Criticism of the OPW Method for Band Structure Calculations

KARL LENDI

Instituto de Física*, Universidade Federal do Rio de Janeiro, Rio de Janeiro

The OPW method is associated with a general eigenvalue problem of type $(A - \lambda B)\vec{x} = 0$, in which the matrix B and in particular its lowest eigenvalue decide upon the stability of the solutions λ and, therefore, upon the applicability of the method which may become very questionable for heavier substances. Analytical proofs as well as explicit numerical estimates for several solids are given.

O método "OPW" (ondas planas ortogonalizadas) relaciona-se com um problema geral de autovalores do tipo $(A - \lambda B)\vec{x} = 0$, onde a matriz B e, particularmente, seu menor autovalor decide sobre a estabilidade das soluções h e, portanto, da aplicabilidade do método, o que poderia tornar-se questionável no caso de substâncias mais pesadas. Apresentamos provas analíticas e avaliações numéricas explícitas para vários sólidos.

1. INTRODUCTION

The method of orthogonalized plane waves (OPW) has been used for the calculation of the electronic band structure of almost all types of solids. It has been successfully applied to metals, mostly in the simplified version of the pseudopotential method[1,2] which neglects nonlocal effects. Most applications have been made to semiconductors[3], whe-

---

* Postal address: Ilha do Fundão. Cidade Universitária, Bloco A-3? e 4?, Rio de Janeiro RJ.

re it is known to be superior to other methods of band theory, particularly in the case of covalent binding where the potential cannot be well approximated by the conventional muffin - tin construction[4]. Finally , even for insulators with large gaps, surprisingly good results could be obtained[5].

Among many attempts to improve the original version of Herring[6] as well as the procedures of practical computation, there is one outstanding example of perfection due to Euwema et $al.$[7] who even got self-consistency of the crystal potential. The approach of these authors provides the most sophisticated state of the method at present.

Summarizing, one can say that the OPW method is a powerful tool of general validity for the calculation of band structures of all kinds of substances. Moreover, it is rather practical and time–saving from the computational point of view since it leads to an eigenvalue problem involving matrix elements which do not depend on the eigenvalues, as in other methods of band theory. Therefore, the eigensolutions can be found easily by conventional methods of linear algebra.

In all the above-mentioned applications, very little has been said about the limitations[8,9] of the methods. For instance, there is a well-known example where the OPW method fails, namely in the case of sd–hybridization in crystals of transition elements[10], whose potentials may exhibit resonances[4]. This failure is clearly due to the structure of the secular problem arising in the OPW formalism which can be related to a Born-series expansion, and it is known in scattering theory[11] that resonances cannot be appropriately accounted for in any order of such an expansion.

On the other hand, when applying the OPW method to heavy substances, it becomes apparent that there must be a basic unreliability inherent in the formalism which gets worse in going from lighter to heavier elements. It is the purpose of the present paper to clarify this particular problem and to establish some criteria to gauge the applicability of the method.

2. THE OPW EIGENVALUE PROBLEM

Assume we know the exact Hamiltonian H including the true crystal poten-
tial, and denote by $\mathcal{H}$ the corresponding Hilbert space spanned by all
exact one-particle eigenstates. The idea of the OPW method is based on
the decomposition of H into the direct sum of three orthogonal comple-
ments:

$$H = H_c \oplus H_P^{(g)} \oplus H_R \quad .$$

H_c is the f-dimensional subspace of the so-called core states, $H_P^{(g)}$ the
g-dimensional subspace of band states, and H_R a rest due to the trunca-
tion of $H_P^{(g)}$ at finite g. The core states will be denoted by $|\alpha\rangle$, where
a denotes the set of relevant quantum numbers of a particular state,
and the plane waves by $|\vec{k} + \vec{k}_j\rangle$, with $\vec{k}$ being the reduced wave vector,
and $\vec{k}_j$ ($j = 1, 2, \ldots, g$) a reciprocal lattice vector. Under the assumption
that the basis of H_c is already known, the interest lies in finding the
basis $|i\rangle$ for $H_P^{(g)}$, starting with ordinary planewaves. This is done
by a partition of the Hamiltonian H into

$$H(\lambda) = H_A + \lambda H_B \ , \qquad H_B = \sum_\alpha |\alpha\rangle\langle\alpha| \qquad ,$$

where H_A contains the kinetic and the attractive as well as non-local
repulsive potential terms, which will not concern us in the following;
H_B incorporates the terms arising from the orthogonalization to H_c (for
details see Refs. 9,12), and λ stands for the set of band eigenvalues.
In this way, the original eigenvalue equation, restricted to $H_P^{(g)}$,

$$H|i\rangle = \lambda_i |i\rangle, \quad |i\rangle \in H_P^{(g)} \ , \text{ for every } i \ ,$$

is transformed into

$$H(\lambda_i) \ |\tilde{i}\rangle = \lambda_i |\tilde{i}\rangle, \text{ or } H_A|\tilde{i}\rangle = \lambda_i (1 - H_B)|\tilde{i}\rangle \qquad ,$$

where $|\tilde{i}>$ is the pure plane wave part of $|i>$ without orthogonality terms. We may write

$$|\tilde{i}> = \sum_{\ell=1}^{g} x_{\ell}^{i}|\vec{k} + \vec{K}_{\ell}> ,$$

and, after taking **matrix elements** of H_{A} and $(1 - H_{B})$, we are left with the general algebraic eigenvalue problem

$$(A - \lambda B)\vec{x} = 0 , \tag{1}$$

$\vec{x} = \{x_{m}\}$, $A = \{a_{mn}\}$, $B = \{b_{mn}\}$, $m,n = 1,\ldots,g$,

$a_{mn} = <\vec{k} + \vec{K}_{m}|H_{A}|\vec{k} + \vec{K}_{n}>$, $b_{mn} = <\vec{k} + \vec{K}_{m}|1 - H_{B}|\vec{k} + \vec{K}_{n}>$.

The additional index i, in h_{i} and $\vec{x}^{i}$, labels the g solutions to (1) . For positive-definite B, the shortest and most stable numerical way to solve (1) is a Cholesky factorization

$$B = R^{T}R ,$$

with R being a regular right-triangular matrix, and a subsequent reduction to a special eigenvalue problem

$$(M - \lambda 1)\vec{v} = 0 \tag{2}$$

by means of the transformations

$$M = (R^{-1})^{T} A R^{-1}, \quad \vec{v} = R \vec{x} .$$

Eq.(2) is then numerically solved by standard techniques, as described in Wilkinson and Reinsch[13]. The use of the inverse of R should make it clear that it is important to know something about the spectrum of B, particularly about its lower limit.

3. THE SPECTRUM OF B

The spectrum of B is of basic importance for reliable solutions of (1).

Since $P \equiv (I - H_B)$ is a projection operator on $H_p^{(g)}$, its spectrum[14] must lie in the closed interval $[0,1]$. In terms of the exact one-particle eigenstates of H the eigenvalues of P would be either 0 or 1, but in our case we take matrix elements of P between normalized plane waves, which always decompose into

$$|\vec{k} + \vec{k}_j> = y|c> + z|p> ,$$

$$|c> \in H_c , \quad |p> \in H_p^{(g)} ,$$

where $|c>$ and $|p>$ are normalized appropriate linear combinations of core states $|a>$ and band states $|i>$, respectively, and $|y|^2 + |z|^2 = 1$. Note in addition that $P|c> = 0$, and as a consequence, we find for the spectrum $\{\lambda_i^B , i = 1,\ldots,g\}$ of B (at least for $g \to \infty$) the whole closed interval

$$0 \leq \lambda_i^B \leq 1 , \text{ for every } i .$$

In fact, in an application to a concrete physical case, the lower limit should not be attained, and hopefully, B does not differ appreciably from the unit matrix, because usually the orthogonality coefficients $<\alpha|\vec{k} + \vec{k}_j>$ are small quantities (compared to 1), and

$$b_{mn} = <\vec{k} + \vec{k}_m|P|\vec{k} + \vec{k}_n> = \delta_{mn} \sum_{\alpha=1}^{f} <\vec{k} + \vec{k}_m|\alpha><\alpha|\vec{k} + \vec{k}_n> . \tag{3}$$

This may be true as long as we deal with lighter elements. However, as can be seen from Eq. (3), the orthogonality terms accumulate with increasing f, and as a consequence, the heavier the elements the more B will differ from unity. The positive diagonal elements become smaller, and the modulus of the off-diagonal elements larger. The diagonal dominance and with it the "good" positive-definiteness are going to be lost[15] and this in turn means that the lowest eigenvalue, hereafter referred to as λ_{min}^B, is approaching its limiting value zero. This is the behavior to be expected when going from light to heavy substances. The qualitative trend will be the same for all kinds of substances but one

expects different **details for monoatomic** and polyatomic crystals as **well** as for mtals and insulators, based on the fact that the energetically highest core states, which give **rise** to the largest orthogonaliy **con**-tributions, may be different for these classes. In the following, we are interested **in** a more quantitative analysis of the situation.

4. QUANTITATIVE NUMERICAL ESTIMATES

First, **it** is important to emphasize that certain eigenvalues **of** $(A-\lambda B)\vec{x}=$ 0 may "explode" for small λ_{min}^{B}, in which case the whole eigenvalue **pro**-blem becomes extremely ill-conditioned with respect to slight changes in the input quantities. In any application to a real physical situation, many appromations and simplifications have to be **made** and consequently, the **matrix elements** of A and B are affected by errors which reappear in the solutions, of course. The central question to be asked concerns the relationship between a relative **error** δ of input quantities and a cor-**responding** error ε of output quantities. **We** would speak of a **well-con**-ditioned eigenvalue **problem,** in fact occurring in the **case** of light **subs**-tances, if ε were roughly equal to 6. Unfortunately, the estimated **be**-havior of λ_{min}^{B} can lead to situations like

$$\varepsilon \sim \frac{1}{\delta}$$

already for low-order matrices, and the situation will get even worse in higher dimensions due to accumulation of numerical errors. Once an instability of the above mentioned strength is present **in** a calculation, the **adjustment** of **parameters** which have to be introduced in every appli-cation (at least the zeroth Fourier coefficient of the crystal **poten**-tial) and their physical interpretation **become** meaningless. Even **achie**-ving a reasonable adjustment, at a particular point in the **Brillouin** zone, would be no real advantage because the dispersion **of** the energy bands is too unreliable.

In the following, we would like to investigate trends of λ_{min}^{B} for **va**-rious substances. Associating a particular a with the main-, angular- and magnetic quantum numbers (n,ℓ,m), respectively, we note for **comple**-

teness that the explicit form of the orthogonality coefficients[12], in (3), is given by

$$<\alpha|\vec{k}+\vec{K}_j> = \left[\frac{4\pi}{\Omega}(2\ell+1)\right]^{1/2}\delta_{m,0}\int_0^R dr\; r j_\ell(\kappa_j r)P_{n\ell}(r)\;;$$

$\Omega = (4/3)\pi R^3$ is the volume per primitive cell, j_ℓ a spherical Bessel function, $P_{n\ell}(r)$ the radial part of a core wavefunction, and $\kappa_j = |\vec{k}+\vec{K}_j|$. Only states with $m=0$ will contribute. For simplicity, we choose substances which crystallize in the fcc- or rocksalt structure. The calculation is based on an expansion into 65 plane waves (unsymmetrized) and a group theoretical reduction of high order to low order matrices, aiming at the greatest possible numerical accuracy. Without loss of generality, we confine ourselves to the center Γ of the Brillouin zone, and consider only the states of symmetry Γ_1 which are the mst critical and interesting ones for our purposes since they are affected by the largest orthogonality contributions. In this case, the matrix dimension is 18.

It is generally true that the spectrum of B is highly degenerate or quasi-degenerate, that is, almost all eigenvalues accumulate in the point 1, and only a very few of them are well separated from the rest, and smaller than 1. The numerical problem is very delicate and, for this reason, we used a Fortran version of a procedure which is the mst stable for such a situation[16].

Table 1 gives a summary of the substance under consideration, their lattice constants a and the number s of completely or partially occupied core shells. To complete the inert gas series, we have included a hypothetical Rn-crystal. The diatomic compounds are labeled by the total nuclear charge per primitive cell.

According to the previous analysis, one would expect λ_{min}^B to be simply a monotonically decreasing function of Z, the total nuclear charge per primitive cell. However, this behavior is obscured by the dependence of the orthogonality coefficient on the lattice constant like $\Lambda^{-1/2}$, where $A = a^3/4$, furthermore by the details of the core wavefunctions for which the overlap with plane waves may be rather sensitive, as well as by the modification of the plane waves due to the change of the lattice cons-

Table 1. Lattice constants and number of core shells

	$a[\text{Å}]$	s		$a[\text{Å}]$		s
^{10}Ne	4.46	1	^{20}MgO	4.21	^{12}Mg	3
					^{8}O	1
^{18}Ar	5.31	3	^{36}KCℓ	6.29	^{19}K	5
^{36}Kr	5.64	6			^{17}Cℓ	3
^{54}Xe	6.13	9	^{54}KBr	6.59	^{35}Br	6
^{86}Rn	6.70	13	^{90}RbI	7.33	^{37}Rb	8
^{13}Aℓ	4.05	·3			^{53}I	9
^{38}Sr	6.08	8	^{71}EuO	5.14	^{63}Eu	12
^{79}Au	4.08	13	^{100}UO	4.92	^{92}U	17

tant. In view of these facts, the following numerical results of the spectral analysis of the matrix B can be taken without much ado. Tables 2, 3 and 4 show a representative contribution to the first diagonal element of B, namely, $\sum_{\alpha}<\vec{0}|\alpha><\alpha|\vec{0}>$, where $\vec{0} = (0,0,0)$ is the zero wavevector, and the lowest three eigenvalues as well as the number N_Δ of eigenvalues which lie in the interval A close to 1. A was chosen to be 5% of the total length of the spectrum, i.e., $0.95 \leq \lambda_\Delta^B \leq 1$.

For the inert gas series, λ_{min}^B is well behaved and will not cause any trouble, not even for the heavy hypothetical Rn-crystal. On the other hand, the ionic and metallic solids show appreciably lower eigenvalues. This can be understood from the fact orthogonalization in the OPW method is required only for those states which do not form the valence states. Usually, the energetically highest states (or the outermost electrons) give the largest orthogonality coefficients. For the inert gases, therefore, the s- and p-electrons of the octed shell do not contribute to the matrix elements of B, whereas for the other elements they do. Irregularities in the trend to decreasing λ_{min}^B, for increasing total ruclear charge, may occur as in the case of KCR. It has already been mentioned that there exists a complicated interplay between the change of lattice

Table 2. Lowest eigenvalues for the inert gases

	$\sum_{\alpha}\langle\vec{0}\|\alpha\rangle\langle\alpha\|\vec{0}\rangle$	λ_{min}^{B}	λ_{2}^{B}	λ_{3}^{B}	N_{Δ}
Ne	0.00162	0.9162	1.0000	1.0000	17
Ar	0.00987	0.6385	0.9483	0.9999	16
Kr	0.01444	0.5306	0.8880	0.9660	16
Xe	0.02501	0.3620	0.7673	0.8888	15
Rn	0.02692	0.3356	0.7344	0.8483	15

Table 3. Lowest eigenvalues for insulators

	$\sum_{\alpha}\langle\vec{0}\|\alpha\rangle\langle\alpha\|\vec{0}\rangle$	λ_{min}^{B}	λ_{2}^{B}	λ_{3}^{B}	N_{Δ}
MgO	0.10427	0.09450	O. 4990	0.8305	15
KCℓ	0.15003	0.03642	0.2331	0.7046	14
KBr	0.13490	0.04875	0.2667	0.6209	14
RbI	0.14343	0.04042	0.2053	0.4913	12

Table 4. Lowest eigenvalues for metals and heavy oxydes

	$\sum_{\alpha}\langle\vec{0}\|\alpha\rangle\langle\alpha\|\vec{0}\rangle$	λ_{min}^{B}	λ_{2}^{B}	λ_{3}^{B}	N_{Δ}
Aℓ	0.07970	0.12513	0.5686	0.9918	16
Sr	0.17444	0.01585	0.1377	0.8792	15
Au	0.21993	0.00626	0.0939	0.1764	14
EuO	0.32958	-0.00494	0.0600	0.6532	11
UO	0.50968	-0.02217	0.0674	0.2355	10

constant connected to the change of nuclear charge when going from one particular substance to another. The most striking situation is found for the heavy oxides EuO and UO in strict contradiction to the exact analytical lower limit for the spectrum.

5. CONCLUSIONS

It was the aim of this paper to display a basic weakness of the OPW method inherent in the formalism itself and rather independent of particular physical approximations used in concrete applications. This weakness is due to the lower limit of the spectrum of the matrix B, in the OPW-eigenvalue problem $(A - \lambda B)\vec{x} = 0$, which approaches zero for heavier substances, a property which in addition is rather sensitive to the ordinary uncertainties in the matrix elements. If one assumes that, in a hypothetical exact calculation, a small lowest eigenvalue results, one knows that already slight errors in B can drive it much closer to zero or even below zero, as is the case for EuO and UO, for instance. Under such circumstances, the matrix elements of A which are much more affected by uncertainties than those of B, because they contain the crystal potential, blow up according to the solution (2). These extreme instabilities will make at least some of the solutions entirely meaningless. The number of stable solutions is rougly equal to the number N_Δ shown in Tables 2,3 and 4. One could argue that a random distribution of errors over the matrix B cannot account for such instabilities, since a coherent change of certain elements is necessary for this mechanism. However, since the matrix elements of B are nothing else than overlap-integrals between orthogonalized plane waves, an error in any of these functions will propagate systematically through rows and columns, causing exactly such a coherent change. For completeness, we would like to mention that $\lambda_{min}^B \to 0$ means that certain rows or columns of B become linearly dependent, and that, in turn, means that by orthogonalizing the plane waves to the core states, quasi-linearly dependent functions have been constructed.

We hope that our criticism will contribute in making future applications of the OPW method more reliable.

REFERENCES

1. WA. Harrison, *Pseudopotentials in the Theory of Metals*, W.Benjamin, New York 1966.

2. M. L. Cohen and V. Heine, in *Solid State Physics 24*, 37 (1970), Academic Press, New York/London.

3. F. Herman *et al.*, in *Methods in Computational Physics 8*, 193 (1968), Academic Press, New York/London.

4. J.M. Ziman, in *Solid State Physics 26*, 1 (1971), Academic Press, New York/London.

5. A.B. Kunz, Phys.Rev. 175, 1147 (1968).

6. C. Herring, Phys. Rev. *57*, 1169 (1940).

7. R.N. Euwema e t *al.*, Phys. Rev. *162*, 710 (1967).

8. J. Callaway, Phys. Rev. 97, 933 (1955).

9. E. Brown, Phys. Rev. *126*, 421 (1962).

10. F.M. Müller, Phys. Rev. *153*, 659 (1967).

11. T.Y. Wu and T. Ohmura, *Quantum Theory of Scattering*, Prentice-Hall, Englewood Cliffs, N.J., 1962.

12. T.O. Woodruff, in *Solid State Physics 4*, 367 (1957), Academic Press, New York/London.

13. J.H. Wilkinson and C. Reinsch, *Linear Algebra*, Springer Verlag , Berlin, Heidelberg, New York, 1971.

14. N. I. Akhiezer and I.M. Glasmann, *Theorie der linearen Operatoren im Hilbert-Raum*, Akademie-Verlag, Berlin, 1968.

15. H. R. Schwarz *et al.*, *Matrizen-Numerik*, Teubner, Stuttgart 1968.

16. H. Rutishauser, Numer. Math, 9 1 (1966).