Hartree-Fock Gibbs States for the Hubbard Model

V. Bach, J. Poelchau
Fachbereich Mathematik MA 7-2
TU Berlin, Str. d. 17.Juni 136
D-10623 Berlin
bach@math.tu-berlin.de, poelchau@math.tu-berlin.de

May 20, 1996

Abstract

The main results of [2] on the Hartree-Fock approximation for the repulsive Hubbard model at half-filling, $\mu = 0$, are reviewed and partly extended away from half-filling, $\mu \neq 0$.

Keywords: Hartree-Fock Theory, Hubbard Model.
AMS subject classification: 82A15, 82A68, 81J05.

1 Introduction

The Hartree-Fock (HF) approximation is one of the oldest in quantum mechanics. It is widely used in quantum chemistry, and it is one of the few approximations in quantum mechanics, whose accuracy can be estimated rigorously [1, 3, 4]. In spite of its popularity, the solution to the corresponding HF variational problem is explicitly known only in a few cases. One of these is the Hubbard model, and the ananlysis of its HF approximation at half-filling can be found in [2]. Our purpose is to review the results of [2] in the repulsive case from a somewhat more general perspective. As a new result (see also [7]) we establish the characterization of the HF approximation by an auxiliary variational problem that merely involves electron densities and spin densities for the case $\mu \neq 0$ (see Theorem 1). A similar result for $\mu = 0$ was the starting point in [2] for proving Theorems 3 (Gap estimate), 4 (Antiferromagnetism of Hartree-Fock Gibbs states) and 5 (Invariance of Hartree-Fock Gibbs states under spatial symmetries). We regard Theorem 1 to be the starting for a similar analysis in the case $\mu \neq 0$, i.e., away from half filling.

The paper is organized as follows. In Section 2, we review the concept of the HF approximation for quantum statistical mechanics of fermion systems. In Section 3, we specialize to the Hubbard model. We construct an auxiliary functional, $\mathcal{R}_\beta$, and characterize the HF Gibbs states by its extremizers. In Section 4, we analyze the limit $\beta \to \infty$ for all $\mu \in \mathbb{R}$. Finally, we review the main results of [2] for the repulsive Hubbard model at half filling in Section 5.

2 Gibbs States and HF-Gibbs States

To simplify the exposition we consider only finite dimensional systems. The one particle Hilbert space is $\mathcal{H} = \mathbb{C}^M$. The fermion Fock space is given by $\mathcal{F} = \bigoplus_{N=0}^{2M} \mathcal{H}^{(N)}$, where $\mathcal{H}^{(N)}$ is the antisymmetric subspace of $\bigotimes_{i=1}^{N} \mathcal{H}$. Let $\{f_i\}_{i=1}^N$ be an ONB of $\mathcal{H}$, and let $c_k^\dagger := c^\dagger(f_k)$ be the usual fermion creation operators on $\mathcal{F}$, i.e., $\{c_i, c_j^\dagger\} = \delta_{ij}$ and $c_j \Omega = 0$, where Ω denotes the vacuum. We consider Hamiltonians which contain a second quantized one body term, $h : \mathcal{H} \to \mathcal{H}$ (typically the kinetic energy), and a (second quantized) two body operator $V : \mathcal{H} \otimes \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$ which represents the pair interaction.

$$H \; = \; \sum_{k,m} \langle f_k \mid h f_m \rangle \; c_k^\dagger c_m \; + \; \frac{1}{2} \sum_{kl,mn} \langle f_k \otimes f_l \mid V(f_m \otimes f_n) \rangle \; c_l^\dagger c_k^\dagger c_m c_n \; . \tag{1}$$

States of such a system are operators ρ on $\mathcal{F}$ with the property $\rho \geq 0$ and $\mathrm{Tr}\rho = 1$ thus constituting a convex set. The expectation value of an observable A on $\mathcal{F}$ in the state ρ is $\rho(A) = \mathrm{Tr}[A\rho]$. The **entropy** of a state is given by

$$S_{qm}(\rho) = -\mathrm{Tr}[\rho \ln \rho] \; .$$

The subscript qm refers to (full) quantum mechanics. In thermodynamics the equilibrium states of a system in contact with a heat and particle reservoir minimize the **grand canonical potential** defined by

$$\Omega_{qm}(\rho) \; = \; \mathrm{Tr}[(H - \mu \mathbf{N})\rho] - \frac{1}{\beta} S_{qm}(\rho) \; , \tag{2}$$

where $\mathbf{N} = \sum_k c_k^\dagger c_k$ is the number operator, $1/\beta$ is the temperature and μ is the chemical potential. The entropy is strictly concave in ρ which implies strict convexity of the grand canonical potential. It attains its minimum since $\mathcal{F}$ is finite dimensional. The corresponding minimizer, $\rho^{(\beta)}$, is unique by strict convexity and is called **Gibbs state**. The Gibbs State is easily seen to be given by

$$\rho^{(\beta)} \; = \; \frac{\exp\big(- \beta(H - \mu \mathbf{N})\big)}{\mathcal{Z}} \; ,$$

where $\mathcal{Z} = \mathrm{Tr}\Big[\exp\big(- \beta(H - \mu\mathbf{N})\big)\Big]$ is the grand canonical partition function. In other words:

$$\Omega_{qm}\big(\rho^{(\beta)}\big) \; = \min_{\rho \geq 0, Tr[\rho]=1} \{\Omega_{qm}(\rho)\} \; = \; \beta^{-1} \ln \mathrm{Tr}[\exp(-\beta H)] \; .$$

Note that we do not display the dependence of $\rho^{(\beta)}$ on the parameter μ and the dependence of Ω_{qm} and $\mathcal{Z}$ on the parameters μ and β. In general it is far out of reach to calculate the value of $\Omega_{qm}\big(\rho^{(\beta)}\big)$ and the properties of the Gibbs state. A more modest aim is to try to describe the properties of the so-called HF-Gibbs states.

HF theory consists of minimizing the function Ω_{qm} on a restricted set of states: the quasi-free states. A state ρ is quasi-free if all truncated correlation functions vanish. For example

$$\rho(c_k^\dagger c_l^\dagger c_m c_n) = \rho(c_k^\dagger c_l^\dagger)\rho(c_m c_n) + \rho(c_k^\dagger c_n)\rho(c_l^\dagger c_m) - \rho(c_k^\dagger c_m)\rho(c_l^\dagger c_n) \; , \tag{3}$$

for a quasi free state ρ. Applying this to the expectation value of the potential $\rho(V)$ the three terms on the right side of (3) yield the pairing , the direct and the exchange energies. In the present paper we only consider repulsive interactions, $V \geq 0$. In this case it suffices to minimize the HF-grand canonical potential on the set of quasi-free states with vanishing pair expectation values $\rho(c_k^\dagger c_m^\dagger) = 0$ (see [2] Theorem 2.11 for a short and simple proof and also for further details about both the attractive and the repulsive case).

We present HF theory as seemingly independent. In Lemma 1 we establish the connection between HF-theory, introduced below, and quantum statistical mechanics, reviewed above.

One particle density matrices (1-pdm) are operators γ on $\mathcal{H}$ with the property $0 \leq \gamma \leq 1$. The **HF-energy functional** is defined on the set of 1-pdm by

$$\mathcal{E}(\gamma) \;=\; \mathrm{Tr}_1[(T-\mu)\gamma] \;+\; \frac{1}{2}\mathrm{Tr}_2[V(1-\mathrm{Ex})(\gamma\otimes\gamma)]\;, \tag{4}$$

where Ex is the exchange operator: $\mathrm{Ex}(f\otimes g) = g\otimes f$, and $\mathrm{Tr}_{1,2}$ are the traces over $\mathcal{H}$ and $\mathcal{H}\otimes\mathcal{H}$, respectively. The **one-particle entropy** is given by

$$S(\gamma) \;=\; -\frac{1}{2}\mathrm{Tr}_1[\gamma\ln\gamma] \;-\; \frac{1}{2}\mathrm{Tr}_1[(1-\gamma)\ln(1-\gamma)]\;. \tag{5}$$

We may now define the **HF-grand canonical potential**:

$$\Omega(\gamma) \;=\; \mathcal{E}(\gamma) - \frac{1}{\beta}S(\gamma)\;. \tag{6}$$

For given values of β and μ the **HF-Gibbs states** $\gamma^{(\beta)}$ are the 1-pdm which minimize the HF-grand canonical potential

$$\Omega\big(\gamma^{(\beta)}\big) \;=\; \min_{0<\gamma<1}\Omega(\gamma) \;=:\; \Omega^{(\beta)}. \tag{7}$$

Note that the minimizers need not be unique, since the HF-energy functional is neither concave nor convex in the 1-pdm. Indeed, as we show in Section 5, the HF-Gibbs states for the Hubbard model are unique only up to certain gauge transformations.

Now we present the connection of HF Gibbs states (operators on $\mathcal{H}$) to Gibbs states (operators on $\mathcal{F}$):

Lemma 1 *To every 1-pdm γ there exists a unique quasi-free state ρ_γ on $\mathcal{F}$ (with vanishing pair expectation values), such that*

$$\Omega(\gamma) \;=\; \Omega_{qm}(\rho_\gamma)\;, \tag{8}$$

and

$$\langle f_m|\gamma f_n\rangle = \rho_\gamma(c_m^\dagger c_n)\;. \tag{9}$$

Lemma 1 implies that

$$\Omega^{(\beta)} \;=\; \Omega\big(\gamma^{(\beta)}\big) \;=\; \Omega_{qm}\big(\rho_{\gamma^{(\beta)}}\big) \;\geq\; \Omega_{qm}\big(\rho^{(\beta)}\big)\;, \tag{10}$$

i.e., HF theory yields an upper bound to the equilibrium value of the grand canonical potential.

To prove Lemma 1, we review the construction of the state ρ_γ in [2]. Let $0 \le \lambda_i \le 1$ be the eigenvalues of γ, let Π_0 be the projection onto the subspace of $\mathcal{F}$ on which $\sum_{i:\lambda_i=0} c_i^\dagger c_i = 0$, and let Π_1 be the projection onto the subspace of $\mathcal{F}$ on which $\sum_{i:\lambda_i=1} c_i^\dagger c_i = 0$. For each i such that $0 < \lambda_i < 1$, choose e_i to satisfy

$$(1+\exp(e_i))^{-1} = \lambda_i, \tag{11}$$

and consider the following operator H on $\mathcal{F}$.

$$H = \sum_{i:\lambda_i\neq 0, \lambda_i \neq 1} e_i c_i^\dagger c_i . \tag{12}$$

Let $G := \Pi_0 \exp(-H)\Pi_1^\perp$, then we define ρ_γ by

$$\rho_\gamma(A) = \mathrm{Tr}[G]^{-1}\mathrm{Tr}[AG] . \tag{13}$$

The operator G looks peculiar but, by introducing the operator $H' = \sum_{i:\lambda_i=0} c_i^\dagger c_i$ which commutes with H, we can write the state ρ as a limit of Gibbs states:

$$\rho_\gamma(A) = \lim_{\beta\to\infty} Z_\beta^{-1}\mathrm{Tr}[A\exp(-\beta H' - H)]. \tag{14}$$

It is well-known (see [5] for a simple proof) that the Gibbs state for an operator of the form $\sum_i e_i c_i^\dagger c_i$ is quasi-free. Hence we conclude, by taking the limit $\beta \to \infty$, that ρ is quasi-free. It follows from (3) that $\rho_\gamma(H) = \mathcal{E}(\gamma)$ and it is a straightforward calculation to show that $S_{qm}(\rho_\gamma) = S(\gamma)$.

The standard way of introducing one-particle density matrices is to start from a state ρ on $\mathcal{F}$, and to define the corresponding 1-pdm γ_ρ by its matrix elements: $(\gamma_\rho)_{mn} := \rho(c_m^\dagger c_n)$. One easyly checks that $\mathrm{Tr}[\rho_\gamma c_m^\dagger c_n] = \delta_{nm}\lambda_m$, indeed.

3 HF-Gibbs States for the Hubbard Model

In this section we characterize the HF-Gibbs States for the Hubbard Model as extremals of a function of electron and spin densities (vector fields on the lattice). The Hubbard model describes itinerant electrons on a lattice, Λ, (a finite set of points) with an on site Coulomb interaction. $|\Lambda|$ denotes the number of lattice sites. The one-particle space is given by the complex-valued functions on $\Lambda \times \{\uparrow,\downarrow\}$.

$$\begin{aligned} \mathcal{H} &= \Big\{ f : \Lambda \times \{\uparrow,\downarrow\} \to \mathbb{C} \Big\} = \mathbb{C}^{|\Lambda|} \otimes \mathbb{C}^2 \\ &= \mathrm{span}\Big\{ \delta_{y,\sigma}(\cdot,\cdot) \;\Big|\; \delta_{y,\sigma}(x,\tau) = \delta_{xy}\cdot\delta_{\sigma\tau} \ ; \ x,y \in \Lambda \ ; \ \sigma,\tau \in \{\uparrow,\downarrow\} \Big\} . \end{aligned} \tag{15}$$

The kinetic energy is described by a $|\Lambda|\times|\Lambda|$-Matrix t with matrix elements $t_{xy} = \bar{t}_{yx}$ $(x,y\in\Lambda)$. A matrix element $t_{xy}\neq 0$ is the transition amplitude of an electron to move from site x to site y or vice versa. Shifting t by its trace, we require without loss of generality that the trace of t vanishes. Let $T := t \otimes 1$, where 1 is the identity on $\mathbb{C}^2$. If g is a complex-valued function on the lattice $(g \in \mathbb{C}^{|\Lambda|})$, we define the projections 1_x on $\mathbb{C}^{|\Lambda|}$ and the projections $\mathbf{1}_x$ on $\mathbb{C}^{|\Lambda|} \otimes \mathbb{C}^2$ by

$$(1_x g)(y) := \delta_{xy}\, g(y) \quad \text{and} \quad \mathbf{1}_x := 1_x \otimes 1 \qquad (x, y \in \Lambda)\ . \tag{16}$$

Let $U_x > 0$ be strictly positive numbers for all $x \in \Lambda$, denoting the strength of the on-site Coulomb repulsion of two electrons located at the site x. With the notation above we may write the Hubbard interaction as

$$V \;=\; \sum_{x \in \Lambda} U_x (\mathbf{1}_x \otimes \mathbf{1}_x)\ . \tag{17}$$

Let

$$U \;=\; \sum_{x \in \Lambda} U_x \mathbf{1}_x\ . \tag{18}$$

In the general setting of Section 2 the Hubbard Hamiltonian assumes the form

$$\sum_{xy,\sigma\tau} (h)_{x\sigma,y\tau}\ c^\dagger_{x,\sigma} c_{y,\tau} \;+\; \frac{1}{2} \sum_{\substack{xy,\tilde{x}\tilde{y} \\ \sigma\tau,\tilde{\sigma}\tilde{\tau}}} V_{x\sigma,y\tau\,;\tilde{x}\tilde{\sigma},\tilde{y}\tilde{\tau}}\ \ c^\dagger_{y,\tau} c^\dagger_{x,\sigma} c_{\tilde{x},\tilde{\sigma}} c_{\tilde{y},\tilde{\tau}}\ , \tag{19}$$

where the one body operator h is given by $h = T - U$. and the creation operators $c^\dagger_{x,\sigma}$ are the ones corresponding to the canonical ONB $\{\delta_{x,\sigma}\}$. An elementary calculation shows that

$$H \;=\; \sum_{\substack{x,y \in \Lambda \\ \sigma \in \{\uparrow,\downarrow\}}} t_{xy} c^\dagger_{x,\sigma} c_{y,\sigma} + \sum_{x \in \Lambda} U_x (n_{x,\uparrow} - \tfrac{1}{2})(n_{x,\downarrow} - \tfrac{1}{2}) \;-\; \frac{1}{4} \sum_{x \in \Lambda} U_x\ , \tag{20}$$

where $n_{x,\sigma} = c^\dagger_{x,\sigma} c_{x,\sigma}$. If $U_x \equiv U > 0$, we get the original Hubbard model, where the chemical potential is shifted by $-U/2$, plus the constant $U/4|\Lambda|$:

$$H = \sum_{x,y;\sigma} t_{xy} c^\dagger_{x,\sigma} c_{y,\sigma} + U \sum_x n_{x,\uparrow} n_{x,\downarrow} - \frac{U}{2}\mathbf{N} + \frac{U}{4}|\Lambda|\ .$$

If U_x is not constant the Hamiltonian (20) differs from the original one by the diagonal operator $\sum_x U_x c^\dagger_{x,\sigma} c_{x,\sigma}$, but the choice of (20) is suggested by the fact that $\mu = 0$ corresponds to the particle number $N = |\Lambda|$ (half filling), see Section 5.

The goal of HF-theory for the Hubbard model is to calculate the value of $\Omega^{(\beta)}$ at temperature $1/\beta$ and for a chemical potential μ and, even more important, to calculate

1. The electron density at the site x: $n^{(\beta)}(x) \;=\; \mathrm{Tr}\big[\mathbf{1}_x \gamma^{(\beta)}\big]$.
2. The spin density at the site x: $\vec{s}^{(\beta)}(x) \;=\; \mathrm{Tr}\big[(1_x \otimes \sigma)\gamma^{(\beta)}\big]$.

Here σ is the vector of Pauli matrices. The following Lemma is a reformulation of the HF-energy functional in the quantities we are interested in:

Lemma 2 *Let* $-\frac{1}{2} \leq \vec{s}_\gamma(x) = \frac{1}{2}\ \mathrm{Tr}_1[(1_x \otimes \sigma)\gamma] \leq \frac{1}{2}$ *be the spin density at site* x *in the state* $0 \leq \gamma \leq 1$, *and let* $0 \leq n_\gamma(x) = \frac{1}{2}\ \mathrm{Tr}_1[\mathbf{1}_x\gamma] \leq 1$ *be the electron density at site* x. *Then*

$$\mathcal{E}(\gamma) = \mathrm{Tr}_1[T_\mu\gamma] + \sum_{x\in\Lambda} U_x\,(n_\gamma(x) - \tfrac{1}{2})^2 - \vec{s}^{\,2}_\gamma(x)\ . \tag{21}$$

To prove Lemma 2 we use $\mathrm{Tr}_2[\mathrm{Ex}(A\otimes B)] = \mathrm{Tr}_1[AB]$ to calculate

$$\begin{aligned}
\mathrm{Tr}_2\big[V(1-\mathrm{Ex})(\gamma\otimes\gamma)\big] &= \sum_{x\in\Lambda} U_x \mathrm{Tr}_2\big[(1-\mathrm{Ex})(\gamma\otimes\gamma)(\mathbf{1}_x\otimes\mathbf{1}_x)\big] \\
&= \sum_{x\in\Lambda} U_x\Big\{\big(\mathrm{Tr}_1[\mathbf{1}_x\gamma]\big)^2 - \mathrm{Tr}_1\big[(\mathbf{1}_x\gamma)^2\big]\Big\} \\
&= \sum_{x\in\Lambda} U_x\Big\{\frac{1}{2}\big(\mathrm{Tr}_1[\mathbf{1}_x\gamma]\big)^2 - \frac{1}{2}\big(\mathrm{Tr}_1[(1_x\otimes\sigma)\gamma]\big)^2\Big\}\ ,
\end{aligned}$$

and inserting this into the HF-energy functional. In the last line we have used that for any operator B on $\mathcal{H}$, it holds

$$\mathrm{Tr}_1\big[(\mathbf{1}_xB)^2\big] = \frac{1}{2}\big(\mathrm{Tr}_1[(1_x\otimes\sigma)B]\big)^2 + \frac{1}{2}\big(\mathrm{Tr}_1[\mathbf{1}_xB]\big)^2\ .$$

Note that the HF-functional is a non-linear function of the 1-pdm contrasting with the linearity of the energy functional $\mathrm{Tr}[H\rho]$ in ρ. Our next goal is to transcribe the minimization of the HF-functional over the 1-pdm into the extremization of an auxiliary functional, $\mathcal{R}_\beta$, over four dimensional vector fields on Λ. Let $\big(d_0,\vec{d}\big)$ be a vector field on Λ:

$$\big(d_0,\vec{d}\big):\ \Lambda \to \mathbb{R}^4\ ,\quad x \mapsto \big(d_0(x),\vec{d}(x)\big)\ .$$

We may identify the 2×2-matrix-valued function $D_x = d_0(x)1 - \vec{d}(x)\cdot\sigma$ on Λ with $\big(d_0,\vec{d}\big)$. Operators of the form $D = \sum_{x\in\Lambda} 1_x\otimes D_x$ are called matrix valued potentials. Let $T_\mu = T-\mu$ and $\mathcal{R}_\beta$ be a function on the vector fields $\big(d_0,\vec{d}\big)$ given by

$$\mathcal{R}_\beta\big(d_0,\vec{d}\big) = -\frac{1}{\beta}\mathrm{Tr}\left[\ln\cosh\left\{\frac{\beta}{2}(T_\mu+UD)\right\}\right] + \sum_{x\in\Lambda} U_x\left(\vec{d}^{\,2}(x) - d_0^2(x)\right)\ , \tag{22}$$

where U is given in (18). A vector field $\big(d_0^{(\beta)},\vec{d}^{(\beta)}\big)$ with the property

$$\mathcal{R}_\beta\big(d_0^{(\beta)},\vec{d}^{(\beta)}\big) = \min_{\vec{d}\in B}\ \max_{d_0\in K}\ \mathcal{R}_\beta\big(d_0,\vec{d}\big)\quad ,$$

is called an **extremizer** of $\mathcal{R}_\beta$. The compact sets K and B are given by

$$\mathrm{K} := \Big\{d_0\ \Big|\ |\,|d_0(x)| \leq 1/2\Big\} = [-\tfrac{1}{2},\tfrac{1}{2}]^{|\Lambda|}$$

and

$$\mathrm{B} := \Big\{\vec{d}\ \Big|\ |\vec{d}(x)| \leq 1/2\Big\} = \big(B_{\mathbb{R}^3}(\tfrac{1}{2})\big)^{|\Lambda|}\ .$$

The zero component of the vector fields considered here plays the role of the density minus 1/2, and the vector component corresponds to the spin density of the HF-Gibbs states.

The following theorem establishes a one-to-one relation between the HF-Gibbs states and the extremizers of $\mathcal{R}_\beta$. Its main advantage is that instead of dealing with abstract one-particle density matrices γ, we are working with vector fields $\left(d_0, \vec{d}\right)$ representing the physical quantities we are interested in.

Theorem 1 [7]

1. *For every* $0 < \beta < \infty$

$$\Omega^{(\beta)} = \min_{\vec{d}\in B} \max_{d_0\in K} \mathcal{R}_\beta(d_0,\vec{d}) - |\Lambda|(\beta^{-1}2\ln 2+\mu) \quad \textit{holds true.} \tag{23}$$

2. *Let* $\left(d_0^{(\beta)}, \vec{d}^{(\beta)}\right)$ *be an extremizer of* $\mathcal{R}_\beta$, *and let*

$$D^{(\beta)} = \sum_{x\in\Lambda} 1_x \otimes D_x^{(\beta)} = \sum_{x\in\Lambda} \mathbf{1}_x \otimes \left(d_0^{(\beta)}(x)\mathbf{1} - \vec{d}^{(\beta)}(x)\cdot\sigma\right)$$

be the associated matrix-valued potential. The 1-pdm $\gamma_\beta\left(D^{(\beta)}\right)$ *defined by*

$$\gamma_\beta(D^{(\beta)}) := \left(1+\exp\left[\beta\left(T_\mu + UD^{(\beta)}\right)\right]\right)^{-1} \tag{24}$$

has the following properties:

(a) $\gamma_\beta\left(D^{(\beta)}\right)$ *is a HF-Gibbs state:*

$$\Omega^{(\beta)} = \Omega\left(\gamma_\beta\left(D^{(\beta)}\right)\right) = \mathcal{R}_\beta\left(d_0^{(\beta)}, \vec{d}^{(\beta)}\right) - |\Lambda|(\beta^{-1}2\ln 2+\mu) \tag{25}$$

(b) $\gamma_\beta\left(D^{(\beta)}\right)$ *is self consistent:*

$$d_0^{(\beta)}(x) = \frac{1}{2}\,\mathrm{Tr}\left[\gamma_\beta\left(D^{(\beta)}\right)\mathbf{1}_x\right] - \frac{1}{2} = n^{(\beta)}(x) - \frac{1}{2} \tag{26}$$

$$\vec{d}^{(\beta)}(x) = \frac{1}{2}\,\mathrm{Tr}\left[\gamma_\beta\left(D^{(\beta)}\right)(1_x\otimes\sigma)\right] = \vec{s}^{(\beta)}(x) \quad . \tag{27}$$

The theorem states that knowing the extremizers of $\mathcal{R}_\beta$ is equivalent to knowing the HF-Gibbs states. The dimensionality of the variational problem is thereby reduced from $\frac{1}{2}(|\Lambda|^2+|\Lambda|)$ to $4|\Lambda|$. This result can be found in [2] for the special case $\mu = 0$. In this case the proof is considerably simpler because there exists a gap of the effective HF-operator $T_0 + UD^{(\beta)}$ for sufficiently low temperatures (see equation (46)), and its spectrum is symmetric about zero (Gap estimate, Lemma 4.2 in [2]). From this, one concludes that the term $\sum_{x\in\Lambda} U_x(n_\gamma(x) - \frac{1}{2})$ of the HF-functional can be discarded in the case $\mu = 0$ (see Theorem 3 and Lemma 4.4 in [2]).

We give the main ideas of the proof of Theorem 1 (a detailed proof is given in [7]). The first step is to use the simple fact that

$$x^2 \geq 2(x\cdot y) - y^2 \quad \text{for all } x, y \in \mathbb{R}^n ,$$

where equality holds if and only if $x = y$. This implies

$$\left(n_\gamma(x) - \tfrac{1}{2}\right)^2 = \max_{|d_0(x)| \le \frac{1}{2}} \left\{2 n_\gamma(x) d_0(x) - d_0^2(x)\right\}$$

and

$$\vec{s}_\gamma^{\,2}(x) = \max_{|\vec{d}(x)| \le \frac{1}{2}} \left\{2 \vec{s}_\gamma \cdot \vec{d}(x) - \vec{d}^{\,2}(x)\right\} .$$

Inserting this into (21), the HF-grand canonical potential assumes the form

$$\Omega(\gamma) \;=\; \min_{\vec{d} \in B} \max_{d_0 \in K} \left\{ f(d_0, \vec{d}, \gamma) \right\} , \tag{28}$$

where f is given by

$$f\big(d_0, \vec{d}, \gamma\big) \;=\; \mathrm{Tr}_1 \left[(T_\mu - UD)\, \gamma \right] - \frac{1}{\beta} S(\gamma) + \sum_{x \in \Lambda} U_x \left(\vec{d}^{\,2}(x) - d_0^2(x) - d_0(x) \right) .$$

Minimizing over all $0 \le \gamma \le 1$ on both sides of (28), we may exchange the two minima on the right side of (28):

$$\Omega^{(\beta)} \;=\; \min_\gamma \Omega(\gamma) \;=\; \min_\gamma \min_{\vec{d} \in B} \max_{d_0 \in K} f\big(d_0, \vec{d}, \gamma\big) \;=\; \min_{\vec{d} \in B} \min_\gamma \max_{d_0 \in K} f\big(d_0, \vec{d}, \gamma\big) .$$

Note that exchanging the minimization over the γ with the maximization over the d_0 yields only a lower bound:

$$\Omega^{(\beta)} \;\ge\; \min_{\vec{d} \in B} \max_{d_0 \in K} \min_\gamma f\big(d_0, \vec{d}, \gamma\big) .$$

It is a straightforward calculation to show that

$$\gamma_\beta(D) = \Big(1 + \exp\left[\beta(T_\mu + UD)\right]\Big)^{-1}$$

minimizes $f\big(d_0, \vec{d}, (\cdot)\big)$ for all $\big(d_0, \vec{d}\,\big)$, and that inserting $\gamma_\beta(D)$ into f yields $\mathcal{R}_\beta\big(d_0, \vec{d}\,\big) - |\Lambda|(\beta^{-1} 2 \ln 2 + \mu)$. So, we have that

$$\Omega^{(\beta)} \ge \min_{\vec{d} \in B} \max_{d_0 \in K} \mathcal{R}_\beta\big(d_0, \vec{d}\,\big) - |\Lambda|(\beta^{-1} 2 \ln 2 + \mu) . \tag{29}$$

To show that equality holds in (29) we use the fact that $\mathcal{R}_\beta$ is strictly concave in d_0 and that $\mathcal{R}_\beta$ attains its unique maximum with respect to d_0 in the interior of K (for all $\vec{d} \in \mathrm{B}$). This leads to the observation that for all $\vec{d} \in \mathrm{B}$

$$\max_{d_0 \in K} f\left(d_0, \vec{d}, \gamma_\beta\big(D^{(\beta)}\big)\right) \le \max_{d_0 \in K} \mathcal{R}_\beta\big(d_0, \vec{d}\,\big) \quad \text{holds true.}$$

From this we conclude that equality holds in Equation (29). Setting the partial derivatives of $\mathcal{R}_\beta$ to zero at $\big(d_0^{(\beta)}, \vec{d}^{(\beta)}\big)$ yields the self-consistency relations (26) and (27).

4 The Limit $\beta \to \infty$

Now we would like to use the results of the previous section to get the HF approximation of the ground state energy $E_{hf} = \min_{0 \leq \gamma \leq 1} \mathcal{E}(\gamma)$ and HF-ground states (minimizers of $\mathcal{E}$). We can not proceed in the same way as in the proof of Theorem 1 for the spectral function $\frac{1}{2}|\lambda| = \lim_{\beta \to \infty} \frac{1}{\beta} \ln \cosh(\frac{\beta}{2}\lambda)$, because the kink of $|\lambda|$ at $\lambda = 0$ causes problems if

$$T_\mu - UD^{(\infty)} := \lim_{\beta \to \infty} T_\mu - UD^{(\beta)}$$

has zero eigenvalues. In this case it is not possible to conclude, as in the proof of Theorem 1, that the partial derivatives of $\mathcal{R}_\beta$ vanish at its extremum. So, we have to check that all the statements of Theorem 1 persist to hold in the zero temperature limit. This is the subject of Theorem 2. Let $\left(d_0^{(\beta)}, \vec{d}^{(\beta)}\right)$ be an extremizer of the function $\mathcal{R}_\beta$ for $\beta < \infty$, and let $D_x^{(\beta)}$ be the associated matrix-valued potentials, $D^{(\beta)} = \sum_{x \in \Lambda} 1_x \otimes D_x^{(\beta)}$. If β tends to infinity, $\gamma_\beta\left(D^{(\beta)}\right)$ is a sequence on the compact set $\{\|\gamma\|_{\mathcal{H}} \leq 1\}$ ($\mathcal{H}$ is finite dimensional). We may choose a converging subsequence of the sequence $\left(d_0^{(\beta)}, \vec{d}^{(\beta)}\right)$, which we denote $\left(d_0^{(\beta)}, \vec{d}^{(\beta)}\right)$ again, such that

$$\lim_{\beta \to \infty} \gamma_\beta\left(D^{(\beta)}\right) =: \gamma_\infty \tag{30}$$

exists. We do not make any statements about the uniqueness of this limit. Let

$$\left(d_0^{(\infty)}, \vec{d}^{(\infty)}\right) := \lim_{\beta \to \infty} \left(d_0^{(\beta)}, \vec{d}^{(\beta)}\right), \tag{31}$$

and let $D^{(\infty)}$ be the associated matrix-valued potential:

$$D^{(\infty)} = \sum_{x \in \Lambda} 1_x \otimes \left(d_0^{(\infty)}(x) 1 - \vec{d}^{(\infty)}(x) \cdot \sigma\right).$$

Theorem 2 [7]

$$\textit{Let} \quad \mathcal{R}_\infty(d_0, \vec{d}) := -\frac{1}{2} \operatorname{Tr}\left[|T_\mu + UD|\right] + \sum_{x \in \Lambda} U_x \left(\vec{d}^{\,2}(x) - d_0^2(x)\right) \tag{32}$$

then it holds:

1. *$\mathcal{R}_\beta$ converges to $\mathcal{R}_\infty$ uniformly in* K $\times$ B*:*

$$\left|\mathcal{R}_\infty(d_0, \vec{d}) - \mathcal{R}(d_0, \vec{d})\right| \leq \frac{2|\Lambda| \ln 2}{\beta}. \tag{33}$$

2. *$\left(d_0^{(\infty)}, \vec{d}^{(\infty)}\right)$ is an extremizer of $\mathcal{R}_\infty$ and it holds*

$$E_{hf}(\mu) = \mathcal{R}_\infty\left(d_0^{(\infty)}, \vec{d}^{(\infty)}\right) - \mu|\Lambda|. \tag{34}$$

3. Let $P_< = \chi\left[T_\mu + UD^{(\infty)} < 0\right]$,

and $P_0 = \chi\left[\left[T_\mu + UD^{(\infty)} = 0\right] = \sum_{j=1}^{m} |\varphi_j\rangle\langle\varphi_j|\right.$.

There exist numbers $0 \le \lambda_1^{(\infty)}, \ldots, \lambda_m^{(\infty)} \le 1$ *such that*

$$\gamma_\infty = P_< + \sum_{j=1}^{m} \lambda_m^{(\infty)} |\varphi_j\rangle\langle\varphi_j| , \tag{35}$$

where $\langle\varphi_i|\varphi_j\rangle = \delta_{ij}$ *and* $|\varphi_j\rangle\langle\varphi_j|$ *denotes the orthogonal projection onto* φ_j.

4. γ_∞ *is self consistent:*

$$d_0^{(\infty)}(x) = \frac{1}{2}\,\mathrm{Tr}\left[\gamma_\infty \mathbf{1}_x\right] - \frac{1}{2} , \tag{36}$$

$$\vec{d}^{(\infty)}(x) = \frac{1}{2}\,\mathrm{Tr}\left[\gamma_\infty (1_x \otimes \sigma)\right] . \tag{37}$$

5. γ_∞ *is a HF-ground state:*

$$E_{hf} = \mathcal{E}\left(\gamma_\infty\right) . \tag{38}$$

We sketch the proof of Theorem 2 and refer the reader to [7] for details. First, the identity

$$\frac{1}{\beta}\ln\cosh\left(\frac{\beta x}{2}\right) - \frac{|x|}{2} = \frac{1}{\beta}\ln\left[\frac{1}{2}\left(1 + e^{-\beta|x|}\right)\right] \tag{39}$$

implies (33) and establishes the continuity of $\beta \mapsto \mathcal{R}_\infty\big(d_0, \vec{d}\big)$ at $\beta = \infty$.

Secondly, one checks that $\mathcal{R}_\infty\big(d_0, \vec{d}\big)$ is differentiable with respect to $d_0(x)$ and $\vec{d}(x)$, for all $x \in \Lambda$. Since this derivative is bounded uniformly in β, the map $\big(\beta, d_0, \vec{d}\big) \mapsto \mathcal{R}_\infty\big(d_0, \vec{d}\big)$ is continuous on (the compact set) $[1,\infty] \times \mathrm{K} \times \mathrm{B}$. Using this continuity, one proves (2.).

We turn to establishing (3.). To this end, we introduce an orthonormal basis, $\{\varphi_j^\infty\}_{j=1}^{2|\Lambda|}$, of eigenvectors of $T + UD^\infty$, corresponding to eigenvalues $\{e_j^\infty\}_{j=1}^{2|\Lambda|}$, i.e.,

$$(T + UD^\infty)\varphi_j^\infty = e_j^\infty \varphi_j^\infty .$$

Since $D \mapsto T + UD$ is continuous with respect to the operator norm and $D^{(\beta)} \to D^\infty$, we may assume without loss of generality that $\varphi_j^\beta \to \varphi_j^\infty$ and $e_j^\beta \to e_j^\infty$, as $\beta \to \infty$. Let

$$f_\beta(x) = \big(1 + \exp[\beta x]\big)^{-1} .$$

Then

$$\gamma_\beta\big(D^\beta\big) = \sum_{j=1}^{2|\Lambda|} f_\beta(e_j^\beta)\; |\varphi_j^\beta\rangle\langle\varphi_j^\beta| , \tag{40}$$

where $|\varphi_j^\beta\rangle\langle\varphi_j^\beta|$ denotes the orthogonal projection onto φ_j^β. Clearly, as $\beta \to \infty$

$$f_\beta(e_j^\beta) \quad \longrightarrow \left\{ \begin{array}{lll} 1 & : & \text{for } e_j^\infty < 0 \,, \\ 0 & : & \text{for } e_j^\infty > 0 \,, \end{array} \right.$$

but for those e_j^β, for which $e_j^\infty = 0$, it is not even clear, a priori, that the limit $\lim_{\beta\to\infty} f_\beta(e_j^\beta)$ exists at all. This is guaranteed only by (30) , where we have selected a convergent subsequence $\gamma_\beta\big(D^{(\beta)}\big) \to \gamma_\infty$. So, knowing the convergence of $f_\beta(e_j^\beta)$ for all j, we set

$$\lambda_j^{(\beta)} \;:=\; \lim_{\beta\to\infty} f_\beta(e_j^\beta) \,, \tag{41}$$

which establishes (3.). The remaining claims (4.) and (5.) follow from (23), (24), (26), (27) and the continuity of $\mathcal{R}_\beta\big(d_0, \vec{d}\big)$.

5 HF-Theory at Half-Filling

In this section, we review the results of Bach, Lieb and Solovej [2] on the characterization of the minimizers γ_β for $\mu = 0$ (half-filling), and bipartite hopping matrix t. The latter means that we can write the lattice Λ as a disjoint union of two sets $A, B \subset \Lambda$, $A \cup B = \Lambda$, $A \cap B = \emptyset$, such that

$$t_{xy} \neq 0 \ \text{ implies } \ (x,y) \in (A \times B \cup B \times A) \,. \tag{42}$$

Moreover, we assume henceforth, without further mentioning, that t is connected, i.e., for any two points $x, y \in \Lambda$ there is a path $\{x, x_1, x_2, \ldots, x_n, y\} \subset \Lambda$ such that $t_{xx_1} t_{x_1x_2} \cdots t_{x_ny} \neq 0$.

The virtue of the bipartiteness of t is that t is unitarily equivalent to $-t$. This unitary is given by the multiplication operator $(-1)^x = [(-1)^x]^* = [(-1)^x]^{-1}$ on $\mathcal{H}$, where $(-1)^x$ denotes the function

$$(-1)^x = \left\{ \begin{array}{rl} -1 & \text{for } x \in A \,, \\ 1 & \text{for } x \in B \,. \end{array} \right.$$

(We use the same notation $(-1)^x$ for the operator $(-1)^x \otimes 1$ on $\mathcal{H} \otimes \mathbb{C}^2$.)

Indeed, we observe that

$$\{(-1)^x, t\} \;=\; (-1)^x t + t(-1)^x \;=\; 0 \,. \tag{43}$$

Additionally using $\mu = 0$, we thus obtain

$$\{(-1)^x, T_{\mu=0}\} \;=\; \{t, (-1)^x\} \otimes 1 \;=\; 0 \,. \tag{44}$$

Equation (44) is the key observation for establishing the following two theorems.

Theorem 3 [2] *Let t be bipartite, and $\mu = 0$. Then*

1.

$$\Omega^{(\beta)} \;=\; \min_{0 \leq \gamma \leq 1} \Omega(\gamma) \;=\; \min_{\vec{d} \in B} \mathcal{R}_\beta \left(d_0 \equiv 0, \vec{d}, \right) - |\Lambda| \beta^{-1} 2 \ln 2 \, . \tag{45}$$

where, for $\beta < \infty$

$$\mathcal{R}_\beta(0, \vec{d}) = -\frac{1}{\beta} \operatorname{Tr} \left[\ln \cosh \left\{ \frac{\beta}{2} (T + UD) \right\} \right] + \sum_{x \in \Lambda} U_x \vec{d}^{\,2}(x) \, ,$$

and for $\beta = \infty$

$$\mathcal{R}_\infty(0, \vec{d}) := -\frac{1}{2} \operatorname{Tr}[|T + UD|] + \sum_{x \in \Lambda} U_x \vec{d}^{\,2}(x)$$

Here

$$D := -\sum_{x \in \Lambda} 1_x \otimes \vec{d}(x) \cdot \sigma \, .$$

2. *If $\mathcal{R}_\beta \left(0, \vec{d}^{(\beta)} \right) = \Omega^{(\beta)}$ and $D^{(\beta)} = -\sum_{x \in \Lambda} 1_x \otimes \left(\vec{d}(x) \cdot \sigma \right)$, then the spectrum, $\sigma \left(T + UD^{(\beta)} \right)$, of $T + UD^{(\beta)}$*

   (a) *is symmetric about 0, i.e., $\sigma \left(T + UD^{(\beta)} \right) = -\sigma \left(T + UD^{(\beta)} \right)$,*

   (b) *has a gap about 0, i.e., any $\lambda \in \sigma \left(T + UD^{(\beta)} \right)$ obeys*

$$|\lambda| \;\geq\; \frac{1}{4|\Lambda|} \min_{x \in \Lambda} \{U_x\} - \frac{2 \ln 2}{\beta} \, . \tag{46}$$

Note that, although (1.) in Theorem 3 does not claim that necessarily $d_0 \equiv 0$ for $\mathcal{R}_\beta(d_0) = \Omega^{(\beta)}$, this follows immediately from the strict concavity of $\mathcal{R}_\beta\big(\cdot, \vec{d}\,\big)$. Thus the self-consistency equation (26) yields

$$n^{(\beta)}(x) \;=\; d_0^{(\beta)} + \frac{1}{2} \;\equiv\; \frac{1}{2} \quad \text{for all } x \in \Lambda \, . \tag{47}$$

In other words, the electron density of the minimizing 1-pdm γ_β is constant and its total particle number equals $|\Lambda|$, hence justifying the term *half filling.*

Theorem 4 [2]*(Antiferromagnetism)*
Let t be bipartite, and $\mu = 0$. Let $\vec{d}^{(\beta)}$ be a minimizer, i.e., $\mathcal{R}_\beta \left(0, \vec{d}^{(\beta)} \right) = \Omega^{(\beta)}$. Then there exists a fixed unit vector, $\vec{e} \in \mathbb{R}^3$, such that

$$\vec{d}^{(\beta)}(x) \;=\; (-1)^x \cdot |\vec{d}^{(\beta)}(x)| \cdot \vec{e} \, . \tag{48}$$

Moreover $|\vec{d}^{(\beta)}(x)| > 0$, for all $x \in \Lambda$ and sufficiently large β.

Using the self-consistency equation (27), (48) tells us that the spin density of the HF Gibbs state is given by

$$\vec{s}^{(\beta)} = (-1)^x \cdot |\vec{d}^{(\beta)}(x)| \cdot \vec{e}\,, \tag{49}$$

with a nonvanishing antiferromagnetic order for sufficiently small temperatures.

Given further information on the structure of the crystal lattice, i.e., symmetries of t and U, we can further specify the magnitude of the spin density, $|\vec{d}^{(\beta)}(x)|$. To this end we introduce a bijective map $\tau : \Lambda \to \Lambda$ and $\alpha_\tau : \Lambda \to \mathbb{R}$ such that

$$t_{\tau(x),\tau(y)} = \exp\big[-i\big(\alpha_\tau(x) - \alpha_\tau(y)\big)\big] t_{xy} \quad \text{and} \quad U_{\tau(x)} = U_x\,. \tag{50}$$

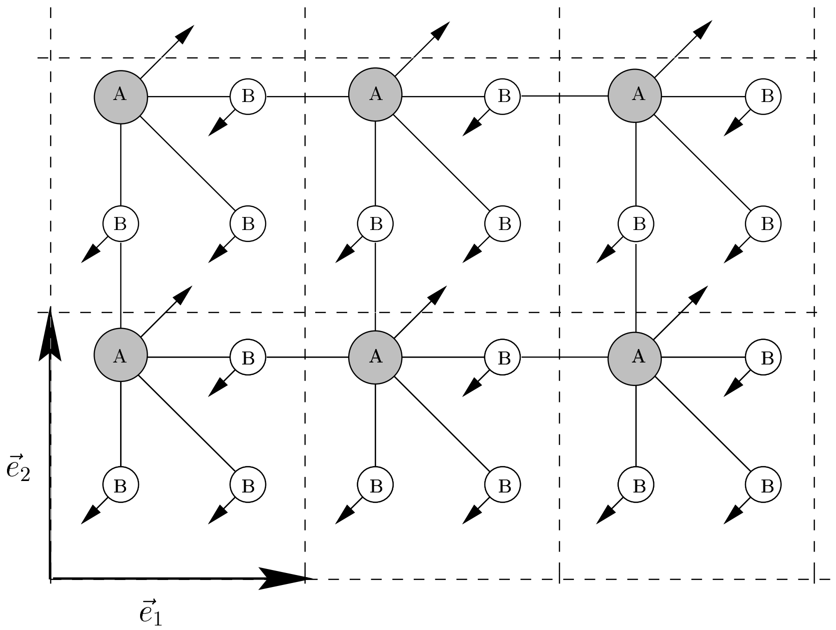
We call τ a **spatial symmetry** of t and U. The set of all spatial symmetries of t and U form a group under composition which we denote by $\mathcal{G}$.

Theorem 5 [2] *Let t be bipartite and $\mu = 0$. Let $\vec{d}^{(\beta)}(x) = (-1)^x \cdot |\vec{d}^{(\beta)}(x)| \cdot \vec{e}$ be a minimizer of $\mathcal{R}_\beta$. Then $|\vec{d}^{(\beta)}(\cdot)|$ is invariant under $\mathcal{G}$, i.e.,*

$$|\vec{d}^{(\beta)}(x)| = |\mathcal{G}|^{-1} \sum_{\tau\in\mathcal{G}} |\vec{d}^{(\beta)}(\tau(x))|\,. \tag{51}$$

We remark that the Theorems 3 - 5 correspond to Lemmata 4.1 - 4.4, Theorem 4.5, and Lemma 4.6 in [2].

To illustrate the contents of Theorems 4 and 5 we consider an example of a lattice with an elementary cell that contains 4 atoms (see figure).



Note that this lattice is invariant under translations by multiples of $\vec{e}_1$ or $\vec{e}_2$. Theorem 4 tells us in which directions the arrows point, and Theorem 5 states that the magnitude of the arrows is invariant under these translations

Finally, we come back to Equation (44), which states that $\{T, (-1)^x 1\} = 0$. To illustrate its importance, we derive Theorems 4 and 5, and (2.) of Theorem 3 from this identity, for the special case of zero temperature and of a translational invariant d-dimensional hypercubic lattice, $\Lambda = (\mathbb{Z}/2L\mathbb{Z})^d$, i.e.,

$$t_{x,y} \;=\; t_{x-y,0} \quad \text{and} \quad U_x \equiv U = \text{const}\,. \tag{52}$$

Denoting $\tau_z := x - z$, the translation by $z \in \Lambda$, with $\alpha_{\tau_z} = 0$, we observe that the group $\mathcal{G}$ of spatial symmetries is simply given by all translations,

$$\mathcal{G} = \{\tau_z | z \in \Lambda\}\;. \tag{53}$$

Thus, Theorems 4 and 5 are contained in the following lemma, for the translation invariant case at zero temperature.

Lemma 3 *Let t be bipartite, $\mu = 0$, $\beta = \infty$, and t, U translation invariant as in (52). Then $\mathcal{R}_\infty(0, \vec{d}) = \Omega^\infty$ implies that, for some unit vector, $\vec{e} \in \mathbb{R}^3$,*

$$\vec{d}(x) \;=\; (-1)^x \cdot \triangle \cdot \vec{e}\,, \tag{54}$$

where $\triangle > 0$ is the unique solution of the BCS-gap equation,

$$\sum_{j=1}^{|\Lambda|} \frac{U}{\sqrt{\nu_j^2 + U^2\Delta^2}} \;=\; 1\;, \tag{55}$$

provided

$$\frac{1}{|\Lambda|}\sum_{j=1}^{|\Lambda|} \frac{1}{|2\nu_j|} \;\geq\; \frac{1}{U}\;, \tag{56}$$

and $\Delta = 0$ in the case that

$$\frac{1}{|\Lambda|}\sum_{j=1}^{|\Lambda|} \frac{1}{|2\nu_j|} \;<\; \frac{1}{U}\;. \tag{57}$$

ν_j denote the eigenvalues of t. Moreover, the spectrum of the HF operator, $T + UD^\infty$, is given by

$$\sigma\left(T + UD^\infty\right) \;=\; \left\{ \pm\sqrt{\nu^2 + U^2\Delta^2} \;\middle|\; \nu \in \sigma(t) \right\}\;. \tag{58}$$

We remark that usually $\frac{1}{|\Lambda|}\sum_{j=1}^{|\Lambda|} |2\nu_j|^{-1}$ diverges in the thermodynamic limit, so Δ is positive (in this limit) even for small coupling U.

Proof: The key idea of the proof can already be found in [6]. We write $|x| = \sqrt{x^2}$ and thus

$$\begin{aligned} \mathcal{R}_\infty(0, \vec{d}) \;&=\; -\frac{1}{2}\mathrm{Tr}\left[\sqrt{(T+UD)^2}\right] + U\sum_{x\in\Lambda}\vec{d}(x)^2 \\ &=\; -\frac{1}{2}\mathrm{Tr}\left[\sqrt{T^2 + U^2D^2 + \{T, UD\}}\right] + U\sum_{x\in\Lambda}\vec{d}(x)^2\;. \end{aligned} \tag{59}$$

Using the unitarity of $(-1)^x$, Identity (44), and $[(-1)^x, UD] = 0$, we observe that

$$\begin{aligned} \mathrm{Tr}\left[\sqrt{T^2 + U^2D^2 + \{T, UD\}}\right] &= \mathrm{Tr}\left[\sqrt{T^2 + U^2D^2 + \{(-1)^x T(-1)^x\,,\, UD\}}\right] \\ &= \mathrm{Tr}\left[\sqrt{T^2 + U^2D^2 - \{T, UD\}}\right] . \end{aligned} \tag{60}$$

Hence, strict concavity of $\sqrt{\cdot}$ on $\mathbb{R}_+$ implies that

$$\begin{aligned} \mathcal{R}_\infty(0, \vec{d}) &= -\frac{1}{4}\mathrm{Tr}\left[\sqrt{T^2 + U^2D^2 + \{T, UD\}}\right] \qquad (61) \\ &\quad -\frac{1}{4}\mathrm{Tr}\left[\sqrt{T^2 + U^2D^2 - \{T, UD\}}\right] + U\sum_{x\in\Lambda}\vec{d}(x)^2 \\ &\geq -\frac{1}{2}\mathrm{Tr}\left[\sqrt{T^2 + U^2D^2}\right] + U\sum_{x\in\Lambda}\vec{d}(x)^2 , \end{aligned}$$

with equality if and only if

$$\{T\,,\,UD\} = 0 . \tag{62}$$

Note that $D^2 = \sum_{x\in\Lambda}\vec{d}(x)^2\,\mathbf{1}_x$, due to $\{\sigma^{(i)}, \sigma^{(j)}\} = 2\delta_{ij}$. Next, we define $\delta > 0$ by

$$|\Lambda|\delta = \sum_{x\in\Lambda}\vec{d}(x)^2$$

and

$$D_z = -\sum_{x\in\Lambda} 1_x \otimes (\vec{\sigma}\cdot\vec{d}(x-z)) ,$$

for any $z \in \Lambda$, so $D = D_0$. Note that D_z represents the spin density $\vec{d}$ shifted by $z \in \Lambda$. Since t and U are translation invariant, we have

$$\mathrm{Tr}\left[\sqrt{T^2 + U^2D^2}\right] = \mathrm{Tr}\left[\sqrt{T^2 + U^2D_z^2}\right] , \tag{63}$$

for any $z \in \Lambda$. So, again using the strict concavity of $\sqrt{\cdot}$, we obtain

$$\begin{aligned} \mathcal{R}_\infty(0, \vec{d}) &\geq -\frac{1}{2|\Lambda|}\sum_{z\in\Lambda}\mathrm{Tr}\left[\sqrt{T^2 + U^2D_z^2}\right] + U\sum_{x\in\Lambda}\vec{d}(x)^2 \\ &\geq -\frac{1}{2}\mathrm{Tr}\left[\sqrt{T^2 + U^2|\Lambda|^{-1}\sum_{z\in\Lambda}D_z^2\,\mathbf{1}}\right] + U\sum_{x\in\Lambda}\vec{d}(x)^2 \\ &= -\frac{1}{2}\mathrm{Tr}\left[\sqrt{T^2 + U^2\delta^2\,\mathbf{1}}\right] + U|\Lambda|\,\delta^2 , \end{aligned} \tag{64}$$

with equality if and only if

$$\{T\,,\,UD\} = 0 \quad \text{and} \quad |\vec{d}(x)| = \delta, \quad \text{for all } x \in \Lambda . \tag{65}$$

It is easily checked that $\vec{d}(x) := (-1)^x\cdot\delta\cdot\vec{e}$, for some fixed unit vector, $\vec{e} \in \mathbb{R}^3$, obeys (65) and is the only choice that does. We conclude that

$$\mathcal{R}_\infty(0, \vec{d}) \geq \mathcal{R}_\infty(0, (-1)^x\cdot\delta\cdot\vec{e}) = -\sum_{j=1}^{|\Lambda|}\sqrt{\nu_j^2 + U^2\delta^2} + U|\Lambda|\,\delta^2 , \tag{66}$$

with equality if and only if $\vec{d}(x) := (-1)^x \cdot \delta \cdot \vec{e}\,'$, for some unit vector $\vec{e}\,' \in \mathbb{R}^3$. The BCS gap equation (55) follows by minimizing the right side of (66) with respect to δ. ■

References

[1] V. Bach. Error bound for the Hartree-Fock energy of atoms and molecules. *Commun. Math. Phys.*,**147**:527-548, 1992.

[2] V. Bach, E.H. Lieb, J.P. Solovej. Generalized Hartree- Fock theory and the Hubbard model. *J.Stat.Phys.*,**76**:3-90, 1994.

[3] V. Bach. Applications for a Correlation Estimate for Fermions. *"Proceedings of the XIth International Congress of Mathematical Physics"; International Press*, 479-489, 1995.

[4] V. Bach, J. Poelchau. Accuracy of the Hartree-Fock Approximation for the Hubbard Model. *submitted to J.Math.Phys.*, 1996.

[5] M. Gaudin. Une démonstration simplifiée du théorème de Wick en méchanique statistique. *Nucl. Phys.*, **15**:89-91, 1960.

[6] T. Kennedy und E.H. Lieb. An itinerant electron model with crystalline or magnetic long range order. *Physica*, **138A**:320-358, 1986.

[7] J. Poelchau. Eine analytische Fehlerabschätzung der Hartree-Fock Näherung für die Grundzustandsenergie des Hubbardmodells. *Diploma Thesis at the Technische Universität Berlin* 1995, (unpublished).