Commun. Math. Phys. 118, 475–494 (1988)

Communications in Mathematical Physics

A New Method for the Thermomodynamics of the BCS Model

N. G. Duffield[1] and J. V. Pulé[1,2]

[1] Department of Mathematical Physics, University College, Belfield, Dublin 4, Ireland
[2] Dublin Institute for Advanced Studies, Dublin, Ireland

Abstract. Using large deviations in combination with the Berezin-Lieb inequalities, we analyse the phase-transition in the BCS model with non-constant energies and interactions.

1. Introduction

N.N. Bogoliubov and his school have made several attempts at solving the full BCS model with hamiltonian

$$H = -\sum_{\mathbf{k}, s=\pm 1} \varepsilon(\mathbf{k}) a^*_{\mathbf{k},s} a_{\mathbf{k},s} - \frac{1}{V} \sum_{\mathbf{k}} \sum_{\mathbf{k}'} a^*_{-\mathbf{k},-1} a^*_{\mathbf{k},1} U(k,k') a_{\mathbf{k}',1} a_{-\mathbf{k}',-1} \tag{1.1}$$

(see [1] and the references therein). The first of these was by a perturbation expansion which can be made rigorous only at zero temperature. Later they developed a mini-max principle which allowed them to treat a class of interactions. In this paper we provide a new method for treating BCS-type models in the quasi-spin formulation

$$H = -\sum_{\mathbf{k}} \varepsilon(\mathbf{k}) \sigma^z_{\mathbf{k}} - \frac{1}{V} \sum_{\mathbf{k}} \sum_{\mathbf{k}'} \sigma^+_{\mathbf{k}} U(\mathbf{k},\mathbf{k}') \sigma^-_{\mathbf{k}'} , \tag{1.2}$$

which can be applied to more general interactions. The result in this paper have already been announced without proof in [2] and an extension of the method to treat this type of model in the original formulation of (1.1) will be given in [3]. The techniques developed here have also been used for the full spin-boson model [4].

Recently Cegła, Lewis, and Raggio [7] have been able to obtain the free energy density for quantum spin systems with homogeneously decomposable hamiltonians. Their methods, amongst other things, allow them to streamline the treatment of the thermodynamics of the BCS model [5, 6] in the strong coupling limit in which $\varepsilon(\mathbf{k})$ and $U(\mathbf{k},\mathbf{k}')$ are replaced by their average value and of other models whose hamiltonians are functions of the total spin operators. They obtained a large deviation principle [8, 9] for the measures arising from the multiplicities of the irreducible representations in the decomposition of the total spin. In some cases the partition function on each irreducible representation can be calculated explicitly. When this is not so the Berezin-Lieb inequalities [10–12] are used to obtain upper and lower bounds for the free energy which coincide in the thermodynamic limit.

The hamiltonian (1.2) which we treat in this paper is not simply a function of the total spin operators. Our method consists of partitioning the system into smaller subsystems and approximating by hamiltonians which are constructed from functions of the total spin operators for these subsystems. The approximating hamiltonians can then be treated by an extension of the techniques in [7]. The approximation is shown to become exact as the number of subsystems is increased and it is proved also that this limit can be interchanged with the thermodynamic limit. This procedure gives a variational formula for the free energy density. This is done in Sect. 2. The variational problem is treated in Sect. 3. The associated Euler-Lagrange equation turns out to be the gap-equation [13]. It is shown that the solution of this equation bifurcates at a certain critical temperature. In Sect. 4 we modify the work Sects. 2 and 3, to calculate the order parameter which is shown to have a non-zero value below a certain critical temperature. In principle other intensive quantities can be calculated in the same way.

Before leaving the introduction we say something about the work of Anderson and Thouless [14, 15] on the full BCS model. In their treatment the hamiltonian is written in terms of the quasi-particle operators through an arbitrary Bogoliubov transformation. In the resulting expression that part which is not a function of the quasi-particle number operator is neglected; this reduces the problem to a classical one in which the free energy can be obtained by minimizing over a function of the quasi-particle densities. Finally the result is minimized over the free rotational parameter in the Bogoliubov transformation. The variational problem arising, which is solved only in the strong-coupling limit, is very similar to the one we solve in Sect. 3, if we identify the quasi-particle densities with the spin densities and the rotational parameter with the spin orientation.

The BCS model has many simplifying features when studied in the thermodynamic limit from the start (see for example [16, 17]). However this approach is not in the same spirit as in this paper since the methods used are algebraic, although in some cases not explicitly so.

2. The Thermodynamic Limit

We consider a slightly more general model than that described in the introduction. Let $\{\Lambda_\ell : \ell = 1, 2, \ldots\}$ be a sequence of regions of Euclidean space $\mathbb{R}^d$ and denote the volume of Λ_ℓ by V_ℓ; we associate with the region Λ_ℓ the sequence of momenta $\{\mathbf{k}_\ell(j) : j = 1, 2, \ldots\}$, where each $\mathbf{k}_\ell(j)$ is in $\mathbb{R}^\nu$. We make the assumption that the sequence of measures $\{\mu_\ell\}$ giving the distribution of momentum states

$$\mu_\ell(B) = \frac{1}{V} \# \{j : \mathbf{k}_\ell(j) \in B\}$$

for Borel subsets B of $\mathbb{R}^\nu$, converges weakly to a measure μ which is absolutely continuous with respect to Lebesgue measure. We shall be considering only those momenta in a *cut-off* region Ω which is a closed bounded region in $\mathbb{R}^\nu$ and we assume that $\mu(\Omega)<\infty$.

The hamiltonian for our model acts on $\mathcal{H}_\ell=\bigotimes_{i=1}^{N_\ell}\mathcal{D}_i$, where $N_\ell=V_\ell\mu_\ell(\Omega)$ and $\mathcal{D}_i$ is a copy of $\mathbb{C}^2$. It is given by

$$H_\ell=-\sum_{i=1}^{N_\ell}\varepsilon(\mathbf{k}_\ell(i))\sigma_i^z-\frac{1}{V_\ell}\sum_{i=1}^{N_\ell}\sum_{j=1}^{N_\ell}\sigma_i^+U(\mathbf{k}_\ell(i),\mathbf{k}_\ell(j))\sigma_j^- \ , \tag{2.1}$$

where $\varepsilon\in C(\Omega)$, $U\in C(\Omega\times\Omega)$ and $\sigma_i^\#=I\otimes\ldots\otimes\sigma^\#\otimes\ldots\otimes I$, the σ's being the usual Pauli matrices. Here the $\varepsilon(\mathbf{k}_\ell(j))$ include the chemical potential.

Let $f_\ell(\beta)$ be the free energy density

$$f_\ell(\beta)=-\frac{1}{\beta V_\ell}\ln\operatorname{trace}e^{-\beta H_\ell} \ . \tag{2.2}$$

We shall prove that $f_\ell(\beta)$ converges as $\ell\to\infty$, and we shall obtain a variational formula for the limit $f(\beta)$. The variational problem is solved in the next section. Our method is to approximate H_ℓ by a hamiltonian for which the method of [7] can be used. Choose $L>0$ such that $[-L,L]^\nu\supset\Omega$ and for each $M\in\mathbb{N}$ partition $[-L,L]^\nu$ into M^ν disjoint cubes of side $2L/M$. Denote these cubes by $\bar{B}_m^M$, $m=1,\ldots,M^\nu$, and let $B_m^M=\bar{B}_m^M\cap\Omega$. We define the approximating hamiltonian by

$$H_\ell^M=-\sum_{m=1}^{M^\nu}\varepsilon_m^M\sum_{\mathbf{k}_\ell(i)\in B_m^M}\sigma_i^z-\frac{1}{V_\ell}\sum_{m,m'=1}^{M^\nu}U_{m,m'}^M\Big(\sum_{\mathbf{k}_\ell(i)\in B_m^M}\sigma_i^+\Big)\Big(\sum_{\mathbf{k}_\ell(j)\in B_{m'}^M}\sigma_j^-\Big)\ , \tag{2.3}$$

where

$$\varepsilon_m^M=\begin{cases}0 & \text{if } \mu(B_m^M)=0\\ \dfrac{1}{\mu(B_m^M)}\displaystyle\int_{B_m^M}\varepsilon(k)\mu(dk) & \text{if } \mu(B_m^M)\neq0\end{cases}$$

and

$$U_{m,m'}^M=\begin{cases}0 & \text{if } \mu(B_m^M)\mu(B_{m'}^M)=0\\ \dfrac{1}{\mu(B_m^M)\mu(B_{m'}^M)}\displaystyle\int_{B_m^M\times B_{m'}^M}U(k,k')\mu(dk)\mu(dk') & \text{if } \mu(B_m^M)\mu(B_{m'}^M)\neq0\ .\end{cases}$$

Let $f_\ell^M(\beta)$ be the corresponding free energy density, then we have:

Theorem 1. *As $M\to\infty$, $f_\ell^M(\beta)$ converges to $f_\ell(\beta)$ uniformly in ℓ.*

Proof. By Bogoliubov's inequality [18] we have

$$|f_\ell(\beta)-f_\ell^M(\beta)|\leqq\|H_\ell^M-H_\ell\|/V_\ell\ .$$

Therefore

$$|f_\ell(\beta)-f_\ell^M(\beta)|\leqq\sum_{m=1}^{M^\nu}\mu_\ell(B_m^M)\sup_{k\in B_m^M}|\varepsilon_m^M-\varepsilon(k)|$$

$$+\sum_{m,m'=1}^{M^\nu}\mu_\ell(B_m^M)\mu_\ell(B_{m'}^M)\sup_{\substack{k\in B_m^M\\ k'\in B_{m'}^M}}|U_{m,m'}^M-U(k,k')|\ .$$

Since ε and U are uniformly continuous given $\varepsilon>0$ we can find M_0 such that for $M>M_0$,

$$\sup_{k\in B_m^M} |\varepsilon_m^M-\varepsilon(k)|<\varepsilon/C$$

for $m=1,\dots,M^\nu$,

$$\sup_{\substack{k\in B_m^M\\ k'\in B_{m'}^M}} |U_{m,m'}^M-U(k,k')|<\varepsilon/C$$

for $m,m'=1,\dots M^\nu$, where C is a number greater than $\mu_\ell(\Omega)+(\mu_\ell(\Omega))^2$ for all ℓ.

Then for $M>M_0$,

$$|f_\ell(\beta)-f_\ell^M(\beta)|<\frac{\varepsilon}{C}\left\{\sum_{m=1}^{M^\nu}\mu_\ell(B_m^M)+\sum_{m,m'=1}^{M^\nu}\mu_\ell(B_m^M)\mu_\ell(B_{m'}^M)\right\}$$

$$=\frac{\varepsilon}{C}\,(\mu_\ell(\Omega)+(\mu_\ell(\Omega))^2<\varepsilon\ . \quad \square$$

We next obtain a variational expression for the limiting approximate free energy density.

Theorem 2. $f^M(\beta)=\lim_{\ell\to\infty} f_\ell^M(\beta)$ *exists and is given by*

$$f^M(\beta)=-\sup\{\mathcal{S}^M(r,\theta,\phi):r\in[0,1]^{M^\nu},\theta\in[0,\pi]^{M^\nu},\phi\in[0,2\pi]^{M^\nu}\} \tag{2.4}$$

where

$$\begin{aligned}\mathcal{S}^M(r,\theta,\phi)=&\frac{1}{2}\sum_{m=1}^{M^\nu}\mu(B_m^M)\varepsilon_m^M r_m\cos\theta_m\\ &+\sum_{m,m'=1}^{M^\nu}\mu(B_m^M)\mu(B_{m'}^M)U_{m,m'}^M r_m r_{m'}\sin\theta_m\sin\theta_{m'}\cos(\phi_m-\phi_{m'})\\ &-\frac{1}{\beta}\sum_{m=1}^{M^\nu}\mu(B_m^M)I(r_m)+\frac{1}{\beta}\ln 2\mu(\Omega)\end{aligned} \tag{2.5}$$

and I is given by

$$I(x)=\tfrac{1}{2}\{(1+x)\ln(1+x)+(1-x)\ln(1-x)\}\ . \tag{2.6}$$

To prove this theorem we shall use the method of large deviations and a result from [7]. For the sake of completeness we shall first describe Varadhan's formulation of the large deviation method and then state the required results from [7].

Let $\{\mathbb{K}_n:n=1,2,\dots\}$ be a sequence of probability measures on the Borel subset of a complete separable metric space E and $\{V_n:n=1,2,\dots\}$ a divergent sequence of positive numbers. We say that $\{\mathbb{K}_n\}$ satisfies the large deviation principle with constants $\{V_n\}$ and rate function $I:E\to[0,\infty]$ if the following conditions hold:

(i) I is lower semicontinuous on E.

(ii) For each $m<\infty$, $\{x:I(x)\leqq m\}$ is compact,

(iii) For each closed subset C of E,

$$\limsup_{n\to\infty} \frac{1}{V_n} \ln \mathbb{K}_n(C) \leqq -\inf_{x\in C} I(x) \ .$$

(iv) For each open subset G of E,

$$\liminf_{n\to\infty} \frac{1}{V_n} \ln \mathbb{K}_n(G) \geqq -\inf_{x\in G} I(x) \ .$$

We shall need the following version of Varadhan's theorem (cf. Theorem 3.4 of [8]).

Varadhan's Theorem. *Suppose that the sequence of probability measures $\{\mathbb{K}_n\}$ on E satisfies the large deviation principle with constants $\{V_n\}$ and rate function I. Let $\{f_n\}$ be a sequence of continuous functions $f_n: E\to\mathbb{R}$ which are uniformly bounded above, and suppose that f_n converges to $f: E\to\mathbb{R}$ uniformly on bounded sets; then*

$$\lim_{n\to\infty} \frac{1}{V_n} \ln \int_E \exp(V_n f_n(x))\mathbb{K}_n(dx) = \sup_E \{f(x)-I(x)\} \ .$$

Let π be the irreducible unitary representation of $SU(2)$ acting on $\mathbb{C}^2$. Let $\mathscr{H}_n$ be the tensor product of $\mathbb{C}^2$ with itself n times and define the unitary representation π_n of $SU(2)$ on $\mathscr{H}_n$ by

$$\pi_n(g) = \underbrace{\pi(g)\otimes\pi(g)\otimes\ldots\otimes\pi(g)}_{n \text{ times}} \ , \qquad g\in SU(2) \ .$$

For $n>1$, π_n is reducible and decomposes into the direct sum

$$\pi_n = \bigoplus_{J\in A_n} \bigoplus_{k=1}^{c(n,J)} \pi^{J,k} \ ,$$

where $A_n = \{0, 1, \ldots, n/2\}$ if n is even and $A_n = \{1/2, 3/2, \ldots, n/2\}$ if n is odd. $\pi^{J,k}$ is a copy of the irreducible representation π^J which acts on $\mathbb{C}^{2J+1}$, and has multiplicity $c(n, J)$. In [7] it is proved that the multiplicities $c(n, J)$ have the following property:

Lemma 1. *Define a probability measure $\mathbb{P}_n$ on the interval* $[0, 1]$ *by*

$$\mathbb{P}_n(B) = \frac{1}{2^n} \sum_{\left\{J:\frac{2J}{n}\in B\right\}} (2J+1)c(n, J) \ ,$$

where B is a Borel subset of $[0, 1]$. *Then the sequence of measures $\{\mathbb{P}_n: n=1, 2, \ldots\}$ satisfies the large deviation principle with constants $\{n\}$ and rate function I, where*

$$I(x) = \tfrac{1}{2}\{(1+x)\ln(1+x) + (1-x)\ln(1-x)\} \ .$$

Proof of Theorem 2. Let $N_\ell^m = V_\ell \mu_\ell(B_m^M)$ and let h_ℓ be the operator on $\bigotimes_{m=1}^{M^\nu} \mathscr{D}_i$ defined by

$$h_\ell = -\sum_{m=1}^{M^\nu} \varepsilon_m^M \sigma_m^z - \frac{1}{V_\ell} \sum_{m=1}^{M^\nu} \sum_{m'=1}^{M^\nu} U_{m,m'}^M \sigma_m^+ \sigma_{m'}^- \ .$$

If p_n and p^J are the representations of the Lie algebra of $SU(2)$ corresponding to π_n and π^J respectively, then

$$H_\ell^M = p_{N_\ell^1} \otimes \ldots \otimes p_{N_\ell^{M^\nu}} h_\ell \ .$$

Thus

$$\operatorname{trace} e^{-\beta H_\ell^M} = \sum_{J \in A_{N_\ell^1} \times \ldots \times A_{N_\ell}^{M^\nu}} c(N_\ell^1, J_1) \ldots c(N_\ell^{M^\nu}, J_{M^\nu}) \operatorname{trace} \exp -\beta \{(p^{J_1} \otimes \ldots \otimes p^{J_{M^\nu}}) h_\ell\} \ . \tag{2.7}$$

Let $f_\ell^M(\beta, \cdot) : [0,1]^{M^\nu} \to \mathbb{R}$ be defined by

$$f_\ell^M(\beta, r) = -\frac{1}{\beta V_\ell} \ln \operatorname{trace} \exp -\beta \{(p^{J_1} \otimes \ldots \otimes p^{J_{M^\nu}}) h_\ell\} \tag{2.8}$$

if $r_j = 2J_j / N_\ell^j, j = 1, \ldots, M^\nu$. By using the Berezin-Lieb inequalities [12] in the Appendix we obtain upper and lower bounds for $f_\ell^M(\beta, r)$:

$$\prod_{m=1}^{M^\nu} \left(1 + \frac{r_m}{N_\ell^m}\right) \int_{(S^2)^{M^\nu}} d\Omega_{M^\nu}(\theta, \phi) e^{-\beta V_\ell \bar{f}_\ell^M(r, \theta, \phi)} \leqq e^{-\beta V_\ell f_\ell^M(\beta, r)}$$
$$\leqq \prod_{m=1}^{M^\nu} \left(1 + \frac{r_m}{N_\ell^m}\right) \int_{(S^2)^{M^\nu}} d\Omega_{M^\nu}(\theta, \phi) e^{-\beta V_\ell \underline{f}_\ell^M(r, \theta, \phi)} \tag{2.9}$$

where $d\Omega_{M^\nu} = \frac{1}{(4\pi)^{M^\nu}} \prod_{m=1}^{M^\nu} \sin\theta_m d\theta_m d\phi_m$, and $\bar{f}_\ell^M(r, \cdot), \underline{f}_\ell^M(r, \cdot)$ are real valued functions on $(S^2)^{M^\nu}$ defined by

$$\bar{f}_\ell^M(r, \theta, \phi) = f_{0,\ell}^M(r, \theta, \phi) - \frac{1}{2V_\ell} \sum_{m=1}^{M^\nu} U_{m,m}^M \mu_\ell(B_m^M) r_m \left(1 - \cos\theta_m - \frac{1}{2}\sin^2\theta_m\right) ,$$

$$\begin{aligned}
\underline{f}_\ell^M(r, \theta, \phi) = f_{0,\ell}^M(r, \theta, \phi) &- \frac{1}{2V_\ell} \sum_{\substack{m, m'=1 \\ m \neq m'}}^{M^\nu} U_{m,m'}^M \\
&\cdot \left(\mu_\ell(B_m^M) r_m + \mu_\ell(B_{m'}^M) r_{m'} + \frac{2}{V_\ell}\right) \sin\theta_m \sin\theta_{m'} \cos(\phi_m - \phi_{m'}) \\
&- \frac{1}{2V_\ell} \sum_{m=1}^{M^\nu} U_{m,m'}^M \left\{ \left(\frac{5}{2} \mu_\ell(B_m^M) r_m + \frac{3}{V_\ell}\right) \sin^2\theta_m \right. \\
&\left. - \left(\mu_\ell(B_m^M) r_m + \frac{2}{V_\ell}\right)(1 + \cos\theta_m) \right\} \\
&- \frac{1}{V_\ell} \sum_{m=1}^{M^\nu} \varepsilon_m^M \cos\theta_m
\end{aligned}$$

and

$$\begin{aligned}
f_{0,\ell}^M(r, \theta, \phi) = &-\frac{1}{2} \sum_{m=1}^{M^\nu} \varepsilon_m^M r_m \mu_\ell(B_m^M) \cos\theta_m \\
&- \frac{1}{4} \sum_{m, m'=1}^{M^\nu} \mu_\ell(B_m^M) \mu_\ell(B_{m'}^M) r_m r_{m'} U_{m,m'}^M \sin\theta_m \sin\theta_{m'} \cos(\phi_m - \phi_{m'}) \ .
\end{aligned}$$

We therefore have

$$-\frac{1}{\beta V_\ell}\ln 2^{N_\ell}\int_{([0,1]\times S^2)^{M^\nu}} e^{-\beta V_\ell \underline{f}_\ell^M(r,\theta,\phi)}d\mathbb{K}_\ell(r,\theta,\phi)\leqq f_\ell^M(\beta)$$

$$\leqq -\frac{1}{\beta V_\ell}\ln 2^{N_\ell}\int_{([0,1]\times S^2)^{M^\nu}} e^{-\beta V_\ell \bar{f}_\ell^M(r,\theta,\phi)}d\mathbb{K}_\ell(r,\theta,\phi)\ , \tag{2.10}$$

where $d\mathbb{K}_\ell(r,\theta,\phi)=d\bar{\mathbb{P}}_\ell(r)d\Omega_{M^\nu}(\theta,\phi)$ and $\bar{\mathbb{P}}_\ell$ is the product measure

$$\mathbb{P}_{N_\ell^1}\times\ldots\times\mathbb{P}_{N_\ell^{M^\nu}}\ .$$

By the above lemma the sequence of measures $\{\mathbb{P}_{N_\ell^m}:\ell=1,2,\ldots\}$ satisfies the large deviation principle with constants $\{N_\ell^M\}$ and rate function I. It is clear that this sequence also satisfies the large deviation principle with constants $\{\beta V_\ell\}$ and rate function $\beta^{-1}\mu(B_m^M)I$. Since with respect to the constants $\{\beta V_\ell\}$ the measure $d\Omega_{M^\nu}$ considered as a constant sequence satisfies the large deviation principle with zero rate function, the product measure $\mathbb{K}_\ell$ satisfies the large deviation principle with rate function $\mathscr{I}_\beta^M:([0,1]\times S^2)^{M^\nu}\to[0,\infty]$, which is the sum of the rate functions of the measures $\{\mathbb{P}_{N_\ell^m}:m=1\ldots M^\nu\}$ (see for example Appendix 1 in [7]); that is

$$\mathscr{I}_\beta^M(r,\theta,\phi)=\frac{1}{\beta}\sum_{m=1}^{M^\nu}\mu(B_m^M)I(r_m)\ . \tag{2.11}$$

Both the functions $\bar{f}_\ell^M$ and $\underline{f}_\ell^M$ converge uniformly on compact subsets of $([0,1]\times S^2)^{M^\nu}$ to f_0^M, where

$$f_0^M(r,\theta,\phi)=-\frac{1}{2}\sum_{m=1}^{M^\nu}\varepsilon_m r_m\mu(B_m^M)\cos\theta_m$$

$$-\frac{1}{4}\sum_{m=1}^{M^\nu}\sum_{m'=1}^{M^\nu}\mu(B_m^M)\mu(B_{m'}^M)r_m r_{m'}U_{m,m'}\sin\theta_m\sin\theta_{m'}\cos(\phi_m-\phi_{m'}) \tag{2.12}$$

Therefore by applying Varadhan's theorem to both sides of the inequality (2.10) we obtain

$$f^M(\beta)=-\sup_{([0,1]\times S^2)^{M^\nu}}\{-f_0^M(r,\theta,\phi)-\mathscr{I}_\beta^M(r,\theta,\phi)\}-\frac{1}{\beta}\ln 2\mu(\Omega)\ .\ \square \tag{2.13}$$

We now obtain a variational formula for $f(\beta)$ by combining Theorems 1 and 2.

Theorem 3. *Let* $\mathscr{M}=\{(r,\theta,\phi):r,\theta,\phi\in L^\infty(\Omega,\mu),0\leqq r(k)\leqq 1,0\leqq\theta(k)\leqq\pi,0\leqq\phi(k)\leqq 2\pi\}$, *and define* $\mathscr{S}:\mathscr{M}\to\mathbb{R}$ *by*

$$\mathscr{S}(r,\theta,\phi)=\frac{1}{2}\int_\Omega\varepsilon(k)r(k)\cos\theta(k)\mu(dk)$$

$$+\frac{1}{4}\int_{\Omega\times\Omega}U(k,k')r(k)r(k')\sin\theta(k)\sin\theta(k')$$

$$\cdot\cos(\phi(k)-\phi(k'))\mu(dk)\mu(dk')$$

$$-\frac{1}{\beta}\int_\Omega I(r(k))\mu(dk)+\frac{1}{\beta}\ln 2\mu(\Omega)\ , \tag{2.14}$$

then

$$f(\beta) = -\sup\{\mathscr{S}(r, \Phi, \phi) : (r, \theta, \phi) \in \mathscr{M}\} \ . \tag{2.15}$$

Proof. Let $\mathscr{S}_{\max} = \sup_{\mathscr{M}} \mathscr{S}(r, \theta, \phi)$ and let $\mathscr{M}^M$ be the set functions in $\mathscr{M}$ of the form

$$r = \sum_{m=1}^{M^\nu} r_m 1_{B_m^M} \ , \qquad \theta = \sum_{m=1}^{M^\nu} \theta_m 1_{B_m^M} \ , \qquad \phi = \sum_{m=1}^{M^\nu} \phi_m 1_{B_m^M} \ .$$

Then clearly

$$f^M(\beta) = -\sup\{\mathscr{S}(r, \theta, \phi) : (r, \theta, \phi) \in \mathscr{M}^M\} \ .$$

Therefore $-f^M(\beta) \leqq \mathscr{S}_{\max}$, and so

$$-\liminf_{M \to \infty} f^M(\beta) \leqq \mathscr{S}_{\max} \ .$$

On the other hand given $\varepsilon > 0$ we can find $(r_0, \theta_0, \phi_0) \in \mathscr{M}$ such that

$$\mathscr{S}(r_0, \theta_0, \phi_0) > \mathscr{S}_{\max} - \varepsilon/2 \ .$$

Since $\bigcup_{k=1}^{\infty} \mathscr{M}^{2^k}$ is dense in $L^\infty(\Omega, \mu) \oplus L^\infty(\Omega, \mu) \oplus L^\infty(\Omega, \mu)$, we can find $(\bar{r}_0, \bar{\theta}_0, \bar{\phi}_0) \in \bigcup_{k=1}^{\infty} \mathscr{M}^{2^k}$ such that

$$\mathscr{S}(\bar{r}_0, \bar{\theta}_0, \bar{\phi}_0) - \mathscr{S}(r_0, \theta_0, \phi_0) > \varepsilon/2 \ .$$

If $(\bar{r}_0, \bar{\theta}_0, \bar{\phi}_0) \in \mathscr{M}^{2^p}$, then $-f^{2^p}(\beta) > \mathscr{S}_{\max} - \varepsilon$. But for $k > k'$, $\mathscr{M}^{2^k} \supset \mathscr{M}^{2^{k'}}$, and so $f^{2^{k'}}(\beta) > f^{2^k}(\beta)$. Therefore for all $k \geqq p$,

$$-f^{2^k}(\beta) \geqq \mathscr{S}_{\max} - \varepsilon \ .$$

Thus

$$-\limsup_{k \to \infty} f^{2^k}(\beta) \geqq \mathscr{S}_{\max} - \varepsilon \ ,$$

and since ε is arbitrary $-\limsup_{k \to \infty} f^{2^k}(\beta) \geqq \mathscr{S}_{\max}$. Therefore $-\lim_{k \to \infty} f^{2^k}(\beta) = \mathscr{S}_{\max}$. By Theorem 1 we conclude that $f(\beta) = \lim_{\ell \to \infty} f_\ell(\beta)$ exists and $-f(\beta) = \mathscr{S}_{\max}$. □

3. The Variational Problem

From now on for simplicity we drop the harmless constant $-\frac{1}{\beta} \ln 2\mu(\Omega)$ in $f(\beta)$. In this section we assume that $U(k, k') > 0$ for all $k, k' \in \Omega$. If U is positive then the supremum in Eq. (2.15) can be restricted to ϕ's for which $\phi(k) = \text{constant}$ and is independent of the constant. Also the supremum need only be taken over those θ's for which $\varepsilon(k) \cos\theta(k) \geqq 0$. We can therefore rewrite the first term as

$$\frac{1}{2} \int_\Omega |\varepsilon(k)| r(k) \cos\theta(k) \mu(dk)$$

and restrict the range of θ to $[0, \pi/2]$. Let

$$\bar{\mathcal{M}} = \{(r, \theta) : r, \theta \in L^\infty(\Omega, \mu), 0 \leqq r(k) \leqq 1, 0 \leqq \theta(k) \leqq \pi/2\} ,$$

and define $\bar{\mathcal{S}} : \bar{\mathcal{M}} \to \mathbb{R}$ by

$$\bar{\mathcal{S}}(r, \theta) = \frac{1}{2} \int_\Omega |\varepsilon(k)| r(k) \cos\theta(k) \mu(dk) - \frac{1}{\beta} \int_\Omega I(r(k)) \mu(dk) + \frac{1}{4} \int_{\Omega \times \Omega} U(k, k') r(k) r(k') \sin\theta(k) \sin\theta(k') \mu(dk) \mu(dk') . \tag{3.1}$$

Then we have seen that for U positive

$$f(\beta) = -\sup\{\bar{\mathcal{S}}(r, \theta) : (r, \theta) \in \bar{\mathcal{M}}\} . \tag{3.2}$$

It is more convenient at this stage to use these variables r and s where $s(k) = r(k) \sin\theta(k)$, than r and θ. We therefore put $\mathcal{N} = \{(r, s) : r, s \in L^\infty(\Omega, \mu), 0 \leqq s(k) \leqq r(k) \leqq 1\}$ and define $\mathcal{J} : \mathcal{N} \to \mathbb{R}$ by

$$\mathcal{J}(r, s) = \frac{1}{2} \int_\Omega \mu(dk) |\varepsilon(k)| (r^2(k) - s^2(k))^{1/2} - \frac{1}{\beta} \int_\Omega I(r(k)) \mu(dk) + \frac{1}{4} \int_{\Omega \times \Omega} \mu(dk) \mu(dk') U(k, k') s(k) s(k') . \tag{3.3}$$

Then $f(\beta) = -\sup\{\mathcal{J}(r, s) : (r, s) \in \mathcal{N}\}$.

For $a \geqq 0$ and $0 \leqq y \leqq x \leqq 1$ let

$$g(x, y; a) = \frac{1}{2} a (x^2 - y^2)^{1/2} - \frac{1}{\beta} I(x) . \tag{3.4}$$

Then $x \mapsto g(x, y; a)$ is concave and its supremum is attained at r_y^a, where

$$r_y^a = \begin{cases} \text{the unique solution of } \dfrac{\tanh^{-1} x}{x} = \dfrac{a\beta}{2} (x^2 - y^2)^{-1/2} , & a > 0 \\ y , & a = 0 . \end{cases} \tag{3.5}$$

Let $\mathcal{L} = \{s \in L^\infty(\Omega, \mu) : 0 \leqq s(k) \leqq 1\}$, and define $\mathcal{V} : \mathcal{L} \to \mathbb{R}$ by $\mathcal{V}(s) = \mathcal{J}(r_{s(\cdot)}^{|\varepsilon(\cdot)|}, s)$, then

$$f(\beta) = -\sup\{\mathcal{V}(s) : s \in \mathcal{L}\} .$$

From now on we shall write simply r_s for $r_{s(\cdot)}^{|\varepsilon(\cdot)|}$. Define the linear operator $\tilde{U}_\beta$ on $L^2(\Omega, \mu)$ by

$$(\tilde{U}_\beta \psi)(k) = \int_\Omega \mu(dk') g_\beta(k) g_\beta(k') U(k, k') \psi(k') ,$$

where

$$g_\beta(k) = \begin{cases} (\beta/2)^{1/2} & \text{if} \quad \varepsilon(k) = 0 , \\ \left(\dfrac{\tanh \beta |\varepsilon(k)|/2}{|\varepsilon(k)|} \right)^{1/2} & \text{if} \quad \varepsilon(k) \neq 0 . \end{cases} \tag{3.6}$$

$\tilde{U}_\beta$ is a compact operator. Let $\lambda(\beta) = \|\tilde{U}_\beta\|$. For $\lambda(\beta) \leqq 1$ we can show directly that the supremum of $\mathscr{V}(s)$ is attained and give the maximizing s.

Theorem 4. *If* $\lambda(\beta) \leqq 1$, *then*

$$f(\beta) = -\frac{1}{\beta} \mu(dk) \ln \cosh\left(\frac{\beta|\varepsilon(k)|}{2}\right) . \tag{3.7}$$

Proof. Let $A = \operatorname{supp} \varepsilon$ and let s be an arbitrary element of $\mathscr{L}$. For $0 \leqq t \leqq 1$ let $F(t) = \mathscr{V}(ts)$. For $0 \leqq t < 1$, $F(t)$ is differentiable and

$$\begin{aligned} F'(t) = & -\frac{t}{2}\left\{\int_A \mu(dk) s^2(k)|\varepsilon(k)|(r_{ts}^2(k) - t^2 s^2(k))^{-1/2}\right. \\ & \left. + \int_{\Omega\times\Omega} \mu(dk)\mu(dk')U(k,k')s(k)s(k') - \frac{2}{\beta}\int_{A^c} \tanh^{-1}(ts(k))s(k)\mu(dk)\right\} \\ \leqq & -\frac{t}{2}\left\{\int_\Omega \mu(dk)\left(\frac{s(k)}{g_\beta(k)}\right)^2 - \int_{\Omega\times\Omega} \mu(dk)\mu(dk')U(k,k')s(k)s(k')\right\} . \end{aligned}$$

To obtain this inequality we have used

$$(r_s^2(k) - s^2(k))^{1/2} \leqq \tanh\left(\frac{\beta|\varepsilon(k)|}{2}\right)$$

for $k \in A$ and

$$\tanh s(k) \leqq s(k)$$

for $k \in A^c$. Let $\hat{s}(k) = s(k)/g_\beta(k)$. Then

$$F'(t) = -\frac{t}{2}(\hat{s}, (I - \tilde{U}_\beta)\hat{s}) \leqq -\frac{t}{2}(1 - \lambda(\beta))\|\hat{s}\|_2^2 \leqq 0 .$$

Therefore $F(t) \leqq F(0)$ for $0 \leqq t < 1$. Since $t \mapsto F(t)$ is continuous $F(t) \leqq F(0)$ for $0 \leqq t \leqq 1$ or $\mathscr{V}(s) \leqq \mathscr{V}(0)$. Now $r_0(k) = \tanh(\beta|\varepsilon(k)|/2)$ for $k \in A$ and $r_0(k) = 0$ for $k \notin A$, and so

$$\begin{aligned} \mathscr{V}(0) &= \frac{1}{2}\int_A \mu(dk)|\varepsilon(k)| \tanh\frac{\beta}{2}|\varepsilon(k)| - \frac{1}{\beta}\int_A \mu(dk) I\left(\tanh\frac{\beta}{2}|\varepsilon(k)|\right) \\ &= \frac{1}{\beta}\int_A \mu(dk) \ln\cosh\left(\frac{\beta}{2}|\varepsilon(k)|\right) . \quad \square \end{aligned}$$

For $\lambda(\beta) > 1$ we have to proceed in a different way. We first prove that the supremum is attained [this proof is valid for all $\lambda(\beta)$] and then show that the maximizer is in the interior of the region, and that it satisfies the corresponding Euler-Lagrange equation which in turn is shown to have a unique solution.

Theorem 5. *There is at least one function* s_* *in* $\mathscr{L}$ *such that* $f(\beta) = -\mathscr{V}(s_*)$.

Proof. We can find a sequence $\{s_n\}$ in $\mathscr{L}$ such that $\lim_{n\to\infty} \mathscr{V}(s_n) = -f(\beta)$. $\mathscr{L}$ is a subset of the closed ball of unit radius in $L^\infty(\Omega, \mu)$.

Now since closed balls are compact in the w^*-topology induced by $L^1(\Omega,\mu)$, there is a subsequence $\{s_{n_k}\}$ which converges in that topology. If $r_n = r_{s_n}$, by the same argument there is a subsequence $\{r_{n_k}\}$ which converges in the w^*-topology on $L^\infty(\Omega,\mu)$. Therefore there are sequences $\{r_n\}$ and $\{s_n\}$ in $\mathscr{L}$ (we have denoted the subsequences by $\{r_n\}$ and $\{s_n\}$ to avoid subscripts) which converge in the w^*-topology and satisfy $r_n = r_{s_n}$. Let $r_* = w^*\text{-}\lim_{n\to\infty} r_n$, and $s_* = w^*\text{-}\lim_{n\to\infty} s_n$. Since $0 \leqq s_n(k) \leqq r_n(k) \leqq 1$, and $r_n(k) \geqq \tanh(\beta|\varepsilon(k)|/2)$, we have also $0 \leqq s_*(k) \leqq r_*(k) \leqq 1$ and $r_*(k) \geqq \tanh(\beta|\varepsilon(k)|/2)$. We shall prove that given $\varepsilon > 0$ for n sufficiently large

$$\mathscr{I}(r_n, s_n) \leqq \mathscr{I}(r_*, s_*) + \varepsilon \ . \tag{3.8}$$

Let $B = \{k : r_*(k) - s_*(k) > 0\}$. Since

$$\int_{B^c} (r_n(k) - s_n(k))\mu(dk) \to \int_{B^c} (r_*(k) - s_*(k))\mu(dk) \ ,$$

the subsequences can be chosen such that $r_n(k) - s_n(k) \to 0$ almost everywhere on B^c, and thus by Lebesgue's dominated convergence theorem

$$\int_{B^c} (r_n^2(k) - s_n^2(k))^{1/2} |\varepsilon(k)| \mu(dk) \to 0 \ .$$

For $\delta > 0$ let $B_\delta = \{k : \delta > r_*(k) - s_*(k) > 0\}$, and let $\bar{B}_\delta = B - B_\delta$. Given $\varepsilon > 0$ choose $\delta > 0$ such that

$$\int_{B_\delta} |\varepsilon(k)| \mu(dk) < \varepsilon/3 \ ,$$

then

$$\int_{B_\delta} (r_n^2(k) - s_n^2(k))^{1/2} |\varepsilon(k)| \mu(dk) < \varepsilon/3 \ .$$

Now the function $(x,y) \mapsto (xy)^{1/2}$ is concave so that

$$(xy)^{1/2} < (x'y')^{1/2} + \frac{1}{2}\left(\frac{y'}{x'}\right)^{1/2}(x - x') + \frac{1}{2}\left(\frac{x'}{y'}\right)^{1/2}(y - y') \ .$$

Therefore:

$$\begin{aligned}(r_n^2(k) - s_n^2(k))^{1/2} &\leqq (r_*^2(k) - s_*^2(k))^{1/2} \\ &+ \frac{1}{2}\left(\frac{r_*(k) - s_*(k)}{r_*(k) + s_*(k)}\right)^{1/2} \{(r_n(k) - r_*(k)) + (s_n(k) - s_*(k))\} \\ &+ \frac{1}{2}\left(\frac{r_*(k) + s_*(k)}{r_*(k) - s_*(k)}\right)^{1/2} \{(r_n(k) - r_*(k)) - (s_n(k) - s_*(k))\} \ .\end{aligned}$$

Since

$$\left(\frac{r_*(k) - s_*(k)}{r_*(k) + s_*(k)}\right)^{1/2} |\varepsilon(k)| 1_{\bar{B}_\delta}(k) \in L^1(\Omega,\mu)$$

and

$$\left(\frac{r_*(k) + s_*(k)}{r_*(k) - s_*(k)}\right)^{1/2} |\varepsilon(k)| 1_{\bar{B}_\delta}(k) \in L^1(\Omega,\mu) \ ,$$

we have

$$\int_{\bar{B}_\delta} (r_n^2(k)-s_n^2(k))^{1/2}|\varepsilon(k)|\mu(dk) < \int_{\bar{B}_\delta} (r_*^2(k)-s_*^2(k))^{1/2}|\varepsilon(k)|\mu(dk)+\varepsilon/3$$

for n sufficiently large. Thus

$$\int_{\Omega} (r_n^2(k)-s_n^2(k))^{1/2}|\varepsilon(k)|\mu(dk) < \int_{\Omega} (r_*^2(k)-s_*^2(k))^{1/2}|\varepsilon(k)|\mu(dk)+\varepsilon$$

for n large enough.

By a similar argument using the convexity of $r\mapsto I(r)$, we have for sufficiently large n,

$$-\frac{1}{\beta}\int_{\Omega} I(r_n(k))\mu(dk) < -\frac{1}{\beta}\int_{\Omega} I(r_*(k))\mu(dk)+\varepsilon .$$

Since $s\mapsto \int_{\Omega\times\Omega} \mu(dk)\mu(dk')U(k,k')s(k)s(k')$

is w^*-continuous we have the inequality (3.8). Therefore $f(\beta)=-\mathscr{J}(r_*,s_*)$. But then we must have $r_*=r_{s_*}$, and so $f(\beta)=-\mathscr{V}(s_*)$. □

In the next lemma we show that s_* cannot be a boundary point of $\mathscr{L}$ unless $s_*\equiv 0$.

Lemma 2. *If for* $s_*\in\mathscr{L}$, $f(\beta)=-\mathscr{V}(s_*)$, *then* $\mu\{k\in\Omega: s_*(k)=1\}=0$, *and either* $\mu\{k\in\Omega: s_*(k)=0\}=0$ *or* $\mu\{k\in\Omega: s_*(k)>0\}=0$.

Proof. Let $C=\{k\in\Omega: s_*(k)=1\}$, and suppose that $\mu(C)\neq 0$. Then for arbitrary $\varepsilon>0$, define $s_\varepsilon\in\mathscr{L}$ by

$$s_\varepsilon(k)=\begin{cases}1-\varepsilon & k\in C\\ s_*(k) & k\notin C .\end{cases}$$

Then

$$\mathscr{V}(s_\varepsilon)-\mathscr{V}(s_*)=\frac{1}{2}\int_C \mu(dk)(r_{1-\varepsilon}^2(k)-(1-\varepsilon)^2)^{1/2}|\varepsilon(k)|$$

$$-\frac{1}{\beta}\int_C \mu(dk)I(r_{1-\varepsilon}(k))+\frac{\mu(C)}{\beta}I(1)+O(\varepsilon)\geqq\frac{\mu(C)}{\beta}(I(1)-I(1-\varepsilon))+O(\varepsilon) ,$$

since $g(r_{1-\varepsilon}(k),1-\varepsilon;|\varepsilon(k)|)\geqq g(1-\varepsilon,1-\varepsilon;|\varepsilon(k)|)$. But since I is a convex function,

$$I(1)-I(1-\varepsilon)\geqq\varepsilon\tanh^{-1}(1-\varepsilon) ,$$

and therefore

$$\mathscr{V}(s_\varepsilon)-\mathscr{V}(s_*)\geqq\frac{1}{\beta}\mu(C)\varepsilon\tanh^{-1}(1-\varepsilon)+O(\varepsilon) ,$$

which means that $\mathscr{V}(s_\varepsilon)-\mathscr{V}(s_*)>0$ for ε sufficiently small, and so s_* is not a maximizer for $\mathscr{V}$. Therefore $\mu(C)=0$.

Now let $D=\{k\in\Omega:s_*(k)>0\}$, and suppose that $\mu(D)\neq 0$ and $\mu(D^c)\neq 0$. For $t\in[0,1]$, let $s_t=s_*+t\,1_{D^c}\in\mathscr{L}$ and put $h(t)=\mathscr{V}(s_t)$. Then

$$h'(0)=\frac{1}{2}\int_{D\times D^c}\mu(dk)\mu(dk')U(k,k')s_*(k)>0\ ,$$

which means that s_* is not a maximizer for $\mathscr{V}$. Therefore either $\mu(D)=0$ or $\mu(D^c)=0$. □

We now concentrate on the case $\lambda(\beta)>1$ and first exclude the possibility that $s_*\equiv 0$.

Lemma 3. *If* $\lambda(\beta)>1$ *and* $f(\beta)=-\mathscr{V}(s_*)$, *then*

$$\mu\{k\in\Omega:s_*(k)=0\}=0\ .$$

Proof. By the Perron-Frobenius theorem [19], we can find $\xi\in L^2(\Omega,\mu)$, such that $\xi(k)\geqq 0$ for all $k\in\Omega$, $\tilde{U}_\beta\xi=\lambda(\beta)\xi$ and $\|\xi\|_2=1$. Define $\xi_n\in L^\infty(\Omega,\mu)$

$$\xi_n(k)=\begin{cases}\xi(k) & \text{if}\quad \xi(k)\leqq n\ ,\\ 0 & \text{if}\quad \xi(k)>n\ .\end{cases}$$

Then by Lebesgue's dominated convergence theorem $(\xi_{n_0},(\tilde{U}_\beta-I)\xi_{n_0})\to\lambda(\beta)-1$ as $n\to\infty$. Choose n_0 such that $(\xi_{n_0},(\tilde{U}_\beta-I)\xi_{n_0})>0$, and let $\hat{s}(k)=\xi_{n_0}(k)g_\beta(k)$. For $t\in[0,\|\hat{s}\|_\infty^{-1})$ let $s_t=t\hat{s}\in\mathscr{L}$, and put $h(t)=\mathscr{V}(s_t)$. Then $h(t)$ is differentiable and

$$h'(t)/t\to\tfrac{1}{2}(\xi_{n_0},(\tilde{U}_\beta-I)\xi_{n_0})>0$$

as $n\to\infty$. Therefore $h'(t)>0$ for $t>0$ sufficiently small and 0 is not a maximizer for $\mathscr{V}$. □

For $s\in\mathscr{L}$ with $s(k)<1$, for all $k\in\Omega$, define $\Phi(s,k)$ by

$$\Phi(s,k)=\begin{cases}\dfrac{|\varepsilon(k)|s(k)}{(r_s^2(k)-s^2(k))^{1/2}}-\int_\Omega\mu(dk')U(k,k')s(k') & \text{if}\quad\varepsilon(k)\neq 0\\ \dfrac{2}{\beta}\tanh^{-1}s(k)-\int_\Omega\mu(dk')U(k,k')s(k) & \text{if}\quad\varepsilon(k)=0\ .\end{cases}\tag{3.9}$$

Then the Euler-Lagrange equation for the variational problem we are studying is $\Phi(s,k)=0$ for almost every k in Ω.

Theorem 6. *If* $\lambda(\beta)>1$ *and* s_* *is a maximizer for* $\mathscr{V}$, *then* s_* *satisfies the Euler-Lagrange equation* $\Phi(s_*,k)=0$ *for almost every* $k\in\Omega$.

Proof. From Lemmas 2 and 3

$$\mu\{k\in\Omega:s_*(k)=1\}=\mu\{k\in\Omega:s_*(k)=0\}=0\ .$$

Let $\delta>0$ and let $\xi\in L^\infty(\Omega,\mu)$ with $\operatorname{supp}\xi\subset\{k\in\Omega:s_*(k)<1-\delta\}$. For $t\in\mathbb{R}$ let $s_t(k)=s_*(k)(1+t\xi(k))$. Then for $|t|$ sufficiently small $s_t\in\mathscr{L}$. Let $h(t)=\mathscr{V}(s_t)$. Since for $|t|$ small enough we have $s_t(k)<1$ for these values of t, h is differentiable and $h'(0)=0$. Now

$$h'(0)=-\frac{1}{2}\int_\Omega\mu(dk)\xi(k)s_*(k)\Phi(s_*,k)\ ,$$

and therefore if we put

$$\xi(k)=\begin{cases}\Phi(s_*,k) & \text{if} \quad s_*(k)<1-\delta \ , \\ 0 & \text{if} \quad s_*(k)>1-\delta \ ,\end{cases}$$

we get

$$\int_{s_*(k)<1-\delta} \mu(dk)s_*(k)[\Phi(s_*,k)]^2=0 \ .$$

Therefore $s_*(k)[\Phi(s_*,k)]^2=0$ for almost every k such that $s_*(k)<1-\delta$. Since $0<s_*(k)<1$ for almost $k\in\Omega$ and δ is arbitrary, $\Phi(s_*,k)=0$ almost everywhere. □

We now change to a new variable κ so that we recover the gap equation [13] from the Euler-Lagrange equation. The variable κ is connected to the order parameter as we shall see later. Let

$$\kappa(k)=\begin{cases} s(k)|\varepsilon(k)|(r_s^2(k)-s^2(k)^{-1/2} \ , & \varepsilon(k)\neq 0 \ , \\ \frac{2}{\beta}\tanh^{-1}s(k) \ , & \varepsilon(k)=0 \ .\end{cases} \tag{3.10}$$

We can now write the equation $\Phi(s,k)=0$ as

$$\kappa(k)=\int_\Omega \mu(dk')U(k,k')\frac{\tanh\left(\frac{\beta}{2}(|\varepsilon(k')|^2+\kappa^2(k'))^{1/2}\right)}{(|\varepsilon(k')|^2+\kappa^2(k'))^{1/2}}\kappa(k') \ . \tag{3.11}$$

This is the usual gap equation. Using this form of the Euler-Lagrange equation we can now show that for $\lambda(\beta)>1$, $\mathscr{V}$ has a unique maximizer.

Theorem 7. *For $\lambda(\beta)>1$ the maximizer of $\mathscr{V}$ is unique.*

Proof. Let

$$h(a,y)=\frac{\tanh(\beta\Phi a^2+y^2)^{1/2}/2)y}{(a^2+y^2)^{1/2}} \ . \tag{3.12}$$

$y\mapsto h(a,y)$ is strictly concave and strictly increasing for $y\in(0,\infty)$. Equation (3.11) can be written as

$$\kappa(k)=\int_\Omega \mu(dk')U(k,k')h(|\varepsilon(k')|,\kappa(k')) \ .$$

Suppose that κ and κ' are two distinct solutions. Let

$$A_1=\{k\in\Omega:\kappa(k)>\kappa'(k)\}$$

and

$$A_2=\{k\in\Omega:\kappa(k)\leqq\kappa'(k)\} \ .$$

Without loss of generality we can assume that $\mu(A_1)\neq 0$. Let $\mu_0=\inf\kappa'(k)/\kappa(k)<1$. Since κ and κ' are continuous, bounded and strictly positive μ_0 is attained. Since

$y \mapsto h(a,y)$ is strictly concave and $h(a,0)=0$,

$$\int_{A_1} \mu(dk')U(k,k')h(|\varepsilon(k')|,\kappa'(k')) > \int_{A_1} \mu(dk')U(k,k')\frac{\kappa'(k)}{\kappa(k)}h(|\varepsilon(k')|,\kappa(k'))$$
$$> \mu_0 \int_{A_1} \mu(dk)U(k,k')h(|\varepsilon(k')|,\kappa(k')) \ .$$

Also since $y \to h(a,y)$ is increasing,

$$\int_{A_2} \mu(dk')U(k,k')h(|\varepsilon(k')|,\kappa(k')) \geqq \int_{A_2} \mu(dk')U(k,k')h(|\varepsilon(k')|,\kappa(k')) \ .$$

Therefore

$$\kappa'(k) > \mu_0 \int_{A_1} \mu(dk)U(k,k')h(|\varepsilon(k')|,\kappa(k'))$$
$$+ \int_{A_2} \mu(dk')U(k,k')h(|\varepsilon(k')|,\kappa(k')) > \mu_0 \kappa(k) \ .$$

Thus $\kappa'(k)/\kappa(k) > \mu_0$ for all $k \in \Omega$ which contradicts the fact that μ_0 is attained. □

We sum up the result of this section in the following theorem. Note that $\lambda \mapsto \lambda(\beta)$ is strictly increasing and as $\beta \to 0$, $\lambda(\beta) \to 0$. Therefore there is a number β_t which can be $+\infty$ such that if $\beta < \beta_t$, then $\lambda(\beta) < 1$ and if $\beta > \beta_t$, $\lambda(\beta) > 1$.

Theorem 8.

$$f(\beta) = -\frac{1}{\beta}\int_{\Omega} \mu(dk)\ln\cosh\left[\frac{\beta}{2}(\varepsilon^2(k)+\kappa_\beta^2(k))\right]^{1/2} + \frac{1}{4}\int_{\Omega} \mu(dk)\kappa_\beta(k)h(|\varepsilon(k)|,\kappa_\beta(k)) \ , \tag{3.13}$$

where if $\beta \leqq \beta_t$, *then* $\kappa_\beta = 0$ *while if* $\beta > \beta_t$, *then* κ_β *is the unique strictly positive solution of the gap-equation*

$$\kappa_\beta(k) = \int_{\Omega} \mu(dk')U(k,k')h(|\varepsilon(k')|,\kappa_\beta(k')) \ . \tag{3.14}$$

If $\beta_t = \infty$, then clearly there is no phase transition. If $\beta_t < \infty$, we see that there is a phase transition in the sense that there is a singularity in $f(\beta)$ at β_t.

4. The Order Parameter

To examine the breaking of rotational symmetry we perturb the hamiltonian H_ℓ in (2.1) by $\alpha \sum_{i=1}^{N_\ell} \sigma_i^x$, let

$$H_\ell(\alpha) = H_\ell + \alpha \sum_{i=1}^{N_\ell} \sigma_i^x \ . \tag{4.1}$$

We want to evaluate the expectation of $(1/V_\ell)\sum_{i=1}^{N_\ell} \sigma_i^x$ with respect to the canonical state corresponding to $H_\ell(\alpha)$.

Theorem 9. *Let*

$$s_\ell^x(\alpha,\beta)=\frac{1}{V_\ell}\frac{\operatorname{trace}\sum_{i=1}^{N_\ell}\sigma_i^x e^{-\beta H_\ell(\alpha)}}{\operatorname{trace} e^{-\beta H_\ell(\alpha)}}, \tag{4.2}$$

then $s^x(\alpha,\beta)=\lim_{\ell\to\infty} s_\ell^x(\alpha,\beta)$ *exists and*

$$\begin{aligned}\lim_{\alpha\downarrow 0} s^x(\alpha,\beta) &= -\lim_{\alpha\uparrow 0} s^x(\alpha,\beta)\\ &=\begin{cases}0\,, & \beta\leqq\beta_t\\ -\int_\Omega \mu(dk)\dfrac{\tanh(\beta/2(|\varepsilon(k)|^2+\kappa^2(k))^{1/2}\kappa(k)}{(|\varepsilon(k)|^2+\kappa^2(k))^{1/2}}\,, & \beta>\beta_t\,,\end{cases}\end{aligned} \tag{4.3}$$

where κ *is as in Theorem 8.*

Proof. Clearly $s_\ell^x(\alpha,\beta)$ is an odd function of α. Therefore it is sufficient to compute $s^x(\alpha,\beta)$ for $\alpha>0$. To do this we use a standard procedure and define

$$f_\ell(\alpha,\beta)=-\frac{1}{\beta V_\ell}\ln \operatorname{trace} e^{-\beta H_\ell(\alpha)}\,. \tag{4.4}$$

Using the techniques of Sect. 2 we can prove $f(\alpha,\beta)=\lim_{\ell\to\infty} f_\ell(\alpha,\beta)$ exists and is given by

$$\begin{aligned}f(\alpha,\beta) &= -\sup\{\mathscr{S}(r,\theta,\phi)-\alpha\int_\Omega r(k)\sin\theta(k)\sin\phi(k)\mu(dk):(r,\theta,\phi)\in\mathscr{M}\}\\ &= -\sup\{\bar{\mathscr{S}}(r,\theta)+\alpha\int_\Omega r(k)\sin\theta(k)\mu(dk):(r,\theta)\in\bar{\mathscr{M}}\}\\ &= -\sup\{\mathscr{V}(s)+\alpha\int_\Omega s(k)\mu(dk):s\in\mathscr{L}\}\,.\end{aligned} \tag{4.5}$$

By using arguments similar to those in Sect. 3 we can prove that for all values of $\lambda(\beta)$ the supremum is attained at a unique element of $\mathscr{L}$, s_α which is the unique non-zero solution of the corresponding Euler-Lagrange equation

$$\Phi(s,k)+\alpha=0 \tag{4.6}$$

for almost every k in Ω. We have $s_\ell^x(\alpha,\beta)=\frac{\partial}{\partial\alpha}f_\ell(\alpha,\beta)$. Since $\alpha\mapsto f_\ell(\alpha,\beta)$ is a convex function if $\alpha\mapsto f(\alpha,\beta)$ is differentiable, then it follows that

$$s^x(\alpha,\beta)=\frac{\partial}{\partial\alpha}f(\alpha,\beta)\,.$$

Now for α, $\alpha'\geqq 0$,

$$\begin{aligned}\mathscr{V}(s_\alpha)+\alpha\int_\Omega s_\alpha(k)\mu(dk) &\geqq \mathscr{V}(s_{\alpha'})+\alpha\int_\Omega s_{\alpha'}(k)\mu(dk)\\ &=\mathscr{V}(s_{\alpha'})+\alpha'\int_\Omega s_{\alpha'}(k)\mu(dk)+(\alpha-\alpha')\int_\Omega s_{\alpha'}(k)\mu(dk)\,.\end{aligned}$$

Therefore

$$f(\alpha,\beta)-f(\alpha',\beta)\leqq -(\alpha-\alpha')\int_\Omega s_{\alpha'}(k)\mu(dk)\ .$$

Thus by interchanging α and α' we get for $\alpha,\alpha'\in[0,\infty)$,

$$-(\alpha-\alpha')\int_\Omega s_\alpha(k)\mu(dk)\leqq f(\alpha,\beta)-f(\alpha',\beta)\leqq -(\alpha-\alpha')\int_\Omega s_{\alpha'}(k)\mu(dk)\ .$$

Therefore $\alpha\mapsto f(\alpha,\beta)$ is differentiable at $\alpha_0>0$ if and only if $\alpha\mapsto\int_\Omega s_\alpha(k)\mu(dk)$ is continuous at α_0, and $\frac{\partial f}{\partial\alpha}(\alpha_0,\beta)=-\int_\Omega s_{\alpha_0}(k)\mu(dk)$. In the next lemma we shall prove that $\alpha\mapsto\int_\Omega s_\alpha(k)\mu(dk)$ with $s_0=s_*$ is continuous on $[0,\infty)$. It then follows that

$$s^\alpha(\alpha,\beta)=-\int_\Omega s_\alpha(k)\mu(dk)$$

and

$$\lim_{\alpha\downarrow 0} s^\alpha(\alpha,\beta)=-\int_\Omega s_0(k)\mu(dk)\ ,$$

which proves the theorem. □

Lemma 4.

$$\alpha\mapsto\|s_\alpha\|_\infty \textit{ is continuous for } \alpha\in[0,\infty)\ .$$

Proof. We shall prove continuity only in a set $[0,\varepsilon)$. Continuity for $\alpha\geqq\varepsilon$, it will be seen from the proof, is easier and follows similarly.

Consider first the case $\lambda(\beta)<1$. Choose $\delta\in(0,1)$, and on $B(0,\delta)$ in $L^\infty(\Omega,\mu)$ define

$$(Fs)(k)=\Phi(s,k)\ . \tag{4.7}$$

F is continuously differentiable on $B(0,\delta)$ and

$$(F'(0)h)(k)=\frac{h(k)}{(g_\beta(k))^2}-\int_\Omega U(k,k')h(k')\mu(dk')\ . \tag{4.8}$$

Therefore $F'(0)$ as a map from $L^2(\Omega,\mu)$ to $L^2(\Omega,\mu)$ is strictly positive and therefore invertible. Now $F'(0)$ maps $L^\infty(\Omega,\mu)$ into $L^\infty(\Omega,\mu)$. But if $h\in L^\infty(\Omega,\mu)$ and $F'(0)\tilde h=h$, then since $\int_\Omega U(\cdot,k')\tilde h(k')\mu(dk')$ and $(g_\beta(k))^2<\beta/2$, $\tilde h$ must also be in $L^\infty(\Omega,\mu)$. Therefore $F'(0)$ is invertible on $L^\infty(\Omega,\mu)$. Since $F(0)=0$, by the inverse function theorem there are neighbourhoods U and V of 0 in $L^\infty(\Omega,\mu)$ such that F is one-to-one on U and $F(U)=V$, and if G is the inverse of F on V, then G is continuously differentiable. But $s_\alpha=G(-\alpha 1)$, and therefore continuous in $[0,\varepsilon)$ for some ε.

For $\lambda(\beta)>1$, $a\equiv\operatorname{ess\,inf} s_*>0$ and $b\equiv\|s_*\|_\infty<1$. Choose $\delta=\max(a,1-b)$ and on $B(s_*,\delta)$ define F as above,

$$(F'(s_*)h)(k)=m(k)h(k)-\int_\Omega U(k,k')h(k')\mu(dk')\ , \tag{4.9}$$

where

$$m(k)=\begin{cases} \dfrac{2}{\beta(1-s_*^2(k))} & \text{when} \quad \varepsilon(k)=0 \ , \\ \dfrac{|\varepsilon(k)|r_{s_*}^2(k)}{(r_{s_*}^2(k)-s_*^2(k))^{3/2}}-\dfrac{\beta}{2}\,\dfrac{|\varepsilon(k)|^2 r_{s_*}^2(k)s_*^2(k)}{(r_{s_*}^2(k)-s_*^2(k))^3} & \\ \cdot\left\{\dfrac{1}{(1-r_{s_*}^2(k))}+\dfrac{\beta}{2}\,\dfrac{|\varepsilon(k)|s_*^2(k)}{(r_{s_*}^2(k)-s_*^2(k))^{3/2}}\right\}^{-1} & \text{when} \quad \varepsilon(k)\neq 0 \ . \end{cases}$$

Using the inequality

$$\inf_{x>\varepsilon>0}\left\{\frac{1}{1-x^2}-\frac{\tanh h^{-1}x}{x}\right\}>0 \ ,$$

we can check that

$$m(k)>\int_{\Omega} U(k,k')\,\frac{s_*(k')}{s_*(k)}\,\mu(dk')+\delta$$

for some $\delta>0$. Therefore

$$\int_{\Omega} h(k)(F'(s_*)h)(k)\mu(dk)>\int_{\Omega\times\Omega}\mu(dk)\mu(dk')U(k,k')\left\{\frac{h^2(k)}{s_*(k)}\,s_*(k')-h(k)h(k')\right\}$$

$$+\delta\|h\|_2^2=\frac{1}{2}\int_{\Omega\times\Omega}\mu(dk)\mu(dk')U(k,k')\left(h(k)\left(\frac{s_*(k')}{s_*(k)}\right)^{1/2}\right.$$

$$\left.-h(k')\left(\frac{s_*(k)}{s_*(k')}\right)^{1/2}\right)^2+\delta\|h\|_2^2>\delta\|h\|_2^2 \ .$$

The rest of the proof then follows as above. □

Remark. What we have calculated above is the thermodynamic limit of the derivative of the free energy density in the direction of the extensive perturbation $\sum_{i=1}^{N_\ell}\sigma_\ell^x$. It is clear that the work of this section can be adapted to calculate the corresponding quantity with $\sum_{i=1}^{N_\ell}\sigma_i^x$ replaced by

$$\frac{1}{V_\ell^{N-1}}\prod_{j=1}^{N}\sum_{i=1}^{N_\ell} f_j(k_\ell(i))\sigma_i^{\#j} \ ,$$

where the f_j are continuous functions and $\sigma_i^{\#j}=I\otimes\ldots\otimes\sigma^{\#j}\otimes\ldots\otimes I$, $\sigma^{\#j}$ being any Pauli matrix.

Appendix: Berezin-Lieb Inequalities

In this appendix we describe the application of the Berezin-Lieb inequalities to Sect. 2. Let Δ^J be the irreducible representation of $SU(2)$ on $D^J=\mathbb{C}^{2J+1}$. Let $\{P(J,\Omega):\Omega\in S^2\}$ denote the family of Bloch coherent projections in $\mathscr{L}(D^J)$. The

$P(J,\Omega)$ satisfy

$$\operatorname{trace} P(J,\Omega)=1 \ , \qquad 2J+1 \int_{S^2} \frac{d\Omega}{4\pi} P(J,\Omega)=\mathbb{1} \ . \tag{A1}$$

Let B be a self-adjoint linear operator on D^J. We call functions B^u and B^L in $L^\infty(S^2, d\Omega)$ the upper and lower symbols respectively of B if

$$B=2J+1 \int_{S^2} \frac{d\Omega}{4\pi} B^u(\Omega)P(J,\Omega) \ , \qquad B^L(\Omega)=\operatorname{trace} BP(J,\Omega) \ . \tag{A2}$$

B^L is determined uniquely by (A2), whereas B^u always exists but is not necessarily unique. Given B, B^L and a B^u can be read off from the table in [12].

Given positive integers $\{n_i : i=1,\dots N\}$, let J_i be in A_{n_i} for $i=1,2,\dots N$. Let $d\boldsymbol{\Omega}$ be the product measure $\prod_{i=1}^{N} d\Omega_i$, $d\Omega_i=d\Omega$, $\mathscr{L}$ be $\bigotimes_{i=1}^{N} \mathscr{L}(D^{J_i})$ and L be $L^\infty((S^2)^N, d\boldsymbol{\Omega})$. Define the map $f:\mathscr{L}\to L$ by

$$f(B)(\Omega_1,\dots,\Omega_N)=\operatorname{trace}_{\mathscr{L}} BP(J_1,\Omega_1)\otimes\dots\otimes P(J_N,\Omega_N) \ , \tag{A3}$$

and the map $F:L\to\mathscr{L}$ by

$$F(b)=\prod_{i=1}^{N} \frac{2J_i+1}{4\pi} \int_{(S^2)^N} d\boldsymbol{\Omega} b(\Omega_1,\dots,\Omega_N)P(J_1,\Omega_1)\otimes\dots\otimes P(J_N,\Omega_N) \ . \tag{A4}$$

By (A1) f and F are both positive and unital; also $x\mapsto e^x$ is convex. We use a generalisation of Theorem B of [20], namely

Proposition A1. *Let M and N be von Neumann algebras and $g:\mathbb{R}\to\mathbb{R}$ be a convex function. Assume that τ is a normal semifinite trace on M and $\alpha:N\to M$ is a positive unital mapping. Then for α self-adjoint in N,*

$$\tau(g(\alpha(a)))\leqq\tau(\alpha(g(a))) \ , \tag{A5}$$

whenever both sides are defined.

By applying Proposition A1 with $\alpha=f$ and $\tau=\prod_{i=1}^{N}(2J_i+1)(4\pi)^{-1}\int_{S^{2N}} d\boldsymbol{\Omega}$, we find using (A1) that

$$\prod_{i=1}^{N} \frac{2J_i+1}{4\pi} \int_{(S^2)^N} d\boldsymbol{\Omega}\, e^{(f(B))(\boldsymbol{\Omega})}\leqq\operatorname{trace} e^B \ , \tag{A6}$$

and similarly with $\alpha=F$ and $\tau=\text{trace}$ that

$$\operatorname{trace} e^{F(b)}\leqq\prod_{i=1}^{N} \frac{2J_i+1}{4\pi} \int_{(S^2)^N} d\boldsymbol{\Omega} e^{b(\boldsymbol{\Omega})} \ . \tag{A7}$$

In the case that B is a tensor product $B_1\otimes B_2\otimes\dots\otimes B_N$ with each B_i self-adjoint, then from (A2) we see that $f(B)=B_1^L B_2^L\dots B_N^J$ and $F(B_1^u B_2^u\dots B_N^u)=B$. The hamiltonian $(p^{J_1}\otimes\dots\otimes p^{J_N})h_\ell$ is a finite linear combination of such terms, so by linearity of the maps f and L, (2.10) follows from (A6) and (A7) using Lieb's table [12]. In the case that $N=1$ these are just the usual Berezin-Lieb inequalities.

Acknowledgements. We would to thank W. Cegła, J.T. Lewis, and G.A. Raggio for making available to us their results and especially the latter for many very useful discussions.

References

1. Bogoliubov, N.N., Jr.: A method for studying model Hamiltonians. International Series of Monographs in Natural Philosophy, Vol 4. Oxford: Pergamon 1972
2. Duffield, N.G., Pulé, J.V.: Lett. Math. Phys. **14**, 329 (1987)
3. Duffield, N.G., Pulé, J.V.: Thermodynamics and phase transitions in the Overhauser model. J. Stat. Phys. (to appear)
4. Raggio, G.A.: The free energy of the full spin-Boson model. Preprint, Dublin Institute for Advanced Studies (1988)
5. Thirring, W.: Commun. Math. Phys. **7**, 181 (1967)
6. Thouless, D.J.: Phys. Rev. **117**, 1256 (1960)
7. Cegła, W., Lewis, J.T., Raggio, G.A.: The free energy of quantum spin systems and large deviations. Commun. Math. Phys. (to appear)
8. Varadhan, S.R.S.: Commun. Pure Appl. Math. **19**, 261 (1966)
9. Ellis, R.S.: Entropy, large deviations, and statistical mechanics. Berlin Heidelberg New York: Springer 1985
10. Simon, B.: Commun. Math. Phys. **71**, 247 (1980)
11. Berezin, F.A.: Math. U.S.S.R. Izv. **6**, 117 (1972)
12. Lieb, E.H.: Commun. Math. Phys. **31**, 327 (1973)
13. Bardeen, J., Cooper, L.N., Schrieffer, J.R.: Phys. Rev. **108**, 1175 (1957)
14. Anderson, P.W.: Phys. Rev. **112**, 1900 (1958)
15. Thouless, D.J.: The quantum mechanics of many-body systems. New York: Academic Press 1961
16. Haag, R.: Il Nuovo Cimento **25**, 287 (1962)
17. Van Hemmen, L.: Fortschr. Physik **26**, 397 (1978)
18. Ginibre, J.: Commun. Math. Phys. **8**, 26 (1968)
19. Jentzsch, R.: Crelles J. **141**, 235 (1912)
20. Petz, D.: Proc. Am. Math. Soc. **99**, 273 (1987)

Communicated by J. Fröhlich

Received February 11, 1988; in revised form April 1, 1988