A Cluster Method for Spectral Properties of Correlated Electrons

David Sénéchal[a]

[a]Département de physique, Université de Sherbroke, Québec, Canada
David.Senechal@USherbrooke.CA

Some technical aspects of Cluster Perturbation Theory (CPT) are presented. CPT is an approximation scheme for calculating the one-particle Green function of the Hubbard model of interacting electrons on a lattice. It proceeds by (i) dividing the lattice into identical clusters and calculating the Green function by exact diagonalization on a given cluster (ii) extending this Green function to the whole lattice by using Dyson's equation. The exact diagonalization technique for the Hubbard model is explained in some detail. Embedding the cluster in the full lattice is then discussed, and two alternate emdedding schemes are briefly described.

Quelques aspects techniques de la théorie des perturbations inter-amas (CPT) sont présentés. La CPT est une méthode approximative pour le calcul de la fonction de Green à une particule du modèle de Hubbard, qui décrit des électrons en interaction sur un réseau cristallin. Elle procède par (i) dallage du réseau en amas identiques et calcul de la fonction de Green par diagonalisation exacte sur un amas (ii) extension de ce résultat au réseau entier via l'équation de Dyson. La méthode de diagonalisation exacte du modèle de Hubbard est expliquée en détail. L'insertion de l'amas dans le réseau est aussi discutée, et deux méthode d'insertion alternatives sont décrites.

1 Introduction

1.1 Context

The quantum theory of electrons in solids is a complicated many-body problem. Even if we accept that core electrons are well localized within atoms, solving the Schrödinger equation for the outer electrons of every atom in a crystal is impossible and practical approximation schemes must be developed. The most common starting point is the independent-electron approximation : each electron is moving in the average potential created by all the others (including the exchange potential), but they are otherwise independent. In principle, the best approximate solution within this scheme is obtained by solving the Hartree-Fock equations. In practice, one uses even simpler approximations, such as the local density approximation (LDA) of Density Functional Theory (DFT, for a review, see [1]). Calculating the band structure of solids within that approximation has become an almost standard procedure, and very efficient programs are available for this task.

However, the very notion of band structure is approximate. It merely provides a basis of states in terms of which the two-body interaction between electrons is, in some sense, minimal. Basic semiconductor Physics is based largely on this picture. So is our understanding of ordinary metals, provided the residual interactions are taken into account in a perturbative way. The Landau theory of Fermi liquids aims precisely at preserving the physical picture of independent electrons while taking into account the main effects of residual electron-electron interactions [2]. However, in a large class of materials – called strongly correlated – those residual interactions are too strong and the picture of independent electrons breaks down completely (see, e.g., [3]). Such are the high-temperature superconductors, organic conductors, colossal magneto-resistance materials, and so on. Strongly correlated materials are typically very anisotropic (quasi two-dimensional or quasi one-dimensional), because in low dimensions quantum or thermal fluctuations suppress the onset of broken symmetry phases that would otherwise replace the high-temperature metallic phase. The latter evolves instead towards a quantum state poorly described in terms of one-electron states.

1.2 The Hubbard model

The most basic model commonly used to study the behavior of electrons with strong residual interactions in the Hubbard model [4]. It is usually expressed in the Wannier basis and assumes that the residual electrostatic repulsion between electrons is purely local, i.e., felt only between electrons states located on the same atom (or lattice site). The Hamiltonian reads

$$H = -\sum_{i,j,\sigma} t_{ij} c^\dagger_{i\sigma} c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} \tag{1}$$

where $c_{j\sigma}$ annihilates an electron of spin σ ($\uparrow$ or $\downarrow$) in a Wannier state centered on site j of the lattice and $n_{j\sigma} = c^\dagger_{j\sigma} c_{j\sigma}$ is the number of electrons in that state. U is the strength of on-site electrostatic repulsion and t_{ij} is the probability amplitude for an electron to hop from site j to site i. If $U = 0$, this model reduces to the tight-binding description of a band of independent electrons, the dispersion relation being

$$\varepsilon(\mathbf{k}) = \frac{1}{N} \sum_{i,j} t_{ij} \mathrm{e}^{i\mathbf{k}\cdot(\mathbf{r}_i - \mathbf{r}_j)} \tag{2}$$

where N is the number of lattice sites ($N \to \infty$) and $\mathbf{r}_i$ the position of site i. On the other hand, when U is large compared to t_{ij}, the strong repulsion suppresses states with more than one electron per site, and this favors an insulating state, at least when the system contains one electron per site on average (density $n = 1$). Thus, the degree of correlations can be tuned by varying the ratio U/t_{ij}.

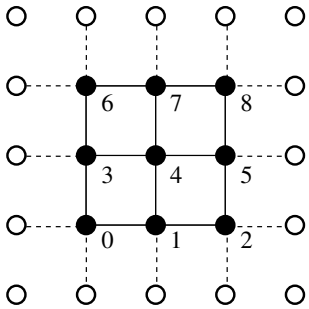

Figure 1. An example of 9-site cluster on a square lattice. The numbering of sites used in practice is illustrated. Full lines indicate intra-cluster hoppings, and dashed-line inter-cluster hoppings.

1.3 Cluster Perturbation Theory

In this paper we will explain in some detail the numerical aspects of an approximate method for calculating the one-particle properties of the Hubbard model. This method, called Cluster Perturbation Theory (or CPT [5,6]) is based on the exact diagonalization of the Hubbard model on a small cluster of sites (typically a dozen), and on a particular way to extend the results of a cluster to the full infinite lattice. We will restrict the discussion to some methodological aspects that are not covered in detail in [5,6].

The quantity we wish to calculate is the one-particle Green function $G(\mathbf{k}, z)$, defined as

$$\begin{aligned} G(\mathbf{k},z) &= G_e(\mathbf{k},z) + G_h(\mathbf{k},z) \\ G_e(\mathbf{k},z) &= \langle\Omega|c(\mathbf{k})\frac{1}{z-H+E_0}c^\dagger(\mathbf{k})|\Omega\rangle \\ G_h(\mathbf{k},z) &= \langle\Omega|c^\dagger(\mathbf{k})\frac{1}{z+H-E_0}c(\mathbf{k})|\Omega\rangle \end{aligned} \quad (3)$$

where $c(\mathbf{k})$ and $c^\dagger(\mathbf{k})$ respectively annihilate and create an electron of wavevector $\mathbf{k}$ (the spin index of these operators is suppressed). E_0 is the ground-state energy. From the Green function many other quantities can be derived, such as the spectral function

$$A(\mathbf{k},\omega) = -2\lim_{\eta\to 0^+} \mathrm{Im}\, G(\mathbf{k},\omega+i\eta) \quad (4)$$

which is the probability for an electron of momentum $\hbar\mathbf{k}$ to have an energy $\hbar\omega$. The negative-frequency part of this function (coming from G_h) is experimentally accessible by angle-resolved photoemission spectroscopy (ARPES).

Cluster perturbation theory proceeds by first tiling the infinite lattice into identical clusters of L sites ($L \leq 16$; see for instance, Fig. 1. The hopping amplitude of the Hubbard model can then be labelled as t_{ab}^{mn}, where the subscripts refer to sites within a cluster, and the superscripts to clusters. The terms with $m = n$, together with the local interaction of the Hubbard model, define a model local to a given cluster, that can be solved numerically by a suitable exact diagonalization technique, as described in the next section. 'Solved' here means that the real-space Green function

$$\begin{aligned} G_{ab}^{(c)}(z) &= G_{ab,e}^{(c)}(z) + G_{ab,h}^{(c)}(z) \\ G_{ab,e}^{(c)}(z) &= \langle\Omega|c_a\frac{1}{z-H+E_0}c_b^\dagger|\Omega\rangle \\ G_{ab,h}^{(c)}(z) &= \langle\Omega|c_b^\dagger\frac{1}{z+H-E_0}c_a|\Omega\rangle \end{aligned} \quad (5)$$

(6)

is calculated. The superscript (c) stands for 'cluster' : this Green function describes the motion of electrons on a small cluster only.

The hopping terms with $m \neq n$, that connect different clusters, are treated approximately: We write

$$t_{ab}^{mn} = t_{ab}^{(c)}\delta_{mn} + V_{ab}^{mn} \quad (7)$$

where $t_{ab}^{(c)}$ is the hopping amplitude within a cluster, which is accounted for exactly, and V_{ab}^{mn} the remainder, corresponding to inter-cluster hopping. In terms of matrices with indices (a, m) and (b, n), our approximation consists in taking the full Green function to be

$$\hat{G}^{-1} = \hat{G}^{(c)-1} - \hat{V} \quad (8)$$

This can be viewed as the lowest order approximation in a formal perturbation expansion in powers of $|V|/U$, where $|V|$ stands for the typical size of an inter-cluster hopping amplitude [6]. It can also be viewed, more heuristically, as an approximation where the full self-energy $\hat{\Sigma}$ of the model is replaced by the cluster self-energy [7]. Recall that the self-energy is defined by Dyson's equation

$$\hat{G}^{-1} = \hat{G}_0^{-1} - \hat{\Sigma} \quad (9)$$

where $\hat{G}_0$ is the noninteracting Green function ($U = 0$):

$$\hat{G}_0^{-1} = \omega - \hat{t} = \omega - \hat{t}^{(c)} - \hat{V} \quad (10)$$

Applying Dyson's equation to the exact, cluster Green function gives

$$\hat{G}^{(c)-1} = \omega - \hat{t}^{(c)} - \hat{\Sigma}^{(c)} \quad (11)$$

and Eq. (8) may then be written as

$$\hat{G}^{-1} = \hat{G}_0^{-1} - \hat{\Sigma}^{(c)} \quad (12)$$

Eq. (8), or its equivalent form (12), is the basic equations of Cluster Perturbation Theory.

2 Exact Diagonalization

The numerically intensive part of the problem, at least as far as memory is concerned, is the exact diagonalization, i.e. the calculation of $G_{ab}^{(c)}$. We will describe the method used in some detail, even though it is fairly straightforward, because the nuts and bolts, at least applied to the Hubbard model, are not necessarily easy to find in the literature.

2.1 Coding of the states

The first step in the exact diagonalization procedure is to define a coding scheme for the quantum basis states. A basis state may be specified by the occupation number $n_{a\sigma}$ ($= 0$ or 1) of electrons spin σ at the site a ($a = 1, \ldots, L$). Thus the string of $n_{a\sigma}$'s forms the binary representation of a 32-bit unsigned integer (if $L \leq 16$) and represents the state

$$(c^\dagger_{1\uparrow})^{n_{1\uparrow}} \cdots (c^\dagger_{L\uparrow})^{n_{L\uparrow}} (c^\dagger_{1\downarrow})^{n_{1\downarrow}} \cdots (c^\dagger_{L\downarrow})^{n_{L\downarrow}} |0\rangle \qquad (13)$$

where the order in which the creation operators are applied is a matter of convention, but important.

There are 2^{2L} such states, but not all are relevant, since the Hamiltonian of the Hubbard model is block-diagonal : the number of electrons of a given spin ($N_\uparrow$ and $N_\downarrow$) are conserved quantities, which commute with the Hamiltonian H. The exact diagonalization is to be performed in a sector (i.e. a subspace) of the total Hilbert space with fixed values of $N_\uparrow$ and $N_\downarrow$. The number of states with these fixed values, i.e. the dimension of the relevant sector, is

$$d = \left(\frac{L!}{N_\uparrow!(L-N_\uparrow)!} \right) \left(\frac{L!}{N_\downarrow!(L-N_\downarrow)!} \right) \qquad (14)$$

Note that the ground state $|\Omega\rangle$ of the Hamiltonian generally belongs to the sector $N_\uparrow = N_\downarrow$. The density $n = (N_\uparrow + N_\downarrow)/L$ is fixed in a given calculation; it can only take a limited number of values (although the use of "superclusters" gives us more flexibility in this matter [6]).

Since the number of states is less than 2^{2L}, one needs a lookup table that gives the index of a given state from its binary representation $(n_{1\uparrow} \cdots n_{L\downarrow})$. This table can be deleted as soon as the Hamiltonian matrix has been built, before allocating the Lanczös vectors.

The next step is to construct the Hamiltonian matrix in the chosen sector. The particular structure of the Hubbard model Hamiltonian brings a considerable simplification. Indeed, the Hilbert space is the tensor product of a space of 'up' electrons with a space of 'down' electrons ($\mathcal{H} = \mathcal{H}_\uparrow \otimes \mathcal{H}_\downarrow$), of dimension $d = d_\uparrow d_\downarrow$. The Hamiltonian has the form

$$H = K_\uparrow \otimes 1 + 1 \otimes K_\downarrow + V_{\text{int.}} \qquad (15)$$

where $K_\uparrow$ only acts on the up electrons and $K_\downarrow$ on the down electrons, and where the Coulomb repulsion term $V_{\text{int.}}$ (the second term of (1)) is diagonal in the basis chosen. Thus storing the Hamiltonian in memory is not a problem : the diagonal $V_{\text{int.}}$ is stored (an array of size d), and the kinetic energy K_σ (a matrix having a small fraction of d_σ^2 elements) is stored in sparse form. Constructing this matrix, formally expressed as

$$K = \sum_{a,b} t_{ab} c^\dagger_a c_b \ , \qquad (16)$$

needs some care with the signs. Basically, two states $|\psi\rangle$ and $|\psi'\rangle$ are connected with this matrix if their binary representations $(n_1, n_2, \ldots, n_L)$ (see Eq. (13)) differ at two positions i and j. The matrix element is then $(-1)^{M_{ab}} t_{ab}$, where M_{ab} is the number of occupied sites between a and b, i.e., assuming $a < b$,

$$M_{ab} = \sum_{c=a+1}^{b-1} n_c \qquad (17)$$

For instance, the two states (10010110) and (10011100) with $L = 8$ are connected with the matrix element $+t_{46}$, where the sites are numbered from 0 to $L-1$.

2.2 The Lanczös algorithm

Next, one must apply the exact diagonalization method per se, using the Lanczös algorithm. Calculating the Green function (5) requires a two-step procedure.

2.2.1 First step: the ground state

In the first step, the ground state $|\Omega\rangle$ is obtained from two consecutive sequences of Lanczös iterations : the first sequence to find the tridiagonal representation of the Hamiltonian, and the second one to find the ground state itself. The sequence of states built during the Lanczös procedure obey a two-step recursion relation:

$$|\phi_{n+1}\rangle = H|\phi_n\rangle - a_n|\phi_n\rangle - b_n^2|\phi_{n-1}\rangle \qquad (18)$$

where

$$a_n = \frac{\langle\phi_n|H|\phi_n\rangle}{\langle\phi_n|\phi_n\rangle} \qquad b_n^2 = \frac{\langle\phi_n|\phi_n\rangle}{\langle\phi_{n-1}|\phi_{n-1}\rangle} \qquad (19)$$

and we set the initial conditions $b_0 = 0$, $|\phi_{-1}\rangle = 0$. At any given step, only three state vectors are kept in memory (ϕ_{n+1}, ϕ_n and ϕ_{n-1}). In the basis of normalized states $|n\rangle = |\phi_n\rangle/\sqrt{\langle\phi_n|\phi_n\rangle}$, the Hamiltonian has the tridiagonal form

$$H = \begin{pmatrix} a_0 & b_1 & 0 & 0 & \cdots & 0 \\ b_1 & a_2 & b_2 & 0 & \cdots & 0 \\ 0 & b_2 & a_3 & b_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & a_N \end{pmatrix} \qquad (20)$$

The initial state $|\phi_0\rangle$ is chosen at random. The iteration sequence terminates when either (i) the lowest eigenvalue of the tridiagonal Hamiltonian converges, (ii) b_n^2 is smaller than a chosen value (e.g. 10^{-14} in double-precision arithmetic) or (iii) the number of iterations has reached a preset maximum (typically a few hundreds). The ground state energy E_0 and the ground state $|\Omega\rangle$ are very well approximated by the lowest eigenvalue and the corresponding eigenvector of that matrix, which are obtained by standard methods. This provides us with the ground state $|\Omega\rangle$ in the reduced basis $\{\phi_n\}$. But we need the ground state in the original basis, and this requires retracing the Lanczös iterations a second time – for the $|\phi_n\rangle$ are not stored in memory – and constructing the ground state progressively at each iteration from the known coefficients $\langle\Omega|\phi_n\rangle$.

2.2.2 Second step: the Green function

Once the ground state is known, the second step of the procedure consists in calculating the Green function (5). Because the Hubbard model conserves spin, we need only to consider the creation and annihilation of up-spin electrons. Consider first the case $a = b$ in $G^{(c)}_{aa,e}$. One first calculates the state $|\phi\rangle = c^\dagger_a|\Omega\rangle$, and then

$$G^{(c)}_{aa,e} = \langle\phi|\frac{1}{z - H + E_0}|\phi\rangle \tag{21}$$

This expression is again calculated using the Lanczös algorithm. A Lanczös sequence is calculated from the initial, normalized state $|\phi_0\rangle = |\phi\rangle/b_0$ where $b_0^2 = \langle\phi|\phi\rangle$. This sequence generates a tridiagonal representation of H, albeit in a different Hilbert space sector : that with $N_\uparrow + 1$ up-spin electrons and $N_\downarrow$ down-spin electrons. Once the preset maximum number of Lanczös steps, or a minimum value of b_n, has been reached, the tridiagonal representation (20) may then be used to calculate (21). This amounts to the matrix element $b_0^2[(z - H + E_0)^{-1}]_{00}$ (the first element of the inverse of a tridiagonal matrix), which has a simple continued fraction form :

$$G^{(c)}_{aa,e} = \cfrac{b_0^2}{z - a_0 - \cfrac{b_1^2}{z - a_1 - \cfrac{b_2^2}{z - a_2 - \cdots}}} \tag{22}$$

Thus, evaluating the Green function, once the arrays $\{a_n\}$ and $\{b_n\}$ have been found, reduces to the calculation of a finite continued fraction.

Consider next the case $a \neq b$. The continued fraction representation applies only to the case where the same state $|\phi\rangle$ appears on the two sides of (21). If $a \neq b$, this is no longer the case, but we may use the following trick : we define the combinations

$$G^{\pm}_{ab,e}(z) = \langle\Omega|(c_a \pm c_b)\frac{1}{z - H + E_0}(c_a \pm c_b)^\dagger|\Omega\rangle \tag{23}$$

Using the symmetry $G_{ab,e}(z) = G_{ba,e}(z)$, this leads to

$$G_{ab,e}(z) = \frac{1}{4}(G^{+}_{ab,e}(z) - G^{-}_{ab,e}(z)) \tag{24}$$

where $G^{\pm}_{ab,e}$ can be calculated in the same way as $G_{aa,e}$, i.e., with a simple continued fraction. We proceed likewise for $G^{\pm}_{ab,h}(z)$.

2.3 Cluster symmetries

One does not need to perform the second Lanczös step for each pair (a, b) of sites on the cluster, since some pairs are related by symmetries of the cluster, such as reflections and rotations. For instance, in a one-dimensional cluster of length L, a reflection with respect to the cluster's center sends site a ($a = 0, 1, \ldots, L-1$) to site $L - a - 1$. The cluster Hamiltonian is invariant under such a reflection, and so $G^{(c)}_{a,b} = G^{(c)}_{L-a-1,L-b-1}$. The number of Lanczös sequences to perform, instead of being L^2, reduces to $L(L+1)/2$. On a two-dimensional, rectangular cluster, reflections about the center along the x and y directions, and their combination, reduces the number of sequences by a factor of 4 in the limit of a large cluster, but in practice, for a 3×4 cluster, from $12^2 = 144$ to 48. On a square cluster, the symmetry is even higher since rotations of $\pi/2$ generate an additional set of transformations that can further reduce the number of pairs, by a factor 8 in the limit of large square clusters, but in practice from $16^2 = 256$ to 55 for a 4×4 cluster, the largest one practically available.

Using these symmetries not only saves times because of the economy in Lanczös sequences, but also later in the calculation, when the cluster Green function is calculated from the continued fraction representation $\{a_n, b_n\}$. That calculation is performed only for pairs of sites (a, b) unrelated by symmetry.

The symmetry of the cluster could also be used to reduce the size of the Hilbert space considered in the Lanczös procedure. Indeed, the Hamiltonian would be block-diagonal in a new basis of eigenfunctions of these symmetry operations. However, this has not been implemented by us, basically because the complexity of the coding involved and the development time needed would, in our opinion, offset the benefits.

2.4 Memory requirements

The memory needed to perform the exact diagonalization comes mainly from storing the diagonal part of the Hamiltonian matrix and three Lanczös vectors, which means $4d$ double precision floating point numbers. For a half-filled, 12 site cluster, $d = 853,776$ and this is modest 26 MB. Had we needed to store the full kinetic energy operator K instead of its tensor product factors, this would have been multiplied by roughly four. But the growth of the Hilbert space is exponential as a function of system size. The same procedure with a half-filled 16 site cluster would require slightly over 5.0 GB of memory. Computing time also explodes, since the Lanczös procedure scales like $d \ln d$ and the number of Lanczös sequences needed for calculating the Green function scales like L^2.

Once the Lanczös sequences are over, one needs to keep in memory the tridiagonal representations $\{a_n, b_n\}$ for each nonequivalent pair (a, b), which takes negligible memory. The calculation then proceeds to the next phase, described in the following section.

3 Embedding the cluster in the lattice

3.1 Restoring translation invariance

Let us go back the basic formulas 8 or 12 of Cluster Perturbation Theory. The indices m, n labelling different clusters can be replaced by wavevector indices in the description of inter-cluster quantities, through a Fourier transform. For instance, the inter-cluster hoppings may be arranged in

a $L \times L$ matrix $V_{ab}(\mathbf{K})$, where $\mathbf{K}$ is a wavevector in the Brillouin zone of the lattice of clusters (the so-called superlattice) :

$$V_{ab}(\mathbf{K}) = \sum_n t_{a,b}^{0,n} \mathrm{e}^{i\mathbf{K}\cdot\mathbf{r}_n} \tag{25}$$

where $\mathbf{r}_n$ is the position of the cluster labelled n with respect to the position of the "base" cluster. We then write Eq. 8 as

$$\hat{G}^{-1}(\mathbf{K},\omega) = \hat{G}^{(c)-1}(\omega) - \hat{V}(\mathbf{K}) \tag{26}$$

where frequency dependence has been explicitly restored, to stress that the matrix $V(\mathbf{K})$ is frequency independent.

Since inter-cluster hoppings V_{ab}^{mn} are treated differently from intra-cluster hoppings ($t_{ab}^{(c)}$), the full translation symmetry of the lattice is broken, and must be restored in some way. The procedure we adopt, explained in Ref. [6], is to perform a residual Fourier transform in order to obtain a fully momentum dependent Green function:

$$G_{\mathrm{CPT}}(\mathbf{k},\omega) = \frac{1}{L}\sum_{a,b=1}^{L} G_{ab}(\mathbf{k},\omega)\mathrm{e}^{-i\mathbf{k}\cdot(\mathbf{r}_a-\mathbf{r}_b)} \tag{27}$$

where $\mathbf{r}_a$ is the position of site a within the cluster. We have replaced the wavevector $\mathbf{K}$, belonging to the reduced Brillouin zone (the Brillouin zone of the superlattice), by the wavevector of interest $\mathbf{k}$ belonging to the full Brillouin zone. The difference $\mathbf{q} = \mathbf{K} - \mathbf{k}$ belongs to the reciprocal superlattice, and so $V_{ab}(\mathbf{K}) = V_{ab}(\mathbf{k})$ since $\mathrm{e}^{i\mathbf{q}\cdot\mathbf{r}_n} = 1$.

To summarize, evaluating the Green function for a given wavevector $\mathbf{k}$ and complex frequency $\omega \to z$ proceeds according to the following steps:

1. For that value of z, the cluster Green function $G_{ab}^{(c)}$ is calculated from the continued fraction representation (22) – whose coefficients were stored in memory – for all pairs (a,b) of sites that are not related by a symmetry of the cluster.

2. For that value of $\mathbf{k}$, one forms the matrix $V_{ab}(\mathbf{k})$ following Eq. (25) and calculates the matrix

$$G_{a,b}(\mathbf{k},z) = \left(\frac{\hat{G}^{(c)}(z)}{1-\hat{V}(\mathbf{k})\hat{G}^{(c)}(z)}\right)_{a,b} \tag{28}$$

   which implies two multiplications and one inversion of a complex-valued matrix of dimension L.

3. One perform the residual Fourier transform (27)

The merits and limitations of this method have been commented on in Ref. [6] and we will avoid repeating what was done in that paper. However, let us stress the most important points, by comparing the spectral function $A(k,\omega)$ obtained by this method to what is obtained by an ordinary exact diagonalization and to the known exact solution in a very special case: that of the one-dimensional Hubbard model with nearest-neighbor hopping t in the $U \to \infty$ limit. In that case, the exact solution is known [8] and can be used to assert the effectiveness of the method we propose. One sees from Fig. 2 the main benefits of CPT :

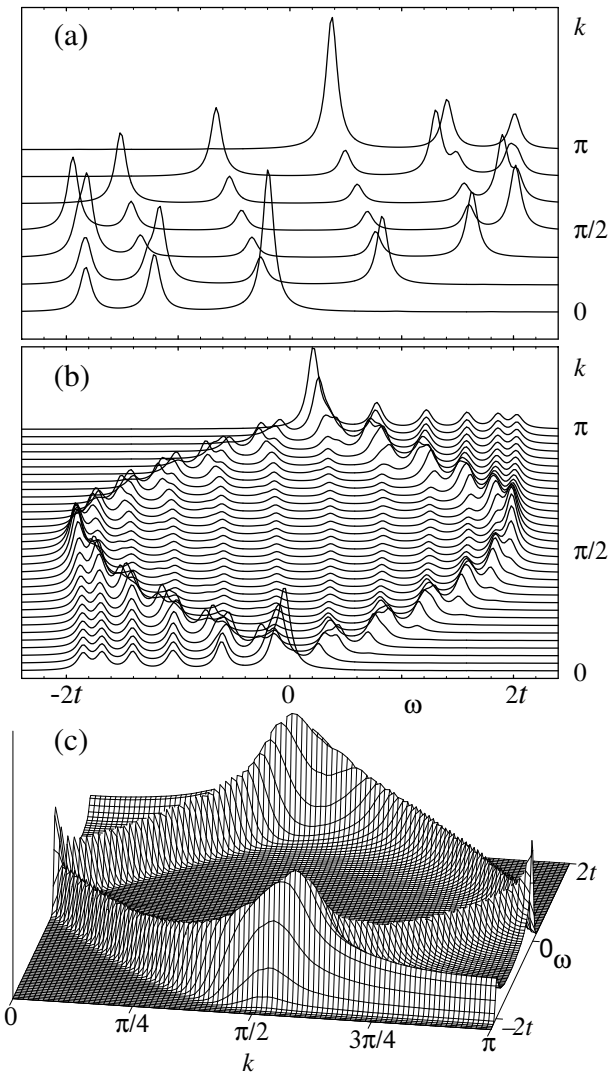

Figure 2. The spectral function of the $U \to \infty$ limit of the one-dimensional Hubbard model, as calculated from (a) an exact diagonalization of the Hubbard model with $U/t = 100$ on a periodic 12-site cluster; (b) the same, but with CPT, on a 12-site cluster with open boundary conditions; (c) the exact solution, taken from [8]; beware: the axes are oriented differently. In (a) and (b) a finite width η (see Eq. (4)) has been given to peaks that would otherwise be Dirac δ-functions.

1. The number of wavevectors is as large as one wishes. the exact diagonalization yields only $L/2+1 = 7$ wavevectors from $k = 0$ to $k = \pi$, in contrast to CPT, for which 32 wavevectors are displayed.

2. The number of poles of the Green function (i.e., peaks in the spectral function) is larger in CPT than in an exact diagonalization, even though both methods proceed through almost identical Lanczös se-

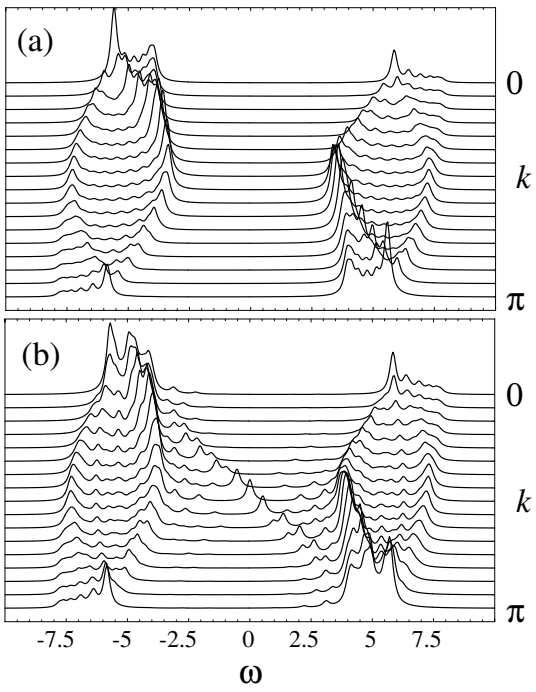

Figure 3. The spectral function of the one-dimensional Hubbard model at $U = 10t$ and $n = 1$ (half-filling), as calculated from (a) the CPT procedure of Sect. 3.1 and (b) from a Fourier transform of the cluster self-energy. One notices in the latter the appearance of spectral weight within the Mott gap.

quences with the same number of steps, and therefore the same number of floors in the continued fraction representation (22). The difference is that the CPT Green function formula (28), like Dyson's equation, rearranges these poles and the corresponding residues to something closer to the thermodynamic limit.

Moreover, the method falls back on the exact result in the $U \to 0$ limit, as well as on the trivial atomic limit in the $t_{ij}/U \to 0$ limit. One thus expects it to be applicable in between, in the intermediate coupling regime ($t_{ij} \sim U$).

3.2 Alternative I : Fourier transform of the self-energy

One possible alternative to the procedure described above would be to restore the translation symmetry by taking the Fourier transform of the self-energy rather than that of the CPT Green function. Explicitly, this amounts to calculating

$$\Sigma(\mathbf{k},\omega) = \frac{1}{L}\sum_{a,b=1}^{L} \Sigma_{ab}^{(c)}(\omega) \mathrm{e}^{-i\mathbf{k}\cdot(\mathbf{r}_a - \mathbf{r}_b)} \qquad (29)$$

where $\hat{\Sigma}^{(c)}$ is defined in Eq. (11), and then to write the full Green function from Dyson's equation:

$$G(\mathbf{k},\omega) = \frac{1}{\omega - \varepsilon(\mathbf{k}) - \Sigma(\mathbf{k},\omega)} \qquad (30)$$

While it is difficult to dismiss this procedure from the start, an explicit calculation shows that it fails (Fig. 3) to reproduce a key feature of the known properties of the one-dimensional Hubbard model : the appearance of a Mott gap for all values of U. Indeed, it is known from the exact solution of the Hubbard model [9] that a gap opens in the spectrum as soon as the interaction U is turned on. This is also the behavior observed in CPT. By contrast, Fourier transforming the cluster self-energy leads to a Mott transition at $U/t = \infty$ instead of $U/t = 0$. Notice, however, that this approach, like CPT, is exact in the two limiting cases $U/t = 0$ and $U/t \to \infty$.

3.3 Alternative II : Minimal self-consistence

CPT is not a self-consistent procedure. This means that the CPT Green function $G_{CPT}(\mathbf{k},\omega)$, when Fourier transformed to calculate the local Green function $G_{ab}^{CPT}(\omega)$, does not yield the cluster Green function $G_{ab}^{(c)}(\omega)$ calculated by exact diagonalization. This is no surprise, since $G_{ab}^{(c)}(\omega)$ feels the effects of the cluster boundary, while $G_{ab}^{CPT}(\omega)$ is translation invariant.

One possible variation of the CPT method would be to make it self-consistent at that level : the local Green function $G_{ab}^{CPT}(\omega)$ calculated from $G_{CPT}(\mathbf{k},\omega)$ would be used to obtain a new CPT Green function through Eq. (28) and then Eq. (27), thus closing the self-consistent loop, which would be iterated until convergence. When doing so, it turns out that the self-consistent solution obtained is identical to what would be found by applying the CPT procedure to a cluster with periodic boundary conditions. Using periodic boundary conditions is possible in CPT provided the hopping terms that need to be added to the cluster to make it periodic are also subtracted in V_{ab}^{mn}, as described in [6]. In other words, the periodic-cluster version of CPT is self-consistent in its spatial dependence. Unfortunately, it is also further from the exact solution than the open-cluster version, as explicitly illustrated in [6]. Thus, it appears that self-consistence is not necessarily a virtue.

4 Relation to other cluster approaches and Outlook

In Cluster Perturbation Theory, a cluster is embedded in its environment by Dyson's equation only. There are other methods currently or recently developed that simulate the cluster's environment, i.e., the rest of the lattice, by a bath of fermionic levels chosen in a self-consistent way. These are the Dynamical Cluster Approximation (DCA) [10] and the Cellular Dynamical Mean Field Theory (CDMFT) [11]. They are both offshoots of Dynamical Mean Field Theory (DMFT), in which a "cluster" of one site is embedded in a self-consistent bath (for a review and additional references, see [12]). Common to DCA and CDMFT is the need to couple the Hubbard model defined on a cluster with an Anderson impurity model (the bath). The whole system (bath + cluster) is then solved, for instance by Quantum Monte

Carlo or exact diagonalization, which implies in the latter case that the physical size of the cluster has to be small (e.g. 4 to 6 sites) to leave space for the bath. The DCA and CDMFT differ in that the former is a momentum-space technique using periodic boundary conditions on the cluster, whereas the latter is a real-space technique using open boundary conditions. CPT does not use a bath, and thus can handle bigger clusters. But it is not self-consistent as the DCA and CDMFT are.

Cluster Perturbation Theory has limitations : In its present state, it cannot handle more than the one-particle Green function, i.e., susceptibilities cannot be calculated. Long-range order is also unaccessible, because of the small cluster sizes involved and the absence of self-consistence.

This being said, CPT passes the test very well when compared to the available exact results in one dimension. It is also capable of accounting for the experimentally observed disappearance of the Fermi surface in high-T_c superconductors [13]. A finite-temperature version of CPT is under development, and already gives interesting results on a 2×2 cluster. Future development will also include treating an ordered state self-consistently. Eventually, it is hoped that some mutual enrichment, if not convergence, will occur between the various cluster techniques for strongly correlated electron systems.

Acknowledgments

Stimulating discussions with A.M. Tremblay, G. Kotliar, M. Jarrell, S. Kancharla, D. Plouffe and X. Barnabé-Thériault are gratefully acknowledged. Calculations were performed on Beowulf clusters whose components were bought thanks to grants from NSERC and CFI. I am grateful to the members of the Centre de Calcul Scientifique (CCS) of the Université de Sherbrooke, who designed and assembled the clusters, in particular to M. Barrette, M. Bozzo-Rey, B. des Ligneris, C. Gauthier, J.-P. Turcotte, P. Vachon and A. Veilleux. Finally, this work is financially supported by NSERC (Canada) and FQRNT (Québec).

References

1. W. Kohn. *Rev. Mod. Phys.*, 71:1253, 1999.
2. A.A. Abrikosov, L.P. Gorkov, and I.E. Dzyaloshinski. Dover, New York, 1963.
3. A.-M. S. Tremblay, C. Bourbonnais, and D. Sénéchal. *Physics in Canada*, 2000.
4. J. Hubbard. *Proc. Royal. Soc. London*, 276:238, 1963.
5. D. Sénéchal, D. Pérez, and M. Pioro-Ladrière. *Phys. Rev. Lett.*, 84:522, 2000.
6. David Sénéchal, Danny Perez, and Dany Plouffe. *Phys. Rev. B*, 66:075129, 2002.
7. Claudius Gros and Roser Valenti. *Annalen der Physik*, 3:460, 1994.
8. J. Favand, S. Haas, K. Penc, F. Mila, and E. Dagotto. *Phys. Rev. B*, 55:R4859, 1997.
9. E.H. Lieb and F.Y. Wu. *Phys. Rev. Lett.*, 20:1445, 1968.
10. Th. Maier, M. Jarrel, Th. Pruschke, and J. Keller. *Eur. Phys. J. B*, 13:613, 2000.
11. C.J. Bolech, S.S. Kancharla, and G. Kotliar. *cond-mat/0206166*, 2002.
12. A. Georges, G. Kotliar, W. Krauth, and M.J. Rozenberg. *Rev. Mod. Phys.*, 68:13, 1996.
13. D. Sénéchal and A.-M. S. Tremblay. *in preparation*, 2003.