

The Free Energy of Quantum Spin Systems and Large Deviations

W. Cegła, J. T. Lewis, and G. A. Raggio

Dublin Institute for Advanced Studies, 10 Burlington Road, Dublin 4, Ireland

Abstract. For a quantum system of n identical spins of magnitude j , we introduce an integrated density of states of definite total spin angular momentum. The underlying sequence $\{\mathbb{K}_n^j: n=1, 2, \dots\}$ of probability measures satisfies Varadhan's large deviation principle, and converges to a degenerate distribution. We use the Berezin-Lieb Inequalities to obtain upper and lower bounds for the limiting specific free-energy of the spins interacting with a second quantum system under specified conditions on the Hamiltonian. The method is illustrated by applications to the BCS model and to the Dicke maser model.

1. Introduction

Large-deviation methods are proving useful in statistical mechanics [1, 2], both for reorganizing existing proofs and for obtaining new results [3–5]. The main purpose of this paper is to prove a large deviation result which has applications to the equilibrium statistical mechanics of spin systems; we illustrate its use by applying it to the calculation of the specific free energy of the BCS model and of the Dicke Maser model. In the case of the BCS model, it provides a short proof of a known result; in the case of the Dicke Maser model, when combined with the Berezin-Lieb inequalities, it provides a substantial improvement on existing treatments in that it requires minimal restrictions on the interaction terms. In both cases, the spins are not required to be of magnitude $1/2$.

The Laplace method (the method of the largest term) is at the heart of equilibrium statistical mechanics, but to make it rigorous in particular instances is usually tedious and sometimes subtle. The large deviation principle, in Varadhan's formulation [6], provides an efficient way of supplying the required proofs; it reduces the tedium and exposes any latent subtleties. The abstract setting is that of a sequence $\{\mathbb{K}_n: n=1, 2, \dots\}$ of probability measures on the Borel subsets of a complete separable metric space E (in the main theorem of this paper, E is the interval $[0, 1]$), and a divergent sequence $\{V_n: n=1, 2, \dots\}$ of positive numbers. We

say that $\{\mathbb{K}_n\}$ satisfies the large deviation principle with constants $\{V_n\}$ and rate-function $I: E \rightarrow [0, \infty]$ if the following conditions hold:

(LD 1) $I(\cdot)$ is lower semicontinuous on E .

(LD 2) For each $m < \infty$, $\{x: I(x) \leq m\}$ is compact.

(LD 3) For each closed subset C of E ,

$$\limsup_{n \rightarrow \infty} V_n^{-1} \log \mathbb{K}_n[C] \leq -\inf_C I(x).$$

(LD 4) For each open subset G of E ,

$$\liminf_{n \rightarrow \infty} V_n^{-1} \log \mathbb{K}_n[G] \geq -\inf_G I(x).$$

The following version of Varadhan’s Theorem is adapted to our requirements; it is readily deduced from Theorem 3.4 of [6]:

Varadhan’s Theorem. Suppose the sequence $\{\mathbb{K}_n\}$ of probability measures on E satisfies the large deviation principle with constants $\{V_n\}$ and rate-function $I: E \rightarrow [0, \infty]$. Let $\{f_n\}$ be a sequence of continuous functions $f_n: E \rightarrow \mathbb{R}$ which are bounded above, and suppose that $\{f_n\}$ converges to $f: E \rightarrow \mathbb{R}$ uniformly on bounded subsets; then

$$\lim_{n \rightarrow \infty} V_n^{-1} \log \int_E \exp(V_n f_n(x)) \mathbb{K}_n[dx] = \sup_E \{f(x) - I(x)\}.$$

Let π^j be an irreducible unitary representation of $SU(2)$ acting on the $(2j + 1)$ -dimensional complex Hilbert space $\mathfrak{D}(j)$. Let \mathfrak{R}_n be the tensor product of $\mathfrak{D}(j)$ with itself n times ($n = 1, 2, 3, \dots$), and define the unitary representation π_n of $SU(2)$ on \mathfrak{R}_n by

$$\pi_n(g) = \pi^j(g) \otimes \pi^j(g) \otimes \dots \otimes \pi^j(g), \quad g \in SU(2). \tag{1.1}$$

n times

By the theory of unitary representations of compact groups, π_n is reducible ($n \geq 2$) and decomposes into the direct sum

$$\pi_n = \bigoplus_{J \geq 0}^{nj} c^j(n, J) \pi^J, \tag{1.2}$$

where the irreducible unitary representation π^J acts on the $(2J + 1)$ -dimensional Hilbert space $\mathfrak{D}(J)$, and has multiplicity $c^j(n, J)$. Here, J runs through $\{0, 1, 2, \dots, nj\}$ if nj is an integer, and through $\{1/2, 3/2, \dots, nj\}$ if nj is a half-integer. In Sect. 2 we show that if X_n is an element of \mathfrak{A}_n , the *-subalgebra of $\mathfrak{B}(\mathfrak{R}_n)$ – the linear operators on \mathfrak{R}_n – generated by π_n :

$$\mathfrak{A}_n = \{\pi_n(g) : g \in SU(2)\}'', \tag{1.3}$$

where $'$ denotes the commutant in $\mathfrak{B}(\mathfrak{R}_n)$, then

$$\text{tr}_{\mathfrak{R}_n}(X_n) = \sum_{J=0}^{nj} c^j(n, J) \text{tr}_{\mathfrak{D}(J)}(\pi^J(X_n)). \tag{1.4}$$

This trace formula can be re-written as follows: Let

$$C^j(n) = \sum_{J=0}^{nj} c^j(n, J), \tag{1.5}$$

and define a probability measure \mathbb{K}_n^j on the interval $[0, 1]$ by

$$\mathbb{K}_n^j[B] = C^j(n)^{-1} \sum_{\{J: J/n_j \in B\}} c^j(n, J); \quad B \subseteq [0, 1]. \tag{1.6}$$

Define

$$f_n(\beta) = (-n\beta)^{-1} \log \operatorname{tr}_{\mathfrak{R}_n}(e^{-\beta H_n}) \tag{1.7}$$

for $H_n \in \mathfrak{R}_n$ self-adjoint, and

$$f_n^j(\beta; J/n_j) = (-n\beta)^{-1} \log \operatorname{tr}_{\mathfrak{R}(J)}(\pi^J(e^{-\beta H_n})). \tag{1.8}$$

Let $f_n^j(\beta; \cdot)$ be any continuous function on $[0, 1]$ satisfying (1.8). We can use (1.4) to write

$$e^{-n\beta f_n(\beta)} = C^j(n) \int_{[0, 1]} e^{-n\beta f_n^j(\beta; u)} \mathbb{K}_n^j[du]. \tag{1.9}$$

In Sect. 3 we show that $\{\mathbb{K}_n^j: n \geq 1\}$ satisfies the large deviation principle with constants $\{n\}$ and rate-function I^j , where

$$I^j(u) = \log(2j + 1) + \sup_{\alpha \geq 0} \left\{ \alpha u - \log \left\{ \frac{\sinh(\alpha(2j + 1)/2j)}{\sinh(\alpha/2j)} \right\} \right\}. \tag{1.10}$$

Suppose now that the sequence $\{f_n^j(\beta; \cdot): n \geq 1\}$ converges uniformly to $f^j(\beta; \cdot)$; then it follows from Varadhan’s theorem that $f(\beta) = \lim_{n \rightarrow \infty} f_n(\beta)$ exists and is given by

$$f(\beta) = \inf_{[0, 1]} \{f^j(\beta; u) - \beta^{-1} I_0^j(u)\}, \tag{1.11}$$

where

$$I_0^j(u) = I^j(1) - I^j(u). \tag{1.12}$$

This is the essence of the method; in order to apply it to the BCS model in Sect. 2, we make a second application of Varadhan’s theorem – this time with the trivial rate-function – in order to prove the convergence of $\{f_n^j(\beta; \cdot)\}$; in order to apply the method to the Dicke Maser model in Sect. 4, we require a generalization of the trace formula (see Sect. 2) and the introduction of the Berezin-Lieb inequalities (see Sect. 4) in order to prove the convergence of the sequence $\{f_n^j(\beta; \cdot)\}$.

To prove that $\{\mathbb{K}_n^j: n \geq 1\}$ satisfies the large deviation principle with rate-function I^j given by (1.10), it is enough to prove that

$$\lim_{n \rightarrow \infty} n^{-1} \log \int_{[0, 1]} e^{njxu} \mathbb{K}_n^j[du] = \log \left\{ \frac{\sinh(\alpha(2j + 1)/2j)}{\sinh(\alpha/2j)} \right\} - \log(2j + 1); \tag{1.13}$$

the rest follows by general principles (see [2], for example). We are unable to prove (1.13) directly, except in the case $j = 1/2$, where the multiplicities $c^j(n, J)$ are known explicitly. Instead, we exploit the fact that $2 \sinh(nx)$ behaves like e^{nx} for large n and that $\int_{[0, 1]} \sinh(nuy) \mathbb{K}_n^j[du]$ can be computed using a special case of the trace formula (1.4). This is done in Sect. 3.

2. A Trace Formula

In this section we determine the class of operators in $\mathfrak{B}(\mathfrak{K}_n)$ for which a trace formula like (1.4) holds. We use nothing more than the theory of finite-dimensional representations of involutive algebras (see Sect. 2.3.5 of [7], for example). The decompositions of π_n into irreducible unitary representations of $SU(2)$ are in one-to-one correspondence with the decompositions of \mathfrak{K}_n into irreducible $*$ -algebras, these are in turn in one-to-one correspondence with the maximal abelian $*$ -subalgebras (masas) of \mathfrak{K}_n . The masas of \mathfrak{K}_n are unitarily equivalent. Given a masa \mathfrak{A}_n of \mathfrak{K}_n , there exists a unitary operator U from \mathfrak{K}_n onto

$$\bigoplus_{J \in \mathbb{J}} \bigoplus_{k \in \mathbb{K}(J)} \mathfrak{D}(J), \tag{2.1}$$

where $\mathbb{J} = \{J : 0 \leq J \leq nj, c^j(n, J) \neq 0\}$, $\mathbb{K}(J) = \{1, 2, \dots, c^j(n, J)\}$, and irreducible $*$ -representations π^J of \mathfrak{K}_n on $\mathfrak{D}(J)$ such that for every $N \in \mathfrak{K}_n$,

$$UNU^* = \bigoplus_{J \in \mathbb{J}} \bigoplus_{k \in \mathbb{K}(J)} \pi^J(N). \tag{2.2}$$

Conversely, if

$$A = \bigoplus_{J \in \mathbb{J}} \bigoplus_{k \in \mathbb{K}(J)} A_0, \quad \text{where } A_0 \in \mathfrak{B}(\mathfrak{D}(J)),$$

then $U^*AU \in \mathfrak{K}_n$. An $X \in \mathfrak{B}(\mathfrak{K}_n)$ is said to be *decomposable* (with respect to \mathfrak{A}_n) if

$$UXU^* = \bigoplus_{J \in \mathbb{J}} \bigoplus_{k \in \mathbb{K}(J)} X(J, k), \quad \text{where } X(J, k) \in \mathfrak{B}(\mathfrak{D}(J)). \tag{2.3}$$

An operator is decomposable if and only if it belongs to \mathfrak{A}'_n . Now suppose that $X \in \mathfrak{B}(\mathfrak{K}_n)$ and there exists a unitary operator $V \in \mathfrak{A}'_n$ such that $V^*XV \in \mathfrak{A}_n$. Then,

$$U(V^*XV)U^* = \bigoplus_{J \in \mathbb{J}} \bigoplus_{k \in \mathbb{K}(J)} \pi^J(V^*XV),$$

$$U^*VU = \bigoplus_{J \in \mathbb{J}} \bigoplus_{k \in \mathbb{K}(J)} V(J, k), \quad \text{with } V(J, k) \in \mathfrak{B}(\mathfrak{D}(J)) \text{ unitary,}$$

and thus

$$UXU^* = \bigoplus_{J \in \mathbb{J}} \bigoplus_{k \in \mathbb{K}(J)} V(J, k)\pi^J(V^*XV)V(J, k)^*. \tag{2.4}$$

The converse is also true; if

$$A = \bigoplus_{J \in \mathbb{J}} \bigoplus_{k \in \mathbb{K}(J)} A(J, k),$$

where for each $J \in \mathbb{J}$, $A(J, k)$ is unitarily equivalent to $A(J, k')$ for all $k, k' \in \mathbb{K}(J)$, then there exists a unitary $V \in \mathfrak{A}'_n$ such that $V^*U^*AUV \in \mathfrak{A}_n$. For these operators, which we call *homogeneously decomposable*, we have the trace formula

$$\text{tr}_{\mathfrak{K}_n}(X) = \sum_{J \in \mathbb{J}} c^j(n, J) \text{tr}_{\mathfrak{D}(J)}(\pi^J(V^*XV)). \tag{2.5}$$

Notice that if $X \in \mathfrak{A}_n$ then (2.5) holds, with $V=1$, for any decomposition.

The above admits a straightforward and useful generalization. Suppose that \mathfrak{h}_n is a second separable complex Hilbert space, and let H_n be a selfadjoint operator on the tensor product $\mathfrak{h}_n \otimes \mathfrak{K}_n$ such that $\exp\{-\beta H_n\}$ is trace-class. Let

$$Z_n = \text{tr}_{\mathfrak{h}_n \otimes \mathfrak{K}_n}(\exp\{-\beta H_n\}), \quad f_n(\beta) = (-n\beta)^{-1} \log Z_n. \tag{2.6}$$

Assume that there exists a masa \mathfrak{A}_n of \mathfrak{R}'_n , and a unitary operator $V \in \mathfrak{A}'_n$ such that $(1 \otimes V)^* H_n(1 \otimes V)$ is associated with $\mathfrak{B}(\mathfrak{h}_n) \otimes \mathfrak{R}_n$. Such operators will be henceforth called *homogeneously decomposable*. Then,

$$Z_n = \sum_{J \in \mathfrak{J}} c^j(n, J) \operatorname{tr}_{\mathfrak{h}_n \otimes \mathfrak{D}(J)}(\exp\{-\beta H_n(J)\}), \tag{2.7}$$

where $H_n(J)$ is the operator on $\mathfrak{h}_n \otimes \mathfrak{D}(J)$ given by

$$H_n(J) = \{\operatorname{Id} \otimes \pi^J\} ([1 \otimes V]^* H_n [1 \otimes V]). \tag{2.8}$$

The trace formula (2.7) can be rewritten as follows: Define

$$Z_n^j(J/nj) = \operatorname{tr}_{\mathfrak{h}_n \otimes \mathfrak{D}(J)}(\exp\{-\beta H_n(J)\}), \tag{2.9}$$

$$f_n^j(\beta; J/nj) = (-n\beta)^{-1} \log Z_n^j(J/nj). \tag{2.10}$$

Then for any function $f_n^j(\beta; \cdot)$ on $[0, 1]$, satisfying (2.10), we can write (2.7) as,

$$Z_n = C^j(n) \int_{[0, 1]} \exp\{-n\beta f_n^j(\beta; u)\} \mathbb{K}_n^j[du], \tag{2.11}$$

and, accordingly,

$$f_n(\beta) = (-n\beta)^{-1} \log C^j(n) - (n\beta)^{-1} \log \left\{ \int_{[0, 1]} \exp\{-n\beta f_n^j(u)\} \mathbb{K}_n^j[du] \right\}. \tag{2.12}$$

We remark that if H_n is associated with $\mathfrak{B}(\mathfrak{h}_n) \otimes \mathfrak{R}_n$, then it is homogeneously decomposable [with $V = 1$ in (2.8)].

In the above, the introduction of the group $SU(2)$ is convenient but irrelevant. We may replace the irreducible representation π^j of $SU(2)$, which we started with, by $\mathfrak{B}(\mathfrak{K})$, where \mathfrak{K} is an arbitrary *finite* dimensional Hilbert space; then define on the n -fold tensor product \mathfrak{K}_n of \mathfrak{K} with itself, the algebra

$$\mathfrak{R}_n = \{U \otimes U \otimes \dots \otimes U \in \mathfrak{B}(\mathfrak{K}_n) : U \text{ a unitary in } \mathfrak{B}(\mathfrak{K})\}''.$$

In the decomposition of \mathfrak{R}_n into irreducible representations of itself, the dimension of the carrier Hilbert spaces as well as the multiplicities are determined only by n and the dimension of \mathfrak{K} . The advantage of introducing the group $SU(2)$ is that polynomials in the generators S_n^x, S_n^y, S_n^z of the representation $d\pi_n$ of the Lie algebra $\mathfrak{su}(2)$ provide a convenient basis for \mathfrak{R}_n since

$$\mathfrak{R}_n = \{S_n^x, S_n^y, S_n^z\}'' . \tag{2.13}$$

The generators of $d\pi_n$ are given by

$$S_n^\alpha = \sum_{k=1}^n ({}^j S_{(k)}^\alpha), \quad \alpha = x, y, z, \tag{2.14}$$

where ${}^j S_{(k)}^\alpha$ is ${}^j S^\alpha$ acting on the k -th component of \mathfrak{K}_n , and $\{{}^j S^\alpha : \alpha = x, y, z\}$ are the generators of the irreducible representation of $\mathfrak{su}(2)$ carried by the Hilbert space $\mathfrak{K} \cong \mathfrak{D}(j)$, $j = (\dim(\mathfrak{K}) - 1)/2$.

3. Large Deviations

In this section we prove the large deviation result announced in Sect. 1. Let $C^J(n)$ be the normalizing factor defined by

$$C^J(n) = \sum_{j=0}^{nj} c^j(n, J). \tag{3.1}$$

Applying the trace formula (2.5) to the identity operator, we have

$$(2j + 1)^n = \sum_{j=0}^{nj} c^j(n, J) (2J + 1). \tag{3.2}$$

Hence

$$(2j + 1)^n \geq C^J(n) \geq (2j + 1)^n / (2nj + 1) \tag{3.3}$$

so that

$$\lim_{n \rightarrow \infty} n^{-1} \log C^J(n) = \log(2j + 1). \tag{3.4}$$

Introduce the function $t : [0, \infty) \times [0, \infty) \rightarrow [0, \infty)$ defined by

$$t(s, v) = \begin{cases} (\sinh 1/2s)^{-1} \sinh(v + 1/2)s, & s > 0, \\ 2v + 1, & s = 0; \end{cases} \tag{3.5}$$

we have

$$\text{tr}_{\mathfrak{D}(J)}(\pi^J(e^{\lambda S^z})) = \sum_{m=-J}^J e^{\lambda m} = t(\lambda, J). \tag{3.6}$$

Lemma 1. *Let $\{\mathbb{K}_n^j : n \geq 1\}$ be the sequence of probability measures on $[0, 1]$ defined by (1.6); then for every $s \geq 0$,*

$$\lim_{n \rightarrow \infty} n^{-1} \log \int_{[0, 1]} e^{njs} \mathbb{K}_n^j[du] = \log t(s, j) - \log t(0, j). \tag{3.7}$$

Proof. For $s = 0$, the result is trivially true. Assume $s > 0$; applying the trace formula (2.5) to the operator $\exp\{\lambda S_n^z\}$, we have

$$t(\lambda, j)^n = \sum_{j=0}^{nj} c^j(n, J) t(\lambda, J) = C^J(n) \int_{[0, 1]} t(\lambda, nju) \mathbb{K}_n^j[du]. \tag{3.8}$$

Using (3.4), we have

$$\lim_{n \rightarrow \infty} n^{-1} \log \int_{[0, 1]} t(s, nju) \mathbb{K}_n^j[du] = \log t(s, j) - \log t(0, j). \tag{3.9}$$

Now, for $x \geq 0$,

$$x(1 + 2x)^{-1} e^x \leq \sinh(x) \leq 1/2e^x \tag{3.10}$$

so that [using the fact that $v \rightarrow (2v + 1)s(1 + (2v + 1)s)^{-1}$ is increasing in $v \geq 0$ for every $s \geq 0$]

$$2(1 + s)s^{-1} e^{-s/2} \sinh(s/2) t(s, nju) \geq e^{njs} \geq 2e^{-s/2} \sinh(s/2) t(s, nju), \tag{3.11}$$

and (3.7) follows from (3.11) and (3.9). \square

Once (3.7) has been established, the proof that $\{\mathbb{K}_n^j\}$ satisfies the large deviation principle is routine (see [2] for an exposition of general principles); to make the paper self-contained, we provide a proof even though it could be deduced, using Lemma 1, from Theorem II.6.1 of [1].

Theorem 1. *The sequence $\{\mathbb{K}_n^j: n \geq j\}$ satisfies the large deviation principle with rate-function $I^j: [0, 1] \rightarrow \mathbb{R}$ given by (1.10).*

Proof. Notice that $I^j(0) = 0$, $I^j(1) = \log(t(0, j))$, and I^j is increasing and convex. The Legendre transform of a function is automatically lower semi-continuous, so (LD 1) holds. This implies that the level sets of I^j are closed; it is easy to check that they are bounded, so (LD 2) holds. It follows from Markov's Inequality that the large deviation upper bound holds for a closed interval $A_y = [y, 1]$, $y > 0$:

$$\mathbb{K}_n^j[A_y] \leq \int_{[0, 1]} e^{n\alpha(u-y)} \mathbb{K}_n^j[du]. \tag{3.13}$$

It follows from (3.7) that

$$\limsup_{n \rightarrow \infty} n^{-1} \log \mathbb{K}_n^j[A_y] \leq -\alpha y + \log t(\alpha/j, j) - \log t(0, j). \tag{3.14}$$

Hence

$$\limsup_{n \rightarrow \infty} n^{-1} \log \mathbb{K}_n^j[A_y] \leq -I^j(y) = -\inf_{A_y} I^j(u). \tag{3.15}$$

Now let C be an arbitrary closed subset of $[0, 1]$; if 0 belongs to C , the bound is trivially true; otherwise, take $y = \inf C$, and consider $A_y \supseteq C$. This establishes (LD 3); to prove (LD 4), we introduce the exponential family $\{\mathbb{K}_n^{j,\alpha}: \alpha \geq 0\}$ defined by

$$\mathbb{K}_n^{j,\alpha}[du] = e^{n\alpha u} \mathbb{K}_n^j[du] / \int_{[0, 1]} e^{n\alpha x} \mathbb{K}_n^j[dx]. \tag{3.16}$$

For $\alpha \geq 0$, let $\varrho(\alpha)$ be given by

$$\varrho(\alpha) = \left\{ \frac{d}{ds} \log t((\alpha + s)/j, j) \right\}_{s=0}. \tag{3.17}$$

Note that $\alpha \rightarrow \varrho(\alpha)$ is an increasing function with $\varrho(0) = 0$ and $\lim_{\alpha \rightarrow \infty} \varrho(\alpha) = 1$.

Lemma 2. *For each $\alpha \geq 0$, the sequence $\{\mathbb{K}_n^{j,\alpha}: n \geq 1\}$ converges weakly to the Dirac measure supported at the point $\varrho(\alpha)$:*

$$\mathbb{K}_n^{j,\alpha} \xrightarrow{w} \delta_{\varrho(\alpha)}.$$

Proof. Let

$$p_n(s) = n^{-1} \log \int_{[0, 1]} e^{nsu} \mathbb{K}_n^j[du], \quad s \geq 0, \tag{3.18}$$

which is convex, differentiable on $(0, \infty)$, and, by Lemma 1, converges to $p(s) = \log t(s/j, j) - \log t(0, j)$. p is differentiable on $(0, \infty)$ with derivative equal to ϱ . Let $s > 0$; one has

$$\log \int_{[0, 1]} e^{su} \mathbb{K}_n^{j,\alpha}[du] = s \{ p_n(\alpha + s/n) - p_n(\alpha) \} / (s/n). \tag{3.19}$$

By convexity, if p'_n denotes the derivative of p_n ,

$$p'_n(\alpha) \leq \{p_n(\alpha + s/n) - p_n(\alpha)\}/(s/n) \leq p'_n(\alpha + s/n). \tag{3.20}$$

Applying the generalized Griffith’s lemma proved in [8], we deduce that both sides of (3.20) converge to $\varrho(\alpha)$. Thus (the case $s = 0$ is trivial),

$$\lim_{n \rightarrow \infty} \log \int_{[0,1]} e^{su} \mathbb{K}_n^{j,\alpha}[du] = s\varrho(\alpha). \tag{3.21}$$

The result follows from the continuity and uniqueness theorems for the Laplace transform. \square

Returning to the proof that (LD 4) holds, let G be an open subset of $[0, 1]$ and let y be an arbitrary point of G . For each $\delta > 0$ such that $B_y^\delta = (y - \delta, y + \delta) \subset G$ we have

$$\begin{aligned} \mathbb{K}_n^j[G] &\geq \mathbb{K}_n^j[B_y^\delta] = \int_{[0,1]} e^{nu\alpha} \mathbb{K}_n^j[du] \int_{B_y^\delta} e^{-nx\alpha} \mathbb{K}_n^{j,\alpha}[dx] \\ &\geq e^{-n(y+\delta)\alpha} \mathbb{K}_n^{j,\alpha}[B_y^\delta] \int_{[0,1]} e^{nu\alpha} \mathbb{K}_n^j[du]. \end{aligned} \tag{3.22}$$

Now choose α such that $y = \varrho(\alpha)$, which is possible since $y \neq 1$ because G is open; then

$$\begin{aligned} \liminf_{n \rightarrow \infty} n^{-1} \log \mathbb{K}_n^j[G] &\geq -(y + \delta)\alpha + \log t(\alpha/j, j) - \log t(0, j) \\ &\quad + \liminf_{n \rightarrow \infty} n^{-1} \log \mathbb{K}_n^{j,\alpha}[B_y^\delta]. \end{aligned} \tag{3.23}$$

By Lemma 2, for n sufficiently large we have $\mathbb{K}_n^{j,\alpha}[B_y^\delta] > 1/2$ so that

$$\lim_{n \rightarrow \infty} n^{-1} \log \mathbb{K}_n^{j,\alpha}[B_y^\delta] = 0, \tag{3.24}$$

and (LD 4) holds since $\delta > 0$ was arbitrary subject to the condition $(y - \delta, y + \delta) \subset G$ and y was an arbitrary point of G . \square

We can now apply Varadhan’s theorem to (2.12) using (3.4) to obtain the following result:

Theorem 2. *Suppose that the self-adjoint operator H_n on $\mathfrak{h}_n \otimes \mathfrak{R}_n$ is homogeneously decomposable and $\exp\{-\beta H_n\}$ is trace-class for every $n \geq 1$. Let $f_n^j(\beta, \cdot)$ be continuous on $[0, 1]$, and satisfy (2.10) for all $n \geq 1$ and $J \in \mathfrak{J}$. Suppose that the sequence $\{f_n^j(\beta, \cdot) : n \geq 1\}$ converges uniformly to $f^j(\beta, \cdot)$, then, $f(\beta) = \lim_{n \rightarrow \infty} f_n(\beta)$ exists and*

$$f(\beta) = \inf \{ f^j(\beta, u) - \beta^{-1} I_0^j(u) : u \in [0, 1] \},$$

where I_0^j is given by (1.12).

Remark. The function $f_n^j(\beta, \cdot)$ is some continuous interpolation of the free-energy (2.10); at first sight, it might be thought that a different choice of continuous interpolation could yield a different limit function. This is not the case under the continuity and convergence hypotheses of Theorem 2.

An Application: The BCS Model. For $j=1/2$, the following Hamiltonian appears in the strong-coupling-limit theory of the BCS model [9–11].

$$H_n = \varepsilon_n \sum_{k=1}^n \{1_{(k)} - {}^jS_{(k)}^z\} + \lambda_n \sum_{k, k'=1}^n {}^jS_{(k)}^+ {}^jS_{(k')}^-;$$

H_n acts on \mathfrak{R}_n and ${}^jS_{(k)}^\pm = {}^jS_{(k)}^x \pm i {}^jS_{(k)}^y$. H_n can be rewritten as

$$H_n = n\varepsilon_n \cdot 1_{\mathfrak{R}_n} + (\lambda_n - \varepsilon_n)S_n^z + \lambda_n(S_n^x)^2 - \lambda_n(S_n^y)^2.$$

By (2.13), H_n is homogeneously decomposable with

$$H_n(J) = (n\varepsilon_n + \lambda_n J(J+1)) \cdot 1_{\mathfrak{D}(J)} + (\lambda_n - \varepsilon_n)J^2S^z - \lambda_n(J^2S^z)^2.$$

Since the eigenvalues of J^2S^z are $\{-J, -J+1, \dots, J-1, J\} = \mathfrak{S}$, we have

$$Z_n^j(J/nj) = \exp\{-\beta[n\varepsilon_n + \lambda_n J(J+1)]\} \sum_{k \in \mathfrak{S}} \exp\{-\beta[(\lambda_n - \varepsilon_n)k - \lambda_n k^2]\}.$$

Let μ be the probability measure on the interval $[-1, 1]$ defined by

$$\mu[E] = (2J+1)^{-1} \{k \in \mathfrak{S} : kJ^{-1} \in E\}.$$

Then,

$$\begin{aligned} f_n^j(\beta; J/nj) &= \varepsilon_n + n^{-1}\lambda_n J(J+1) - (n\beta)^{-1} \log(2J+1) \\ &\quad - (n\beta)^{-1} \log \left\{ \int_{[-1, 1]} \exp\{\beta[J^2\lambda_n x^2 + J(\varepsilon_n - \lambda_n)x]\} \mu[dx] \right\}. \end{aligned}$$

Thus,

$$\begin{aligned} f_n^j(\beta; u) &= \varepsilon_n + j^2 n \lambda_n u(u + (nj)^{-1}) - (n\beta)^{-1} \log(2nju + 1) \\ &\quad - (n\beta)^{-1} \log \left\{ \int_{[-1, 1]} e^{n\beta \varphi_n^j(u, x)} \mu[dx] \right\}, \end{aligned}$$

where we have defined $\varphi_n^j(\cdot, \cdot)$ on $[0, 1] \times [-1, 1]$ by

$$\varphi_n^j(u, x) = j^2 n \lambda_n u^2 x^2 + j(\varepsilon_n - \lambda_n)ux.$$

Suppose now that

$$\lim_{n \rightarrow \infty} \varepsilon_n = \varepsilon, \quad \text{and} \quad \lim_{n \rightarrow \infty} n \lambda_n = \lambda$$

both exist. The sum of the first three summands in the formula for $f_n^j(\beta; \cdot)$ converges uniformly to $\varepsilon + j^2 \lambda u^2$. Moreover,

$$\lim_{n \rightarrow \infty} \varphi_n^j(u, \cdot) = \varphi^j(u, \cdot) \text{ uniformly,}$$

where

$$\varphi^j(u, x) = j\varepsilon ux + j^2 \lambda u^2 x^2.$$

Applying Varadhan's theorem to the constant sequence $\{\mu_n \equiv \mu : n \geq 1\}$, we have

$$\lim_{n \rightarrow \infty} (n\beta)^{-1} \log \left\{ \int_{[-1, 1]} e^{n\beta \varphi_n^j(u, x)} \mu[dx] \right\} = \sup \{ \varphi^j(u, x) : x \in [-1, 1] \}.$$

We conclude from Theorem 2 (using compactness of $[0, 1]$), that $f(\beta)$ exists, and

$$f(\beta) = \varepsilon + \inf \{ -j\epsilon u x + j^2 \lambda u^2 (1 - x^2) - \beta^{-1} I_0^j(u) : u : u \in [0, 1], x \in [-1, 1] \}.$$

Large deviation methods are used in [5] to discuss the general BCS model (see [12]).

4. The Berezin-Lieb Inequalities for Z_n^j

In this section we use the Berezin-Lieb inequalities to obtain upper and lower bounds on the partition functions $Z_n^j(J/nj)$. This then gives us bounds on Z_n which can be treated by Large Deviations.

Consider a fixed J in $\{0, 1/2, 1, 3/2, 2, \dots\}$, and let $\{|J, \hat{e}\rangle : \hat{e} \in S^2\}$ be the family of Bloch-coherent unit vectors in $\mathfrak{D}(J)$ parametrized by the unit sphere S^2 in \mathbb{R}^3 (we use the conventions of [13]). We let m be the probability measure on S^2 given by $m[d\hat{e}] = (4\pi)^{-1} dx dy dz$. Consider $P(J, \hat{e})$, the orthogonal projection of $\mathfrak{D}(J)$ onto the subspace spanned by $|J, \hat{e}\rangle$; then

$$(2J + 1) \int_{S^2} P(J, \hat{e}) m[d\hat{e}] = 1_{\mathfrak{D}(J)}. \tag{4.1}$$

Given any selfadjoint $A \in \mathfrak{B}(\mathfrak{D}(J))$, let

$$A^1(J, \hat{e}) = \text{tr}_{\mathfrak{D}(J)}(AP(J, \hat{e})). \tag{4.2}$$

There exists a function $A^u(J, \cdot)$ on S^2 (see [13, 14]), such that (in the sense of matrix elements)

$$A = (2J + 1) \int_{S^2} A^u(J, \hat{e}) P(J, \hat{e}) m[d\hat{e}]. \tag{4.3}$$

The Berezin-Lieb inequalities [13–15] state that

$$(2J + 1) \int_{S^2} e^{A^1(J, \hat{e})} m[d\hat{e}] \leq \text{tr}_{\mathfrak{D}(J)} e^A \leq (2J + 1) \int_{S^2} e^{A^u(J, \hat{e})} m[d\hat{e}]. \tag{4.4}$$

More generally, if \mathfrak{h} is a separable complex Hilbert space and H is a selfadjoint operator on $\mathfrak{h} \otimes \mathfrak{D}(J)$ such that $\exp\{-\beta H\}$ is trace-class, then ([13]; see Appendix 2 for a proof)

$$(2J + 1) \int_{S^2} \text{tr}_{\mathfrak{h}} e^{-\beta H^1(J, \hat{e})} m[d\hat{e}] \leq \text{tr}_{\mathfrak{h} \otimes \mathfrak{D}(J)} e^{-\beta H} \leq (2J + 1) \int_{S^2} \text{tr}_{\mathfrak{h}} e^{-\beta H^u(J, \hat{e})} m[d\hat{e}], \tag{4.5}$$

where $H^1(J, \cdot)$ is the \mathfrak{h} -operator valued function on S^2 defined by

$$H^1(J, \hat{e}) = \text{tr}_{\mathfrak{D}(J)}(H(1 \otimes P(J, \hat{e}))), \tag{4.6}$$

and $H^u(J, \cdot)$ is any \mathfrak{h} -operator valued function S^2 such that, in the weak sense on domains,

$$H = (2J + 1) \int_{S^2} H^u(J, \hat{e}) (1 \otimes P(J, \hat{e})) m[d\hat{e}]. \tag{4.7}$$

We now apply the inequality (4.5) to the situation described in Sect. 2. Let H_n acting on $\mathfrak{h}_n \otimes \mathfrak{R}_n$ be homogeneously decomposable and $\exp\{-\beta H_n\}$ trace-class.

Let Z_n , and $Z_n^j(J/nj)$ be defined as in Sect. 2. Let

$$f_n^j(\beta; J/nj, \hat{\varrho}) = (-n\beta)^{-1} \log \text{tr}_{\mathfrak{h}_n} \{ \exp \{ -\beta H_n^j(J, \hat{\varrho}) \} \}; \tag{4.8}$$

$$\tilde{f}_n^j(\beta; J/nj, \hat{\varrho}) = (-n\beta)^{-1} \log \text{tr}_{\mathfrak{h}_n} \{ \exp \{ -\beta H_n^j(J, \hat{\varrho}) \} \}. \tag{4.9}$$

Then, by (2.11), and (4.5) applied to $Z_n^j(J/nj)$,

$$\begin{aligned} & \sum_{J \geq 0}^n c^j(n, J) (2J + 1) \int_{S^2} \exp \{ -n\beta f_n^j(\beta; J/nj, \hat{\varrho}) \} m[d\hat{\varrho}] \leq Z_n \\ & \leq \sum_{J \geq 0}^n c^j(n, J) (2J + 1) \int_{S^2} \exp \{ -n\beta \tilde{f}_n^j(\beta; J/nj, \hat{\varrho}) \} m[d\hat{\varrho}]. \end{aligned} \tag{4.10}$$

Define a probability measure \mathbb{P}_n^j on $[0, 1]$ by [see (3.2)]

$$\mathbb{P}_n^j[E] = (2j + 1)^{-n} \sum_{\substack{J \in \mathbb{J} \\ \{J/nj \in E\}}} (2J + 1) c^j(n, J), \tag{4.11}$$

and consider the product measure $\mathbb{M}_n^j = \mathbb{P}_n^j \times m$ on $\mathbb{T} = [0, 1] \times S^2$. Let the functions $f_n^j(\beta; \cdot, \hat{\varrho})$ and $\tilde{f}_n^j(\beta; \cdot, \hat{\varrho})$ on $[0, 1]$ satisfy (4.8), respectively (4.9), for every $n \geq 1$, and $J \in \mathbb{J}$; we can rewrite (4.10) as

$$\begin{aligned} & C^j(n) (2j + 1)^n \int_{\mathbb{T}} \exp \{ -n\beta f_n^j(\beta; t) \} \mathbb{M}_n^j[dt] \leq Z_n \\ & \leq C^j(n) (2j + 1)^n \int_{\mathbb{T}} \exp \{ -n\beta \tilde{f}_n^j(\beta; t) \} \mathbb{M}_n^j[dt], \end{aligned} \tag{4.12}$$

where $t = (u, \hat{\varrho})$.

Recalling the definition (1.6) of \mathbb{K}_n^j , and the inequality $0 \leq J \leq nj$ for $J \in \mathbb{J}$, we have

$$C^j(n) (2j + 1)^{-n} \mathbb{K}_n^j[B] \leq \mathbb{P}_n^j[B] \leq (2nj + 1) C^j(n) (2j + 1)^{-n} \mathbb{K}_n^j[B], \tag{4.13}$$

for every $B \subseteq [0, 1]$, and every $n \geq 1$. The following result is then a consequence of Theorem 1.

Corollary. *The sequence $\{\mathbb{P}_n^j : n \geq 1\}$ satisfies the large deviation principle with rate-function I^j .*

Using the proposition in Appendix 1, we can apply Varadhan’s theorem to both sides of (4.12) to get the following result:

Theorem 3. *Suppose that the self-adjoint operator H_n on $\mathfrak{h}_n \otimes \mathfrak{S}_n$ is homogeneously decomposable and $\exp \{ -\beta H_n \}$ is trace-class for every $n \geq 1$. Let the sequences $\{f_n^j(\beta; \cdot) : n \geq 1\}$, and $\{\tilde{f}_n^j(\beta; \cdot) : n \geq 1\}$ of continuous functions on \mathbb{T} satisfy (4.8), respectively (4.9), for all $n \geq 1$, and $J \in \mathbb{J}$, and converge uniformly on \mathbb{T} to $f^j(\beta; \cdot)$, respectively $\tilde{f}^j(\beta; \cdot)$; then*

$$\begin{aligned} & \inf \{ f^j(\beta; u, \hat{\varrho}) - \beta^{-1} I_0^j(u) : (u, \hat{\varrho}) \in \mathbb{T} \} \leq \liminf_{n \rightarrow \infty} f_n(\beta) \leq \limsup_{n \rightarrow \infty} f_n(\beta) \\ & \leq \inf \{ \tilde{f}^j(\beta; u, \hat{\varrho}) - \beta^{-1} I_0^j(u) : (u, \hat{\varrho}) \in \mathbb{T} \}. \end{aligned}$$

First Application: A General Quadratic Spin-Hamiltonian. On \mathfrak{S}_n , let

$$H_n = c_n \cdot 1_{\mathfrak{S}_n} + r_n \cdot \underline{S}_n + \sum_{\mu, \nu \in \{x, y, z\}} Q_n(\mu, \nu) S_n^\mu S_n^\nu,$$

where $c_n \in \mathbb{R}$, $r_n \in \mathbb{R}^3$, and $Q_n = (Q_n(\mu, \nu))$ is a complex hermitian (3×3) matrix. Then H_n is homogeneously decomposable and

$$H_n(J) = c_n \cdot 1_{\mathfrak{D}(J)} + r_n \cdot J \mathcal{S} + (J \mathcal{S} | Q_n^J \mathcal{S}).$$

Using Lieb’s table (reproduced in Appendix 3) we have

$$\begin{aligned} H_n^l(J, \hat{\varrho}) &= c_n + J r_n \cdot \hat{\varrho} + J(J - 1/2)(\hat{\varrho} | Q_n \hat{\varrho}) + 1/2 J \operatorname{tr}(Q_n) + J k_n \cdot \hat{\varrho}; \\ H_n^u(J, \hat{\varrho}) &= c_n + (J + 1)r_n \cdot \hat{\varrho} + (J + 1)(J + 3/2)(\hat{\varrho} | Q_n \hat{\varrho}) \\ &\quad - 1/2(J + 1) \operatorname{tr}(Q_n) + (J + 1)k_n \cdot \hat{\varrho}; \end{aligned}$$

where $k_n = \operatorname{Im}(Q_n(z, y), Q_n(x, z), Q_n(y, x))$.

Now, $\tilde{f}_n^j(\beta; u, \hat{\varrho}) = n^{-1} H_n^j(nju, \hat{\varrho})$, and $f_n^j(\beta; u, \hat{\varrho}) = n^{-1} H_n^l(nju, \hat{\varrho})$. Thus,

$$\begin{aligned} f_n^j(\beta; u, \hat{\varrho}) &= n^{-1} c_n + ju r_n \cdot \hat{\varrho} + j^2 nu(u - 1/2nj)(\hat{\varrho} | Q_n \hat{\varrho}) \\ &\quad + 1/2ju \operatorname{tr}(Q_n) + ju k_n \cdot \hat{\varrho}, \end{aligned}$$

and similarly for \tilde{f}_n^j . Suppose that

$$\lim_{n \rightarrow \infty} n^{-1} c_n = c, \quad \lim_{n \rightarrow \infty} r_n = r, \quad \text{and} \quad \lim_{n \rightarrow \infty} n Q_n = Q \text{ (elementwise),}$$

then $k_n \rightarrow 0$, and $\operatorname{tr}(Q_n) \rightarrow 0$, so that $\{f_n^j(\beta; \cdot) : n \geq 1\}$, and $\{\tilde{f}_n^j(\beta, \cdot) : n \geq 1\}$ both converge uniformly on \mathbb{T} to

$$f^j(\beta; u, \hat{\varrho}) = c + ju r \cdot \hat{\varrho} + j^2 u^2 (\hat{\varrho} | Q \hat{\varrho}).$$

Hence, by Theorem 3,

$$f(\beta) = c + \inf \{ ju r \cdot \hat{\varrho} + j^2 u^2 (\hat{\varrho} | Q \hat{\varrho}) - \beta^{-1} I_0^j(u) : (u, \hat{\varrho}) \in \mathbb{T} \}.$$

Second Application: The Spin-Boson Model. Consider $n(2j + 1)$ -level atoms or j -spins interacting with a quantal radiation field in a region $\mathcal{A}_n \subset \mathbb{R}^d$ of finite “volume” V_n . The one-particle Hamiltonian \mathfrak{h}_n is assumed to be a positive injective self-adjoint operator on $L^2(\mathcal{A}_n)$ such that $\exp\{-\beta \mathfrak{h}_n\}$ is trace-class. It follows that \mathfrak{h}_n has a bounded inverse. Let \mathfrak{F}_n be the symmetric Fock space over the Hilbert space $L^2(\mathcal{A}_n)$. The Hamiltonian for the full system is taken to be (see [16])

$$H_n = d\Gamma(\mathfrak{h}_n) \otimes 1 + \varepsilon \sum_{k=1}^n (1 \otimes J S_{(k)}^z) + (V_n)^{-1/2} \sum_{k=1}^n \{ a^*(\lambda_n) + a(\lambda_n) \} \otimes J S_{(k)}^x,$$

acting on $\mathfrak{F}_n \otimes \mathfrak{R}_n$, where ε is real, $\lambda_n \in L^2(\mathcal{A}_n)$, $d\Gamma$ denotes second quantization, and $a(f)$, $f \in L^2(\mathcal{A}_n)$, denotes the usual annihilation operator (see e.g. [17]). Let

$$A_n = \|(\mathfrak{h}_n)^{-1/2} \lambda_n\|^2.$$

We consider the limit $V_n \rightarrow \infty$ as $n \rightarrow \infty$, with $\varrho = n/V_n$, the density of particles, fixed. H_n is homogeneously decomposable since

$$H_n = d\Gamma(\mathfrak{h}_n) \otimes 1 + \varepsilon(1 \otimes S_n^z) + (V_n)^{-1/2} \{ a^*(\lambda_n) + a(\lambda_n) \} S_n^x.$$

A glance at Lieb’s table (see Appendix 3) shows that

$$\begin{aligned} H_n^l(J, \hat{\varrho}) &= d\Gamma(\mathfrak{h}_n) + \varepsilon J z \cdot 1_{\mathfrak{b}} + (\varrho/n)^{1/2} J x \{ a^*(\lambda_n) + a(\lambda_n) \}, \\ H_n^u(J, \hat{\varrho}) &= H_n^l(J + 1, \hat{\varrho}). \end{aligned}$$

If $W[f] = \exp\{a(f) - a^*(f)\}$, $f \in L^2(A_n)$, denotes the unitary Weyl operator on \mathfrak{F}_n , then [18],

$$H_n^l(J, \hat{e}) = W[(\varrho/n)^{1/2} Jx(\mathfrak{h}_n)^{-1} \lambda_n]^* d\Gamma(\mathfrak{h}_n) W[(\varrho/n)^{1/2} Jx(\mathfrak{h}_n)^{-1} \lambda_n] + \varepsilon Jz \cdot 1_{\mathfrak{h}} - (\varrho/n) J^2 x^2 A_n \cdot 1_{\mathfrak{h}},$$

and similarly for H_n^u . The traces in (4.8) and (4.9) can be done easily, to obtain

$$f_n^j(\beta; u, \hat{e}) = \varepsilon j u z - \varrho(j u x)^2 A_n + \varrho^{-1} f_n^0(\beta),$$

$$\tilde{f}_n^j(\beta; u, \hat{e}) = \varepsilon j(u + (1/nj))z - \varrho(j(u + (1/nj))x)^2 A_n + \varrho^{-1} f_n^0(\beta),$$

where $f_n^0(\beta)$ is the specific free energy of the free field:

$$f_n^0(\beta) = (-\beta V_n)^{-1} \log \text{tr}_{\mathfrak{F}_n}(\exp\{-\beta d\Gamma(\mathfrak{h}_n)\}).$$

Suppose that

$$A = \lim_{n \rightarrow \infty} A_n, \quad \text{and} \quad f^0(\beta) = \lim_{n \rightarrow \infty} f_n^0(\beta),$$

both exist; then both $f_n^j(\beta; \cdot)$, and $\tilde{f}_n^j(\beta; \cdot)$ converge uniformly to

$$f^j(\beta; u, \hat{e}) = \varepsilon j u z - \varrho A(j u x)^2 + \varrho^{-1} f^0(\beta).$$

It follows from Theorem 3, that

$$f(\beta) = f^0(\beta) + \varrho \inf\{\varepsilon j u x - \varrho A(j u)^2(1 - x^2) - \beta^{-1} I_0^j(u) : u \in [0, 1], x \in [-1, 1]\}.$$

A more complete discussion of the statistical thermodynamics of this model will be given in [19]. We note that our result seems to be optimal in the sense that the conditions imposed on the coupling constants are minimal. The full details of the steepest-descent argument of Hepp and Lieb [16] were given in [20] for $j = 1/2$, where $c^j(n, J)$ is known explicitly. The best previous result appears to be that of Zagrebnov [21], who also gives a very detailed bibliography.

Appendix 1

The following is a double application of Theorem 3.1 of [6].

Lemma. *Let X and Y be complete separable metric spaces. Suppose $\{\mathbb{K}_n : n \geq 1\}$ is a sequence of Borel probability measures on X satisfying the large deviation principle with constants $\{V_n\}$ and rate-function $H : X \rightarrow [0, \infty]$, and that $\{\mathbb{L}_n : n \geq 1\}$ is a Borel probability measure on Y satisfying the large deviation principle with the same constants $\{V_n\}$ and rate-function $I : Y \rightarrow [0, \infty]$. Let F be an extended real valued upper semicontinuous function on $X \times Y$ which is bounded above; then*

$$\limsup_{n \rightarrow \infty} V_n^{-1} \log \int_{X \times Y} \exp(V_n F(x, y)) \{\mathbb{K}_n \times \mathbb{L}_n\} [d(x, y)] \leq \sup_{X \times Y} \{F(x, y) - H(x) - I(y)\}.$$

Proof. Define the function $G_n : X \rightarrow \mathbb{R}$ by

$$G_n(x) = V_n^{-1} \log \int \exp(V_n F(x, y)) \mathbb{L}_n[dy],$$

by using Fubini’s theorem. It follows that G_n is bounded above with the same bound as F . Let $\{x_n : n \geq 1\}$ be a sequence in X converging to $x \in X$. Define $F_n : Y \rightarrow \mathbb{R}$ by $F_n(y) = F(x_n, y)$. Then F_n is bounded above, and if $\{y_n : n \geq 1\}$ is a sequence in Y converging to $y \in Y$, we have $\limsup_{n \rightarrow \infty} F_n(y_n) \leq F(x, y)$ by upper semicontinuity. Thus, the hypothesis of Theorem 3.1 of [6] is met (see the remark on p. 278 of [6]), and we conclude that

$$\begin{aligned} \limsup_{n \rightarrow \infty} G_n(x_n) &= \limsup_{n \rightarrow \infty} V_n^{-1} \log \int_Y \exp(V_n F_n(y)) \mathbb{L}_n[dy] \\ &\leq \sup_Y \{F(x, y) - I(y)\}. \end{aligned} \tag{A1.1}$$

By Fubini’s theorem,

$$\int_{X \times Y} \exp(V_n F(x, y)) \{\mathbb{K}_n \times \mathbb{L}_n\}[d(x, y)] = \int \exp(V_n G_n(x)) \mathbb{K}_n[dx]. \tag{A1.2}$$

Again, (A1.1) says that the conditions of Theorem 3.1 of [6] are satisfied, so that

$$\limsup_{n \rightarrow \infty} V_n^{-1} \log \int_X \exp(V_n G_n(x)) \mathbb{K}_n[dx] \leq \sup_X \left\{ \sup_Y \{F(x, y) - I(y)\} - H(x) \right\},$$

and the claim follows from (A1.2). \square

The following result is heuristically obvious.

Proposition. *Take the assumptions and notation of the lemma. The sequence $\{\mathbb{M}_n : n \geq 1\}$ of Borel probability measures $\mathbb{M}_n = \mathbb{K}_n \times \mathbb{L}_n$ on $X \times Y$ satisfies the large deviation principle with the same constants $\{V_n\}$ and rate-function $J : X \times Y \rightarrow [0, \infty]$ given by*

$$J(x, y) = H(x) + I(y).$$

Proof. J satisfies (LD 1) and (LD 2). A rectangle is a subset of $X \times Y$ of the form $A \times B$, where $A \subseteq X$ and $B \subseteq Y$. For rectangles, $\log \mathbb{M}_n[A \times B] = \log \mathbb{K}_n[A] + \log \mathbb{L}_n[B]$. Thus (LD 3) and (LD 4) follow directly from the assumptions.

To prove (LD 4), let z be a point in an arbitrary open subset G of $Z \equiv X \times Y$. There exist open subsets G_1 of X , and G_2 of Y such that $z \in G_1 \times G_2 \subseteq G$. It follows that $\mathbb{M}_n[G] \geq \mathbb{M}_n[G_1 \times G_2]$; dividing this inequality by V_n , taking logarithms, the inferior limit, and invoking (LD 4) for rectangles, we obtain:

$$\liminf_{n \rightarrow \infty} V_n^{-1} \log \mathbb{M}_n[G] \geq - \inf_{G_1 \times G_2} \{J(x, y)\} \geq -J(z); \tag{A1.3}$$

the right inequality is obvious since $z \in G_1 \times G_2$. Since z was arbitrary in G , (LD 4) follows from (A1.3) by taking the supremum with respect to G .

Let C be closed in Z and define F_C on Z by $F_C(x, y) = 0$ if $(x, y) \in C$, and $F_C = -\infty$ otherwise. F_C is upper semicontinuous and

$$\mathbb{M}_n[C] = \int_Z \exp(V_n F_C(x, y)) \mathbb{M}_n[d(x, y)]. \tag{A1.4}$$

Dividing (A1.4) by V_n , taking logarithms and the superior limit, (LD 3) follows from the lemma. This completes the proof. \square

Appendix 2

Let (X, \mathfrak{S}, μ) be a measure space; a family of *coherent projections* on (X, \mathfrak{S}, μ) is a weakly measurable map $x \rightarrow P(x)$ from X to $\text{proj}(\mathfrak{R})$, the set of orthogonal projections on a separable Hilbert space \mathfrak{R} , satisfying:

$$\text{tr}(P(x)) = 1, \tag{CP 1}$$

$$\int_X P(x)\mu(dx) = 1_{\mathfrak{R}}. \tag{CP 2}$$

It follows [14] from (CP 1) and (CP 2) that $\mu[X] = \text{dim}(\mathfrak{R})$.

Theorem. (Berezin-Lieb Inequalities). *Let $\mathfrak{h}, \mathfrak{R}$ be separable Hilbert spaces; let A be a self-adjoint operator on $\mathfrak{h} \otimes \mathfrak{R}$ such that $\text{tr}_{\mathfrak{h} \otimes \mathfrak{R}}(e^A)$ is finite. Let $P : X \rightarrow \text{proj}(\mathfrak{R})$ be a family of coherent projections on the measure space (X, \mathfrak{S}, μ) and define the map A^l from X to the self-adjoint operators on \mathfrak{h} by*

$$A^l(x) = \text{tr}_{\mathfrak{R}}(A(1 \otimes P(x))). \tag{A.2.1}$$

Then

$$\int_X \text{tr}_{\mathfrak{h}}(e^{A^l(x)})\mu(dx) \leq \text{tr}_{\mathfrak{h} \otimes \mathfrak{R}}(e^A). \tag{BL 1}$$

Suppose, in addition, that there exists a map A^u from X to the self-adjoint operators on \mathfrak{h} such that

$$\int_X A^u(x) \otimes P(x)\mu(dx) = A, \tag{A.2.2}$$

in a weak sense; then

$$\text{tr}_{\mathfrak{h} \otimes \mathfrak{R}}(e^A) \leq \int_X \text{tr}_{\mathfrak{h}}(e^{A^u(x)})\mu(dx). \tag{BL 2}$$

Proof. (Adapted from [14] where it is proved for the case $\mathfrak{h} = \mathbf{C}$). Since A is self-adjoint and e^A is trace class, the spectrum of A consists entirely of isolated eigenvalues of finite multiplicity; let $\{\xi(j) : j \in \mathbb{J}\}$ be an orthonormal basis for $\mathfrak{h} \otimes \mathfrak{R}$ consisting of eigenvectors of A :

$$A\xi(j) = \lambda(j)\xi(j), \tag{A.2.3}$$

where the eigenvalues $\{\lambda(j) : j \in \mathbb{J}\}$ are labelled according to their multiplicities. Let $Q : \mathbb{J} \rightarrow \text{proj}(\mathfrak{h} \otimes \mathfrak{R})$ be the map defined by

$$Q(j)\eta = \langle \xi(j), \eta \rangle \xi(j) \tag{A.2.4}$$

for each vector η in $\mathfrak{h} \otimes \mathfrak{R}$, then the operator

$$Q(j, x) = \text{tr}_{\mathfrak{R}}(Q(j)(1 \otimes P(x)))$$

satisfies

$$0 \leq Q(j, x) \leq 1_{\mathfrak{h}} \quad \text{for all } j, x. \tag{A.2.5}$$

By (CP 1),

$$\sum_{j \in \mathbb{J}} Q(j, x) = 1_{\mathfrak{h}} \quad \text{for all } x; \tag{A.2.6}$$

by (A.2.3),

$$A^l(x) = \sum_{j \in \mathbb{J}} \lambda(j) Q(j, x) \quad \text{for all } x, \tag{A.2.7}$$

and by (CP 2),

$$\int_X \text{tr}_{\mathfrak{h}}(Q(j, x)) \mu(dx) = 1 \quad \text{for all } j. \tag{A.2.8}$$

Let ψ be a unit vector in \mathfrak{h} ; then $j \rightarrow Q_\psi(j, x) = \langle \psi, Q(j, x)\psi \rangle_{\mathfrak{h}}$ defines a probability measure on the countable set \mathbb{J} by (A.2.5) and (A.2.6). Using Jensen's Inequality and (A.2.7), we have

$$\exp\{\langle \psi, A^l(x)\psi \rangle_{\mathfrak{h}}\} \leq \sum_{j \in \mathbb{J}} e^{\lambda(j)} Q_\psi(j, x) \quad \text{for all } x. \tag{A.2.9}$$

Choosing an orthonormal basis for \mathfrak{h} consisting of eigenvectors of $A^l(x)$, we have by (A.2.9):

$$\text{tr}_{\mathfrak{h}}(e^{A^l(x)}) \leq \sum_{j \in \mathbb{J}} e^{\lambda(j)} \text{tr}_{\mathfrak{h}}(Q(j, x)) \quad \text{for all } x. \tag{A.2.10}$$

Integrating (A.2.10) with respect to μ and using (A.2.8) we get (BL 1).

For each x , let $\{\psi(x, m) : m \in \mathbb{M}\}$ be an orthonormal basis for \mathfrak{h} consisting of eigenvalues of $A^u(x)$:

$$A^u(x)\psi(x, m) = a(x, m)\psi(x, m), \tag{A.2.11}$$

where the eigenvalues $\{a(x, m) : m \in \mathbb{M}\}$ are labelled according to their multiplicities. Let $R : X \times \mathbb{M} \rightarrow \text{proj}(\mathfrak{h})$ be the map defined by

$$R(x, m)\phi = \langle \psi(x, m), \phi \rangle_{\mathfrak{h}} \psi(x, m), \tag{A.2.12}$$

for each vector ϕ in \mathfrak{h} ; then, by (A.2.11), we have

$$A^u(x) = \sum_{m \in \mathbb{M}} a(x, m) R(x, m) \quad \text{for each } x, \tag{A.2.13}$$

and

$$\sum_{m \in \mathbb{M}} R(x, m) = 1_{\mathfrak{h}} \quad \text{for each } x. \tag{A.2.14}$$

For each Borel subset E of X and each m in \mathbb{M} , put

$$\sigma_j(E, m) = \int_E \text{tr}_{\mathfrak{h}}(R(x, m)Q(j, x)) \mu(dx); \tag{A.2.15}$$

then, by (CP 2) and (A.2.14), $E \times \{m\} \rightarrow \sigma_j(E, m)$ determines a probability measure on the Borel subsets of $X \times \mathbb{M}$. By Jensen's Inequality, we have

$$\exp\left\{\sum_{m \in \mathbb{M}} \int_X a(x, m) \sigma_j(dx, m)\right\} \leq \sum_{m \in \mathbb{M}} \int_X e^{a(x, m)} \sigma_j(dx, m) \tag{A.2.16}$$

but, by (A.2.4) and (A.2.14),

$$\sum_{m \in \mathbb{M}} \int_X a(x, m) \sigma_j(dx, m) = \lambda(j). \tag{A.2.17}$$

Summing up (A.2.16) over j and using (A.2.8), we get (BL 2). \square

Appendix 3

The following is a completion of Lieb’s table (p. 330 of [13]) computed by the method expounded in Appendix A of [13].

S^α , $\alpha = x, y, z$, are the components of $^J\mathcal{S}$, where J is fixed in $\{0, 1/2, 1, 3/2, 2, \dots\}$.

	Lower symbol $\{\cdot\}^l(J, \hat{e})$	Upper symbol $\{\cdot\}^u(J, \hat{e})$
S	$J\hat{e}$	$(J+1)\hat{e}$
$(S^x)^2$	$J(J-1/2)x^2 + 1/2J$	$(J+1)(J+3/2)x^2 - 1/2(J+1)$
$(S^y)^2$	$J(J-1/2)y^2 + 1/2J$	$(J+1)(J+3/2)y^2 - 1/2(J+1)$
$(S^z)^2$	$J(J-1/2)z^2 + 1/2J$	$(J+1)(J+3/2)z^2 - 1/2(J+1)$
S^xS^y	$J(J-1/2)xy + i1/2Jz$	$(J+1)(J+3/2)xy + i1/2(J+1)z$
S^zS^x	$J(J-1/2)xz + i1/2Jy$	$(J+1)(J+3/2)xz + i1/2(J+1)y$
S^yS^z	$J(J-1/2)yz + i1/2Jx$	$(J+1)(J+3/2)yz + i1/2(J+1)x$

The complex conjugate of an upper symbol (respectively the lower symbol) of $S^\beta S^\alpha$ is an upper (respectively the lower symbol) of $S^\alpha S^\beta$. We remark that in Lieb’s table (p. 330 of [13]) “ $\cos \varphi$ ” should be replaced by “ $\sin \varphi$ ” in the row corresponding to $(S^y)^2$.

References

1. Ellis, R.S.: Entropy, large deviations, and statistical mechanics. Berlin, Heidelberg, New York: Springer 1985
2. Lewis, J.T.: The large deviation principle in statistical mechanics: an expository account. In: Stochastic mechanics and stochastic processes, Swansea Conference Proceedings 1986. Davies, I., Truman, A. (eds.). Lecture Notes in Mathematics. Berlin, Heidelberg, New York: Springer (in press)
3. van den Berg, M., Lewis, J.T., Pulè, J.V.: The large deviation principle and some models of an interacting boson gas. Commun. Math. Phys. **118**, 61–85 (1988)
4. van den Berg, M., Lewis, J.T., Pulè, J.V.: Large deviations and the boson gas. In: Stochastic mechanics and stochastic processes, Swansea Conference Proceedings 1986. Davies, I., Truman, A. (eds.). Lecture Notes in Mathematics. Berlin, Heidelberg, New York: Springer (in press)
5. Duffield, N.G., Pulè, J.V.: Thermodynamics of the B.C.S. model through large deviations. Lett. Math. Phys. **14**, 329–331 (1987)
6. Varadhan, S.R.S.: Asymptotic probabilities and differential equations. Commun. Pure Appl. Math. **19**, 261 (1966)
7. Dixmier, J.: C^* -algebras. Amsterdam, New York, Oxford: North-Holland 1977
8. Lewis, J.T., Pulè, J.V.: The equivalence of ensembles in statistical mechanics. In: Stochastic analysis and its applications, Swansea Conference Proceedings 1983. Truman, A., Williams, D. (eds.). Lecture Notes in Mathematics, Vol. 1095. Berlin, Heidelberg, New York: Springer 1984
9. Wada, Y., Takano, F., Fukuda, N.: Exact treatment of Bardeen’s theory of superconductivity in the strong coupling limit. Progr. Theoret. Phys. (Kyoto) **19**, 597 (1958)
10. Thouless, D.J.: Strong-coupling limit in the theory of superconductivity. Phys. Rev. **117**, 1256 (1960)
11. Thirring, W., Wehrl, A.: On the mathematical structure of the B.C.S. model. Commun. Math. Phys. **4**, 303 (1967)

12. Thouless, D.J.: The quantum mechanics of many-body systems. New York: Academic Press 1961
13. Lieb, E.H.: The classical limit of quantum spin systems. *Commun. Math. Phys.* **31**, 327 (1973)
14. Simon, B.: The classical limit of quantum partition functions. *Commun. Math. Phys.* **71**, 247 (1980)
15. Berezin, F.A.: Covariant and contravariant symbols of operators. (Russian) *Izv. Akad. Nauk SSSR Ser. Mat.* **36**, 1134 (1972) [English transl.: *Math. USSR Izv.* **6**, 117 (1972)]
16. Hepp, K., Lieb, E.H.: Equilibrium statistical mechanics of matter interacting with the quantized radiation field. *Phys. Rev.* **A8**, 2517 (1973)
17. Bratteli, O., Robinson, D.W.: *Operator algebras and quantum statistical mechanics. II.* Berlin, Heidelberg, New York: Springer 1981
18. Cook, J.M.: Asymptotic properties of a boson field with given source. *J. Math. Phys.* **2**, 33 (1961)
19. Lewis, J.T., Raggio, G.A.: Equilibrium thermodynamics of a spin-boson model. *J. Stat. Phys.* **50**, 1201–1220 (1988)
20. Fannes, M., Sisson, P.N.M., Verbeure, A., Wolfe, J.C.: Equilibrium state and free energy of an infinite mode Dicke maser model. *Ann. Phys.* **98**, 38 (1976)
21. Zagrebnov, V.A.: The approximating Hamiltonian method for an infinite-mode Dicke maser model with A^2 -term. *Z. Phys.* **B55**, 75 (1984)

Communicated by J. Fröhlich

Received September 28, 1987; in revised form January 15, 1988