal{H}^2(\Omega) \) is a symmetric operator with an existing self-adjoint extension, hence it conserves the probability current\textsuperscript{28}. General theory of elliptic partial differential equations characterizes the problem \( H_0 \) in terms of the corresponding Fredholm type theorems (e.g. \textsuperscript{29}), so that the problem \( H_0 \) has a countable set of eigenvalues \( E_i \), smallest eigenvalue \( E_0 \) is simple and corresponding eigenstate \( \psi_0 \) is of constant sign in \( \Omega \), furthermore

\[
E_i \in [\rho_0, \infty), \quad \rho_0 = E_0^r + \min\{0, \inf_{\Omega} V(x)\}, \quad i = 0, \ldots, n,
\]

where \( V(x) \) is real.

If \( V(x) \) is a gently varying function over the unit cell in the sense of \textsuperscript{31} the original operator \( H_0 \) can be approximated by another operator \( H \) (using Bloch theorem), determined by the projection \( P \) of \( H_0 \) on the considered eigenspace and Lowding perturbation theory.\textsuperscript{13} The last step in this approximation procedure accounts for the influence of the elements from the space complement to the eigenspace by the formula

\[
H = PH_0 + \sum_{i=1}^{r} \lambda^i H^{(r)}
\]

up to the order \( r \). Setting \( \lambda = 1 \) leads one to the final approximation, under the assumption that the series \( \sum \) is convergent for such \( \lambda \). Despite a wide applicability of such approximations, the intrinsic ellipticity requirements for the realizations of \( H \) have not been explicitly verified in a systematic manner (see \textsuperscript{11,14,22} as well as more recent works).

As an example, let us consider two classical Hamiltonians for ZB\textsuperscript{13} and WZ\textsuperscript{21} type of materials with more scrutiny. In what follows we use the parameter notation identical to the works where corresponding Hamiltonians were obtained. When necessary, the parameters will be converted from one notation to another by using the formulas from \textsuperscript{1,30}.

First consider the Luttinger-Kohn (LK) Hamiltonian from \textsuperscript{3}

\[
H^{LK} = \begin{pmatrix}
\frac{1}{2} P & L & M & 0 & \frac{i}{\sqrt{2}} L \\
L^* & \frac{1}{2} P + \frac{2}{3} Q & 0 & M^* & -\frac{1}{\sqrt{2}} M \\
M^* & 0 & \frac{1}{2} P + \frac{2}{3} Q & L^* & -\frac{1}{\sqrt{2}} M \\
0 & M^* & -L^* & \frac{1}{2} P & \frac{i}{\sqrt{2}} L \\
-\frac{i}{\sqrt{2}} L^* & \frac{i}{3\sqrt{2}} (P - 2Q) & \frac{i}{\sqrt{2}} L & \frac{i}{3\sqrt{2}} (P - 2Q) & \frac{1}{3} (P + Q)
\end{pmatrix}
\]

where each of the \( P, Q, L, M \) is a second order position dependent differential operator or equivalently second order polynomial in the momentum representation\textsuperscript{\textsuperscript{3}}

\[
P = (A + B)k_+ k_- + 2Bk_z^2, \quad Q = B k_+ k_- + Ak_z^2, \quad M = \frac{1}{2\sqrt{3}}((A - B)(k_x^2 - k_y^2) - 2iCk_xk_y), \quad L = -i \frac{C}{2\sqrt{3}} k_z, \quad k_\pm = k_x \pm ik_y.
\]
Our aim is to check the type (elliptic, hyperbolic or essentially hyperbolic) of the $H^{LK}$ as a partial-differential operator (PDO) on $\mathcal{H}^1(\Omega)$ as we know that the given Schrödinger operator from [1] is elliptic. Only the second order derivative terms are playing the dominant role in the following analysis because contributions from the terms linear in $k$ as well as from the potential, are bounded in the domain $D(H^{LK})$. It means that the results for more complicated physical models with potential contributions from additional fields (e.g. strain, magnetic field, etc.) will stay the same as for the original $H^{LK}$, analyzed here. The fact that the Hamiltonian is a linear operator guarantees that it is also true for any other representation of $H^{LK}$ obtained by linear (basis) transformations.

In a more general sense, for any $m$-dimensional matrix PDO $H = \{h_{ij}\}_{i,j=1}^{m}$, where each element $h_{ij}$ is a second order one dimensional PDO

$$h_{ij} = \sum_{k,l=0}^{n} h_{ij}^{kl} \frac{\partial^2}{\partial x_k \partial x_l},$$

(4)

the associated quadratic form is defined by

$$G(\xi_1, \ldots, \xi_{nm}) = vMv^T, \quad v = (\xi_1, \ldots, \xi_{nm}),$$

(5)

where $M$ is an $mn \times mn$ matrix composed from the elements $h_{ij}^{kl}$. The $k \cdot p$ Hamiltonians in $\mathbb{R}^3$ are a special case of (4) with the real eigenvalues $\alpha_i$ and $n = 3$ (e.g. [14]).

Using these notations, the procedure of obtaining the ellipticity condition for $H$ reduces to the question about the sign of $\lambda_i$ for the associated $M$. More precisely, the matrix differential operator $H$ will be elliptic if and only if all eigenvalues of the corresponding Hermitian $M$ will have the same sign [27,29].

In general, it is a challenging task to calculate the eigenvalues of $M$ explicitly, even for such small as $3 \times 3$ Hamiltonians, but we recall that we deal here with usually sparse, band structure operators.

Taking into account the fact that the sequence of eigenenergies of $H_0$ is semi-bounded, for an approximation $H^{LK}$ (in the momentum representation), we obtain

$$\lambda_i < 0, \quad \forall \ i = 0, 1, \ldots, nm.$$  

(6)

The last constrains guarantee the ellipticity (in strong sense [20]) of Hamiltonian $H$. The operator $H$ possesses a self-adjoint extension in $D(H) \subset \mathcal{H}^{k+2}(\Omega)$, $k > 0$, provided that the domain $\Omega$ is sufficiently smooth (piecewise Lipschitz). Then it can be extended to a Hermitian operator by closure in the norm [p. 113, 29] or via Lax-Miligram procedure [27]. From the physical point of view the smoothness characteristics of $D(H)$ fulfill the natural assumption of quantum theory that the state of the system must be a continuous function of spatial variables even when some coefficients of $H$ have finite jumps [30] like in the heterostructures consisting of different materials [27,29].

| $H^{LK}$ as a partial-differential operator (PDO) on $\mathcal{H}^1(\Omega)$ | $H^{LK}$ as a partial-differential operator (PDO) on $\mathcal{H}^1(\Omega)$ |
|---|---|
| $\lambda_1 = -\frac{C}{T} + A$ | $\lambda_{2/3} = (\frac{1}{4} \pm \sqrt{3}) C + A$ |
| $\lambda_{4/5} = \pm \frac{C}{T} + B$ | $\lambda_{0/7} = \pm \frac{C}{T} + B,$ |

where $\lambda_1$, $\lambda_{2/3}$, $\lambda_{4/5}$ have the multiplicity 2 and $\lambda_{0/7}$ have the multiplicity 4, respectively, and $A, B, C$ are usual material parameters. By substituting (7) into (6), we receive the system of linear inequalities with respect to $A$, $B$ and $C$. They describe the feasibility region in the $A, B, C$ space, when $H^{LK}$ is an elliptic partial differential operator with the discrete and decreasing sequence of eigenvalues. One can use similar reasoning to obtain corresponding inequalities for other common representations of $H^{LK}$ through Luttinger parameters [20]. Evidently, any solution of (6) for (7) would have a unique corresponding solution in $\gamma_i$ notation. Note next that the solution domain $\Lambda_-(H^{LK})$ of (7) is symmetric with respect to the sign of $C$. It follows from the form of the $H^{LK}$ which is dependent on $|C|$ only. $\Lambda_-(H^{LK})$ comprises an unbounded pyramid in $\mathbb{R}^3$ (cf. Fig. 1) with the following rays as its edges: 

$$(-\frac{1}{2}, -\frac{1}{2}, -1), \quad \left(\frac{1}{4} + \sqrt{3} \right) t, -\frac{1}{2} t, t,$$

and

$$(-t, 0, 0), \quad \left(0, -\left(\frac{3}{4} + \sqrt{3} \right) t, 0\right), \quad \text{for} \ t \in [0, \infty).$$

A B C d | A B C d |
|---|---|---|---|---|---|---|---|
| $\text{AlAs}^a$ | -7.5 | -2 | 8.4 | 1.97 | Ge$^b$ | -30 | -4.6 | 33 | 10.64 |
| $\text{AlP}^a$ | -3.8 | -3.4 | 7.2 | 0.92 | Ge$^{ce}$ | -30 | -4.6 | 33 | 10.64 |
| $\text{AlS}^e$ | -3.7 | -3.4 | 7.2 | 1.01 | Ge$^a$ | -30 | -4.4 | 36 | 11.99 |
| $\text{GaS}^d$ | -7.5 | -3.8 | 6.1 | In $\text{InP}$ | -11 | -1.8 | 10 | 2.86 |
| $\text{C}^g$ | -4 | -3.4 | 6.6 | 0.42 | InP$^{ce}$ | -15 | -2.1 | 17 | 5.72 |
| $\text{C}^h$ | -2.9 | -2.3 | 3.9 | In $\text{InP}$ | -9 | -3.3 | 9.6 | 1.34 |
| $\text{C}^i$ | -6 | -3.8 | 6 | In $\text{InSb}$ | -100 | -4 | 96 | 39.35 |
| $\text{C}^j$ | -3.1 | -1.7 | 0.9 | In $\text{InSb}$ | -100 | -3 | 96 | 40.25 |
| $\text{GaS}^k$ | -16 | -2 | 6 | 0.89 | Si$^k$ | -6 | -3.5 | 9.6 | 1.18 |
| $\text{GaS}^l$ | -17 | -2.2 | 6.6 | 0.98 | Si$^k$ | -5.5 | -3.6 | 8.4 | 0.54 |
| $\text{GaS}^m$ | -15 | -3.1 | 17 | 4.83 | Si$^k$ | -6 | -3.4 | 8.4 | 0.72 |
| $\text{GaP}^n$ | -6 | -3 | 7.2 | 0.54 | SiC$^a$ | -2.6 | -1.7 | 4.2 | 0.38 |
| $\text{GaP}^o$ | -8 | -2.2 | 10 | 2.50 | SiC$^a$ | -4.8 | -1.8 | 5.1 | 0.67 |

$^a$ Measured under $T = 4.2K$

$^b$ Obtained by extrapolations from $14 \times 14 k \cdot p$ model

$^c$ Measured under $T = 300K$

$^d$ Ref. [20]

$^e$ Ref. [19]

$^f$ Ref. [20]

$^g$ Most probable value (set 5 from [30])

$^h$ Sets 1, 2, 3 from [31]

$^i$ Sets 7 (T=50K), 8 (T=70K), 2 from [19]

$^j$ Sets 7 (T=60..300K), 2, 1 from [19]

$^k$ Sets 2 (T=1,26K), 3, 6 from [19]

$^l$ Sets 4 (T=1.6K), 3, 6 from [19]
The ellipticity analysis for ZB, can be applied to a WZ Hamiltonian. These conditions are also violated in most of practically important materials, among which, we would like to mention GaN, AlN and ZnO. Here, the distances to the WZ ellipticity region defined by (9) are approximately equal: 0.804, 0.862, 0.606 for GaN parameter sets, 1.132, 1.27, 1.007 for AlN; 1.067 for ZnO. The distances have the order of terms in the unperturbed part of the WZ Hamiltonian (which in the dimensionless Luttinger-like notation equal to 1), and thus are considerably high.

Let us return back to the feasible parameters for two materials C and GaN. For carbon, the parameter values were analyzed in [37] where authors showed that they don’t agree with the Hall effect experimental measurements. In the same paper the authors suggested another, more consistent (in term of the measurements), set of parameters (row 5, Table I). Observe, however, that the latter set does not belong to the ellipticity region \( \Lambda_{-}(H^{LK}) \). In terms of the distance to \( \Lambda_{-}(H^{LK}) \), we can also classify other mostly large band gap materials, such as Si, SiC, AIP and GaN, as those belonging to the same group. For GaN we have the set of \( A, B, C \) lying inside the region and for other three materials sets lie relatively close to this region. Such small deviations are within the reported order of measurement accuracy (0.25%, 0.3% 0.6% for Si, SiC and AIP, respectively). They can be eliminated by direct adjustments. The fact that the Si belongs to that category in spite of its smaller band gap of \( \approx 1.11 \) can be easily explained. Indeed, it is one component diamond crystal with highly regular parabolic main valence and conduction bands diagrams, and additionally its structure follows the time reversal symmetry at \( \Gamma \) point.

The rest of the materials from Table I have more complicated structure, e.g. the InSb is a small band gap, big effective mass material. It is known [39] that by accounting for the valence bands only, LK approximation would be insufficient for InSb like materials, and presented analysis support this fact theoretically. Concerning the Ge and GaAs they have anisotropic lower conduction bands without time reversal symmetry and high coupling between the \( p \)-bonding topmost valence band and \( p \)-antibonding conduction band states [24]. Inclusion these conduction band states leads to more precise 8 × 8 and 14 × 14 models [38]. The setup described above is still applicable for these models but with a few minor modifications.
1. G. Bir and G. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors* (Wiley, New York, 1974).

2. S. L. Chuang, *Physics of optoelectronic devices* (Wiley, 1995).

3. J. M. Luttinger and W. Kohn, Phys. Rev. 97, 869 (1955).

4. E. O. Kane, J. Phys. Chem. Solids 1, 249 (1957).

5. J.-B. Xia, Phys. Rev. B 43, 9856 (1991).

6. S. R. Patil and R. V. N. Melnik, Nanotechnology 20, 125402 (2009).

7. F. J. de Abajo, Reviews of Modern Physics 79, 1267 (2007).

8. S. L. Chuang, *Physics of photonic devices* (John Wiley & Sons, 2009).

9. D. L. Smith and C. Mailhiot, Phys. Rev. B 33, 8345 (1986).

10. B. A. Foreman, Phys. Rev. B 75, 235331 (2007).

11. W. Yang and K. Chang, Phys. Rev. B 72, 233309 (2005).

12. B. Lassen, R. Melnik, and M. Willatzen, Commun. Comput. Phys. 6, 699 (2009).

13. R. G. Veprek, S. Steiger, and B. Witzigmann, Phys. Rev. B 54, 11539 (1996).

14. F. Mireles and S. E. Ulloa, Phys. Rev. B 62, 2562 (2000).

15. J.-B. Jeon, Y. Sirenko, K. Kim, M. Littlejohn, and M. Stroscio, Solid State Communications 99, 423 (1996).

16. W. J. Fan, J. B. Xia, P. A. Agus, S. T. Tan, S. F. Yu, and X. W. Sun, Journal of Applied Physics 99, 013702 (2006).

17. L. Reggiani, D. Waeber, and S. Zukotynski, Phys. Rev. B 28, 3550 (1983).

18. In this case the smoothness coefficient $k$ in the definition of domain and image of $H$ is $k = 1$. 

---

* Electronic address: sytnikd@gmail.com
† Electronic address: spatil@wlu.ca
‡ Electronic address: rmelnik@wlu.ca