EXACT SOLUTION OF THE SIMPLIFIED HUBBARD MODEL

PAVOL FARKAŠOVSKÝ

Institute of Experimental Physics, Slovak Acad. Sci.,
Solovjevova 47, 043 53 Košice, Slovak Republic

Received 6 May 1992

We present the exact solution of the simplified Hubbard model in which only one kind of electrons can hop and this quantum mechanical hopping of electrons is assumed to be unconstrained. It is shown that the model still behaves nontrivially, although it no longer depends on the lattice structure and the dimensionality of the system. For this case we find: (i) a gap in the ground state energy always exists at the half-filled band point \( n = 1 \), (ii) a preferred magnetic state at \( n = 1 \) and large \( U \) is a total spin singlet, (iii) \( U \)-dependence of the ground state energy has qualitatively the same form as one of the conventional Hubbard model with the \( (t^2/U) \)-behavior at large \( U \). A phase diagram of the model is discussed.

1. Introduction

Since its introduction in 1963 by Hubbard [1], Kanamori [2] and Gutzwiller [3], the Hubbard model (HM) has become an important standard model for a description of correlated fermions on a lattice. It combines the kinetic energy of itinerant electrons with the on-site Coulomb interaction between two electrons of opposite spin

\[
H = \sum_{ij} t_{ij} c_{i\sigma}^+ c_{j\bar{\sigma}} + U \sum_i c_{i\uparrow}^+ c_{i\downarrow} c_{i\downarrow}^+ c_{i\uparrow}.
\] (1)

The operators \( c_{i\sigma}^+ (c_{i\sigma}) \) create (annihilate) \( \sigma \)-spin fermions on site \( i \) on \( d \)-dimensional lattice, \( U > 0 \) is the on-site interaction parameter, and \( t_{ij} \) is the general matrix element for hopping between sites \( i \) and \( j \). For the conventional HM, it is usually assumed that \( t_{ij} = -t \) if \( i \) and \( j \) are the nearest neighbors and \( t_{ij} = 0 \) otherwise, but it will not be our case, so we leave it arbitrary for the moment.

In spite of its simplicity and enormous research activity of physicists in the past, an exact solution exists only in one dimension [4]. In this situation, it is particularly valuable to investigate the model in yet perhaps unrealistic limits, since it can yield an important insight into the more physical regimes. Three interesting, although little realistic limits were introduced for the HM: the limit of infinite dimension [5], the limit of unconstrained hopping [6,7] and the static limit [8], which is more often referred to as the Falicov-Kimball model [9], or directly as the modified (simplified) version of the HM [10,11].

The simplified version of the Hubbard Hamiltonian can be written as

\[
H = \sum_{ij} t_{ij} c_i^+ c_j + U \sum_i w_i c_i^+ c_i.
\] (2)
It means that only one kind of electrons (say with $\sigma = 1$) may hop $(c_i^+ c_i = c_i^+ c_i)$ and $\sigma = -1$ electrons are infinitely massive, with $w_i = 1$ if a fixed electron is at $i$ and $w_i = 0$ otherwise.

The Hamiltonian (2) has various physical interpretations. If one considers localized f-electrons, or classical ions instead of down-spin electrons, one obtains the model convenient for the description of mixed-valence states in rare-earth compounds [9], or for the study of the crystallization in a system of classical ions interacting with itinerant electrons through the on-site potential [8].

Concerning the simplified version of the HM, the following exact results are known: (1) In the symmetric case, when the number of up-spin and down-spin electrons per site are equal to $1/2$, down-spin electrons (ions) form a checkerboard pattern [8,12]. (2) Outside the symmetry point there are three down-spin electrons configurations: the checkerboard configuration, the completely empty or fully occupied configurations may be ground states [13,14]. The basic assumption underlying these results is that the lattice has AB-structure. From this point of view, it is particularly valuable to study some special limits of the Hamiltonian (2). Both above mentioned limits: the limit of high dimension and the limit of unconstrained hopping are clearly unrealistic, but as it was shown by Metzner and Vollhardt [5] they are still useful, because they can yield important insight into the more physical regimes. For example, many essential features of systems in $d = 3$ (and even lower) dimension are well described by results in $d \rightarrow \infty$ [5]. Although the most interesting goal, the exact solution of the HM in $d \rightarrow \infty$, has not yet been reached, using the fact that in $d \rightarrow \infty$ the proper self-energy becomes site diagonal [5] the Falicov-Kimball model was solved exactly in $d \rightarrow \infty$ by Brandt and Mielisch [15]. In this article we will investigate another limit of the Hamiltonian (2): the infinite-range hopping limit, when $t_{ij} = -t$ for all $i \neq j$. The full HM with infinite-range hopping was investigated by van Dongen and Vollhardt [8] using standard methods of many-body theory. Their solution is based on the fact that the entire kinetic energy of the system is carried only by two electrons in state $k = 0$, which does not play a preferred role in the interaction term and therefore may be neglected in thermodynamic limit as a term of relative order $1/L$. The model thus becomes trivial and postulates that the system has a ground state with the least possible number of doubly occupied sites independent of the strength of the interaction. On the other hand, Long [12] pointed out that the model exhibits some non-trivial behavior when the interaction energy $U$ scales with the volume of the system. Apart from some unphysical consequences of this choice, the author studied the model with regard to its magnetic content. For the case less then a quarter band filling, he found that the lowest energy solution is paramagnetic if $(U/Lt) \rightarrow 0$ or $(Lt/U) \rightarrow 0$, i.e., the preferred magnetic state is a total spin singlet for both cases. Unlike the exact solution of the infinite-range HM model [8], which works only in the thermodynamic limit, our solution of the simplified HM is applicable to the finite systems too and it is not necessary to scale $U$ with the volume of the system. The fact that our results are applicable to the finite systems has another advantage. For some small lattices with $L = 2, 3$ or 4 sites, where unconstrained hopping can
be interpreted as the nearest neighbors (nn) hopping, we directly obtain as the bonus the exact solution of the conventional Falicov-Kimball model. The four-site problem e.g. can be considered to represent a regular tetrahedron of sites with nn coupling. Similarly, the three-site case can be considered as an equilateral triangle, and the two-site case as a dumbbell, both cases with nn coupling.

2. Solution of the model

Our starting point in the next investigation will be the Hamiltonian

\[ H = -t \sum_{ij} c_i^+ c_j + u \sum_i w_i c_i^+ c_i, \]  

where the prime on the sum means that the term with \( i = j \) was omitted. This Hamiltonian can be exactly diagonalized by the canonical transformation

\[ \begin{pmatrix} a_1^+ \\ \vdots \\ a_L^+ \end{pmatrix} = U \begin{pmatrix} c_1^+ \\ \vdots \\ c_L^+ \end{pmatrix}, \]

where \( U \) is an unitary matrix (see Appendix A), into the simple form

\[ H = \sum_i E_i a_i^+ a_i. \]

Here \( E_i \) are eigenvalues of the matrix

\[ \begin{pmatrix} U & w_1 & -t & \cdots & -t \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ -t & -t & \cdots & U & w_L \end{pmatrix}, \]

which do not depend on the real configuration of the particles with spin down, but only on their number

\[ N_1 = \sum_i w_i. \]

Solving the secular equation, we find the following eigenvalues and degeneracies

\[ E = E_1 = \frac{1}{2} \left( U - Lt + 2t - \sqrt{(U + Lt)^2 - 4tUN_1} \right), \]

\( E = t, \quad (L - N_1 - 1)\)-times,

\[ E = E_2 = \frac{1}{2} \left( U - Lt + 2t + \sqrt{(U + Lt)^2 - 4tUN_1} \right), \]

\( E = t + U, \quad (N_1 - 1)\)-times,
where $N_1 = 1, 2 \ldots L - 1$. The problem is trivial for $N_1 = 0$, when

$$E = \begin{cases} -t(L-1), \\ t, \end{cases} \text{ (L-1)-times,}$$

and for $N_1 = L$, when

$$E = \begin{cases} -t(L-1) + U, \\ t + U, \end{cases} \text{ (L-1)-times.}$$

From these expressions it follows that the considered model is not one of the independent particles (as might seem at first sight) because $w_i$ are allowed to vary and, in the ground state (GS), must be chosen to minimize the ground state energy (GSE). To calculate the GSE, we adopt the following definition [8]. If the total particle number $N = N_T + N_1$ is fixed, then the energy of the GS is

$$E_G(N) = \min \{E(N_T, N_1) | N_T + N_1 = N \}.$$  

Here $E(N_T, N_1)$ is the GS for fixed $N_T$ and $N_1$. It is defined by taking the GS of $H$ with respect to $N_T$.

Let us first investigate the case $N < L$, where $L$ is the number of lattice sites. Since $\sigma = -1$ electrons are localized at $E = 0$ (symmetric case), the ground state $E(N_T, N_1)$ for fixed $N_T$ and $N_1$ is obtained such that all $N_T$ electrons are placed on the lowest energy levels available for occupation: $E = E_1$ and $E = t$, which is $(L - N_1 - 1)$-times degenerated. This implies that

$$E(0, N) = 0, \quad \text{for } N_1 = N$$

and

$$E(N - N_1, N_1) = E(N_1 = (N - N_1)t + \frac{1}{2} \left(U - Lt - \sqrt{(U + Lt)^2 - 4tU N_1} \right),$$

for $0 \leq N_1 \leq N - 1$. Using the identities

$$E(1, N - 1) < E(0, N),$$

$$E(N - 1, 1) < E(N, 0),$$

the GSE may be expressed only in terms of $E(N_1)$ ($0 \leq N_1 \leq N - 1$)

$$E_G(N) = \min \{E(N_1) | N_1 + N_1 = N \}.$$  

Thus, to determine $E_G(N)$, we must find a minimum of $E(N_1)$ for $0 \leq N_1 \leq N - 1$. Analysing (13) one can show straightforwardly that $E(N_1)$ has a minimum for some $N_1 = N_0$ and therefore $E_G(N) = E(N_0)$ for $N > N_0$ and $E_G(N) = E(N - 1)$ for $N \leq N_0$. Combining this fact with the identities (14) and (15), we arrive that the following conclusion: the GS of (3) can not be purely ferromagnetic for any interaction strength $U$ and any total number of particles $N \leq L$ (certainly that for
Exact solution of the simplified Hubbard model

$N > L$ too). Here by ferromagnetic we mean a maximum-spin GS and analogously by antiferromagnetic we will mean a minimum-spin GS. It is convenient for a system to have at least one electron with spin up, because as follows from (12) and (13) only in such a way it can reduce its own energy. The analytic calculation for the GSE yields

$$E_G(N, U) = \begin{cases} E(N - 1) & \text{for } 0 \leq (U/t) \leq U_{N-1}, \\ E(N_0) & \text{for } U_{N_0+1} \leq (U/t) \leq U_{N_0}, \\ E(L/2) & \text{for } U_{L/2} \leq (U/t), \end{cases}$$

(17)

where

$$U_{N_0} = \frac{L^2 - 1}{2(2N_0 - L + 1)}$$

(18)

and $N_0$ may be $N - 2$, $N - 3$, ..., $\frac{1}{2}L + 1$. It should be mentioned that the expression (17) is valid for arbitrary $L$ and therefore can be used to investigate finite systems too. For example, the four-site problem (a regular tetrahedron of sites with nn coupling) has for the total electron number $N = 4$ the following solution: the state with $N_1 = 3$ is stable for $0 < U/t < 7.5$ and the total spin singlet is stable for $7.5 < U/t$.

For large lattices (17) is reduced to

$$E_G(N) = \begin{cases} \frac{1}{2} \left( U - Lt + 2t - \sqrt{(U + Lt)^2 - 4U(N - 1)} \right), & \text{for } N \leq N_0 \\ -(L - N)t - \frac{L^2 t^2}{4U}, & \text{for } N > N_0 \end{cases}$$

(19)

where

$$N_0 = \frac{L}{2} + \frac{L^2 t}{4U}.$$ 

(20)

Two important results follow from (17)–(20).

1. If $U \to \infty$ then $N_1 \to L/2$. Thus, for $n = N/L = 1$ (half-filled band) we also have $N_1 \to L/2$, and in this case the GS is a total spin singlet.

2. For both cases $N \leq N_0$ and $N > N_0$, $U$-dependence of the GSE has $(t^2/U)$-behavior at large $U$, which is well known for the Hubbard-type models in this limit [16,17]. The total $U$-dependence of the GSE is shown in Fig. 1. It reveals that $E_G(U)$ has qualitatively the same form as $E_G(U)$ of the ordinary HM in one dimension [4].

Thus, we see that at first sight the trivial model retains some characteristic features of the original HM. Moreover, the model exhibits another important characteristic of the HM: an energy gap at $n = 1$. To prove it, let us consider the case $N > L$. Because on the lowest energy levels ($E_1$, $t$ for the up-spin electrons and $E = 0$ for the down spin electrons) we can put always only $L$ electrons, additional electrons must go on $E_2$ and $U + t$. But, as follows from (8), the simultaneously

occupied energy levels $E_1$ and $E_2$ give to the GSE a contribution $E_1 + E_2$, which is independent of $N_1$ and one can obtain for $E(N_1)$

$$E(N_1) = MU - (N_1 - M)t,$$

(21)

where $M = N - L$. $E(N_1)$ reaches the minimum at $N_1 = L$ and with respect to (11) the GSE for $N > L$ is given by

$$E_G(N) = MU - (L - M)t.$$

(22)

Now, combining (17) or (19) with (22) we have an analytical expression for the GSE, for an arbitrary total number of electrons. Using these results, and noting that the chemical potential $\mu(N)$ is defined as

$$\mu(N) = E_G(N + 1) - E_G(N),$$

(23)

one can show straightforwardly that a chemical potential is discontinuous at $n = 1$ for all $U > 0$. Thus, the model predicts a metal-insulator transition in the half-filled band point, which occurs also in the GSE of the conventional HM in the static limit [8] and even in the exact GS of the ordinary HM in $d = 1$ [4]. Besides the prediction of the Mott-transition at $n = 1$, the $n$-dependence of the GSE exhibits good agreement with exact results in $d = 1$ when $n$ is less than and nearly equal to 1. Of course, for $n \to 0$ these behaviors are strongly different, because in our case only one electron carries all kinetic energy so that $E_G(N)$ is nonvanishing in this limit, while $E_G(N)$ approaches zero for the HM with unconstrained hopping, as was found by Long [18] for the case of less than quarter band filling.

Finally, let us sketch a phase diagram of the model. There are three different regions with respect to the values of $u = U/Lt$ and $n$, denoted in Fig. 2 as A, B, C. For $u$ and $n$ from region A, the GS is characterized by $N_1 = 1$ and $N_L = N - 1$. This state is mainly preferred for small $u$ and $n$. One electron with spin up carries all kinetic energy and additional electrons have spin down and energy $E = 0$. In spite of the fact that down-spin electrons are localized at $E = 0$, the energy of the

Fig. 1. The $u$-dependence of the ground state energy for two different densities.
state increases as the total particle number increases, because the energy of up-spin electron increases too and the state becomes unstable. It becomes, for critical value of the interaction energy,

$$u^* = \frac{1}{2(2n - 1)}.$$  \hfill (24)

When the interaction parameter $u$ for some $n$ is greater then $u^*$, the system is in the GS with $N_1 = N_0$ and $N_\uparrow = N - N_0$ (region B). The case $n > 1$ (region C) is obvious; $N_1 = L$ and $N_\uparrow = N - L$ for all $n$ and $u$. It can be compared with some exact results obtained for 2-d Falicov-Kimball model with $nn$ coupling. The corresponding phase diagram is presented in Ref. 14. The authors investigated three ion configurations: checkerboard configuration, the completely empty configuration $s_-$ ($w_i = 0$ for all $i$) and fully occupied configuration $s_+$ ($w_i = 1$ for all $i$). Using a procedure based on Tchebycheff-Markov inequalities, they found domains in the plane of chemical potentials of electrons and ions, where these configurations are ground states. The phase C in Fig. 2 is identical with their phase $s_+$. The completely empty configuration is not present in our phase diagram, but it occurs in the case when down-spin electrons are localized at energy $E < 0$ [19]. The ground state with $N_\uparrow = N_1 = L/2$ is stable only if $u \to \infty$ and does not depend on the actual configuration of down-spin electrons, thus checkerboard configuration is not preferred for unconstrained hopping.

In summary, we have presented the exact solution of the simplified Hubbard model. It is shown that, although the model no longer depends on the actual lattice structure and the dimensionality of the system, it still yields some nontrivial conclusions. We summarize the main ones: (i) a gap in the ground state always exists at the half-filled band, (ii) the $U$-dependence of the ground state energy has qualitatively the same form as that of the exact solution of the ordinary Hubbard model in one dimension, (iii) for large $U (n \leq 1)$, $E_G(U)$ goes as $(t^2/U)$ and the ground state in this limit at $n = 1$ is the total spin singlet.
Appendix A: Exact diagonalization

The canonical transformation (4), which diagonalizes the Hamiltonian (3), is given by unitary matrix

\[ U = \begin{pmatrix}
  b_1(Uw_1 + t + E_1)^{-1} & b_1(Uw_2 + t + E_1)^{-1} & \cdots & b_1(Uw_L + t + E_1)^{-1} \\
  (1 - w_1)p_{12} & (1 - w_2)p_{22} & \cdots & (1 - w_L)p_{LL} \\
  \vdots & \vdots & \ddots & \vdots \\
  (1 - w_1)p_{1L-N,1} & (1 - w_2)p_{2L-N,1} & \cdots & (1 - w_L)p_{LL-N,1} \\
  b_2(Uw_1 + t - E_2)^{-1} & b_2(Uw_2 + t - E_2)^{-1} & \cdots & b_2(Uw_L + t - E_2)^{-1} \\
  w_1q_1L-N,1+2 & w_2q_2L-N,1+2 & \cdots & w_Lq_LL-N,1+2 \\
  \vdots & \vdots & \ddots & \vdots \\
  w_1q_1L & w_2q_2L & \cdots & w_Lq_LL 
\end{pmatrix},
\]

where

\[ b_i = \frac{1}{\sqrt{\frac{N_1}{(U + t - E_i)^2} + \frac{L-N_1}{(t - E_i)^2}}}, \quad i = 1, 2 \]

and the eigenvalues \( E_1, E_2 \) are given by (8). For the coefficients \( p_{ij} \) and \( q_{ij} \), which determine eigenvectors of (6) we obtained

\[ p_{ij} = \begin{cases} 
0 & \text{for } w_i = 1 \\
\Pi_{ij} & \text{for } w_i = 0 
\end{cases} \]

\[ q_{ij} = \begin{cases} 
0 & \text{for } w_i = 0 \\
Q_{ij} & \text{for } w_i = 1 
\end{cases} \]

where \( P_{ij}(Q_{ij}) \) are elements of the matrix \( P(Q) \) of the form

\[ P(Q) = \begin{pmatrix}
  \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & \cdots & 0 & \cdots \\
  \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & 0 & \cdots & 0 & \cdots \\
  \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \\
  \frac{1}{\sqrt{t^2+i^2}} & \frac{1}{\sqrt{t^2+i^2}} & \frac{1}{\sqrt{t^2+i^2}} & \frac{1}{\sqrt{t^2+i^2}} & \cdots & \frac{1}{\sqrt{t^2+i^2}} & \cdots \\
  \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots 
\end{pmatrix}, \quad i = 1, 2, \ldots
\]

The matrix \( P \) is of the type \((L - N_1 - 1) \times (L - N_1)\) and the matrix \( Q \) is of the type \((N_1 - 1) \times N_1\). A physical meaning of the new operators \( a_i^\dagger \) is now obvious, e.g., \( a_i^\dagger \) creates an electron delocalized on all sites with the probability \( b_i^2/(t-E_1)(b_i^2/(U+t-E_1)) \) on sites occupied (unoccupied) by a down spin electron etc.
Appendix B: Exact diagonalization

If we regard the down-spin electrons as fixed and their configuration is \( \{w_i\} = (w_1, w_2, \ldots, w_L) \), then the motion of the up-spin electrons is similar to the motion of electrons in a random alloy model. A relation to this model, in particular to the coherent potential approximation (CPA), as well as to the exact solution of the Falicov-Kimball model in large dimensions [15] may be established through Green’s functions. Before doing it, let us recall that CPA is exact for the random alloy in large dimensions [20] and it yields the same form for self-energy functionals of the Falicov-Kimball model in \( D \to \infty \) as the exact solution [15], but there the probability of finding the site occupied by an f-electron still depends on the site diagonal Green’s function, in contrast to the CPA where the probability is constant and it becomes important. Using the standard technique [1], we find that Green’s functions satisfy the equation

\[
EG_{ij} = \frac{\delta_{ij}}{2\pi} - t \sum_m G_{mj} + (t + U w_i)G_{ij},
\]

which has the solution

\[
G = \frac{1}{2\pi} A^{-1}.
\]

Here \( A \) is given by (6) and \( G \) is the matrix of the up-spin electron Green’s functions \( G_{ij} \) corresponding to \( \{w_i\} \)

\[
G_{ii}(E) = \frac{1}{2\pi} \left[ \frac{1}{(E - U w_i - t)} - \frac{t}{1 + tx (E - U w_i - t)^2} \right], \quad (B3)
\]

\[
G_{ij}(E) = -\frac{t}{2\pi 1 + tx (E - U w_i - t)(E - U w_j - t)} \frac{1}{(E - U w_j - t)}, \quad (B4)
\]

where \( x = N_1/(E - U - t) + (L - N_1)/(E - t) \). The density of (pseudo-particle) states per atom is defined in terms of \( G_{ii}(E) \) [1] and can be directly calculated

\[
\rho(E) = \frac{i}{L} \lim_{\varepsilon \to 0^+} \sum_j [G_{jj}(E + i\varepsilon) - G_{jj}(E - i\varepsilon)] = \frac{1}{L} \left[ (N_1 - 1)\delta(E - U - t) + (L - N_1 - 1)\delta(E - t) + \delta(E - E_1) + \delta(E - E_2) \right].
\]

On the other hand, solving equations of the CPA [21,22] one directly finds the self-energy, the entire averaged Green’s function and hence the density of states per atom

\[
\rho(E) = \frac{1}{L} \left\{ (N_1 - n_1)\delta(E - U - t) + (L - N_1 - 1 + n_1)\delta(E - t) + \delta(E - E_1) + \delta(E - E_2) \right\}.
\]

We see that the CPA gives the correct energy spectrum, but degeneracies of energy levels are slightly different from exact ones. This fact is not important in the weak coupling limit, while it becomes crucial in the strong coupling limit as one should expect for the CPA.

References