THE RECURSIVE SOLUTION OF THE SCHRÖDINGER EQUATION

Roger HAYDOCK

Cavendish Laboratory, Madingley Road, Cambridge, CB3 0HE, England

A new approach to the computational solution of the Schrödinger equation is based on the partial transformation of the Hamiltonian to a tridiagonal matrix. The method is especially suited to tight-binding Hamiltonians encountered in solid state physics and permits of the order of 10^4 degrees of freedom to be included in a calculation. Independent particle Green functions are calculated naturally from the partially tridiagonalized Hamiltonian. These lead to simple computation of small energy differences, binding energies, transition matrix-elements and other useful quantities.

1. Introduction

Computation of the solutions of the time-independent Schrödinger equation,

$$\mathcal{H}\psi = E\psi, \tag{1}$$

where $\mathcal{H}$ is the Hamiltonian operator and ψ is a wavefunction, is one of the major uses of the computer in physics. In every area of physics where quantum mechanics describes the behaviour of systems, eq. (1) must be solved in order to compare theories about interactions with experimental observations.

In solid state physics the Schrödinger equation is needed to describe the motion of electrons, traditionally, in crystals with various cellular potentials. Using the translational symmetry, Bloch's theorem reduces eq. (1) to the problem of diagonalizing 'small' matrices. The result of these calculations is the electronic band-structure: an efficient summary of the solutions of the Schrödinger equations in a crystal.

The success of quantum mechanics in explaining the electronic properties of simple crystals has turned interest toward more complicated crystals, surfaces, defects and amorphous materials, systems where Bloch's theorem does not reduce eq. (1) to a small matrix. In these cases the problem is made computable by the use of some discrete basis set, such as a tight-binding basis, for representation of the electronic wavefunctions. These approaches often lead to representation of eq. (1) as an infinite matrix equation which is solved approximately in a subspace of vectors with dimensionality of the order 10^4.

The subject of this article is a new approach to the solution of the Schrödinger equation when it may be represented as a large sparse matrix equation.

Before solving the Schrödinger equation, there is the problem of how to present the results. Although an enumeration of eigenvalues and wavefunctions is possible in principle, one wants a far more economical summary in a form which bears directly on the physical questions. Such a form is the independent-particle Green function or various elements of the resolvent matrix of $\mathcal{H}$ [1]. It is the calculation of this object toward which the method is oriented.

The Lanczos method [2] is one of several methods for transforming a finite matrix to tridiagonal or Jacobi form. It is based on the work of many mathematicians on the solution of the classical moment problem [3], and computationally it has the advantage that it requires storage of only two vectors rather than whole matrices. Its usefulness as a computational method was doubted until 1972 when numerical investigations of the symmetric version by Paige [4] showed remarkable stability.

The recursion method [5] is related to the solution of the moment problem and the Lanczos method in that elements of a Jacobi matrix are calculated from a three term recurrence relation. It differs from the Lanczos method, and this distinction is not just a nicety of academic priority, in that the Hamiltonian is only partially tridiagonalized and an infinite Jacobi matrix is either implicitly or explicitly constructed. The Green function is then computed in the spirit of the solution to the moment problem. There are many

other variants of the Lanczos and recursion methods which have been used in other branches of physics, among which are statistical mechanics [6] and nuclear physics [7].

The remainder of this paper is devoted to a brief description of the main steps in the recursive solution of the Schrödinger equation. The next section presents the concept of a chain model in quantum mechanics and shows how any quantum mechanical model can be numerically transformed to a chain model. Section 3 discusses Green functions and shows how these may be computed from the chain model. The fourth section mentions a number of other computations which are possible within the recursion method.

For a fuller account of these methods as well as a discussion of their applications in solid state physics, the reader is referred to volume 35 of the Solid State Physics series [8] and references therein.

2. Construction of the chain model

In this section I show how an arbitrary quantum mechanical model can be reduced to a special model, namely the chain model. Although the form of this model is the same for all physical systems, of course the parameters are different and reflect the physics of the system being modelled.

The chain model consists of a set of basis states or orbitals $\{u_n\}$, which spans the states of the system which are accessible from u_0. In terms of these orbitals the Hamiltonian is described by parameters $\{a_n\}$ and $\{b_n\}$ so that

$$u_n = a_n u_n + b_{n+1} u_{n+1} + b_n u_{n-1}. \tag{2}$$

There is no state u_{-1}, so eq. (2) is truncated when $n = 0$. This model is displayed pictorially in fig. 1. The above model is symmetric in that each of the parameters b_n appears in two relations, namely the ones for u_{n-1} and u_n, so that the parameter connecting u_n and u_{n-1} is the same in both directions.

The chain model expresses a hierarchical relationship between its basis orbitals. A system, initially in state u_0, cannot reach state u_n without first passing through the intervening states on the chain, $u_1, u_2, ..., u_{n-1}$. Also the chain is complete in the sense that its basis orbitals span all the states accessible to the
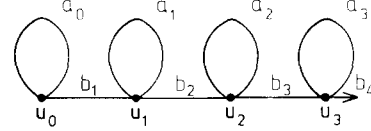

Fig. 1. A graphic representation of the chain model including the states $\{u_n\}$ and the recurrence parameters $\{a_n\}$ and $\{b_n\}$.

system from u_0. Taken together these properties imply the picture of a system initially in state u_0 which, depending on the parameters $\{a_n\}$ and $\{b_n\}$ decays into a superposition of states including u_n, with n increasing with time. In short, the system diffuses along the chain from u_0.

Given a quantum mechanical model expressed in terms of a linearly independent basis $\{\phi_\alpha\}$, by a Hamiltonian matrix H and an overlap matrix S, we can construct a chain model from an initial orbital u_0. The Hamiltonian matrix H is determined from the Hamiltonian operator by the relation

$$\mathcal{H}\phi_\alpha = \sum_\beta [H]_{\beta\alpha}\phi_\beta. \tag{3}$$

Note that, when the basis is not orthogonal, the elements of H are not what are usually called matrix-elements of the Hamiltonian. The overlap matrix S is just the matrix of overlap integrals,

$$S_{\alpha\beta} = \int \phi_\alpha^* \phi_\beta \, d\tau, \tag{4}$$

where the integral is over the configuration space of the system.

The choice of initial orbital u_0 determines what aspect of the physics will be described by the chain model. Each choice of u_0 leads to a distinct chain so by the choice of u_0 we limit the scope of the resulting calculation. The initial orbital u_0 must satisfy the mathematical requirement that $u_0^\dagger S H^n u_0$, the diagonal matrix element of H^n for u_0, must be finite for all finite n. The requirement translates into the physical condition that u_0 must not contain too large components of too high energy states.

The second element of the chain is the normalized component of $H u_0$ which is orthogonal to u_0. The projection of $H u_0$ on u_0,

$$a_0 = u_0^\dagger S H u_0 \tag{5}$$

is the first parameter of the model, and the magnitude

of the remainder is the second parameter,

$$b_1 = |(H - a_0)u_0|.$$

The second orbital is just the normalized remainder,

$$u_1 = (H - a_n)u_0/b_1 . \tag{6}$$

This completes the first step of the construction.

The general step of the transformation is similar to the first step except that u_{n+1} is the normalized component of Hu_n which is orthogonal to all the previous orbitals. Suppose we have already found $u_0, u_1, ...,$ $u_n, a_0, a_1, ..., a_{n-1}$, and $b_1, b_2, ..., b_n$, then we can show that $Hu_n - b_n u_{n-1}$ is orthogonal to $u_0, u_1, ...,$ u_{n-1}. The projection of Hu_n on u_n is

$$a_n = u_n^\dagger SHu_n , \tag{7}$$

the next parameter of the model, and the magnitude of the component of Hu_n orthogonal to the known orbitals is the parameter,

$$b_{n+1} = |(H - a_n)u_n - b_n u_{n-1}| . \tag{8}$$

The next orbital is thus,

$$u_{n+1} = ((H - a_n)u_n - b_n u_{n-1})/b_{n+1} . \tag{9}$$

This procedure continues until b_N is found to be zero, or until one has included enough degrees of freedom in the chain to answer the physical question.

The above procedure has been programmed in FORTRAN by Nex [9].

When H can be stored in some condensed form or when Hx can be computed directly, the above procedure has a number of computational advantages. Apart from H, it requires store for only two vectors at a time. This makes it possible to use a Hamiltonian which effectively includes tens of thousands of degrees of freedom. Although (among other problems) this procedure is a slow way of reducing the whole of H to Jacobi form, the physics is determined by only a fraction of the whole chain and so it is an extremely fast way of calculating physical properties to a given accuracy.

A perplexing property of this transformation is its extreme numerical stability. It can be observed from a computation that the rounding error leads to a cumulative loss of orthogonality among the $\{u_n\}$. However, when applied in the Lanczos method, the same computation leads to eigenvalues correct to the machine accuracy [4]. Thus it seems cumulative loss

of orthogonality does not produce a cumulative loss of accuracy in the eigenvalues.

The explanation of this fortunate property is that the orthogonality of the chain basis has nothing to do with the accuracy of the model. The eigenvalues and eigenvectors of an operator do not depend on the scalar product. Hence as long as the basis orbitals constructed for the chain are linearly independent, there exists a scalar product under which they are orthogonal. So unless at some stage the basis orbitals cease to be linearly independent, the chain has the full machine accuracy.

Because of rounding error the chain basis will not be orthogonal with respect to the scalar product under which the Hamiltonian is self-adjoint, so, in general, the chain model cannot be fully symmetric. At the stage where, for a finite model, the degrees of freedom coupled to the chain are exhausted, the next chain orbital will be linearly dependent on the preceding ones. The chain can be terminated by an un-symmetric link (where the parameter connecting u_{n-1} to u_n is different from that connecting u_n to u_{n-1}), to make b_{n+1} zero. This terminating procedure is described in ref. [8] and makes the chain model exact to machine precision.

Other methods of correcting for rounding error, such as reorthogonalization, corrupt the basic three-term recurrence, thereby either introducing new errors, if the resulting model is interpreted as a chain, or simply producing a lower triangular representation.

3. Green's functions

The Green functions or matrix-elements of the resolvent are an economical and convenient way of describing the solution of the Schrödinger equation for electrons in solids with little or no symmetry. High symmetry, such as in simple crystals, generates quantum numbers, related to the symmetry, which almost exhaustively classify the eigenstates. This is the origin of band structure. However, in systems with less symmetry, quantum numbers become more complicated or there are more states with the same quantum number and this approach becomes useless. The Green function is an entirely different way of describing eigenvalues and eigenstates. It combines in one function of energy, E, a magnitude which tells

the overlap of eigenstates at E with a particular state and an imaginary part which describes the density of eigenvalues at E.

An example of a Green function is the local density of states $n_0(e)$. It is related to the diagonal Green function or resolvent element for u_0 of the Hamiltonian,

$$G_0(E) = u_0^\dagger S[E - H]^{-1} u_0 . \tag{10}$$

The local density of states on u_0 is,

$$n_0(E) = - \lim_{\epsilon \to 0^+} \text{Im}\{G_0(E + i\epsilon)\}/\pi . \tag{11}$$

Its physical relevance can be better seen from the spectral representation,

$$n_0(E) = \sum_\alpha |u_0^\dagger S \psi_\alpha|^2 \, \delta(E - E_\alpha) , \tag{12}$$

where the sum is over an orthonormal set of eigenvectors of H, ψ_α, and E_α is the eigenvalue of ψ_α. Thus $n_0(E)$ is the intensity of the eigenvector of H at energy E on u_0. This is related directly to experimental physical quantities such as photoemission and indirectly to binding energies, etc:

The Green function is also directly related to the chain model [8]. In terms of the chain parameters,

$$G_0(E) = 1/(E - a_0 - b_1^2/(E - a_1$$
$$- b_2^2/(E - a_2 - ...))) . \tag{13}$$

This is the continued-fraction representation and it can be used to compute the local density of states by, in the finite case, introducing a small imaginary part to E, or in the infinite case, replacing an infinite part of the fraction by an analytic approximation. An example of a Green function is that of the inifnite chain model with $a_n = 0$ and $b_n = b$. The analytic expression is

$$G_0 = \begin{cases} (E - (E^2 - 4b^2)^{1/2}/2b^2, & \text{Re}\{E\} > 0, \\ (E + (E^2 - 4b^2)^{1/2}/2b^2, & \text{Re}\{E\} < 0, \end{cases} \tag{14}$$

for E in the upper half of the complex plane and for the usual definition of the square root. The real and imaginary parts of this function are sketched in fig. 2.

In many physical applications one wishes to calculate a chain-model for some finite approximation to the Hamiltonian and then calculate from that an approximation to the Green function for the infinite system. This procedure poses the question: in what
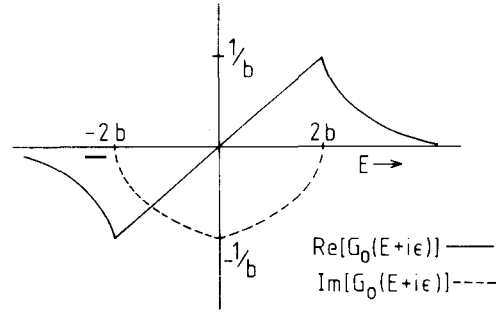

Fig. 2. The real and imaginary parts of $G_0(E + i\epsilon)$ for a constant chain.

sense does the Green function for a finite system approach that for an infinite system? The answer is that the indefinite integrals of the local density of states times functions, all of whose derivatives have the same sign, converge in the usual sense while along the real energy axis, the Green function does not converge. Thus the Green function only converges in the mean along the real energy axis.

For many applications the local density of states is the quantity of particular physical interest. For any finite system, the local density of states is a collection of delta functions whereas for infinite systems and especially metallic ones, the local density of states is not only continuous but smooth. One wishes to produce from the chain model for the finite system an approximation to the local density of states of the infinite system. There are two basic approaches to this problem.

The first approach is that of fitting the asymptotic part of an infinite chain to the finite part which can be calculated exactly. When H is sparse as for a tight-binding model where u_0 is localized in the centre of a finite cluster of atoms, a number of chain parameters can be calculated exactly for the infinite cluster. The rest of the chain model is approximated by a chain whose Green function is known. The difficulty of this is that for good results the infinite approximate chain must have Van Hove singularities in exactly the right places. An example of this method is terminating a computed chain with a constant chain whose band limits have been fitted to the system.

The second approach to this problem is to compute a number of levels of the chain for the finite system and smooth the resulting local density of

states. A particularly attractive way of doing this, which takes advantage of convergence in the mean, is to compute the bounds of the indefinite integral of the local density of states, take the mean at each energy and differentiate the result [8].

The densities of states calculated with increasing numbers of levels from the chain show behaviour typical of an asymptotic approximation. At first the density of states shows increasing amounts of smooth structure characterized by densities of states which are little changed by the increase of one level, however, after a certain point the change in density of states with number of levels begins to diverge. The most accurate density of states is the one for which the addition of one level makes the least difference.

4. Further applications

There are a wide variety of applications of the chain model beyond the calculation of Green functions and local densities of states, mentioned in the previous section. In this section I will introduce briefly a number of other uses of the model. For further information, the reader is referred to the volume of Solid State Physics [8].

The position and nature of Van Hove singularities in the Green function are related to the asymptotic behaviour of the chain model, the variation of chain parameters with n for large n. Many of the singularities are determined by the first-order variation of the chain parameters from their limiting values. The theory of this was originated by Hodges [10] and leads to the result that the change in the local density of states is the Fourier series whose coefficients are the change in chain parameters.

An important application of the chain model is the determination of small energy differences. The first example of this is to determine the realtive stability of very similar structures. This often depends on a subtle relation between geometry, the orbital symmetry of the valence electrons, and the electron-per-atom ratio. Comparison of structural energy involves integrals of the eigenstate energy over the total density of states. This can be done very accurately from a chain model using the Gaussian quadrature generated by the continued fraction.

The second application of the chain model to cal-

culation of small energy differences is in the calculation of binding energies. When an atom adsorbs onto the surface of a metal, it shifts the infinitely many electronic levels by infinitesimal amounts adding up to a finite binding energy. This can be calculated from the logarithm of the ratio of Green functions before and after adsorption.

Another useful technique is a perturbation theory of the chain model itself. If a Hamiltonian differs slightly from one whose chain model is known, then the chain parameters of the new Hamiltonian differ only slightly from those already known. These differences can be computed directly from the known chain and its basis in analogy with other perturbation theories. One particularly useful application of this is to the asymptotic properties of the chain. By this means, changes in the singularities of the Green function can be related directly to changes in the Hamiltonian. Another application of this is to the calculation of transition matrix-elements. A small perturbation of the system causes it to make transitions between various stationary states. This perturbation theory enables one to calculate the probability of such transitions.

5. Conclusion

The recursive solution of the Schrödinger equation provides a framework for the efficient computation of a wide variety of physical quantities in quantum mechanics. The idea of transforming all quantum models to a standard model allows programs to be very general in their application. The large body of useful mathematical properties of the chain model permit us to be precise about errors and to derive from the chain model the many physical quantities measured by experiment.

As well as being a computational tool, the recursion method gives insight into the properties of quantum models. Expressed in terms of a chain, one can see how the parameters affect the rate of diffusion of the system among the chain basis states. Knowledge of the basis states in turn allows one to relate diffusion on the chain to the diffusion of the system among the states of the original model. The intuitive understanding gained from this is valuable.

There are a number of directions in which the

recursive solution can still progress. Most applications have so far been to systems modelled by tight-binding Hamiltonians. In principle the methods also apply to nearly-free-electron models, and models which hybridize the two. Much work is needed to extend the method to these systems.

Another area for further work is the investigation of the singularity structure resulting from chains with randomness in their parameters. Amorphous and glassy materials are homogeneously random. This means that the chain models for electrons in these materials are also random. There is still much to be discovered in the analysis of random chains.

References

[1] J. Friedel, Advan. Phys. 3 (1954) 446.
[2] C. Lanczos, J. Res. Natl. Bur. Std. 45 (1950) 255.
[3] N.I. Akhiezer, The classical moment problem (Oliver and Boyd, London, 1965);
J.A. Shohat and J.D. Tamarkin, The problem of moments, Am. Math. Soc., Mathematical Surveys I, Rev. ed., Providence, RI (1950).
[4] C.C. Paige, J. Inst. Math. Appl. 10 (1972) 373.
[5] R. Haydock, V. Heine and M.J. Kelly, J. Phys. C8 (1975) 2591.
[6] H. Mori, Japan. Prog. Theor. Phys. 34 (1965) 399.
[7] R.R. Whitehead, A. Watt, B.J. Cole and D. Morrison, in: Adv. in nucl. phys., Vol. 9, eds. M. Baranger and E. Vogt, (Academic Press, New York, 1977).
[8] D.W. Bullett, R. Haydock, V. Heine and M.J. Kelly, Solid state physics, vol. 35 (Academic Press, New York, 1980).
[9] C.M.M. Nex, The Cambridge Recursion Library.
[10] C.L. Hodges, J. de Phys. Lett. 38 (1977) 187.