

Calculation of electronic transition matrix elements based on local atomic environment

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1977 J. Phys. A: Math. Gen. 10 461

(<http://iopscience.iop.org/0305-4470/10/4/009>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 129.252.86.83

The article was downloaded on 30/05/2010 at 13:55

Please note that [terms and conditions apply](#).

Calculation of electronic transition matrix elements based on local atomic environment

Roger Haydock

Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, UK

Received 9 November 1976

Abstract. In a wide variety of solids the electronic structure is well described in terms of local atomic environment by the recursion method. A theory of perturbations, directed towards the calculation of transition matrix elements, is developed in terms of the local environment using the recursion method as the unperturbed solution. As well as embodying this physical concept, the theory permits calculations on systems with many more degrees of freedom than other methods.

1. Introduction

The calculation of transition matrix elements is a vital step in solid state physics for the comparison of theory and experiment. Such techniques as photo-emission, Auger spectroscopy, soft x-ray emission, and so forth, all measure transition probabilities rather than spectra of electronic states. These experiments could only be compared qualitatively with theory in situations where the matrix elements are not constant.

In the bulk of simple crystals, matrix elements can be calculated using eigenstates with particular crystal momentum. However, there is much interest in surfaces, dislocations, and amorphous structures where band theory is difficult to apply. It was for these systems that the recursion method (Haydock *et al* 1975, Haydock 1974) was devised. This theory is the generalization of band theory to low-symmetry structures. However, like band theory it produces eigenstates and eigenvalue spectra rather than the dynamic behaviour of the system.

The probes used in many experimental methods may be treated as perturbations of the system's Hamiltonian. The perturbation is time dependent so it gives rise to transitions between states of different energy. For example the perturbation in photo-emission is the dipole potential of the electric field of the photon. For photons of a particular frequency, this dipole perturbation gives rise to electronic transitions between the ground state and excited states with the photon energy.

The philosophy of the recursion method is to express the stationary electronic states in terms of a sequence of states first near a single atom and then successively, linear combinations of states near other atoms. This embodies the physical principle that the electronic structure in a region is mainly determined by the atomic environment of that region and secondly by the effect on that region of the neighbouring atomic environment. This approach has been successful in producing quick and accurate densities of states for localized orbital models.

The motivation of the theory presented here is to apply this philosophy to the problem of electronic transitions. There are two aspects of this. In the case of

photo-emission, the probe is extended in nature so that while emission from a particular region is related to the local environment, emission takes place from a number of regions simultaneously. The second case is that of a localized probe such as a single atom ionized by some process. In this case both the probe and the electronic structure are determined mainly by the local environment.

As the recursion method expresses electronic structure in terms of successive neighbours of an atom, so the perturbation theory describes the changes in the neighbouring states. This comes about in two ways. Not only does a perturbation alter the energy of and coupling between neighbouring states, but it also alters the order of importance of different neighbouring environments. The theory accounts for both effects.

Although this theory was developed for the purpose of calculating transition matrix elements, it has great generality. Its most important property is that it converges for continuous spectra—as must a perturbation theory for use in the solid state. Approximants for physical quantities converge even when there are singularities in the coupling constant. For a perturbation theory to possess this quality, physical quantities must not be expressed as power series in the coupling constant, but rather, as is the case, the coupling constant enters in a way similar to a Padé approximant. The approximation includes the presence of singularities in the coupling constant.

The recursion method expresses the Hamiltonian for a system as a three-term recurrence relation, equivalently as a tridiagonal matrix. Once in this form there are simple expressions for the eigenvalues, eigenvectors, Green functions and other physical quantities. This theory is based on the idea of a perturbed recurrence relation. Thus the recurrence coefficients or tridiagonal matrix elements are expressed as a power series in the coupling constant. Likewise the basis is also dependent on the coupling constant. This means that the perturbed Hamiltonian is tridiagonal in the new basis. The changes in the coefficients express the changes in the properties of the environment while the change in basis expresses the movement of the environment.

Computationally the recursion method and this theory have great advantages. Systems with a large number of degrees of freedom can be studied because the method requires only the storage for a few states (vectors) of the system rather than the storage of matrices. Typical calculations with recursion theories involve 10 000 degrees of freedom while storage of matrices for such a system would require 10^8 storage locations. It is also found that the approximants converge rapidly so that one gets a good result far faster than with other methods.

The body of the paper is organized in two sections. The first is a description of the theory. It is completely general although I point out where simplifications can be made in particular cases. There are no proofs although the method of proof is outlined. There are also certain matters such as precise limits of convergence which I leave to future work. The section is in two parts. The first shows how to obtain the perturbed recurrence relation and the second part shows how to use the perturbed recurrence to calculate physical quantities. The second section is a discussion of an application of the theory to interpretation of optical absorption experiments.

2. New perturbation theory

This section will deal mainly with the mathematics of the perturbation theory. The notation is summarized in table 1. However, the first point is to define the relationship

Table 1. A table of notation used in presenting the recursion perturbation theory.

$[X]_{uw}$	the \mathbf{u}, \mathbf{w} component of \mathbf{X}	$\mathbf{J}_n^{(l)}$	coefficient of λ^l in \mathbf{J}_n
\mathbf{X}^\dagger	complex conjugate transpose of \mathbf{X}	λ	scalar strength of perturbation
a_n, b_n, c_n	matrix elements of \mathbf{J}	$P_n(E)$	orthogonal polynomial n
$a_n^{(l)}, b_n^{(l)}, c_n^{(l)}$	coefficients of λ^l in a_n, b_n, c_n	$Q_n^{(l,m)}(E)$	orthogonal polynomial related to the l, m element of $\hat{\mathbf{G}}_n^{(0)}(E)$
\mathbf{e}_m	unit column vector with one in component m	$\mathcal{R}^n\{f(x), x_0\}$	the n th order residue of $f(x)$ at x_0
$\mathbf{G}(E)$	energy resolvent operator	\mathbf{S}	matrix of overlaps of $\{z_m\}$
$\mathbf{G}(E)$	resolvent matrix in the $\{z_m\}$ basis	\mathbf{U}	matrix made up of columns \mathbf{u}_n
$\mathbf{G}^{(l)}(E)$	coefficients of λ^l in $\mathbf{G}(E)$	\mathbf{u}_n	elements of the left-hand tridiagonal basis
$\hat{\mathbf{G}}(E)$	resolvent matrix in the $\{\mathbf{u}_n\}, \{\mathbf{w}_n\}$ basis	$\mathbf{u}_n^{(l)}$	coefficient of λ^l in \mathbf{u}_n
$\hat{\mathbf{G}}^{(l)}(E)$	coefficient of λ^l in $\hat{\mathbf{G}}(E)$	\mathbf{V}	perturbing potential operator
$\hat{\mathbf{G}}_n(E)$	$(E\mathbf{I}_n - \mathbf{J}_n)^{-1}$	\mathbf{V}	matrix of \mathbf{V} in $\{z_m\}$ basis
$\hat{\mathbf{G}}_n^{(l)}(E)$	coefficient of λ^l in $\hat{\mathbf{G}}_n(E)$	\mathbf{W}	matrix made up of columns \mathbf{w}_n
\mathbf{H}	Hamiltonian operator	\mathbf{w}_n	elements of the right-hand tridiagonal basis
\mathbf{H}	hopping matrix among the $\{z_m\}$	$\mathbf{w}_n^{(l)}$	coefficient of λ^l in \mathbf{w}_n
\mathbf{I}	identity matrix	\mathbf{x}_n	intermediate left-hand vector in recursion
\mathbf{I}_n	$n \times n$ identity matrix	$\mathbf{x}_n^{(l)}$	coefficient of λ^l in \mathbf{x}_n
\mathbf{J}	Jacobi or tridiagonal matrix for $\mathbf{H} + \lambda \mathbf{V}$	\mathbf{y}_n	intermediate right-hand vector in recursion
$\mathbf{J}^{(l)}$	coefficient of λ^l in \mathbf{J}	$\mathbf{y}_n^{(l)}$	coefficient of λ^l in \mathbf{y}_n
\mathbf{J}_n	first n columns and rows of \mathbf{J}	z_m	basis states of \mathbf{H} and \mathbf{V}

between the physical description of a system and the mathematical entities involved in this theory. Let us think of the physics in terms of some solid. The electrons behave quantum mechanically and suppose further that they are well described by an independent-particle model, although the theory applies just as well to a many-body interaction.

Select a set of states $\{z_m\}$ to use as a basis for the electronic Hamiltonian. Let \mathbf{H} be the unperturbed Hamiltonian and \mathbf{V} the perturbation. We may express the operators \mathbf{H} and \mathbf{V} as matrices \mathbf{H} and \mathbf{V} in terms of the basis by means of

$$\mathbf{H}z_m = \sum_n [\mathbf{H}]_{nm}z_n \tag{1}$$

and,

$$\mathbf{V}z_m = \sum_n [\mathbf{V}]_{nm}z_n \tag{2}$$

where $[\mathbf{H}]_{nm}$ means the element of matrix \mathbf{H} in the n th row and m th column. These definitions of \mathbf{H} and \mathbf{V} do not require $\{z_m\}$ to be orthogonal, nor indeed do they require any specification of a scalar product (Haydock 1974). The basis must simply be linearly independent for unique specification of \mathbf{H} and \mathbf{V} .

The scalar product is now defined independently of \mathbf{H} and \mathbf{V} . This is by means of another matrix \mathbf{S} ,

$$[\mathbf{S}]_{nm} = z_n \cdot z_m. \tag{3}$$

Thus the matrix element of \mathbf{H} between two states z_n and z_m is

$$z_n \cdot \mathbf{H}z_m = \sum_l [\mathbf{S}]_{nl}[\mathbf{H}]_{lm}. \tag{4}$$

From the above definitions, the physics has been expressed in terms of three matrices \mathbf{H} ,

\mathbf{V} , and \mathbf{S} . The results of this calculation are the matrix elements of \mathbf{V} between eigenstates of \mathbf{H} . A compact way of representing this information is by means of the resolvent,

$$\mathbf{G}(E) = (E - \mathbf{H} - \lambda \mathbf{V})^{-1}, \quad (5)$$

where λ is a scalar expressing the strength of the perturbation \mathbf{V} . In terms of the basis, the matrix for the resolvent is

$$\mathbf{G}(E) = (E\mathbf{I} - \mathbf{H} - \lambda \mathbf{V})^{-1}, \quad (6)$$

where \mathbf{I} is the identity matrix.

Let us construct a similarity transformation for the Hamiltonian, expressed as new left- and right-hand basis sets, whose elements in terms of $\{\mathbf{z}_n\}$, are given by the columns of \mathbf{U} for the left-hand basis and \mathbf{W} for the right-hand basis. They are to be biorthonormal, so

$$\mathbf{U}^\dagger \mathbf{W} = \mathbf{I} \quad (7)$$

where the dagger means the complex conjugate transpose of a matrix. A further condition is that

$$\mathbf{U}^\dagger (\mathbf{H} + \lambda \mathbf{V}) \mathbf{W} = \mathbf{J}, \quad (8)$$

where \mathbf{J} is a tridiagonal matrix, one which has non-zero elements only on the main diagonal and the two adjacent diagonals so that in each row or column there are at most three non-zero elements.

Suppose we are interested in the transitions between two states \mathbf{u} and \mathbf{w} . The matrix element of the resolvent

$$[\mathbf{G}(E)]_{uw} = \mathbf{u}^\dagger \mathbf{S} \mathbf{G}(E) \mathbf{w}, \quad (9)$$

contains all the information about the transitions between the eigenstates of \mathbf{H} comprising \mathbf{u} and \mathbf{w} . We now construct the first element of each new basis by finding the smallest non-negative power, m , of $\mathbf{H} + \lambda \mathbf{V}$ such that there are non-zero b_0 and c_0 for some λ ,

$$b_0 c_0 = \mathbf{u}^\dagger \mathbf{S} (\mathbf{H} + \lambda \mathbf{V})^m \mathbf{w}. \quad (10)$$

If no such m exists, then $[\mathbf{G}(E)]_{uw}$ is zero. The choice of b_0 and c_0 is arbitrary except that their product is fixed. Define the first column of \mathbf{U} as

$$\mathbf{u}_0^\dagger = \mathbf{u}^\dagger \mathbf{S} / b_0, \quad (11)$$

and the first column of \mathbf{W} as

$$\mathbf{w}_0 = (\mathbf{H} + \lambda \mathbf{V})^m \mathbf{w} / c_0. \quad (12)$$

This choice of \mathbf{u}_0^\dagger and \mathbf{w}_0 is simple but not unique. Using the expansion of $\mathbf{G}(E)$ in powers of $1/E$, and the fact that m is the smallest non-negative power giving $b_0 c_0$ non-zero, then

$$[\mathbf{G}(E)]_{uw} = b_0 c_0 \mathbf{u}_0^\dagger \mathbf{G}(E) \mathbf{w}_0 / E^m. \quad (13)$$

Having defined \mathbf{u}_0 and \mathbf{w}_0 , we proceed with the construction of the remainder of \mathbf{U} and \mathbf{W} as well as \mathbf{J} . Define both \mathbf{u}_{-1} and \mathbf{w}_{-1} to be zero. Given $\{\mathbf{u}_n, \dots, \mathbf{u}_0, \mathbf{u}_{-1}\}$, $\{\mathbf{w}_n, \dots, \mathbf{w}_0, \mathbf{w}_{-1}\}$, $\{a_{n-1}, \dots, a_0\}$, $\{b_n, \dots, b_0\}$, and $\{c_n, \dots, c_0\}$, the iterative construction of the vectors \mathbf{u}_{n+1}^\dagger and \mathbf{w}_{n+1} , as well as the matrix elements, a_n , b_{n+1} and c_{n+1} , of \mathbf{J} ,

proceeds from

$$b_{n+1} \mathbf{u}_{n+1}^\dagger = \mathbf{u}_n^\dagger (\mathbf{H} + \lambda \mathbf{V} - a_n) - c_n \mathbf{u}_{n-1}^\dagger, \quad (14)$$

and

$$c_{n+1} \mathbf{w}_{n+1} = (\mathbf{H} + \lambda \mathbf{V} - a_n) \mathbf{w}_n - b_n \mathbf{w}_{n-1}, \quad (15)$$

where,

$$a_n = \mathbf{u}_n^\dagger (\mathbf{H} + \lambda \mathbf{V}) \mathbf{w}_n, \quad (16)$$

$$b_{n+1} c_{n+1} = [\mathbf{u}_n^\dagger (\mathbf{H} + \lambda \mathbf{V} - a_n) - c_n \mathbf{u}_{n-1}^\dagger][(\mathbf{H} + \lambda \mathbf{V} - a_n) \mathbf{w}_n - b_n \mathbf{w}_{n-1}]. \quad (17)$$

The \mathbf{u}_{n+1} and \mathbf{w}_{n+1} constitute successive columns of \mathbf{U} and \mathbf{W} respectively. The choice of b_{n+1} and c_{n+1} is arbitrary except that their product is fixed. The procedure terminates if $b_{n+1} c_{n+1}$ is zero. The matrix \mathbf{J} is defined as follows:

$$[\mathbf{J}]_{nm} = \begin{cases} a_n & \text{if } n = m \\ b_{n+1} & \text{if } n + 1 = m \\ c_n & \text{if } n = m + 1 \\ 0 & \text{otherwise} \end{cases}, \quad n, m \geq 0. \quad (18)$$

The properties in equations (7) and (8) follow simply from the construction.

Following this transformation, equation (13) takes a much simpler form. If \mathbf{e}_n represents a column vector with one in component n and zero for all other components,

$$[\mathbf{G}(E)]_{uw} = b_0 c_0 \mathbf{e}_0^\dagger (\mathbf{E} \mathbf{I} - \mathbf{J})^{-1} \mathbf{e}_0 / E^m. \quad (19)$$

So far we have defined \mathbf{U} , \mathbf{W} , and \mathbf{J} as functions of λ to yield a simple form of $[\mathbf{G}(E)]_{uw}$. Note that most of \mathbf{U} and \mathbf{W} are not needed for $[\mathbf{G}(E)]_{uw}$; this is the recursion method of Haydock (1974). For a perturbation theory we need to express \mathbf{U} , \mathbf{W} , and \mathbf{J} as power series in λ and then show that the calculation separates in such a way that evaluating a given quantity to a given order never involves any higher order.

Before going on to a perturbation theory, let us consider briefly the analytic properties of $[\mathbf{G}(E)]_{uw}$ as a function of λ . For a perturbation theory we must know over what regions of λ , \mathbf{U} , \mathbf{W} and \mathbf{J} are analytic. If, for some N , $b_{N+1} \mathbf{u}_{N+1}^\dagger$ or $c_{N+1} \mathbf{w}_{N+1}$ are zero for all λ , then the recursion terminates. If, however $b_{N+1} \mathbf{u}_{N+1}^\dagger$ or $c_{N+1} \mathbf{w}_{N+1}$ is zero for only some values of λ , this introduces singularities in λ into \mathbf{U} , \mathbf{W} and \mathbf{J} at those values. The singularities may be of any type including: (i) poles of arbitrary order, which are associated with the breaking of a finite-order degeneracy by \mathbf{V} ; and (ii) essential singularities, which are associated with breaking a degeneracy of infinite order such as phase change in a many-body system. It is hoped to discuss singularities in λ , i.e. degenerate perturbation theory, in future papers. For the remainder of this paper, we will consider only the case where \mathbf{U} , \mathbf{W} , and \mathbf{J} are analytic in λ at zero, that is, non-degenerate perturbation theory. For some removable singularities in λ it is easy to see how to apply the theory for λ analytic at zero.

Let us now define $\mathbf{u}_n^{(l)}$, $\mathbf{w}_n^{(l)}$, $a_n^{(l)}$, $b_n^{(l)}$, $c_n^{(l)}$ and $\mathbf{J}^{(l)}$ as the coefficients of λ^l in the power series for \mathbf{u}_n , \mathbf{w}_n , a_n , b_n , c_n , and \mathbf{J} respectively. The initial conditions are that

$$\mathbf{u}_{-1}^{(l)} = \mathbf{w}_{-1}^{(l)} = 0. \quad (20)$$

The coefficients $b_n^{(0)}$ and $c_n^{(0)}$ are both non-zero by the assumption that \mathbf{U} , \mathbf{W} , and \mathbf{J} are analytic at $\lambda = 0$. Define $\mathbf{u}_n^{(-1)}$ and $\mathbf{w}_n^{(-1)}$ to be zero. The choice of $b_0^{(l)}$ and $c_0^{(l)}$ is arbitrary except for this and the condition in equation (10) which defines the product of their

power series. These $b_0^{(l)}$ and $c_0^{(l)}$ in turn determine $\mathbf{u}_0^{(l)}$ and $\mathbf{w}_0^{(l)}$ through equations (11) and (12).

Having specified the initial conditions, we may define a recursion for $\mathbf{u}_{n+1}^{(l)}$, $\mathbf{w}_{n+1}^{(l)}$, $a_n^{(l)}$, $b_{n+1}^{(l)}$, and $c_{n+1}^{(l)}$ from equations (14), (15), (16) and (17). Expanding both sides of equation (16) in powers of λ and equating coefficients of the same power gives

$$a_n^{(l)} = \sum_{m=0}^l \mathbf{u}_n^{(m)\dagger} \mathbf{H} \mathbf{w}_n^{(l-m)} + \sum_{m=0}^{l-1} \mathbf{u}_n^{(m)\dagger} \mathbf{V} \mathbf{w}_n^{(l-m-1)}. \tag{21}$$

Define a sum to be zero if the upper limit is less than the lower. The calculation of $a_n^{(l)}$ involves \mathbf{u}_n^\dagger and \mathbf{w}_n to order l . Having calculated a_n to order l and knowing b_n , c_n , \mathbf{u}_n^\dagger and \mathbf{w}_n to order l , define

$$\mathbf{x}_n^{(l)\dagger} = \mathbf{u}_n^{(l)\dagger} \mathbf{H} + \mathbf{u}_n^{(l-1)\dagger} \mathbf{V} - \sum_{m=0}^l (\mathbf{u}_n^{(m)\dagger} a_n^{(l-m)} + \mathbf{u}_{n-1}^{(m)\dagger} c_n^{(l-m)}). \tag{22}$$

Thus \mathbf{x}_n^\dagger is defined to order l . Similarly, define \mathbf{y}_n to order l as

$$\mathbf{y}_n^{(l)} = \mathbf{H} \mathbf{w}_n^{(l)} + \mathbf{V} \mathbf{w}_n^{(l-1)} - \sum_{m=0}^l (a_n^{(m)} \mathbf{w}_n^{(l-m)} + b_n^{(m)} \mathbf{w}_{n-1}^{(l-m)}). \tag{23}$$

Applying this to equation (17) and equating coefficients of λ^l on both sides gives

$$\sum_{m=0}^l b_{n+1}^{(m)} c_{n+1}^{(l-m)} = \sum_{m=0}^l \mathbf{x}_n^{(m)\dagger} \mathbf{y}_n^{(l-m)}. \tag{24}$$

Starting with $l = 0$, increasing l by one gives one constraint for $b_{n+1}^{(l)}$ and $c_{n+1}^{(l)}$. However, they can be chosen to order l from \mathbf{x}_n and \mathbf{y}_n to order l . Now doing the same to equations (14) and (15) gives

$$\sum_{m=0}^l \mathbf{u}_{n+1}^{(m)\dagger} b_{n+1}^{(l-m)} = \mathbf{x}_n^{(l)\dagger}, \tag{25}$$

$$\sum_{m=0}^l c_{n+1}^{(m)} \mathbf{w}_{n+1}^{(l-m)} = \mathbf{y}_n^{(l)}. \tag{26}$$

Each of these equations can be solved to give

$$\mathbf{u}_{n+1}^{(l)\dagger} = \left(\mathbf{x}_n^{(l)\dagger} - \sum_{m=0}^{l-1} \mathbf{u}_{n+1}^{(m)\dagger} b_{n+1}^{(l-m)} \right) (b_{n+1}^{(0)})^{-1}, \tag{27}$$

$$\mathbf{w}_{n+1}^{(l)} = \left(\mathbf{y}_n^{(l)} - \sum_{m=0}^{l-1} c_{n+1}^{(l-m)} \mathbf{w}_{n+1}^{(m)} \right) (c_{n+1}^{(0)})^{-1}. \tag{28}$$

Thus equations (21), (22), (23), (24), (27), and (28) define the recursion for the perturbation and because one can calculate exactly to order l without using any higher order, the system fits the main requirement for a perturbation theory. The recursion terminates when $b_{n+1}^{(0)} c_{n+1}^{(0)}$ is zero. This is due to our assumption that \mathbf{J} is analytic at $\lambda = 0$.

The alteration to the above which takes into account poles of finite order at $\lambda = 0$ involves the introduction of similar poles at $\lambda = 0$ in \mathbf{U} and \mathbf{W} and the replacement of $b_{n+1}^{(0)}$ and $c_{n+1}^{(0)}$ by the lowest-order non-zero coefficient in b_{n+1} and c_{n+1} . Otherwise the equations are the same.

This perturbation theory of \mathbf{J} has the same properties that make the recursion method so useful. In any computation only a small number of vectors need be stored at

any one time. If the matrices \mathbf{H} and \mathbf{V} can be stored in some condensed form (which is usually the case with a local orbital basis for the electronic structure), then the theory can be applied to systems with many thousands of degrees of freedom. A further saving is made because it is not necessary to calculate the entire matrix \mathbf{J} . As with the recursion method, one obtains good results with only a 15-dimensional submatrix of \mathbf{J} . This completes the non-degenerate perturbation theory of \mathbf{U} , \mathbf{W} , and \mathbf{J} . The next part of this section deals with the calculation of the transition matrix elements from \mathbf{J} .

The result of the above calculation is a power series in λ for the bases $\{\mathbf{u}_n\}$, $\{\mathbf{w}_n\}$, and the matrix \mathbf{J} . The resolvent in these bases is simply $(\mathbf{E}\mathbf{I} - \mathbf{J})^{-1}$. As λ varies, the resolvent varies in two ways. The first is due to the variation of \mathbf{J} with λ and the second is due to the variation of the new bases vectors with λ . In order to calculate the matrix elements of \mathbf{V} between eigenstates of \mathbf{H} we can now concentrate on \mathbf{J} . Because $\mathbf{u}_0^{(0)}$ and $\mathbf{w}_0^{(0)}$ each contain the same eigenvectors of \mathbf{H} as \mathbf{u} and \mathbf{w} , we can be sure that these are all eigenvectors of $\mathbf{J}^{(0)}$. Furthermore to order l in λ , the eigenvectors of $\mathbf{H} + \lambda \mathbf{V}$ contained in \mathbf{u} and \mathbf{w} are also eigenvectors of \mathbf{J} .

Let us see how the perturbation theory of \mathbf{J} can be applied to calculating the matrix elements we require. Equating the derivatives with respect to λ (at $\lambda = 0$) of equations (13) and (19) gives

$$\begin{aligned} \mathbf{e}_0^\dagger (\mathbf{E}\mathbf{I} - \mathbf{J}^{(0)})^{-1} \mathbf{J}^{(1)} (\mathbf{E}\mathbf{I} - \mathbf{J}^{(0)})^{-1} \mathbf{e}_0 \\ = \mathbf{u}_0^{(0)\dagger} (\mathbf{E}\mathbf{I} - \mathbf{H})^{-1} \mathbf{V} (\mathbf{E}\mathbf{I} - \mathbf{H})^{-1} \mathbf{w}_0^{(0)} + \mathbf{u}_0^{(1)\dagger} (\mathbf{E}\mathbf{I} - \mathbf{H})^{-1} \mathbf{w}_0^{(0)} + \mathbf{u}_0^{(0)\dagger} (\mathbf{E}\mathbf{I} - \mathbf{H})^{-1} \mathbf{w}_0^{(1)}. \end{aligned} \quad (29)$$

If we now expand equation (29) in terms of the eigenvectors of \mathbf{H} and \mathbf{J} we find that the residues of the second-order poles at energy E of the left-hand side are just the matrix elements of \mathbf{V} between eigenvectors of \mathbf{H} with energy E , weighted by the projections of the eigenvectors of \mathbf{H} on $\mathbf{u}_0^{(0)\dagger}$ and $\mathbf{w}_0^{(0)}$. An attempt to derive a generalization of equation (29), which produces the matrix element between eigenstates with different energies, gives an extra term, proportional to the difference in energy involving a matrix which is not tridiagonal like $\mathbf{J}^{(1)}$:

$$\begin{aligned} \mathbf{e}_0^\dagger (\mathbf{E}_0\mathbf{I} - \mathbf{J}^{(0)})^{-1} [\mathbf{J}^{(1)} + (\mathbf{E}_1 - \mathbf{E}_0) \mathbf{U}^{(0)\dagger} \mathbf{W}^{(1)}] (\mathbf{E}_1\mathbf{I} - \mathbf{J}^{(0)})^{-1} \mathbf{e}_0 \\ = \mathbf{u}_0^{(0)\dagger} (\mathbf{E}_0\mathbf{I} - \mathbf{H})^{-1} \mathbf{V} (\mathbf{E}_1\mathbf{I} - \mathbf{H})^{-1} \mathbf{w}_0^{(0)} + \mathbf{u}_0^{(0)\dagger} (\mathbf{E}_0\mathbf{I} - \mathbf{H})^{-1} \mathbf{w}_0^{(1)} \\ + \mathbf{u}_0^{(1)\dagger} (\mathbf{E}_1\mathbf{I} - \mathbf{H})^{-1} \mathbf{w}_0^{(0)}. \end{aligned} \quad (30)$$

In general the matrix $\mathbf{U}^{(0)\dagger} \mathbf{W}^{(1)}$ is not tridiagonal and so it requires the full storage of a matrix.

A time-dependent perturbation such as the electric dipole in photo-emission involves transitions between states of different energies and so equation (30) would seem appropriate. However, if we put the problem in a different way, we can obtain the desired result using equation (29). The way of doing this is to include the perturbing force in the unperturbed Hamiltonian as a separate, non-interacting entity, and then make the perturbation the interaction between them.

For example, suppose we wish to calculate matrix elements for photo-emission. Rather than make \mathbf{V} the time-dependent dipole potential, include in the Hamiltonian a photon of energy $\hbar\omega$. There is no electron-photon interaction in \mathbf{H} but \mathbf{V} is now the dipole potential. \mathbf{V} couples states consisting of an electron and photon with states having only an electron. Hence, the transitions conserve the total energy of the system and equation (29) applies. This approach doubles the size of the basis since states for

the electron and photon must be included as well as the purely electronic states. A careful choice of \mathbf{u} and \mathbf{w} avoids introducing degeneracies which \mathbf{V} splits.

Having now established what we wish to calculate, let us proceed to express

$$[\mathbf{G}^{(1)}(E)]_{uw} = \mathbf{e}_0^\dagger (\mathbf{E}\mathbf{I} - \mathbf{J}^{(0)})^{-1} \mathbf{J}^{(1)} (\mathbf{E}\mathbf{I} - \mathbf{J}^{(0)})^{-1} \mathbf{e}_0, \tag{31}$$

in terms of polynomials in E . We can define a sequence of approximants to

$$\check{\mathbf{G}}^{(0)}(E) = (\mathbf{E}\mathbf{I} - \mathbf{J}^{(0)})^{-1}. \tag{32}$$

These are the $\check{\mathbf{G}}_n^{(0)}(E)$ obtained by using $\mathbf{J}_n^{(0)}$ which consists of the first n rows and columns of $\mathbf{J}^{(0)}$, instead of $\mathbf{J}^{(0)}$ in equation (32). The various elements of $\check{\mathbf{G}}_n^{(0)}(E)$ can be calculated by taking the ratio of the cofactor and the determinant of $(\mathbf{E}\mathbf{I}_n - \mathbf{J}_n^{(0)})$. Each of these objects is a polynomial in E , and because $\mathbf{J}^{(0)}$ is tridiagonal, there are simple recurrence relations for these polynomials. I will not prove these relations here, but they all follow from the tridiagonal form of $\mathbf{J}^{(0)}$ and the determinantal formula for elements of the inverse of a matrix.

First we define the inverse elements of $(\mathbf{E}\mathbf{I}_n - \mathbf{J}_n^{(0)})$ as

$$[\check{\mathbf{G}}_n^{(0)}(E)]_{lm} = \mathbf{e}_l^\dagger (\mathbf{E}\mathbf{I}_n - \mathbf{J}_n^{(0)})^{-1} \mathbf{e}_m = Q_n^{(l,m)}(E) / P_n(E). \tag{33}$$

The recurrence relations for $P_n(E)$, which is proportional to the determinant of $(\mathbf{E}\mathbf{I}_n - \mathbf{J}_n^{(0)})$, are:

$$\begin{aligned} P_{-1}(E) &= 0, & P_0(E) &= b_0^{(0)}, \\ P_{n+1}(E) &= (E - a_n^{(0)})P_n(E) / b_n^{(0)} - c_n^{(0)}P_{n-1}(E) / b_{n-1}^{(0)}. \end{aligned} \tag{34}$$

The $Q_n^{(l,m)}(E)$ are proportional to the cofactor of the l, m element of $(\mathbf{E}\mathbf{I}_n - \mathbf{J}_n^{(0)})$. Their recurrence relations are:

$$Q_n^{(l,m)}(E) = 0, \quad n \leq \max(l, m), \tag{35}$$

$$Q_n^{(l,n-1)}(E) = P_l(E) / b_l^{(0)}, \quad l \leq n, \tag{36}$$

$$Q_n^{(n-1,m)}(E) = c_{n-1}^{(0)}Q_{n-1}^{(n-2,m)}(E) / b_{n-1}^{(0)}, \quad m < n - 1 \tag{37}$$

and,

$$Q_{n+1}^{(l,m)}(E) = (E - a_n^{(0)})Q_n^{(l,m)}(E) / b_n^{(0)} - c_n^{(0)}Q_{n-1}^{(l,m)}(E) / b_{n-1}^{(0)}, \quad \max(l, m) < n - 1. \tag{38}$$

There are also relations for $Q_n^{(l,m)}(E)$ at constant n :

$$P_n(E)\delta_{lm} = (E - a_l^{(0)})Q_n^{(l,m)}(E) - b_{l+1}^{(0)}Q_n^{(l+1,m)}(E) - c_l^{(0)}Q_n^{(l-1,m)}(E), \tag{39}$$

$$P_n(E)\delta_{lm} = (E - a_m^{(0)})Q_n^{(l,m)}(E) - c_{m+1}^{(0)}Q_n^{(l,m+1)}(E) - b_m^{(0)}Q_n^{(l,m-1)}(E). \tag{40}$$

The calculation of $[\check{\mathbf{G}}_n^{(1)}(E)]_{00}$ involves elements $[\check{\mathbf{G}}_n^{(0)}(E)]_{l0}$ and $[\check{\mathbf{G}}_n^{(0)}(E)]_{0m}$ only. Because $\mathbf{J}^{(1)}$ is tridiagonal, one needs very few of the products. The formula for $[\check{\mathbf{G}}_n^{(1)}(E)]_{00}$ becomes

$$[\check{\mathbf{G}}_n^{(1)}(E)]_{00} = \left(\sum_{l=0}^{n-1} (a_l^{(1)}Q_n^{(0,l)}(E) + b_l^{(1)}Q_n^{(0,l-1)}(E) + c_{l+1}^{(1)}Q_n^{(0,l+1)}(E))Q_n^{(l,0)}(E) \right) P_n(E)^{-2}. \tag{41}$$

The matrix elements of \mathbf{V} between eigenstates of total energy E_0 are given by the second-order residue at E_0 of $[\check{\mathbf{G}}_n^{(1)}(E)]_{00}$. This is given by

$$\begin{aligned} & \mathcal{R}^{(2)}\{[\check{\mathbf{G}}_n^{(1)}(E)]_{00}, E_0\} \\ &= \left(\sum_{l=0}^{n-1} (a_l^{(1)} Q_n^{(0,l)}(E_0) + b_l^{(1)} Q_n^{(0,l-1)}(E_0) + c_{l+1}^{(1)} Q_n^{(0,l+1)}(E_0)) Q_n^{(l,0)}(E_0) \right) (P'(E_0))^{-2} \quad (42) \end{aligned}$$

where $P_n(E_0) = 0$. The n th approximant has n eigenstates to approximate those of \mathbf{H} . $[\check{\mathbf{G}}_n^{(1)}(E)]_{00}$ approximates $[\check{\mathbf{G}}^{(1)}(E)]_{00}$ in the sense that the contour integrals around the spectral points of each, of any polynomial up to degree $2n - 1$ are the same. However when calculating densities in energy, the discrete spectrum of the n th approximant is often too spikey. Some method of smoothing must be used. Any method which does not alter the property of the integrals over polynomials is fine.

In practice \mathbf{J}_n can be calculated where n is less than several hundred. The matrix elements of \mathbf{J} fall into two regions. Those for small n which give information about the distribution of spectral weight with energy and for large n which determine band limits and van Hove singularities. There is no sharp boundary between regions and the distinction is a matter of judgment. \mathbf{J}_n must be calculated for n sufficiently large to reach the asymptotic region. Once this is done, a smooth spectrum may be obtained by a continuation of \mathbf{J}_n based on knowledge of the singularities, or by obtaining bounds on the indefinite integral of $[\check{\mathbf{G}}_n^{(0)}(E)]_{00}$ or $[\check{\mathbf{G}}_n^{(1)}(E)]_{00}$ by varying a_{n+1} , b_{n+1} , or c_{n+1} and then differentiating the bound to give a smooth density. I hope to discuss this in a future paper.

A number of simplifications are possible in special cases. Firstly, b_n and c_n can always be chosen to be the same requiring one-third fewer numbers to store. When \mathbf{H} and \mathbf{V} are self-adjoint and \mathbf{u}_0 and \mathbf{w}_0 are identical, then \mathbf{u}_n and \mathbf{w}_n are all identical and hence the store required is halved. Also, a basis may be chosen so that \mathbf{H} and \mathbf{V} are real. If \mathbf{u}_0 and \mathbf{w}_0 are real, this also saves store. It should be noted that even without the savings mentioned above, at most a small number of vectors need be stored at any time. Hence, this perturbation theory retains the computational advantages of the recursion method.

The perturbation theory in the first part of this section developed \mathbf{J} to arbitrary order in λ . For the calculation of matrix elements, we require \mathbf{J} only to first order in λ . However, it is conceivable that one would wish higher-order terms in the series for the resolvent

$$\check{\mathbf{G}}(E) = (E\mathbf{I} - \mathbf{J})^{-1}. \quad (43)$$

These are easy to calculate by a generalization of the method for $[\check{\mathbf{G}}^{(1)}(E)]_{00}$. It only involves more products of the $[\check{\mathbf{G}}^{(0)}(E)]_{lm}$ which are expressed in terms of the $P_n(E)$ and the $Q_n^{(l,m)}(E)$. These terms may be related to terms in $\mathbf{G}(E)$ by the transformation $\mathbf{U}^{(0)\dagger}$ and $\mathbf{W}^{(0)}$.

3. Conclusion

Unfortunately there are no non-trivial systems where this perturbation theory can be worked analytically. The simplest system which absorbs by means of electrons making interband transitions is an extended polymer or linear chain. Although the recursion method can be performed to give the density of states analytically for this case,

calculation of the optical absorption cannot be done analytically. Since the computations involved in doing the optical absorption of the linear chain are the same as for optical absorption of any other system, the application of this perturbation theory will have to wait for a future paper where there is sufficient space.

The recursion method for calculating the density of states has been applied to a wide variety of systems and large numbers of atoms. This perturbation theory is more complicated than the recursion method, however, to calculate matrix elements for say, photo-emission, requires only four times the store necessary for the density of states. The factor of four comes from the necessity of including in the Hamiltonian electronic states both with and without the photon and the storing of vectors to both zeroth and first order in the perturbation. Although a factor of four is lost in the number of atoms in the cluster, one only loses a factor of about two in the time for the calculation.

The advantages of this perturbation theory are similar to those of the recursion method. First, the arrangement of atoms is arbitrary, so that the method applies where there are no bands. Second, the method does not involve complete diagonalization of the Hamiltonian. The physical information is extracted from what might be called a 'partial diagonalization'. Thirdly, the results converge rapidly with increasingly complete 'partial diagonalization' but always before complete diagonalization, and there is an error theory.

Acknowledgments

The author would like to gratefully acknowledge the support for this work of the Stokes Research Fellowship at Pembroke College, Cambridge, and thank Dr V Drchal for a critical reading of the manuscript.

References

- Haydock R 1974 *J. Phys. A: Math., Nucl. Gen.* **7** 2120-4
Haydock R, Heine V and Kelly M 1975 *J. Phys. C: Solid St. Phys.* **8** 2591-605