Wehrl's entropy of spin states and Lieb's conjecture

1988 J. Phys. A: Math. Gen. 21 3749

(http://iopscience.iop.org/0305-4470/21/19/013)

J. Phys. A: Math. Gen. 21 (1988) 3749-3761. Printed in the UK

Wehrl's entropy of spin states and Lieb's conjecture

Ching Tsung Lee

Department of Physics, Alabama A & M University, Normal, AL 35762, USA

Received 28 March 1988

Abstract. Wehrl's entropy is the entropy of the probability distribution in the phase space corresponding to the Q representation (antinormal ordering of operators) of a quantum state in terms of coherent states. The Hilbert space of a system of N spins (or N two-level atoms) is of 2^N dimensions. The subspace with maximum total spin $N/2$ is of $N+1$ dimensions. All the discussions in this paper are restricted within this subspace. The probability amplitude for an arbitrary pure state in this subspace is a polynomial of degree N which can be factorised into the product of N linear factors and each root can be identified with one point on the unit sphere. Hence, an arbitrary state in the subspace can be represented geometrically by N unit vectors. The general expression for the Wehrl entropy is obtained as a finite series expansion in terms of some symmetric functions of $\sin^2(\omega_{ij}/2)$, where ω_{ij} is the angle between a pair of unit vectors. The special cases of coherent and almost coherent spin states are then considered. It was pointed out by Lieb that the Wehrl entropy of a coherent spin state of an N-spin system is $N/(N+1)$ and it was conjectured by him that this is the absolute minimum. Based on the fact that, in the geometric representation of a coherent spin state, the N points on the unit sphere condense into a single point, it is shown that the Wehrl entropy of a coherent state is a local minimum; thus Lieb's conjecture is partially confirmed. It is also conjectured that the maximum Wehrl entropy is attained when the N points form a possible regular polyhedron.

1. Introduction

Two kinds of coherent states are widely used in quantum optics to describe radiation-matter interaction; namely, the Glauber coherent states for a quantum harmonic oscillator (Schrödinger 1926, Glauber 1963, Sudarshan 1963) and the coherent spin states (also called the atomic coherent states) for a system of spins or a system of two-level atoms (Radcliffe 1971, Arecchi *et al* 1972). Intuitively, coherence can be considered as the opposite of chaos; thus one important characteristic of coherent states is that they have minimum uncertainties measured by standard deviation. However, some serious defects in using standard deviation as the measure of uncertainties have been discussed by Uffink and Hilgevoord (1985) and, on the other hand, Deutsch (1983) has proposed the use of entropy as an alternative measure of uncertainties.

Quantum mechanical entropy has been defined by von Neumann as $-\mathrm{Tr}(\hat{\rho}\ln\hat{\rho})$, where $\hat{\rho}$ is the density matrix or density operator. The trouble with this definition is that a pure quantum state, coherent or not, always has a minimum entropy of 0. This certainly cannot display the unique character of coherent states. Wehrl (1979) introduced a new definition of 'classical' entropy in terms of the Glauber coherent state $|z\rangle$

as

$$S \equiv -\int \frac{dz}{\pi} \rho(z) \ln \rho(z) \tag{1.1}$$

where

$$\rho(z) \equiv \langle z|\hat{\rho}|z\rangle \tag{1.2}$$

is the diagonal element of the density matrix and, hence, must be non-negative. It was conjectured by Wehrl (1979) and proved by Lieb (1978) that a Glauber coherent state has a minimum entropy of 1 as defined by (1.1).

We believe that this new definition of entropy can best characterise the uniqueness of a coherent state. We also recognise that $\rho(z)$ defined by (1.2) is exactly the distribution function corresponding to antinormal ordering of operators, which was first discussed by Husimi (1940) and was advocated again by Kano (1965); in the terminology of phase space representation of coherent states (see, for example, Haken 1970) it is called Q representation, in contrast with Wigner's (1932) W representation and Glauber's (1963) P representation.

The Q representation and its extension for spin coherent states is finding increasing use in a number of areas such as quantum chaos by Takahashi and Saitô (1987) and by Toda and Ikeda (1987). In its original definition, the Wehrl entropy is relevant to a special class of simultaneous measurement of position and momentum (see, for example, Davies 1976, Milburn 1985). In view of these developments, it should be of current interest to explore the extension of Wehrl's definition of entropy to spin coherent states for possible use in quantum chaos and possible special interpretation in measurement theory.

In the same paper by Lieb (1978), it was also conjectured that the extension of Wehrl's definition of entropy for coherent spin states will yield a minimum entropy of $N/(N+1)$, where N is the number of spins in the system.

The original motive of this work was to prove Lieb's conjecture. But we will first try to lay the foundation, keeping in mind also the possible future exploration of other aspects of spin states. In § 2 we will use a geometric representation for the distribution function of an arbitrary pure state in the $(N+1)$-dimensional subspace with maximum total spin characterised by N unit vectors. In § 3 we will derive the general expression for the Wehrl entropy of an arbitrary spin state with maximum total spin in terms of $\sin^2(\omega_{ij}/2)$, where ω_{ij} is the angle between a pair of two unit vectors. In § 4 we will focus our attention on the coherent spin states and show that the N unit vectors representing a coherent spin state must be identical, i.e. all the ω_{ij} are zero. Then for an almost coherent state considered in § 5, we have $\sin^2(\omega_{ij}/2) \ll 1$ for any pair of unit vectors. This will be the basis in our approach to prove that the Wehrl entropy of a coherent spin state is a local minimum. Unfortunately, we are unable to prove that it is also a global minimum.

It is interesting that the proof of Wehrl's conjecture was published one year before the conjecture itself was formally published. In contrast a decade has passed since the publication of Lieb's conjecture without its confirmation being seen in the literature. This is, perhaps, an indication of the degree of difficulty we face in finding the proof. In view of this fact our partial solution might be a worthwhile first step towards the final solution.

2. Basic formulation

2.1. Q representation of a maximum total spin state

It is well known that, as far as mathematical formulation is concerned, a system of N two-level atoms is exactly the same as a system of N spins. This equivalence was first used by Dicke (1954) to develop his theory of superradiance. The basic quantum states are denoted by $|r, s\rangle$, which have since been called Dicke states, where r is called the cooperation number and s is half of the difference of the number of atoms in the excited state and that in the ground state. But for a spin system the physical meanings are different; r is the total angular momentum of the system and s is the eigenvalue of the z component of the angular momentum.

In this paper, we will consider only the case with $r = N/2$; so there are $N+1$ possible values for s, ranging from $-N/2$ to $+N/2$. For simplicity we will use $|n\rangle$ to denote the basic eigenstates with $n \equiv N/2+s$ so that n runs from 0 to N.

In lemma 2 of Lieb's (1978) paper, it is proved that the state that minimises the Wehrl entropy must be a pure state. Because of this we will restrict ourselves to considering pure states only.

The density matrix of an arbitrary pure state in the $(N+1)$-dimensional subspace can be written in terms of the Dicke states $|n\rangle$ as

$$\hat{\rho}_N \equiv \sum_{n=0}^{N} \sum_{m=0}^{N} C_m^* C_n |n\rangle\langle m| \tag{2.1}$$

with the normalisation condition

$$\sum_{n=0}^{N} C_n^* C_n = 1. \tag{2.2}$$

A coherent spin state can then be expressed as

$$|\theta, \phi\rangle_N \equiv \sum_{n=0}^{N} |n\rangle \binom{N}{n}^{1/2} (\cos(\theta/2))^{N-n} (\sin(\theta/2)\, \mathrm{e}^{-\mathrm{i}\phi})^n \tag{2.3}$$

where θ and ϕ are two angles in the standard spherical coordinate system. The probability density function over the spherical surface in the Q representation is defined as

$$Q_N(\theta, \phi) \equiv \langle \theta, \phi | \hat{\rho} | \theta, \phi \rangle_N \equiv P_N^*(\theta, \phi) P_N(\theta, \phi) \tag{2.4}$$

where

$$P_N(\theta, \phi) \equiv \sum_{n=0}^{N} \binom{N}{n}^{1/2} C_n (\cos(\theta/2))^{N-n} (\sin(\theta/2)\, \mathrm{e}^{\mathrm{i}\phi})^n \tag{2.5}$$

is a polynomial of degree N, and the Wehrl entropy is defined as

$$S_N \equiv -\frac{N+1}{4\pi} \int_0^{\pi} \mathrm{d}\theta \sin\theta \int_0^{2\pi} \mathrm{d}\phi\, Q_N(\theta, \phi) \ln Q_N(\theta, \phi). \tag{2.6}$$

From (2.4) we can see clearly that $Q_N(\theta, \phi)$ is always non-negative, which is absolutely necessary to ensure that S_N is well defined.

2.2. Factorisation and geometric representation

To carry out the integration in (2.6) it is convenient to reduce $P_N(\theta, \phi)$ of (2.5) to its N linear factors. It is always possible to express $P_N(\theta, \phi)$ as the following product:

$$P_N(\theta, \phi) \equiv K_N^{1/2} \prod_{i=1}^{N} p(\theta, \phi; \theta_i, \phi_i) \tag{2.7}$$

where K_N is the normalisation factor and

$$p(\theta, \phi; \theta_i, \phi_i) \equiv \cos(\theta/2)\cos(\theta_i/2) + \sin(\theta/2)\sin(\theta_i/2)\, \mathrm{e}^{\mathrm{i}(\phi-\phi_i)}. \tag{2.8}$$

Substitution of (2.7) into (2.4) gives

$$Q_N(\theta, \phi) = K_N \prod_{i=1}^{N} q(\theta, \phi; \theta_i, \phi_i) \tag{2.9}$$

with

$$q(\theta, \phi; \theta_i, \phi_i) \equiv |p(\theta, \phi; \theta_i, \phi_i)|^2 = \cos^2(\omega_i/2) \tag{2.10}$$

where ω_i is the angle between (θ, ϕ) and (θ_i, ϕ_i).

Expressing $P_N(\theta, \phi)$ in this way will provide a simple geometric representation and hence a clear picture of the distribution function $Q_N(\theta, \phi)$. Consider (θ_i, ϕ_i) as the direction of a unit vector or as the coordinate of a point on the surface of the unit sphere; then an arbitrary pure state with maximum total spin of an N-spin system can be characterised by the locations of N points on the surface of the unit sphere. From (2.9) and (2.10) we see that the distribution function is the product of N individual distributions, each being proportional to $\cos^2(\omega_i/2)$, which peaks at the point (θ_i, ϕ_i) and vanishes at the diagonally opposite point.

2.3. Rotation of the coordinate system

The integration in (2.6) is over the whole surface of the unit sphere. Therefore S_N is invariant under any rotation of the coordinate system. We will use this fact to simplify its calculation. Consider

$$|\theta, \phi\rangle \equiv \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2)\, \mathrm{e}^{\mathrm{i}\phi} \end{pmatrix} \tag{2.11}$$

as the state vector of a one-particle (spin or two-level atom) state so that (2.8) and (2.10) can be written as

$$p(\theta, \phi; \theta_i, \phi_i) = \langle \theta_i, \phi_i | \theta, \phi \rangle \tag{2.12}$$

and

$$q(\theta, \phi; \theta_i, \phi_i) = \langle \theta, \phi | \theta_i, \phi_i \rangle \langle \theta_i, \phi_i | \theta, \phi \rangle. \tag{2.13}$$

Then consider a rotation represented by the matrix

$$T(\theta_n, \phi_n) \equiv \begin{pmatrix} \cos(\theta_n/2) & \sin(\theta_n/2)\, \mathrm{e}^{-\mathrm{i}\phi_n} \\ \sin(\theta_n/2)\, \mathrm{e}^{\mathrm{i}\phi_n} & -\cos(\theta_n/2) \end{pmatrix}. \tag{2.14}$$

Under this rotation we have

$$p(\theta, \phi; \theta_i, \phi_i) \xrightarrow{T(\theta_n, \phi_n)} p_{in}(\theta, \phi) \tag{2.15}$$

where

$$\begin{aligned} p_{in}(\theta, \phi) \equiv \langle \theta_i, \phi_i | T(\theta_n, \phi_n) | \theta, \phi \rangle \\ = \cos(\theta_i/2) \cos(\theta_n/2) \cos(\theta/2) + \sin(\theta_i/2) \sin(\theta_n/2) \cos(\theta/2) \, e^{i(\phi_n - \phi_i)} \\ + \cos(\theta_i/2) \sin(\theta_n/2) \sin(\theta/2) \, e^{i(\phi - \phi_n)} \\ - \sin(\theta_i/2) \cos(\theta_n/2) \sin(\theta/2) \, e^{i(\phi - \phi_i)}. \end{aligned} \tag{2.16}$$

The significance of $T(\theta_n, \phi_n)$ can be seen in the special case when (θ_n, ϕ_n) coincides with (θ_i, ϕ_i); then (2.16) reduces to

$$p_{ii}(\theta, \phi) = \cos(\theta/2). \tag{2.17}$$

Under this rotation $q(\theta, \phi; \theta_i, \phi_i)$ becomes

$$\begin{aligned} q_{in}(\theta, \phi) \equiv |p_{in}(\theta, \phi)|^2 \\ = \tfrac{1}{2}[1 + \cos \omega_{in} \cos \theta + (f_{in}^* \, e^{i\phi} + f_{in} \, e^{-i\phi}) \sin \theta] \end{aligned} \tag{2.18}$$

where

$$f_{in} \equiv \tfrac{1}{2}[\cos \theta_i \sin \theta_n - \sin \theta_i \cos \theta_n \cos(\phi_n - \phi_i) + i \sin \theta_n \sin(\phi_n - \phi_i)] \, e^{i\phi_n}. \tag{2.19}$$

3. General expression for the normalisation factor and the entropy

In this section we will derive formulae for the normalisation constant K_N and the Wehrl entropy S_N for an arbitrary pure spin state with maximum total spin. They will be expressed as finite series expansions in terms of some symmetric functions of $\sin^2(\omega_{ij}/2)$. We need two kinds of symmetric functions; they are defined as follows:

$$D_m^N \equiv \sum_{i_1=1}^{N} \sum_{j_1>i_1}^{N} \sum_{i_2>i_1}^{N}{}^{*} \sum_{j_2>i_2}^{N}{}^{*} \cdots \sum_{i_m>i_{m-1}}^{N}{}^{*} \sum_{j_m>i_m}^{N}{}^{*} \sigma_{i_1 j_1} \sigma_{i_2 j_2} \cdots \sigma_{i_m j_m} \tag{3.1}$$

$$\begin{aligned} E_{m,n}^N \equiv \sum_{i=1}^{N} \sum_{j_1=1}^{N}{}^{*} \sum_{j_2>j_1}^{N}{}^{*} \cdots \sum_{j_n>j_{n-1}}^{N}{}^{*} \sum_{k_1=1}^{N}{}^{*} \sum_{l_1>k_1}^{N}{}^{*} \sum_{k_2>k_1}^{N}{}^{*} \sum_{l_2>k_2}^{N}{}^{*} \cdots \sum_{k_m>k_{m-1}}^{N}{}^{*} \sum_{l_m>k_m}^{N}{}^{*} \\ \times \sigma_{ij_1} \sigma_{ij_2} \cdots \sigma_{ij_n} \sigma_{k_1 l_1} \sigma_{k_2 l_2} \cdots \sigma_{k_m l_m} \end{aligned} \tag{3.2}$$

where

$$\sigma_{ij} \equiv \sin^2(\omega_{ij}/2) \tag{3.3}$$

and the * indicates a restriction on the summations so that all non-repeated indices in each term take different values. A particular case is the definition in (2.1) when $m = 0$ should be specified as $D_0^N \equiv 1$.

The general definitions of (3.1) and (3.2) seem to be somewhat complicated. So it might be helpful to look at some specific examples as illustration:

$$\begin{aligned} D_2^5 \equiv \sigma_{12}\sigma_{34} + \sigma_{12}\sigma_{35} + \sigma_{12}\sigma_{45} + \sigma_{13}\sigma_{24} + \sigma_{13}\sigma_{25} + \sigma_{13}\sigma_{45} + \sigma_{14}\sigma_{23} \\ + \sigma_{14}\sigma_{25} + \sigma_{14}\sigma_{35} + \sigma_{15}\sigma_{23} + \sigma_{15}\sigma_{24} + \sigma_{15}\sigma_{34} + \sigma_{23}\sigma_{45} + \sigma_{24}\sigma_{35} + \sigma_{25}\sigma_{34} \end{aligned}$$

$$E^5_{1,2} \equiv \sigma_{12}\sigma_{13}\sigma_{45} + \sigma_{12}\sigma_{14}\sigma_{35} + \sigma_{12}\sigma_{15}\sigma_{34} + \sigma_{13}\sigma_{14}\sigma_{25} + \sigma_{13}\sigma_{15}\sigma_{24} + \sigma_{14}\sigma_{15}\sigma_{23}$$
$$+ \sigma_{21}\sigma_{23}\sigma_{45} + \sigma_{21}\sigma_{24}\sigma_{35} + \sigma_{21}\sigma_{25}\sigma_{34} + \sigma_{23}\sigma_{24}\sigma_{15}$$
$$+ \sigma_{23}\sigma_{25}\sigma_{14} + \sigma_{24}\sigma_{25}\sigma_{13} + \dots.$$

There are 30 terms in the last expression.

3.1. *General expression for K_N*

The normalisation factor K_N appears in the following expression for the distribution function:

$$Q_N(\theta, \phi) \equiv K_N \prod_{i=1}^{N} q_{in}(\theta, \phi)$$
$$= K_N \prod_{i=1}^{N} [\cos^2(\theta/2) - \sigma_{in} \cos\theta + (f^*_{in} e^{i\phi} + f_{in} e^{-i\phi}) \sin(\theta/2)\cos(\theta/2)] \tag{3.4}$$

where f_{in} is defined in (2.18), which implies the very useful relation

$$f^*_{in} f_{jn} + f_{in} f^*_{jn} = \sigma_{in} + \sigma_{jn} - \sigma_{ij} - 2\sigma_{in}\sigma_{jn} \tag{3.5}$$

where n can be arbitrary. K_N is then to be determined by the following normalisation condition:

$$\frac{N+1}{4\pi} \int_0^{\pi} d\theta \sin\theta \int_0^{2\pi} d\phi\, Q_N(\theta, \phi) = 1. \tag{3.6}$$

Using (3.4) and (3.5) in (3.6), and some thorough investigation, we can convince ourselves that the inverse of K_N can be expressed as a finite series expansion in terms of the symmetric functions defined by (3.1) and (3.2) in the form

$$K_N^{-1} \equiv \sum_{m=0}^{[N/2]} d^N_m D^N_m + \sum_{m=0}^{[N/2]} \sum_{n=2}^{N-2m-1} e^N_{m,n} E^N_{m,n} \tag{3.7}$$

where

$$[N/2] \equiv \begin{cases} N/2 & \text{if } N \text{ is even} \\ (N-1)/2 & \text{if } N \text{ is odd} \end{cases} \tag{3.8}$$

and the expansion coefficients can be calculated by the following formulae:

$$d^N_m = \frac{(-1)^m (N+1)}{m!} \int_0^{\pi} d\theta [\sin(\theta/2)]^{2m+1} [\cos(\theta/2)]^{2(N-m)+1} = (-1)^m \frac{(N-m)!}{N!} \tag{3.9}$$

$$e^N_{m,n} = \sum_{i=0}^{[n/2]} \sum_{j=0}^{M} (-1)^{m+n+i+j} \frac{(N-2m-n-1)!\,n!}{(N-2m-n-j-1)!\,(m+i+j)!\,(n-2i-j)!\,i!\,j!}$$
$$\times \int_0^{\pi} d\theta [\sin(\theta/2)]^{2(m+i+j)+1} [\cos(\theta/2)]^{2(N-m-n+i)+1} (\cos\theta)^{n-2i-j}$$
$$= 0 \tag{3.10}$$

where

$$M \equiv \min\{N-2m-n-1,\, n-2i\}. \tag{3.11}$$

The result in (3.10) will be derived in appendix 1.

Using (3.9) and (3.10) in (3.7) we obtain

$$K_N^{-1} = \sum_{m=0}^{[N/2]} (-1)^m \frac{(N-m)!}{N!} D_m^N . \tag{3.12}$$

3.2. General expression for S_N

The major difficulty in using (2.6) to calculate the Wehrl entropy is caused by the logarithmic function. Our strategy to overcome this difficulty is the following; we use (3.4) in (2.6) and break down the whole integral into the summation

$$S_N = -\ln K_N - \sum_{i=1}^{N} I_i \tag{3.13}$$

where

$$I_i \equiv K_N \frac{N+1}{4\pi} \int_0^{\pi} \mathrm{d}\theta \sin\theta \int_0^{2\pi} \mathrm{d}\phi \ln q_{in}(\theta, \phi) \prod_{j=1}^{N} q_{jn}(\theta, \phi). \tag{3.14}$$

Each of the integrals I_i is still invariant, by itself, under any rotation of the coordinate system. So we have the freedom to adopt a suitable rotation to simplify the calculation of each of the I_i individually. For a particular I_i we choose the rotation $T(\theta_i, \phi_i)$ as defined by (2.14); then we can have

$$\ln q_{in}(\theta, \phi) = \ln q_{ii}(\theta, \phi) = \ln \cos^2(\theta/2). \tag{3.15}$$

Using this strategy and some thorough investigation, we find that the general expression for the Wehrl entropy can be written as

$$S_N = K_N \left(\sum_{m=0}^{[N/2]} \delta_m^N D_m^N + \sum_{m=0}^{[N/2]} \sum_{n=2}^{N-2m-1} \varepsilon_{m,n}^N E_{m,n}^M \right) - \ln K_N \tag{3.16}$$

where

$$\begin{aligned} \delta_m^N &= (-1)^m \frac{(N+1)}{m!} \int_0^{\pi} \mathrm{d}\theta \{(N-2m)(2m-1)[\sin(\theta/2)]^{2m+1}[\cos(\theta/2)]^{2(N-m)+1} \\ &\quad - 2m^2 \cos\theta [\sin(\theta/2)]^{2m-1}[\cos(\theta/2)]^{2(N-m)+1}\} \ln \cos^2(\theta/2) \\ &= (-1)^m \frac{(N-m)!}{N!} \left(\sum_{l=N-m+1}^{N+1} \frac{N}{l} - 2m \right) \end{aligned} \tag{3.17}$$

$$\begin{aligned} \varepsilon_{m,n}^N &= \sum_{i=0}^{[n/2]} \sum_{j=0}^{M} (-1)^{m+n+i+j} \frac{(N-2m-n-1)!\,n!}{(N-2m-n-j-1)!(m+i+j)!(n-2i-j)!\,i!\,j!} \\ &\quad \times \int_0^{\pi} \mathrm{d}\theta [\sin(\theta/2)]^{2(m+i+j)+1} \\ &\quad \times [\cos(\theta/2)]^{2(N-m-n+i)+1} (\cos\theta)^{n-2i-j} \ln \cos^2(\theta/2) \\ &= (-1)^{m+1} \frac{(N-m-n)!(n-2)!}{N!}. \end{aligned} \tag{3.18}$$

The last result will be derived in appendix 2.

Substitution of (3.17) and (3.18) into (3.16) yields

$$S_N = K_N\left[\sum_{m=1}^{[N/2]}(-1)^m \frac{(N-m)!}{N!}\left(\sum_{l=N-m+1}^{N}\frac{N}{l}-2m\right)D_m^N - \sum_{m=0}^{[N/2]}\sum_{n=2}^{N-2m-1}(-1)^m\frac{(N-m-n)!(n-2)!}{N!}E_{m,n}^N\right]+\frac{N}{N+1}-\ln K_N. \tag{3.19}$$

3.3. Some examples

Normalisation factors:

$$K_2^{-1} = 1-\tfrac{1}{2}\sigma_{12}$$
$$K_3^{-1} = 1-\tfrac{1}{3}(\sigma_{12}+\sigma_{13}+\sigma_{23})$$
$$K_4^{-1} = 1-\tfrac{1}{4}(\sigma_{12}+\sigma_{13}+\sigma_{14}+\sigma_{23}+\sigma_{24}+\sigma_{34})$$
$$K_5^{-1} = 1-\tfrac{1}{5}(\sigma_{12}+\ldots)+\tfrac{1}{20}(\sigma_{12}\sigma_{34}+\ldots).$$

Wehrl entropy:

$$S_2 = K_2(\tfrac{2}{3}+\tfrac{1}{6}\sigma_{12})-\ln K_2$$
$$S_3 = K_3[\tfrac{3}{4}+\tfrac{1}{12}(\sigma_{12}+\sigma_{13}+\sigma_{23})-\tfrac{1}{6}(\sigma_{12}\sigma_{13}+\sigma_{12}\sigma_{23}+\sigma_{13}\sigma_{23})]-\ln K_3$$
$$S_4 = K_4[\tfrac{4}{5}+\tfrac{1}{20}(\sigma_{12}+\sigma_{13}+\sigma_{14}+\sigma_{23}+\sigma_{24}+\sigma_{34})-\tfrac{13}{180}(\sigma_{12}\sigma_{34}+\sigma_{13}\sigma_{24}+\sigma_{14}\sigma_{23}) -\tfrac{1}{12}(\sigma_{12}\sigma_{13}+\ldots)-\tfrac{1}{24}(\sigma_{12}\sigma_{13}\sigma_{14}+\ldots)]-\ln K_4$$
$$S_5 = K_5[\tfrac{5}{6}+\tfrac{1}{30}(\sigma_{12}+\ldots)-\tfrac{11}{240}(\sigma_{12}\sigma_{34}+\ldots)-\tfrac{1}{20}(\sigma_{12}\sigma_{13}+\ldots) +\tfrac{1}{60}(\sigma_{12}\sigma_{13}\sigma_{14}+\ldots)-\tfrac{1}{60}(\sigma_{12}\sigma_{13}\sigma_{14}\sigma_{15}+\ldots)]-\ln K_5.$$

4. Coherent spin states

From the definition given in (2.3) we can easily obtain the probability density function in the Q representation for a certain coherent spin state of an N-spin system $|\theta_0, \phi_0\rangle_N$ as

$$\begin{aligned} Q_N(\theta,\phi;\theta_0,\phi_0) &\equiv \langle\theta,\phi|\hat{\rho}(\theta_0,\phi_0)|\theta,\phi\rangle_N = |\langle\theta,\phi|\theta_0,\phi_0\rangle_N|^2 \\ &= |\cos(\theta/2)\cos(\theta_0/2)+\mathrm{e}^{\mathrm{i}(\phi-\phi_0)}\sin(\theta/2)\sin(\theta_0/2)|^{2N} \\ &= [\cos(\omega_0/2)]^{2N} \end{aligned} \tag{4.1}$$

where ω_0 is the angle between (θ, ϕ) and (θ_0, ϕ_0).

Equation (4.1) implies that the N points on the unit sphere in the geometric representation of a spin state reduce to a single point for the particular case of a coherent spin state; this in turn implies that

$$\sigma_{ij} = 0 \tag{4.2}$$

for any pair of unit vectors. Using (4.2) in (3.1) and (3.2), we have

$$D_m^N = \delta_{m,0} \tag{4.3}$$

$$E_{m,n}^N = 0. \tag{4.4}$$

Now we can use (4.3) and (4.4) in (3.12) and (3.19) to obtain

$$K_N = 1 \qquad S_N = N/(N+1). \tag{4.5}$$

5. Almost coherent spin states

If the N points on the unit sphere in the geometric representation of a spin state are all very close to one another, it is considered as an almost coherent spin state. Let

$$\theta_0 \equiv \frac{1}{N}\sum_{i=1}^{N}\theta_i \qquad \phi_0 \equiv \frac{1}{N}\sum_{i=1}^{N}\phi_i \tag{5.1a}$$

$$\Delta\theta_i \equiv \theta_i - \theta_0 \qquad \Delta\phi_i \equiv \phi_i - \phi_0 \tag{5.1b}$$

then, for an almost coherent state, we have

$$\Delta\theta_i \ll 1 \qquad \Delta\phi_i \ll 1 \tag{5.2}$$

for $i = 1, 2, \ldots, N$.

The task of this section is to derive the Taylor series expansion for the Wehrl entropy of an almost coherent spin state in terms of the $\Delta\theta_i$ and the $\Delta\phi_i$. We will see later that the first-, second- and third-order terms in the expansion all vanish identically; hence, the series expansion will be carried to the fourth order. As will be seen later, to reach the fourth-order terms for S_N, we need only to use the leading second-order terms of σ_{ij}.

The Taylor series expansion for $\cos\omega_{ij}$ is

$$\begin{aligned}\cos\omega_{ij} &\equiv \cos\theta_i\cos\theta_j + \sin\theta_i\sin\theta_j\cos(\phi_i-\phi_j)\\ &\approx 1-\tfrac{1}{2}(\Delta\theta_i-\Delta\theta_j)^2-\tfrac{1}{2}\sin^2\theta_0(\Delta\phi_i-\Delta\phi_j)^2.\end{aligned} \tag{5.3}$$

Thus the Taylor series for σ_{ij} is

$$\sigma_{ij} \approx \tfrac{1}{4}(\Delta\theta_i-\Delta\theta_j)^2+\tfrac{1}{4}\sin^2\theta_0(\Delta\phi_i-\Delta\phi_j)^2. \tag{5.4}$$

From (3.12) and (3.19), we have

$$K_N^{-1} \approx 1-\frac{1}{N}D_1^N+\frac{1}{N(N-1)}D_2^N \tag{5.5a}$$

$$S_N \approx \frac{N}{N+1}+K_N\left(\frac{1}{N}D_1^N-\frac{2N-3}{N(N-1)^2}D_2^N-\frac{1}{N(N-1)}E_{0,2}^N\right)-\ln K_N. \tag{5.5b}$$

The few terms shown in (5.5a) and (5.5b) are all we need in order to obtain the Taylor series up to the fourth order.

Substitution of (5.5a) into (5.5b) gives

$$S_N \approx \frac{N}{N+1}+\frac{1}{2N^2}(D_1^N)^2-\frac{N-2}{N(N-1)^2}D_2^N-\frac{1}{N(N-1)}E_{0,2}^N \tag{5.6}$$

and, by definition, we have

$$D_1^N \equiv \sum_i\sum_{j>i}\sigma_{ij} \tag{5.7a}$$

$$D_2^N \equiv \sum_i\sum_{j>i}\sum_{k>i}{}^{*}\sum_{l>k}{}^{*}\sigma_{ij}\sigma_{kl} \tag{5.7b}$$

$$E_{0,2}^N \equiv \sum_i\sum_j{}^{*}\sum_{k>j}{}^{*}\sigma_{ij}\sigma_{ik} \tag{5.7c}$$

where the * again indicates the restriction that the indices i, j, k and l must all be different in the same term.

Using (5.4) and (5.7) we obtain

$$D_1^N \approx \frac{N}{4}(F_1 + F_3) \tag{5.8a}$$

$$D_2^N \approx \frac{N^2-3N+3}{32}(F_1+F_3)^2 - \frac{N(N-1)}{32}G + \tfrac{1}{8}(F_2^2 - F_1F_3) \tag{5.8b}$$

$$E_{0,2}^N \approx \frac{3N-6}{32}(F_1+F_3)^2 + \frac{N(N-2)}{32}G - \tfrac{1}{4}(F_2^2 - F_1F_3) \tag{5.8c}$$

where we have introduced some new symmetric functions defined as follows:

$$F_1 \equiv \sum_i (\Delta\theta_i)^2 \tag{5.9a}$$

$$F_2 \equiv \sin\theta_0 \sum_i (\Delta\theta_i)(\Delta\phi_i) \tag{5.9b}$$

$$F_3 \equiv \sin^2\theta_0 \sum_i (\Delta\phi_i)^2 \tag{5.9c}$$

$$G \equiv \sum_i [(\Delta\theta_i)^2 + \sin^2\theta_0(\Delta\phi_i)^2]^2. \tag{5.9d}$$

Now, substituting (5.8) into (5.6), we have

$$S_N \approx \frac{N}{N+1} + \frac{1}{32(N-1)^2}(F_1 - F_3)^2 + \frac{1}{8(N-1)^2}(F_2)^2. \tag{5.10}$$

The first term on the right-hand side of (5.10) is the entropy of the coherent spin state. It is obvious that the rest of the expression is positive semidefinite. This proves that the Wehrl entropy of a coherent spin state is a local minimum. This minimum should be very flat because the first non-vanishing higher-order terms are of the fourth order.

6. Conjecture on maximum entropy states

We have seen in this paper that the Wehrl entropy of a spin state with maximum total spin attains a minimum when all the N roots of its probability amplitude function in the Q representation are identical, i.e. the N points on the surface of the unit sphere in the geometric representation all coincide at a single point. Equation (5.10) indicates that, as these points spread out a little, the Wehrl entropy increases. It is natural to speculate that as these points spread further apart from one another the Wehrl entropy will continue to increase until these points are as far as possible from one another on the unit sphere. Therefore, we conjecture that the maximum Wehrl entropy is attained by a pure spin state with maximum total spin when the N points in its geometric representation are located as follows. For $N=2$, the two points are diametrically opposite to each other; for $N=3$, the three points form an equilateral triangle on a large circle; and for $N=4$, 6, 8, 12, 20, etc, the N points form regular polyhedra.

7. Summary

In this paper we have considered the probability distribution over the surface of a unit sphere for an arbitrary pure spin state with maximum total spin of an N-spin system in terms of the Q representation of coherent spin states or atomic coherent states. We have used a geometric representation of the distribution function as N points on the surface of the unit sphere or N unit vectors. We have introduced two kinds of symmetric functions D_m^N and $E_{m,n}^N$ of $\sin^2(\omega_{ij}/2)$, where ω_{ij} is the angle between a pair of unit vectors. We have obtained a general expression for the Wehrl entropy of an arbitrary maximum total spin state as a finite series expansion in terms of D_m^N and $E_{m,n}^N$. We have shown that the N unit vectors characterising a spin state all join together in the special case of a coherent spin state. Based on this fact we have obtained a Taylor series expansion for spin states in the neighbourhood of a coherent spin state. This series expansion establishes that the Wehrl entropy of a coherent spin state is a local minimum. Therefore, Lieb's conjecture is partially confirmed.

Acknowledgment

This work was supported by the US Navy Office of Naval Research under grant no N00014-82-K-0521.

Appendix 1. Evaluation of $e_{m,n}^N$

We first evaluate the integral in the formula for $e_{m,n}^N$ given in (3.10) as follows:

$$\int_0^\pi d\theta[\sin(\theta/2)]^{2(m+i+j)+1}[\cos(\theta/2)]^{2(N-m-n+i)+1}(\cos\theta)^{n-2i-j}$$

$$= \sum_{k=0}^{n-2i-j}(-1)^k\binom{n-2i-j}{k}B(m+i+j+k+1,\, N-m-i-j-k+1)$$

$$= \frac{(n-2i-j)!(m+i+j)!(N-m-n+i)!}{(N+1)!}$$

$$\times\frac{1}{2\pi i}\oint\frac{dz}{z^{n-2i-j+1}}(1+z)^{-(m+i+j+1)}(1-z)^{-(N-m-n+i+1)} \tag{A1.1}$$

where the contour integral should be taken around the origin of the complex plane. Substituting (A1.1) into (3.10) we obtain

$$e_{m,n}^N = \sum_{i=0}^{[n/2]}(-1)^{m+n+i}\frac{(N-m-n+i)!\,n!}{(N+1)!\,i!}$$

$$\times\frac{1}{2\pi i}\oint\frac{dz}{z^{n-2i+1}}(1+z)^{-(m+i+1)}(1-z)^{-(N-m-n+i+1)}$$

$$\times\sum_{j=0}^{N-2m-n-1}(-1)^j\binom{N-2m-n-1}{j}\left(\frac{z}{1+z}\right)^j \tag{A1.2}$$

where we have fixed the upper limit of j to be $N-2m-n-1$ because, in case the latter is greater than $n-2i$, those terms corresponding to $j>n-2i$ cannot survive the contour integral. The summation over j in (A1.2) can be easily carried out to yield

$$\begin{aligned} e^N_{m,n} &= (-1)^{m+n}\frac{(N-m-n)!\,n!}{(N+1)!}\frac{1}{2\pi i}\oint\frac{dz}{z^{n+1}}(1+z)(1-z^2)^{-(N-m-n+1)} \\ &\quad\times\sum_{i=0}^{\infty}(-1)^i\binom{N-m-n+i}{i}\left(\frac{z^2}{1-z^2}\right)^i \\ &= (-1)^{m+n}\frac{(N-m-n)!\,n!}{(N+1)!}\frac{1}{2\pi i}\oint\frac{dz}{z^{n+1}}(1+z)=0 \end{aligned} \tag{A1.3}$$

where we let i run up to ∞ because those terms corresponding to $i>[n/2]$ cannot survive the contour integral and the last result was obtained since $n\geqslant 2$.

Appendix 2. Evaluation of $\varepsilon^N_{m,n}$

We first evalute the integral in the formula for $\varepsilon^N_{m,n}$ given in (3.18) as follows:

$$\begin{aligned} &\int_0^\pi d\theta[\sin(\theta/2)]^{2(m+i+j)+1}[\cos(\theta/2)]^{2(N-m-n+i)+1}(\cos\theta)^{n-2i-j}\ln\cos^2(\theta/2) \\ &\quad= \sum_{k=0}^{n-2i-j}(-1)^k\binom{n-2i-j}{k}B(m+i+j+k+1,N-m-i-j-k+1) \\ &\quad\quad\times\int_0^1 dt\frac{t^{N-m-i-j-k}-t^{N+1}}{1-t} \\ &\quad= \frac{(n-2i-j)!(m+i+j)!(N-m-n+i)!}{(N+1)!} \\ &\quad\quad\times\int_0^1\frac{dt}{t(1-t)}\frac{1}{2\pi i}\oint\frac{dz}{z^{n-2i-j+1}}(1+z)^{-(m+i+j+1)}\left(\frac{t}{1-zt}\right)^{N-m-n+i+1} \end{aligned} \tag{A2.1}$$

where we have dropped the term $\int_0^1 dt\, t^{N+1}/(1-t)$ because its contribution to $\varepsilon^N_{m,n}$ has the factor $e^N_{m,n}=0$, as shown in appendix 1. Substituting (A2.1) into (3.18) and carrying out the summation over j, we obtain

$$\begin{aligned} \varepsilon^N_{m,n} &= (-1)^{m+n}\frac{(N-m-n)!\,n!}{(N+1)}\int_0^1\frac{dt}{t(1-t)}\frac{1}{2\pi i}\oint\frac{dz}{z^{n+1}}(1+z)\left(\frac{t}{(1+z)(1-zt)}\right)^{N-m-n+1} \\ &\quad\times\sum_{i=0}^{\infty}(-1)^i\binom{N-m-n+i}{i}\left(\frac{z^2t}{(1+z)(1-zt)}\right)^i \\ &= (-1)^{m+n}\frac{(N-m-n)!\,n!}{(N+1)!}\int_0^1\frac{dt}{t(1-t)}\frac{1}{2\pi i}\oint\frac{dz}{z^{n+1}}(1+z) \\ &\quad\times\left(\frac{t}{1+z(1-t)}\right)^{N-m-n+1} \\ &= (-1)^m\frac{(N-m-1)!}{(N+1)!}\int_0^1 dt\, t^{N-m-n}[(N-m)(1-t)^{n-1}-n(1-t)^{n-2}] \\ &= (-1)^{m+1}\frac{(N-m-n)!(n-2)!}{N!}. \end{aligned} \tag{A2.2}$$

References

Arecchi F T, Courtens E, Gilmore R and Thomas H 1972 *Phys. Rev.* A **6** 2211
Davies E B 1976 *The Quantum Theory of Open Systems* (New York: Academic)
Deutsch D 1983 *Phys. Rev. Lett.* **50** 631
Dicke R H 1954 *Phys. Rev.* **93** 99
Glauber R J 1963 *Phys. Rev. Lett.* **10** 84
Haken H 1970 *Handbuch der Physik* **XXV/2c** (Berlin: Springer) p 62
Husimi K 1940 *Proc. Phys. Math. Soc. Japan* **22** 264
Kano Y 1965 *J. Math. Phys.* **6** 1913
Lieb E H 1978 *Commun. Math. Phys.* **62** 35
Milburn G J 1985 *Phys. Rev.* A **33** 674
Radcliffe J M 1971 *J. Phys. A: Gen. Phys.* **4** 313
Schrödinger E 1926 *Naturw.* **14** 664
Sudarshan E C G 1963 *Phys. Rev. Lett.* **10** 210
Takahashi K and Saitô N 1987 *Phys. Rev. Lett.* **55** 645
Toda M and Ikeda K 1987 *Phys. Lett.* **124A** 165
Uffink J B M and Hilgevoord J 1985 *Found. Phys.* **15** 925
Wehrl A 1979 *Rep. Math. Phys.* **16** 353
Wigner E P 1932 *Phys. Rev.* **40** 749