THE ROYAL SOCIETY

# Geometrization of statistical mechanics

By Dorje C. Brody[1] and Lane P. Hughston[2]

[1]*Department of Applied Mathematics and Theoretical Physics,*
*Silver Street, Cambridge CB3 9EW, UK*
*and Churchill College, Cambridge CB3 0DS, UK*
[2]*Department of Mathematics, King's College London,*
*Strand, London WC2R 2LS, UK*

Classical and quantum statistical mechanics are cast here in the language of projective geometry to provide a unified geometrical framework for statistical physics. After reviewing the Hilbert-space formulation of classical statistical thermodynamics, we show that the specification of a canonical polarity on the real projective space $RP^n$ induces a Riemannian metric on the state space of statistical mechanics. In the case of the canonical ensemble, equilibrium thermal states are determined by a Hamiltonian gradient flow with respect to this metric. This flow is characterized by the property that it induces a projective automorphism on the state manifold. The measurement problem for thermal systems is studied by the introduction of the concept of a random state. The general methodology is extended to establish a new framework for the quantum-mechanical dynamics of equilibrium thermal states. In this case, the relevant phase space is the complex projective space $CP^n$, here regarded as a real manifold $\Gamma$ endowed with the Fubini–Study metric and a compatible symplectic structure. A distinguishing feature of quantum thermal dynamics is the inherent multiplicity of thermal trajectories in the state space, associated with the non-uniqueness of the infinite-temperature state. We are then led to formulate a geometric characterization of the standard KMS relation often considered in the context of $C^*$ algebras. Finally, we develop a theory of the quantum microcanonical and canonical ensembles, based on the geometry of the quantum phase space $\Gamma$. The example of a quantum spin-1/2 particle in a heat bath is studied in detail.

**Keywords: Hilbert-space geometry; projective geometry;**
**equilibrium statistical mechanics; quantum dynamics**

## 1. Introduction

One of the most fascinating advances in the application of differential geometry is its use in statistical physics, including quantum mechanics and statistical mechanics. The purpose of this paper is to develop a unified framework that allows for a geometric characterization of both of these aspects of statistical physics.

In quantum mechanics, one typically works with square-integrable wave functions, i.e. elements of a complex Hilbert space $\mathcal{H}$. This space possesses natural geometrical structures induced by its norm. However, in order to seek a compelling axiomatic formulation of quantum mechanics, it is reasonable to consider a space of more direct physical relevance (Haag & Kastler 1964; Landsman 1998). This is not the Hilbert space $\mathcal{H}$ itself, but rather the manifold $\Gamma$ of 'instantaneous pure states' (Kibble

1979), which has the structure of a complex projective space $\mathrm{CP}^n$, possibly infinite dimensional, enriched with a Hermitian correspondence, i.e. a complex conjugation operation that maps points to hyperplanes in $\mathrm{CP}^n$, and vice versa. Equivalently, we can think of $\mathrm{CP}^n$ as being endowed with a Riemannian structure given by the Fubini–Study metric.

The space $\varGamma$ is the quantum analogue of the classical phase space of mechanical systems. Hence, one can interpret the Schrödinger equation as Hamilton's equations on $\mathrm{CP}^n$, and the equation of motion for a general density matrix can be identified with the Liouville equation (Gibbons 1992). An advantage of working with the manifold $\varGamma$, rather than the Hilbert space of state vectors, is that it can readily accommodate generalizations of quantum mechanics (Hughston 1995). Furthermore, the structure of $\varGamma$ allows for the possibility of a natural probabilistic interpretation, even if the standard linear quantum theory is modified.

As we discuss elsewhere (Brody & Hughston 1996, 1997, 1998$a, c$), the statistical aspects of quantum measurement can be clarified if we shift our view slightly, and regard the Hilbert space $\mathcal{H}$ of quantum mechanics not as a complex Hilbert space, but rather a real Hilbert space endowed with a real metric and a compatible complex structure (cf. Gibbons & Pohle 1993). Once quantum theory is formulated this way, its relation to other aspects of statistical physics becomes more apparent.

Statistical mechanics, in particular, can also be formulated concisely (Brody & Hughston 1998$b$) in terms of the geometry of a real Hilbert space $\mathcal{H}$. This can be seen by taking the square-root of the Gibbs density function, which maps the space of probability distributions to vectors in a convex cone $\mathcal{H}_+$ in $\mathcal{H}$. In this way, the various probabilistic and statistical operations of statistical mechanics can be given a transparent geometric meaning in $\mathcal{H}$ (cf. Gibbons 1997). However, it can be argued that even at the classical level of statistical mechanics the 'true' state space is obtained by identifying pure states along a given ray through the origin of $\mathcal{H}$. In this case, the space obtained is the real projective space $\mathrm{RP}^n$. This is the view taken here, and we shall study properties of thermal states that become apparent when the theory is developed in a fully geometric context.

The present paper is organized as follows. In §2 we review basics of the Hilbert-space formulation of statistical mechanics. This section paves the way to the approach in terms of projective geometry presented later. We begin with a brief review of statistical geometry, including the theory of the Fisher–Rao metric on the parameter space of a family of probability distributions. The Gibbs distribution when viewed in this way can be seen as a curve in Hilbert space, parametrized by inverse temperature, and there is a striking formal resemblance to the Schrödinger equation, even though we work at a classical level.

A measurement theory for thermal states is developed by analogy with the theory of density matrices in quantum mechanics. We are then led to a set of uncertainty relations for the measurement of thermodynamic conjugate variables such as energy and inverse temperature. We also introduce an alternative approach to measurement theory that is not based on a density matrix description. Our approach, based on the introduction of random states, extends naturally also to quantum mechanics, where it can be seen to be more appealing in a probabilistic context than standard treatments, but reduces to the conventional density matrix approach in special cases.

In §3 we introduce a projective geometric framework for the probabilistic representation of the canonical thermal state associated with the standard Gibbs measure.

Thermal states are shown to lie on a trajectory in $\mathrm{RP}^n$, which is endowed with the natural 'spherical' metric. In this connection, we find it convenient to develop some useful differential geometric results characterizing projective transformations on the state space $\mathrm{RP}^n$. We find that the equilibrium thermal trajectories, which are shown to be given by a Hamiltonian gradient flow, generate projective automorphisms of the state manifold.

In §4 we synthesize the approaches outlined in §§2 and 3, and consider the relationship of the classical thermal state space, $\mathrm{RP}^n$, and the quantum phase space, $\mathrm{CP}^n$, in a study of the quantum-mechanical dynamics of equilibrium thermal states. First we examine the quantum state space from the viewpoint of complex algebraic geometry, which shows that this space is endowed with a Riemannian geometry given by the Fubini–Study metric, along with a natural symplectic structure. For thermal physics it is instructive to look at quantum mechanics from an entirely 'real' point of view as well, as developed in §4 *b*.

Our formulation is then compared to the standard KMS construction (Kubo 1957; Martin & Schwinger 1959; Haag *et al*. 1967) for equilibrium states. In particular, once we pass to the mixed-state description we recover the KMS state. However, our quantum-mechanical pure thermal state, which does not obey the KMS condition, can be viewed as a more fundamental construction.

In §4 *d* we develop a theory of the quantum microcanonical and canonical ensembles, formulated in terms of the geometry of the quantum phase space $\Gamma$. This is set up in such a way as to admit the possibility of generalizations to nonlinear quantum theories. Finally, we study more explicitly the case of a quantum spin-1/2 particle in a heat bath.

## 2. Statistical states in Hilbert space

### (*a*) *Hilbert-space geometry*

Let us begin by demonstrating how classical statistical mechanics can be formulated in an appealing way by the use of a Hilbert-space formalism. Consider a real Hilbert space $\mathcal{H}$ equipped with an inner product $g_{ab}$. We have in mind in particular the case $\mathcal{H} = L^2(\mathcal{R})$ where $\mathcal{R} = \mathbb{R}^{2m}$ is the phase space of a classical dynamical system, though much of what we discuss here can be generalized to incorporate discrete phase spaces, such as the configuration spaces of lattice spin systems. A probability density function $p(x)$ on $\mathcal{R}$ can be mapped into $\mathcal{H}$ by taking the square root $\psi(x) = (p(x))^{1/2}$, denoted by a vector $\psi^a$ in $\mathcal{H}$. The normalization condition

$$\int (\psi(x))^2 \, \mathrm{d}x = 1 \tag{2.1}$$

is written $g_{ab}\psi^a\psi^b = 1$, indicating that $\psi^a$ lies on the unit sphere $\mathcal{S}$ in $\mathcal{H}$. Since $p(x)$ is non-negative, the image of the map $f : p(x) \to \psi(x)$ is the intersection $\mathcal{S}_+ = \mathcal{S} \cap \mathcal{H}_+$ of $\mathcal{S}$ with the convex cone $\mathcal{H}_+$ formed by the totality of quadratically integrable non-negative functions. If we consider the space of all probability distributions as a metric space relative to the Hellinger distance (Xia 1972), then $f$ is an isometric embedding in $\mathcal{H}$. We call $\psi^a$ the state vector of the corresponding probability density $p(x)$. Our calculations here are facilitated by the introduction of a standard 'singular'

basis $u^a(x)$ characterized by the relations

$$g_{ab}u^a(x)u^b(y) = \delta(x,y) \quad \text{and} \quad \int u^a(x)u^b(x)\,\mathrm{d}x = g^{ab},$$

where $\mathrm{d}x$ is the volume element on $\mathcal{R}$. Then, for any function $\psi(x) \in L^2(\mathcal{R})$, we can write

$$\psi^a = \int \psi(x)u^a(x)\,\mathrm{d}x \tag{2.2}$$

for the corresponding abstract vector, and $\psi(x) = g_{ab}\psi^a u^b(x)$ for the reverse operation.

A typical random variable is represented on $\mathcal{H}$ by a symmetric tensor $X_{ab}$, whose expectation in a normalized state $\psi^a$ is

$$E_\psi[X] = X_{ab}\psi^a\psi^b. \tag{2.3}$$

Similarly, the expectation of its square is $X_{ac}X^c_b\psi^a\psi^b$. The variance of $X_{ab}$ in the state $\psi^a$ is, therefore,

$$\mathrm{Var}_\psi[X] = \tilde{X}_{ac}\tilde{X}^c_b\psi^a\psi^b, \tag{2.4}$$

where $\tilde{X}_{ab} = X_{ab} - g_{ab}E_\psi[X]$ represents the deviation of $X_{ab}$ from its mean in the state $\psi^a$. If $X(x)$ denotes the given random variable as a function on $\mathcal{R}$, with density function $p(x)$, then in terms of the basis $u^a(x)$ we can write

$$X_{ab} = \int X(x)u_a(x)u_b(x)\,\mathrm{d}x. \tag{2.5}$$

Then it is not difficult to see that the formulae above reduce to the familiar relations

$$E_p[X] = \int X(x)p(x)\,\mathrm{d}x \quad \text{and} \quad \mathrm{Var}_p[X] = \int (X(x))^2 p(x)\,\mathrm{d}x - (E_p[X])^2.$$

We note that the *general* random variable on $\mathcal{H}$ corresponds to a symmetric two-point function $X(x,y)$ in $L^2(\mathcal{R})\otimes L^2(\mathcal{R})$. However, in this section we consider *classical* (i.e. non-quantum) random variables, which correspond to simultaneously diagonalizable functions of the form $X(x,y) = X(x)\delta(x,y)$. The metric then corresponds to the delta function, which is consistent with the general relations

$$X_{ab}u^a(x)u^b(y) = X(x,y) \quad \text{and} \quad \iint X(x,y)u_a(x)u_b(y)\,\mathrm{d}x\mathrm{d}y = X_{ab}.$$

We consider now the unit sphere $\mathcal{S}$ in $\mathcal{H}$, and within this sphere a submanifold $\mathcal{M}$ given parametrically by $\psi^a(\theta)$, where $\theta^i$ $(i = 1, \ldots, r)$ are local coordinates. In particular, later on we have in mind the case where the parameter space spanned by $\theta^i$ represents the space of coupling constants associated with a given physical system in statistical mechanics. In the case of the Gibbs measure, there is a single such parameter, corresponding to the inverse temperature variable $\beta = 1/k_\mathrm{B}T$. We write $\partial_i$ for $\partial/\partial\theta^i$. Then there is a natural Riemannian metric $\mathcal{G}_{ij}$ on the parameter space $\mathcal{M}$, induced by $g_{ab}$, given in local coordinates by $\mathcal{G}_{ij} = g_{ab}\partial_i\psi^a\partial_j\psi^b$.

This can be seen by noting that the squared distance between the endpoints of two vectors $\psi^a$ and $\eta^a$ in $\mathcal{H}$ is $g_{ab}(\psi^a - \eta^a)(\psi^b - \eta^b)$. If both endpoints lie on $\mathcal{M}$, and $\eta^a$ is obtained by infinitesimally displacing $\psi^a$ in $\mathcal{M}$, i.e. $\eta^a = \psi^a + \partial_i\psi^a\,\mathrm{d}\theta^i$, then the separation $\mathrm{d}s$ between the two endpoints on $\mathcal{M}$ is given by $\mathrm{d}s^2 = \mathcal{G}_{ij}\,\mathrm{d}\theta^i\mathrm{d}\theta^j$,

where $\mathcal{G}_{ij}$ is given as above. The metric $\mathcal{G}_{ij}$ is, up to a conventional factor of four, the so-called Fisher–Rao metric on the space of the given family of distributions. The Fisher–Rao metric is usually defined in terms of a complicated expression involving the covariance matrix of the gradient of the log-likelihood function; but here we have a transparent geometrical construction. The Fisher–Rao metric is important since it provides a geometrical basis for the key links between the statistical and physical aspects of the systems under consideration.

### (*b*) *Thermal trajectories*

Now suppose we consider the canonical ensemble of classical statistical mechanics, in the case for which the system is characterized by a phase space and an assignment of an energy value for each point in the phase space. The parametrized family of probability distributions then takes the form of the Gibbs measure

$$p_\beta(x) = q(x)\exp[-\beta H(x) - W(\beta)]. \tag{2.6}$$

Here, the variable $x$ ranges over the phase space, $H(x)$ represents the energy, $W(\beta)$ is a normalization factor, and $q(x)$ determines the distribution at $\beta = 0$, where $\beta$ is the inverse temperature.

We now formulate a Hilbert-space characterization of this distribution. Taking the square root of $p_\beta(x)$, we find that the state vector $\psi^a(\beta)$ in $\mathcal{H}$, corresponding to the Gibbs measure (2.6), satisfies the differential equation

$$\frac{\partial \psi^a}{\partial \beta} = -\tfrac{1}{2}\tilde{H}^a_b \psi^b, \tag{2.7}$$

where $\tilde{H}_{ab} = H_{ab} - g_{ab}E_\psi[H]$. Here, the Hilbert-space operator $H_{ab}$ corresponds to the specified Hamiltonian function $H(x)$ in (2.6). The solution of this equation can be represented as follows:

$$\psi^a(\beta) = \exp[-\tfrac{1}{2}(\beta H^a_b + \tilde{W}(\beta)\delta^a_b)]q^b, \tag{2.8}$$

where $\tilde{W}(\beta) = W(\beta) - W(0)$ and $q^a = \psi^a(0)$ is the prescribed distribution at $\beta = 0$. Since equation (2.7) respects the normalization $g_{ab}\psi^a\psi^b = 1$, it follows that for each value of the temperature, equation (2.8) determines a point on $\mathcal{M}$ in $\mathcal{S}_+$.

Now let $u^a(x)$ denote a singular basis in $\mathcal{H}$. Here, the variable $x$ labels the points in the phase space of the system, each of which has a definite energy $H(x)$. In other words, for each point $x$ in phase space we have a corresponding basis vector $u^a(x)$ in $\mathcal{H}$. With this choice of basis, a classical thermal state $\psi^a(\beta)$ can be expressed as a superposition

$$\psi^a(\beta) = \mathrm{e}^{-(1/2)W(\beta)}\int_x \mathrm{e}^{-(1/2)\beta H(x)}u^a(x)\,\mathrm{d}x, \tag{2.9}$$

where

$$\exp[W(\beta)] = \int_x \exp(-\beta H(x))\,\mathrm{d}x \tag{2.10}$$

is the partition function. We note that the states $u^a(x)$ are the energy eigenstates of the system, with eigenvalues $H(x)$. That is to say, $H^a_b u^b(x) = H(x)u^a(x)$. The

variable $x$ in these formulae is formal in the sense that the integration may, if appropriate, be replaced by a summation. By comparing equations (2.8) and (2.9), we find that the initial ($\beta = 0$) thermal state $q^a$ is

$$q^a = \mathrm{e}^{-(1/2)W(0)} \int_x u^a(x) \, \mathrm{d}x, \tag{2.11}$$

which corresponds to the 'centre' point in $\mathcal{S}_+$ determined by the system of energy eigenstates. This relation reflects the fact that at infinite temperature all configurations are equally likely to occur.

Viewed as a function of $\beta$, the state trajectory $\psi^a(\beta)$ thus starts at the centre point $q^a$, and follows a curve on $\mathcal{S}$ generated by the Hamiltonian $H_{ab}$ according to (2.7). It is interesting to note that the curvature of this trajectory, given by

$$K_\psi(\beta) = \frac{\langle \tilde{H}^4 \rangle}{\langle \tilde{H}^2 \rangle^2} - \frac{\langle \tilde{H}^3 \rangle^2}{\langle \tilde{H}^2 \rangle^3} - 1, \tag{2.12}$$

arises naturally in a physical characterization of the accuracy bounds for temperature measurements. This point is pursued further in §2 $d$ below, and in Brody & Hughston (1998$b$). In equation (2.12) the expression $\langle \tilde{H}^n \rangle$ denotes the $n$th central moment of the observable $H_{ab}$. The curvature of the curve $\psi^a(\beta)$, which is necessarily positive, is defined by the square of the 'acceleration' vector, normalized by the square of the velocity vector.

### (c) *Measurement for thermal states*

Given the thermal trajectory $\psi^a(\beta)$, we propose, in the first instance, to consider measurement and estimation by analogy with the von Neumann approach in quantum mechanics. According to this scheme the *general* state of a thermodynamic system is represented by a 'density matrix' $\rho^{ab}$, which in the present context should be understood to be a symmetric semi-definite matrix with trace unity; that is to say, $\rho^{ab}\xi_a\xi_b \geqslant 0$ for any covector $\xi_a$, and $\rho^{ab}g_{ab} = 1$. Then, for example, we can write

$$E_\rho[X] = X_{ab}\rho^{ab}, \tag{2.13}$$

for the expectation of a random variable $X_{ab}$ in the state $\rho^{ab}$, and

$$\mathrm{Var}_\rho[X] = X_{ab}X_c^b\rho^{ac} - (X_{ab}\rho^{ab})^2, \tag{2.14}$$

for the variance of $X_{ab}$ in that state. It should be evident that in the case of a pure state, for which $\rho^{ab}$ is of the form $\rho^{ab} = \psi^a\psi^b$ for some state vector $\psi^a$, these formulae reduce to expressions considered in §2 $a$.

In particular, let us consider measurements made on a pure equilibrium state $\psi^a(\beta)$. Such measurements are characterized by projecting the prescribed state onto the ray in the Hilbert space corresponding to a specified point in the phase space. Hence, the probability density of observing the state $u^a(x)$ corresponding to the energy level $E(x)$, when the system is in the pure state $\psi^a$, is given by the Boltzmann weight

$$p(x) = (g_{ab}\psi^a u^b(x))^2. \tag{2.15}$$

In terms of the density matrix description, the state before measurement is given by the degenerate pure state matrix $\rho^{ab} = \psi^a\psi^b$, for which the thermal development is

$$\frac{\mathrm{d}\rho^{ab}}{\mathrm{d}\beta} = -\tfrac{1}{2}(\tilde{H}_c^a\rho^{bc} + \tilde{H}_c^b\rho^{ac}), \tag{2.16}$$

or equivalently $\mathrm{d}\rho/\mathrm{d}\beta = -\{\tilde{H}, \rho\}$, where $\{A, B\}$ denotes the symmetric product between the operators $A$ and $B$. After a measurement, $\rho^{ab}$ takes the form of a mixed state, characterized by a non-degenerate diagonal density operator for which the diagonal elements are the Boltzmann weights $p(x)$. As a consequence, the von Neumann entropy $- \mathrm{Tr}[\rho \ln \rho]$ changes from 0 to its maximum value $S = \beta\langle H\rangle + W_\beta$, which can be viewed as the quantity of information gained from the observation.

More generally, suppose we consider the measurement of an arbitrary observable $X_{ab}$ in the state $\psi^a(\beta)$, in the situation when the spectrum of $X_{ab}$ admits a continuous component. In this case, we require the spectral measure associated with the random variable $X_{ab}$, according to which the probability density for the measurement outcome $x$ is given by the expectation $p_\beta(x) = \Pi_{ab}(X, x)\psi^a\psi^b$ of the projection operator

$$\Pi_b^a(X, x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp[\mathrm{i}\lambda(X_b^a - x\delta_b^a)]\, \mathrm{d}\lambda. \tag{2.17}$$

In other words, we assign a projection-valued measure $\Pi_X(x)$ on the real line associated with each symmetric operator $X$, so that for a given unit vector $\psi^a$, the mapping $x \in \mathbb{R} \mapsto E_\psi[\Pi_X(x)]$ is a probability measure. This measure determines the distribution of values obtained when the observable $X$ is measured while the system is in the state $\psi$.

For a more refined view of the measurement problem we need to take into account some ideas from statistical estimation theory. Suppose we wish to make a measurement or series of measurements to estimate the value of the parameter characterizing a given thermal equilibrium state. In this situation the observable that we measure is called an 'estimator' for the given parameter. We are interested in the case for which the estimator is unbiased in the sense that its expectation gives the value of the required parameter. To be specific, we consider the case when we estimate the value of the temperature. Let $B_{ab}$ be an unbiased estimator for $\beta$, so that along the trajectory $\psi^a(\beta)$ we have

$$\frac{B_{ab}\psi^a\psi^b}{g_{cd}\psi^c\psi^d} = \beta. \tag{2.18}$$

As a consequence of this relation and the thermal state equation (2.7), we observe that the inverse temperature estimator $B$ and the Hamiltonian $H$ satisfy the 'weak' anti-commutation relation $E_\psi[\{B, \tilde{H}\}] = -1$ along the state trajectory. In statistical terms, this implies that these conjugate variables satisfy the covariance relation $E_\psi[BH] - E_\psi[B]E_\psi[H] = -1$ along this trajectory.

### (d) *Thermodynamic uncertainty relations*

Equipped with the above definitions, one can verify that (Brody & Hughston 1998*b*) the variance in estimating the inverse temperature parameter $\beta$ can be expressed by the geometrical relation

$$\mathrm{Var}_\psi[B] = \tfrac{1}{4}g^{ab}\nabla_a\beta\nabla_b\beta, \tag{2.19}$$

on the unit sphere $\mathcal{S}$, where $\nabla_a\beta = \partial\beta/\partial\psi^a$ is the gradient of the temperature estimate $\beta$. The essence of formula (2.19) can be understood as follows. First, recall that $\beta$ is the expectation of the estimator $B_{ab}$ in the state $\psi^a(\beta)$. Suppose that the

state changes rapidly as $\beta$ changes. Then, the variance in estimating $\beta$ is small— indeed, this is given by the squared magnitude of the 'functional derivative' of $\beta$ with respect to the state $\psi^a$. On the other hand, if the state does not change significantly as $\beta$ changes, then the measurement outcome of an observable is less conclusive in determining the value of $\beta$.

The squared length of the gradient vector $\nabla_a \beta$ can be expressed as a sum of squares of orthogonal components. To this end, we choose a new set of orthogonal basis vectors given by the state $\psi^a$ and its higher derivatives. If we let $\psi_n^a$ denote $\psi^a$ for $n = 0$, and for $n > 0$ the component of the derivative $\partial^n \psi^a / \partial \beta^n$ orthogonal to the state $\psi^a$ and its lower-order derivatives, then our orthonormal vectors are given by

$$\hat{\psi}_n^a = \psi_n^a (g_{bc} \psi_n^b \psi_n^c)^{-1/2}, \quad \text{for } n = 0, 1, 2, \ldots. \tag{2.20}$$

With this choice of orthonormal vectors, we find that the variance of the estimator $B$ satisfies

$$\text{Var}_\psi[B] \geqslant \sum_n \frac{(\tilde{B}_{ab} \psi_n^a \psi^b)^2}{g_{cd} \psi_n^c \psi_n^d}, \tag{2.21}$$

for any range of the index $n$. This follows because the squared magnitude of the vector $\frac{1}{2} \nabla_a \beta = \tilde{B}_{ab} \psi^b$ is greater than or equal to the sum of the squares of its projections onto the basis vectors given by $\hat{\psi}_n^a$ for the specified range of $n$.

In particular, for $n = 1$ we have $B_{ab} \psi_1^a \psi^b = \frac{1}{2}$ on account of the relation $B_{ab} \psi^a \psi^b = \beta$, and $g_{ab} \psi_1^a \psi_1^b = \frac{1}{4} \Delta H^2$, which follows from the thermal equation (2.7). Therefore, if we write $\text{Var}_\psi[B] = \Delta \beta^2$, we find that (2.21) implies the following thermodynamic uncertainty relation:

$$\Delta \beta^2 \Delta H^2 \geqslant 1, \tag{2.22}$$

valid along the trajectory of the thermal equilibrium states. The variance $\Delta \beta^2$ is to be understood in the sense of estimation theory. That is, although $\beta$ does not actually fluctuate, as should be clear from the definition of the canonical ensemble, there is, nonetheless, an inevitable lower bound for the variance of the measurement, given by (2.22), if we wish to estimate the heat-bath temperature. It is worth pointing out that the exposition given here is consistent with the view put forward by Mandelbrot (1962, 1989), who should perhaps be credited with first having introduced an element of modern statistical reasoning into the long-standing debate on the status of temperature fluctuations (cf. Feshbach 1987; Kittel 1988).

Note that, although we consider the variance $\langle (B - \langle B \rangle)^2 \rangle$ here, the higher-order central moments $\mu_n = \langle (B - \langle B \rangle)^n \rangle$ can also be expressed geometrically. First, recall that for any observable $F_{ab}$ with $E_\psi[F] = f$, we have $\nabla_a f = 2 \tilde{F}_{ab} \psi^b$ on the unit sphere $\mathcal{S}$. Therefore, letting $F = \tilde{B}^n$ we construct the higher central moments in terms of the cosines of the angles between certain gradient vectors, e.g.

$$4 \mu_3 = g^{ab} \nabla_a \mu_2 \nabla_b \beta, \qquad 4 \mu_4 = g^{ab} \nabla_a \mu_2 \nabla_b \mu_2 - 4 \mu_2^2, \tag{2.23}$$

and so on. In particular, the even-order moments are expressible in terms of combinations of the squared lengths of normal vectors to the surfaces of constant central moments of lower order.

### (*e*) *Random states*

Let us return to the consideration of measurements on thermal states, which we now pursue in greater depth. In doing so, we shall introduce the idea of a 'random state', a concept that is applicable both in clarifying the measurement problem in statistical physics, as well as in providing a useful tool when we consider ensembles. It also turns out that the idea of a random state is helpful in the analysis of conceptual problems in quantum mechanics. We shall have more to say on this later on, when we consider quantum statistical mechanics.

Suppose we consider a pure thermal state $\psi^a(\beta)$ for some value $\beta$ of the inverse temperature. We know that this state is given by

$$\psi^a(\beta) = \int_x \sqrt{p(x)} u^a(x) \, \mathrm{d}x, \tag{2.24}$$

where $u^a(x)$ is a normalized energy eigenstate with eigenvalue $H(x)$, and $p(x)$ is the associated Gibbs probability. It is natural to model the outcome of a measurement by a *random state* $\boldsymbol{\Psi}^a$. Thus we consider $\boldsymbol{\Psi}^a$ to be a random variable (indicated by use of a bold font) such that the probability density for $\boldsymbol{\Psi}^a$ to take on value $u^a$ is $p(x)$. This way of thinking about the measurement process is to some extent complementary to the density matrix approach, though in what follows we shall make it clearer just what the relationship is.

In particular, the expectation of an observable $X_{ab}$ in the random state $\boldsymbol{\Psi}^a$ is given by averaging over the random states, that is,

$$E_{\boldsymbol{\Psi}}[X] = X_{ab} \boldsymbol{\Psi}^a \boldsymbol{\Psi}^b. \tag{2.25}$$

This relation should be interpreted as a specification of the *conditional* expectation of $X_{ab}$ in the random state $\boldsymbol{\Psi}^a$. Then the associated *unconditional* expectation $E[X] = E[E_{\boldsymbol{\Psi}}[X]]$ is

$$E[X] = X_{ab} E[\boldsymbol{\Psi}^a \boldsymbol{\Psi}^b]. \tag{2.26}$$

However, it should be evident that

$$E[\boldsymbol{\Psi}^a \boldsymbol{\Psi}^b] = \rho^{ab}, \tag{2.27}$$

where the density matrix $\rho^{ab}$ is defined by

$$\rho^{ab} = \int_x p(x) u^a(x) u^b(x) \, \mathrm{d}x. \tag{2.28}$$

Thus, the unconditional expectation of the random variable $X_{ab}$ is given, as noted earlier, by $E[X] = X_{ab}\rho^{ab}$. It should be observed, however, that here we are not emphasizing the role of the density matrix $\rho^{ab}$ as representing a 'state', but rather its role in summarizing information relating to the random state $\boldsymbol{\Psi}^a$.

The feature that distinguishes the density matrix in this analysis is that it is fully sufficient for the characterization of unconditional statistics relating to the observables and states under consideration. This point is clearly illustrated when we calculate the variance of a random variable $X_{ab}$ in a random state, $\boldsymbol{\Psi}^a$. Such a situation arises if we want to discuss the uncertainties arising in the measurement of an observable $X_{ab}$ for an ensemble. In this case, the system we have in mind consists of a large number of identical independent particles, each of which has a definite energy, where the distribution of the energy is given according to the Gibbs distribution. One

might take this as an elementary model for a classical gas. Then the ensemble can be described by a random state $\boldsymbol{\Psi}^a$. Note that here, the interpretation is slightly different from what we had considered before (the random outcome of a measurement for an isolated system), though it will be appreciated that the relation of these two distinct interpretations is of considerable interest for physics and statistical theory alike.

The conditional variance of the observable $X_{ab}$ in the random state $\boldsymbol{\Psi}^a$ is given by

$$\mathrm{Var}_{\boldsymbol{\Psi}}[X] = X_{ab}X_c^b\boldsymbol{\Psi}^a\boldsymbol{\Psi}^c - (X_{ab}\boldsymbol{\Psi}^a\boldsymbol{\Psi}^b)^2. \tag{2.29}$$

The average over the different values of $\boldsymbol{\Psi}^a$ then gives us

$$E[\mathrm{Var}_{\boldsymbol{\Psi}}[X]] = X_{ab}X_c^b\rho^{ac} - X_{ab}X_{cd}\rho^{abcd}, \tag{2.30}$$

where $\rho^{ab}$ is, as before, the density matrix (2.27), and $\rho^{abcd}$ is a certain higher moment of $\boldsymbol{\Psi}^a$, defined by

$$\rho^{abcd} = E[\boldsymbol{\Psi}^a\boldsymbol{\Psi}^b\boldsymbol{\Psi}^c\boldsymbol{\Psi}^d]. \tag{2.31}$$

The appearance of this higher-order analogue of the density matrix may be surprising, though it is indeed a characteristic feature of conditional probability. However, the unconditional variance of $X$ is not given by the expectation $E[\mathrm{Var}_{\boldsymbol{\Psi}}[X]]$, but rather (see, for example, Ross 1996) by the *conditional variance formula*:

$$\mathrm{Var}[X] = E[\mathrm{Var}_{\boldsymbol{\Psi}}[X]] + \mathrm{Var}[E_{\boldsymbol{\Psi}}[X]]. \tag{2.32}$$

For the second term we have

$$\begin{aligned}
\mathrm{Var}[E_{\boldsymbol{\Psi}}[X]] &= E[(X_{ab}\boldsymbol{\Psi}^a\boldsymbol{\Psi}^b)^2] - (E[X_{ab}\boldsymbol{\Psi}^a\boldsymbol{\Psi}^b])^2 \\
&= X_{ab}X_{cd}\rho^{abcd} - (X_{ab}\rho^{ab})^2,
\end{aligned} \tag{2.33}$$

which also involves the higher moment $\rho^{abcd}$. The terms in (2.32) involving $\rho^{abcd}$ then cancel, and we are left with $\mathrm{Var}[X] = X_{ab}X_c^b\rho^{ac} - (X_{ab}\rho^{ab})^2$ for the unconditional variance, which, as indicated earlier, only involves the density matrix. It follows that the random-state approach does indeed reproduce the earlier density matrix formulation of our theory, though the status of the density matrix is diminished. In other words, whenever conditioning is involved, it is the set of totally symmetric tensors,

$$\rho^{ab\cdots cd} = E[\boldsymbol{\Psi}^a\boldsymbol{\Psi}^b\cdots\boldsymbol{\Psi}^c\boldsymbol{\Psi}^d], \tag{2.34}$$

that plays the fundamental role, though it is sufficient to consider the standard density matrix $\rho^{ab}$ when conditioning is removed.

All this is worth bearing in mind when we turn to quantum statistical mechanics, where the considerations developed here in a classical context reappear in a new light. We want to de-emphasize the role of the density matrix, not because there is anything wrong *per se* with the use of the density matrix in an appropriate context, but rather for two practical reasons. First, when we consider conditioning, focusing attention exclusively on the density matrix hampers our thinking, since, as we have indicated, higher moments of the random-state vector also have a role to play. Second, when we consider generalizations of quantum mechanics, such as nonlinear theories of the Kibble–Weinberg type (Kibble 1978; Weinberg 1989*a*, *b*), or stochastic theories of the type considered by Gisin (1989), Pearle (1993), Percival (1994) and others (cf. Hughston 1996; Brody & Hughston 1999, and references cited therein), the density matrix is either an ill-formulated concept, or plays a diminished role. We shall return to this point when we consider quantum statistical mechanics in § 4.

## 3. Statistical phase space

### (*a*) *Projective space and probabilities*

To proceed further, it will be useful to develop a formalism for the algebraic treatment of real projective geometry, with a view to its probabilistic interpretation in the context of classical statistical mechanics. Let $Z^a$ be coordinates for an $(n+1)$-dimensional real Hilbert space $\mathcal{H}^{n+1}$. Later, when we consider quantum theory from a real point of view we shall double this dimension. In the Hilbert-space description of classical probabilities, the normalization condition is written $g_{ab}Z^aZ^b = 1$. However, this normalization is physically irrelevant since the expectation of an arbitrary operator $F_{ab}$ is defined by the ratio

$$\langle F \rangle = \frac{F_{ab}Z^aZ^b}{g_{cd}Z^cZ^d}. \tag{3.1}$$

Therefore, the physical state space is not the Hilbert space $\mathcal{H}$, but the space of equivalence classes obtained by identifying the points $\{Z^a\}$ and $\{\lambda Z^a\}$, for all $\lambda \in \mathbb{R} - \{0\}$. In this way, we 'gauge away' the irrelevant degree of freedom. The resulting space is the real projective $n$-space $\mathrm{RP}^n$, the space of rays through the origin in $\mathcal{H}^{n+1}$. Thus, two points, $X^a$ and $Y^a$, in $\mathcal{H}^{n+1}$ are equivalent in $\mathrm{RP}^n$ if they are proportional, i.e. $X^{[a}Y^{b]} = 0$.

The coordinates $Z^a$ (excluding $Z^a = 0$) can be used as homogeneous coordinates for points of $\mathrm{RP}^n$. Clearly $Z^a$ and $\lambda Z^a$ represent the same point in $\mathrm{RP}^n$. In practice, one treats the homogeneous coordinates as though they define points of $\mathcal{H}^{n+1}$, with the stipulation that the allowable operations of projective geometry are those that transform homogeneously under the rescaling $Z^a \to \lambda Z^a$.

A *prime* or $(n-1)$-plane in $\mathrm{RP}^n$ consists of a set of points $Z^a$ that satisfy a linear equation $P_aZ^a = 0$, where we call $P_a$ the homogeneous coordinates of the prime. Clearly, $P_a$ and $\lambda P_a$ determine the same prime. Therefore, a prime in $\mathrm{RP}^n$ is an $\mathrm{RP}^{n-1}$, and the set of all primes in $\mathrm{RP}^n$ is itself an $\mathrm{RP}^n$ (the 'dual' projective space). In particular, the metric $g_{ab}$ on $\mathcal{H}^{n+1}$ can be interpreted in the projective space as giving rise to a non-singular *polarity*, that is, an invertible map from points to hyperplanes in $\mathrm{RP}^n$ of codimension one. This map is given by $P^a \to P_a := g_{ab}P^b$. See Hughston & Hurd (1983) for further discussion of some of the geometric operations employed here.

If a point $P^a$ in $\mathrm{RP}^n$ corresponds to a probability state, then its negation $\neg P^a$ is the hyperplane $P_aZ^a = 0$. To be more precise, we take $P^a$ as describing the probability state for a set of events. Then, all the probability states corresponding to the complementary events lie in the prime $P_aZ^a = 0$. Thus, the points $Z^a$ on this plane are precisely the states that are orthogonal to the original state $P^a$. The intuition behind this is as follows. Two states $P^a$ and $Q^a$ are orthogonal if and only if any event that, in the state $P^a$ (resp. $Q^a$), has a positive probability is assigned zero probability by the state $Q^a$ (resp. $P^a$). This is the sense in which orthogonal states are 'complementary'. For any state $P^a$, the plane consisting of all points $Z^a$ such that $P_aZ^a = 0$ is the set of states complementary to $P^a$ in this sense.

Distinct states $X^a$ and $Y^a$ are joined by a real projective line represented by the skew tensor $L^{ab} = X^{[a}Y^{b]}$. The points on this line are the various real superpositions of the original two states. The intersection of the line $L^{ab}$ with the plane $R_a$ is given by $S^a = L^{ab}R_b$. Clearly, $S^a$ lies on the plane $R_a$, since $R_aS^a = 0$ on account of the antisymmetry of $L^{ab}$.

The hyperplanes that are the negations (polar planes) of two points $P^a$ and $Q^a$ intersect the joining line $L^{ab} = P^{[a}Q^{b]}$ at a pair of points $\tilde{P}^a$ and $\tilde{Q}^a$, respectively. That is, if $P_a = g_{ab}P^b$ is the coordinate of a plane and $L^{ab}$ represents a line in $\mathrm{RP}^n$, then the point of intersection is given by $\tilde{P}^a = L^{ab}P_b$, and similarly $\tilde{Q}^a = L^{ab}Q_b$. The projective cross ratio between these four points $P^a$, $Q^a$, $\tilde{P}^a = P^a(Q^b P_b) - Q^a(P^b P_b)$ and $\tilde{Q}^a = P^a(Q^b Q_b) - Q^a(P^b Q_b)$, given by $P^a\tilde{Q}_a Q^b\tilde{P}_b / P^c\tilde{P}_c Q^d\tilde{Q}_d$, reduces, after some algebra, to the following simple expression:

$$\kappa = \frac{(P^a Q_a)^2}{P^b P_b Q^c Q_c}, \tag{3.2}$$

which can be interpreted as the *transition probability* between $P^a$ and $Q^a$. It is interesting to note that this formula has an analogue in quantum mechanics (Hughston 1995).

The projective line $L^{ab}$ can also be viewed as a circle, with $\tilde{P}^a$ and $\tilde{Q}^a$ antipodal to the points $P^a$ and $Q^a$. In that case, the cross ratio $\kappa$ is $\frac{1}{2}(1 + \cos\theta) = \cos^2(\theta/2)$ where $\theta$ defines the angular distance between $P^a$ and $Q^a$, in the geometry of $\mathrm{RP}^n$. We note that $\theta$ is, in fact, *twice* the angle made between the corresponding Hilbert-space vectors, so orthogonal states are maximally distant from one another.

Now suppose we let the two states $P^a$ and $Q^a$ approach one another. In the limit, the resulting formula for their second-order infinitesimal separation determines the natural line element on real projective space. This can be obtained by setting $P^a = Z^a$ and $Q^a = Z^a + \mathrm{d}Z^a$ in (3.2), while replacing $\theta$ with the small angle $\mathrm{d}s$ in the expression $\frac{1}{2}(1 + \cos\theta)$, retaining terms of the second order in $\mathrm{d}s$. Explicitly, we obtain

$$\mathrm{d}s^2 = 4\left[\frac{\mathrm{d}Z^a\,\mathrm{d}Z_a}{Z^b Z_b} - \frac{(Z^a\,\mathrm{d}Z_a)^2}{(Z^b Z_b)^2}\right]. \tag{3.3}$$

Note that this metric (Kobayashi & Nomizu 1963, 1969) is locally equivalent to the metric on the sphere $\mathcal{S}^n$ in $\mathcal{H}^{n+1}$, except that in the latter case one does not identify opposite points.

We now consider the case where the real projective space is the state space of classical statistical mechanics. If we write $\psi^a(\beta)$ for the trajectory of thermal state vectors, as discussed in §2, then we can regard $\psi^a(\beta)$, for each value of $\beta$, as representing homogeneous coordinates for points in the state space $\mathrm{RP}^n$. Since $\psi^a(\beta)$ satisfies (2.7), it follows that the line element along the curve $\psi^a(\beta)$ is given by $\mathrm{d}s^2 = \langle\tilde{H}^2\rangle\,\mathrm{d}\beta^2$, which can be identified with the Fisher–Rao metric induced on the thermal trajectory by virtue of its embedding in $\mathrm{RP}^n$. This follows by insertion of the thermal equation (2.7) into expression (3.3) for the natural spherical metric on $\mathrm{RP}^n$.

### (b) *The projective thermal equations*

Let us write $H_{ab}$ for the symmetric Hamiltonian operator, and $\psi^a(\beta)$ for the one-parameter family of thermal states. Then, in our notation, the thermal equation is $\mathrm{d}\psi^b = -\frac{1}{2}\tilde{H}^b_c\psi^c\mathrm{d}\beta$. However, we are concerned with this equation only inasmuch as it supplies information about the evolution of the state of the system in $\mathrm{RP}^n$. We are interested therefore primarily in the *projective* thermal equation

$$\psi^{[a}\,\mathrm{d}\psi^{b]} = -\tfrac{1}{2}\psi^{[a}\tilde{H}^{b]}_c\psi^c\,\mathrm{d}\beta, \tag{3.4}$$

obtained by skew-symmetrizing the thermal equation (2.7) with $\psi^a$. Equation (3.4) defines the equilibrium thermal trajectory of a statistical mechanical system in the proper state space.

In fact, the thermal equation generates a Hamiltonian gradient flow on the state manifold. This can be seen as follows. First, recall that a physical observable $F$ is associated with a symmetric operator $F_{ab}$, and the set of such observables forms a vector space of dimension $\frac{1}{2}(n+1)(n+2)$. Such observables are determined by their diagonal matrix elements, which are real valued functions on $\mathrm{RP}^n$ of the form $F(\psi^a) = F_{ab}\psi^a\psi^b/\psi^c\psi_c$. In particular, we are interested in the Hamiltonian function $H(\psi^a)$. Then, by a direct substitution, we find that the vector field $H^a = g^{ab}\partial H/\partial\psi^b$, associated with the Hamiltonian function $H$, takes the form $H^a = 2\tilde{H}^a_b\psi^b/\psi^c\psi_c$. Therefore, we can write the differential equation for the thermal state trajectory in $\mathcal{H}^{n+1}$ in the form

$$\frac{\mathrm{d}\psi^a}{\mathrm{d}\beta} = -\tfrac{1}{4}g^{ab}\nabla_b H. \tag{3.5}$$

By projecting this down to $\mathrm{RP}^n$, we find

$$\psi^{[a}\mathrm{d}\psi^{b]} = -\tfrac{1}{4}\psi^{[a}\nabla^{b]}H\,\mathrm{d}\beta, \tag{3.6}$$

where $\nabla^a H = g^{ab}\nabla_b H$. From this, we can then calculate the line element to obtain

$$\begin{aligned}
\mathrm{d}s^2 &= 8\psi^{[a}\mathrm{d}\psi^{b]}\psi_{[a}\mathrm{d}\psi_{b]}/(\psi^c\psi_c)^2 \\
&= \tfrac{1}{2}\psi^{[a}\nabla^{b]}H\psi_{[a}\nabla_{b]}H\,\mathrm{d}\beta^2 \\
&= \tfrac{1}{4}\nabla^b H\nabla_b H\psi^a\psi_a\,\mathrm{d}\beta^2 \\
&= \langle\tilde{H}^2\rangle\,\mathrm{d}\beta^2,
\end{aligned} \tag{3.7}$$

which establishes the result we noted earlier.

The critical points of the Hamiltonian function are the fixed points in the state space associated with the gradient vector field $g^{ab}\nabla_b H$. In the case of Hamilton's equations, the fixed points are called stationary states. In the Hilbert space $\mathcal{H}^{n+1}$, these are the points corresponding to the energy eigenstates $u^a_k$ given by $H^a_b u^b_k = E_k u^a_k$ (in the general situation with distinct energy eigenvalues $E_k$, $k = 0, 1, \ldots, n$). Note that, for the sake of intuition, we write the singular basis $u^a(x)$ in a discrete form $u^a_k$ in this section. Therefore, in the projective space $\mathrm{RP}^n$, we have a set of fixed points corresponding to the states $u^a_k$, and the thermal states are obtained by superposing these points with an appropriate set of coefficients given by the Boltzmann weights. Since these coefficients are non-zero at finite temperature, it should be clear that the thermal trajectories do not intersect any of these fixed points.

In particular, the infinite-temperature ($\beta = 0$) thermal state $\psi^a(0)$ is located at the centre point of $\mathcal{S}_+$ in $\mathcal{H}^{n+1}$ determined by $H_{ab}$. The distances from $\psi^a(0)$ to the various energy eigenstates are all equal. This implies that in $\mathrm{RP}^n$ the cross ratios between the fixed points and the state $\psi^a(0)$ are equal. Therefore, we can single out the point $\psi^a(0)$ as an initial point, and form a geodesic hypersphere in $\mathrm{RP}^n$. All the fixed points of the Hamiltonian gradient vector field $H^a$ thus lie on this sphere. Since the cross ratios between the fixed points are also equal (i.e. maximal), these fixed points form a regular simplex on the sphere. The thermal trajectory thus starts at $\psi^a(0)$, and asymptotically approaches a fixed point associated with the lowest energy eigenvalue $E_0$, as $\beta \to \infty$.

If we take the orthogonal prime of the thermal state for any finite $\beta$, i.e. $\neg\psi^a(\beta)$, then the resulting hyperplane clearly does not contain any of the fixed points. On the other hand, if we take the orthogonal prime of any one of the fixed points $u_k^a$, then the resulting hyperplane includes a sphere of codimension one, where all the other fixed points lie. This sphere is given by the intersection of the original hypersphere surrounding $\psi^a(0)$ with the orthogonal prime of the given excluded fixed point. There is a unique prime containing $n$ general points in $\mathrm{RP}^n$. It is worth noting, therefore, that if we choose $n$ general points given by $u_j^a$ ($j = 0, 1, \ldots, n$; $j \neq k$), then there is a unique solution, up to proportionality, of the $n$ linear equations $X_a u_0^a = 0, \ldots, X_a u_n^a = 0$. The solution is then given by $X^a = u_k^a$, where $u_k^a = \epsilon^a_{bc\cdots d} u_0^b u_1^c \cdots u_n^d$. Here $\epsilon_{ab\cdots c} = \epsilon_{[ab\cdots c]}$, with $n + 1$ indices, is the totally skew tensor determined up to proportionality.

It would be interesting to explore whether the orientability characteristics of $\mathrm{RP}^n$ lead to any physical consequences. There may be a kind of purely classical 'spin-statistics' relation in the sense that the state space for half-integral spin is associated with a topological invariant, while the state space for even spin is not.

### (c) Hamiltonian flows and projective transformations

We have seen in §3 b how the thermal trajectory of a statistical mechanical system is generated by the gradient flow associated with the Hamiltonian function. In other words, for each point on the state manifold we form the expectation of the Hamiltonian in that state. This gives us a global function on the state manifold, which we call the Hamiltonian function. Next, we take the gradient of this function, and raise the index by use of the natural metric to obtain a vector field. This vector field is the generator of the thermal trajectories.

It is interesting to note that there is a relation between the geometry of such vector fields and the global symmetries of the state manifold. In particular, we show below that gradient vector fields generated by observables on the state manifold can be interpreted as the generators of *projective transformations*. A projective transformation on a Riemannian manifold is an automorphism that maps geodesics onto geodesics. In the case of a real projective space endowed with the natural metric, the general projective automorphism is generated, as we shall demonstrate, by a vector field that is expressible in the form of a sum of a Killing vector and a gradient flow associated with an observable function.

To pursue this point further, we develop some differential geometric aspects of the state manifold. We consider $\mathrm{RP}^n$ now to be a differential manifold endowed with the natural spherical metric $g_{ab}$. Here, bold indices signify local tensorial operations in the tangent space of this manifold. Thus, we write $\nabla_a$ for the covariant derivative associated with $g_{ab}$, and for an arbitrary vector field $V^a$ we define the Riemann tensor $R_{abc}{}^d$ according to the convention

$$\nabla_{[a}\nabla_{b]}V^c = \tfrac{1}{2}R_{abd}{}^c V^d. \tag{3.8}$$

It follows that $R_{abcd} := R_{abc}{}^e g_{de}$ satisfies $R_{abcd} = R_{[ab][cd]}$, $R_{abcd} = R_{cdab}$, and $R_{[abc]d} = 0$. In the case of $\mathrm{RP}^n$ with the natural metric

$$\mathrm{d}s^2 = \frac{8Z^{[a}\mathrm{d}Z^{b]}Z_{[a}\mathrm{d}Z_{b]}}{\lambda(Z^c Z_c)^2}, \tag{3.9}$$

where $Z^a$ are homogeneous coordinates and $\lambda$ is a scale factor (set to unity in the preceding analysis), the Riemann tensor is given by

$$R_{abcd} = -\tfrac{1}{4}\lambda(g_{ac}g_{bd} - g_{bc}g_{ad}). \tag{3.10}$$

Now we consider projective transformations on $\mathrm{RP}^n$. First, we make a few general remarks about projective transformations on Riemannian manifolds (see, for example, Tomonaga 1970). Suppose we have a Riemannian manifold with metric $g_{ab}$, and we consider the effect of dragging the metric along the integral curves of a vector field $\xi^a$. For an infinitesimal transformation we have

$$g_{ab} \to \hat{g}_{ab} = g_{ab} + \epsilon\mathcal{L}_\xi g_{ab}, \tag{3.11}$$

where $\mathcal{L}_\xi g_{ab} = 2\nabla_{(a}\xi_{b)}$ is the Lie derivative of the metric ($\epsilon \ll 1$). The Levi–Civita connection $\hat{\nabla}_a$ associated with $\hat{g}_{ab}$ is defined by a tensor $Q^c_{ab}$, symmetric on its lower indices, such that

$$\hat{\nabla}_a V^c = \nabla_a V^c + \epsilon Q^c_{ab} V^b, \tag{3.12}$$

for an arbitrary vector field $V^a$, and $\hat{\nabla}_a \hat{g}_{bc} = 0$. A familiar line of argument based on the fundamental theorem of Riemannian geometry shows that

$$Q^c_{ab} = \tfrac{1}{2}\hat{g}^{cd}(\nabla_a\hat{g}_{bd} + \nabla_b\hat{g}_{ad} - \nabla_d\hat{g}_{ab}), \tag{3.13}$$

where $\hat{g}^{ab}$ is the inverse of $\hat{g}_{ab}$, given to first order in $\epsilon$ here by $\hat{g}^{ab} = g^{ab} - 2\epsilon\nabla^{(a}\xi^{b)}$. A calculation making use of (3.8) shows that to first order we have

$$Q^c_{ab} = \nabla_{(a}\nabla_{b)}\xi^c + R^c{}_{(ab)d}\xi^d. \tag{3.14}$$

Now, suppose the vector field $U^a$ satisfies the geodesic equation, which we write in the form $(U^a\nabla_a U^{[b]})U^{c]} = 0$. It should be apparent that $U^a$ is geodesic with respect to the transformed metric $\hat{g}_{ab}$ if and only if $(U^a\hat{\nabla}_a U^{[b]})U^{c]} = 0$, or equivalently, $U^a U^b Q^{[c}_{ab} U^{d]} = 0$. However, this relation is satisfied for *all* geodesic vector fields if and only if there exists a vector field $\phi_a$ such that $Q^c_{ab} = \delta^c_{(a}\phi_{b)}$. Hence, we conclude that the vector field $\xi^a$ is the generator of a projective transformation, mapping geodesics to geodesics, if and only if there exists a vector field $\phi_a$ such that

$$\nabla_{(a}\nabla_{b)}\xi^c + R^c{}_{(ab)d}\xi^d = \delta^c_{(a}\phi_{b)}. \tag{3.15}$$

On the other hand, it follows as a direct consequence of (3.8), and the identity $R_{[abc]d} = 0$ that for *any* vector field $\xi^a$, we have

$$\nabla_{[a}\nabla_{b]}\xi^c - R^c{}_{[ab]d}\xi^d = 0. \tag{3.16}$$

Hence, combining (3.15) and (3.16) we obtain

$$\nabla_a\nabla_b\xi^c + R^c{}_{bad}\xi^d = \delta^c_{(a}\phi_{b)}, \tag{3.17}$$

as a condition equivalent to (3.15) for the existence of a projective transformation on the given manifold. This form of the condition is especially useful for calculations (cf. Tomonaga 1970).

In particular, if $\xi^a$ is a Killing vector, so $\nabla_{(a}\xi_{b)} = 0$, then (3.17) is automatically satisfied, with $\phi_a = 0$, since, as a consequence of (3.8), a Killing vector necessarily satisfies $\nabla_a\nabla_b\xi^c + R^c{}_{bad}\xi^d = 0$. Thus, a Killing symmetry generates a projective transformation. For example, in the case of $\mathrm{RP}^n$ we have the symmetries of the projective orthogonal group.

Now we derive an integrability condition for (3.17) that will be helpful in our analysis of the remaining projective transformations on $\mathrm{RP}^n$, apart from those associated with Killing symmetries. Returning to the transformation (3.11), we consider the curvature $\hat{R}_{abc}{}^d$ associated with the metric $\hat{g}_{ab}$. Clearly, to first order in $\epsilon$ this is given by

$$\hat{R}_{abc}{}^d = R_{abc}{}^d + \epsilon \mathcal{L}_\xi R_{abc}{}^d. \tag{3.18}$$

On the other hand, it is known (see, for example, Hughston & Tod 1990) that for a change of connection given by (3.12) the corresponding change in curvature is

$$\tfrac{1}{2}\hat{R}_{abc}{}^d = \tfrac{1}{2}R_{abc}{}^d + \epsilon \nabla_{[a}Q_{b]c}{}^d + \epsilon^2 Q_{r[a}{}^d Q_{b]c}{}^r. \tag{3.19}$$

Thus, by consideration of terms to first order in $\epsilon$ we deduce that

$$\tfrac{1}{2}\mathcal{L}_\xi R_{abc}{}^d = \nabla_{[a}Q_{b]c}{}^d, \tag{3.20}$$

with $Q_{ab}^c$ given as in equation (3.14). This relation holds for any infinitesimal transformation of the form (3.11) generated by a given vector field $\xi^a$. In the case of a projective transformation, for which $Q_{ab}^c = \delta_{(a}^c \phi_{b)}$, it follows, therefore, after some rearrangement, that

$$\mathcal{L}_\xi R_{abc}{}^d + \delta_{[a}^d \nabla_{b]}\phi_c - \nabla_{[a}\phi_{b]}\delta_c^d = 0. \tag{3.21}$$

This is the desired integrability condition. In the case of a space of constant curvature we have (3.10) and, hence,

$$\mathcal{L}_\xi R_{abc}{}^d = -\tfrac{1}{2}\lambda[\nabla_{(a}\xi_{c)}\delta_b^d - \nabla_{(b}\xi_{c)}\delta_a^d]. \tag{3.22}$$

Inserting this relation into (3.21), we obtain

$$-\tfrac{1}{2}\lambda\nabla_{(a}\xi_{c)}g_{bd} + \tfrac{1}{2}\lambda\nabla_{(b}\xi_{c)}g_{ad} + g_{d[a}\nabla_{b]}\phi_c - \nabla_{[a}\phi_{b]}g_{cd} = 0, \tag{3.23}$$

which implies, after contraction over the indices $c$ and $d$, that $\nabla_{[a}\phi_{b]} = 0$, i.e. that $\phi_a = -\lambda\nabla_a H$ for some function $H(x)$ on the state manifold. Inserting this relation back into (3.23), we obtain $\nabla_{(a}\xi_{b)} = \nabla_a\nabla_b H$, which implies that $\xi^a$ is necessarily of the form

$$\xi^a = \eta^a + g^{ab}\nabla_b H, \tag{3.24}$$

where $\eta^a$ is a Killing vector. Returning this information to equation (3.15), we conclude that

$$\nabla_{(a}\nabla_b\nabla_{c)}H + \lambda g_{(ab}\nabla_{c)}H = 0. \tag{3.25}$$

In other words, $K_{ab} := \nabla_a\nabla_b H + \lambda g_{ab}H$ is a Killing tensor: $\nabla_{(a}K_{bc)} = 0$. However, this is the defining equation on $\mathrm{RP}^n$ for observable functions, i.e. functions of the form

$$H(x) = \frac{H_{ab}Z^a Z^b}{g_{cd}Z^c Z^d}, \tag{3.26}$$

in terms of homogeneous coordinates on $\mathrm{RP}^n$. Thus, we conclude that the generator of a projective transformation on the state space is given as in (3.24), by the sum of a Killing vector and the gradient of an observable function.

We remark, incidentally, that if the dimension $n$ of the manifold is finite, then we can form the trace of (3.17), which gives $\nabla_a(\nabla^2 H) + \tfrac{1}{4}\lambda(n+1)\nabla_a H = 0$, from which we conclude that, up to an additive constant, $H(x)$ is an eigenfunction of the

Laplacian operator, i.e. $\nabla^2 H + \frac{1}{4}\lambda(n+1)H = 0$. The only *global* solutions of this equation on $\mathrm{RP}^n$ are necessarily of the form (3.26). Such functions also satisfy (3.25). In the case of infinite dimension, the trace is not necessarily a valid operation, and we lose the representation of observable functions as eigenfunctions of the Laplacian. On the other hand, our characterization of projective transformations as vector fields of the form (3.24), given by the sum of a Killing vector and the gradient of a scalar $H(x)$ satisfying (3.25) and (3.26), is entirely general, valid in both finite and infinite dimensions. Setting $\lambda = 1$, we see, therefore, that thermal states are characterized by vector fields on $\mathrm{RP}^n$ of the form $\xi^{\boldsymbol{a}} = g^{\boldsymbol{ab}}\nabla_{\boldsymbol{b}}H$, where the Hamiltonian function $H(x)$ satisfies

$$\nabla_{(\boldsymbol{a}}\nabla_{\boldsymbol{b}}\nabla_{\boldsymbol{c})}H + g_{(\boldsymbol{ab}}\nabla_{\boldsymbol{c})}H = 0. \tag{3.27}$$

It is worthwhile noting that, among the trajectories generated by the Hamiltonian gradient flow, the thermal trajectory is singled out on account of the special initial state corresponding to $\beta = 0$. This state is determined by the specification of the Hamiltonian function. In particular, for a given Hamiltonian $H_{ab}$, there are in general $n + 1$ fixed points in the state manifold. The initial state for the thermal trajectory is uniquely determined by the condition that it is equidistant from these $n + 1$ fixed points.

## 4. Quantum statistical mechanics

### (*a*) *The quantum phase space*

In the algebraic approach to quantum theory, one generally works with the totality of mixed states, while regarding pure states as extremal elements of this convex set (Haag *et al.* 1967), and one works in the Heisenberg picture for dynamics. The approach we adopt here, however, will focus primarily on the system of pure states, and for dynamics we shall work in the Schrödinger picture. This is consistent with the point of view we put forward earlier in §2 in connection with the discussion of classical thermal states.

Given a complex Hilbert space $\mathcal{H}_{\mathbb{C}}^{n+1}$ with complex coordinates $Z^\alpha$ representing pure quantum states, we identify the state vector $Z^\alpha$ with its complex multiples $\lambda Z^\alpha$, $\lambda \in \mathbb{C} - \{0\}$, to obtain the complex projective space $\mathrm{CP}^n$. Here we use Greek indices ($\alpha = 0, 1, \ldots, n$) for vectors in $\mathcal{H}_{\mathbb{C}}^{n+1}$. Following the line of argument indicated earlier in the case of the thermal state space, we regard $\mathrm{CP}^n$ as the true 'state space' of quantum mechanics. The status of density matrices will be discussed shortly. Our goal now is to build up the necessary projective geometric machinery appropriate for the consideration of quantum mechanics.

Suppose we regard the state vector $Z^\alpha$ as representing homogeneous coordinates on the projective Hilbert space. The complex conjugate of the state vector $Z^\alpha$ is written $\bar{Z}_\alpha$, with the corresponding Hermitian inner product $Z^\alpha\bar{Z}_\alpha$. The complex conjugate of a point $P^\alpha$ in $\mathrm{CP}^n$ is the hyperplane (prime) $\bar{P}_\alpha Z^\alpha = 0$. The points on this plane are the states that are orthogonal to the original state $P^\alpha$, and we denote this $\neg P^\alpha$ in the sense of the probability rules discussed above. Thus, $\mathrm{CP}^n$ is equipped with a Hermitian correlation, i.e. a complex conjugation operation that maps points to hyperplanes of codimension one.

Distinct points $X^\alpha$ and $Y^\alpha$ are joined by a complex projective line $L^{\alpha\beta} = X^{[\alpha}Y^{\beta]}$, representing the various complex superpositions of the two states. The quantum-

mechanical transition probability between two states $X^\alpha$ and $Y^\alpha$ is then given by the cross ratio

$$\kappa = \frac{X^\alpha \bar{Y}_\alpha Y^\beta \bar{X}_\beta}{X^\gamma \bar{X}_\gamma Y^\delta \bar{Y}_\delta}. \tag{4.1}$$

More precisely, we recall that if the system is in the state $Y^\alpha$ and a measurement is made to see if the system is in the state $X^\alpha$ (corresponding to a measurement of the observable represented by the projection operator $X^\alpha \bar{X}_\beta / X^\gamma \bar{X}_\gamma$), then $\kappa$ is the probability that the result of the measurement is affirmative. If we set $\kappa = \cos^2(\theta/2)$, and set $X^\alpha = Z^\alpha$ and $Y^\alpha = Z^\alpha + \mathrm{d}Z^\alpha$, retaining terms to second order, we recover the natural unitary-invariant metric on complex projective space known as the Fubini–Study metric (Kobayashi & Nomizu 1963, 1969):

$$\mathrm{d}s^2 = 8(Z^\alpha \bar{Z}_\alpha)^{-2} Z^{[\alpha} \mathrm{d}Z^{\beta]} \bar{Z}_{[\alpha} \mathrm{d}\bar{Z}_{\beta]}. \tag{4.2}$$

Now suppose we write $H^\alpha_\beta$ for the Hamiltonian operator, assumed Hermitian. Then for the Schrödinger equation we have

$$\frac{\mathrm{d}Z^\alpha}{\mathrm{d}t} = \mathrm{i}H^\alpha_\beta Z^\beta. \tag{4.3}$$

However, we are interested in the *projective* Schrödinger equation, given by

$$Z^{[\alpha} \mathrm{d}Z^{\beta]} = \mathrm{i}Z^{[\alpha} H^{\beta]}_\gamma Z^\gamma \mathrm{d}t, \tag{4.4}$$

which eliminates the superfluous degree of freedom associated with the direction of $Z^\alpha$. This equation is well defined on the projective state space. Insertion of (4.4) into (4.2) gives

$$\mathrm{d}s^2 = 4\langle \tilde{H}^2 \rangle \mathrm{d}t^2, \tag{4.5}$$

which verifies that the velocity of a state along its trajectory in state space is $\mathrm{d}s/\mathrm{d}t = 2\Delta H$, where $\Delta H = \langle \tilde{H}^2 \rangle^{1/2}$ is the energy uncertainty (Anandan & Aharonov 1990).

An alternative way of looking at this structure is to regard the state manifold $\mathrm{CP}^n$ as a real manifold $\Gamma$ of dimension $2n$, equipped with a Riemannian metric $g_{ab}$ and a compatible symplectic structure $\Omega_{ab}$. Here again we use bold indices for tensorial operations in the tangent space of $\Gamma$. In this formulation the Schrödinger equation takes the form of a flow $\xi^a$ on $\Gamma$ given by

$$\xi^a = 2\Omega^{ab}\nabla_b H. \tag{4.6}$$

Here $H$ is the real function on $\Gamma$, given at each point by the expectation of the Hamiltonian operator at that state. It is interesting to note that in the quantum-mechanical case, the dynamical trajectory is given by a Hamiltonian *symplectic* flow, in contrast to the case of statistical mechanics where the relevant flow is given by a Hamiltonian *gradient* flow for thermal trajectories. Furthermore, we note that in the case of quantum mechanics the flow is necessarily a Killing field, satisfying $\nabla_{(a}\xi_{b)} = 0$, where $\xi_a = g_{ab}\xi^b$. In other words, the isometries of the Fubini–Study metric on $\Gamma$ can be lifted to $\mathcal{H}_\mathbb{C}^{n+1}$ to yield unitary transformations. In the statistical mechanical case, on the other hand, the general projective transformation of the manifold includes both Killing vectors and Hamiltonian gradient flows, but we exclude the former. In both quantum mechanics and statistical mechanics the fixed points of the Hamiltonian operator play a pivotal role.

## (*b*) *The quantum thermal states*

In §4*a* above, we have indicated two useful ways of looking at the quantum state space. The approach via complex homogeneous coordinates is appropriate when the *global* algebraic geometry of the state manifold is of interest. The differential geometric approach is useful when *local* properties of the state manifold are considered (Hughston 1995, 1996). There is a third, equally important approach, however, which we shall call the 'real' approach (Gibbons 1992; Gibbons & Pohle 1993; Brody & Hughston 1996, 1998*a*; Ashtekar & Schilling 1998) to quantum mechanics. In this case, we regard the Hilbert space of quantum mechanics as a *real* vector space, of dimension $2n + 2$. The importance of the real approach is that it links up directly with modern notions from probability theory and statistics. It should not, therefore, be surprising that this approach is also useful in the quantum dynamics of thermal states, and indeed in considering the relation between classical and quantum statistical mechanics.

Let us write $\xi^a$ ($a = 1, 2, \ldots, 2n + 2$) for the real Hilbert-space coordinates corresponding to a quantum state vector. The other ingredients we have at our disposal are the metric $g_{ab}$ on $\mathcal{H}^{2n+2}$ and a compatible complex structure tensor $J_a^b$. We proceed in two stages. First, we regard the state vector $\xi^a$ as homogeneous coordinates for a real projective space $\mathrm{RP}^{2n+1}$ of one dimension lower. This corresponds to factoring out by the freedom $\xi^a \to \lambda \xi^a$ with $\lambda$ real and non-vanishing. Then there is another more subtle freedom, corresponding to factoring out by the 'phase', given by the transformation

$$\xi^a \to \tfrac{1}{2}\mathrm{e}^{\mathrm{i}\phi}(\xi^a + \mathrm{i}J_b^a\xi^b) + \tfrac{1}{2}\mathrm{e}^{-\mathrm{i}\phi}(\xi^a - \mathrm{i}J_b^a\xi^b), \tag{4.7}$$

where $J_b^a$ is the complex structure, and $\phi$ is a phase factor. In this way we obtain a map $\mathrm{RP}^{2n+1} \to \mathrm{CP}^n$, which is related to the Hopf map $S^{2n+1} \to \mathrm{CP}^n$, and we have the diagram

$$
\begin{array}{ccc}
\mathcal{H}^{2n+2} \longrightarrow & S^{2n+1} & \\
\searrow & \downarrow & \searrow \\
& \mathrm{RP}^{2n+1} \longrightarrow & \mathrm{CP}^n
\end{array}
\tag{4.8}
$$

showing the relation between these various spaces. The point of this line of reasoning is that if we take the classical statistical argument based on the state space $\mathrm{RP}^{2n+1}$, and introduce a complex structure on the real Hilbert space $\mathcal{H}^{2n+2}$, then the analogous relations for quantum statistical mechanics can be developed by studying the relation between $\mathrm{RP}^{2n+1}$ and $\mathrm{CP}^n$.

Now we consider more explicitly the characterization of thermal states in the context of quantum 'thermodynamics'. There are a number of features of the Hilbert-space description of conventional statistical mechanics that carry through to the quantum regime, though there are new features as well. In particular, if we start with the real Hilbert space $\mathcal{H}^{2n+2}$ and the associated state space $\mathrm{RP}^{2n+1}$ obtained when we neglect the complex structure on $\mathcal{H}^{2n+2}$ given by quantum theory, then we can apply the theory developed earlier to define thermal trajectories in the space $\mathrm{RP}^{2n+1}$. The main difference arising between the classical and quantum cases is that in quantum theory the Hamiltonian operator $H_{ab}$ satisfies the Hermitian property

$$H_{ab}J_c^a J_d^b = H_{cd}, \tag{4.9}$$

which ensures that it has at most $n + 1$ distinct eigenvalues, unlike the classical situation where there could be $2n + 2$. The thermal equation, giving the development of $\psi^a(\beta)$ as the inverse temperature $\beta$ is changed, is again given by

$$\frac{\mathrm{d}\psi^a(\beta)}{\mathrm{d}\beta} = -\tfrac{1}{2}\tilde{H}^a_b\psi^b(\beta), \tag{4.10}$$

where

$$\tilde{H}_{ab} = H_{ab} - g_{ab}H_{cd}\psi^c\psi^d / g_{ef}\psi^e\psi^f. \tag{4.11}$$

Another feature that distinguishes the quantum theory is that, in general, we find a multiplicity of infinite-temperature ($\beta = 0$) states, corresponding to different values of the 'phases' associated with the state. Topologically, this implies that there is a torus $T^n = (S^1)^n$ in $\mathrm{CP}^n$, the points of which correspond to the $\beta = 0$ state. Thus, for a specified infinite-temperature state, i.e. for a given point on $T^n$, we shall call the solution of (4.10) a 'primitive' thermal state parametrized by a set of phases. In other words, there is a multiplicity of primitive thermal trajectories, each corresponding to a choice of infinite-temperature state. In the case of a two-state system (e.g. a spin-1/2 particle), the two energy eigenstates correspond to the north and south poles of the state space $\mathrm{CP}^1 \sim S^2$, and the infinite-temperature states are given by points on the equator. The thermal trajectories are then given by the geodesic segments that join the equator to one of the poles. This example is studied in more detail in § 4 *e*.

For many applications we have to consider ensembles of particles for which the starting point of the thermal trajectory is effectively random. In other words, although each particle in the ensemble has a definite thermal trajectory, when we form expectations based on the behaviour of the ensemble we have to average over the manifold of infinite-temperature states, i.e. average over the random phases. This leads to the well-known density-matrix-based description of thermal states, as we discuss below. Thus, once again, the density matrix formulation emerges as an essentially secondary construction, although it remains essential for many practical applications.

Another characteristic of quantum thermal dynamics is that the thermal trajectories have a time evolution, given by the Schrödinger equation. Written in terms of real state vectors this evolution is given by

$$\psi^a_\phi(t, \beta) = \exp(-tJ^a_b\tilde{H}^b_c)\psi^c_\phi(0, \beta), \tag{4.12}$$

for each thermal trajectory, where the variables collectively denoted $\phi$ label the $\beta = 0$ states. Thus the primitive thermal states are pure states, subject to the usual evolutionary laws.

In our discussion of classical thermal systems, we introduced the idea of a random state. This notion is also useful in a quantum-mechanical context. For example, a random state $\boldsymbol{\Psi}^a$ can be used to represent the outcome of a measurement process, or to describe the statistics of an ensemble. In particular, in the case of a quantum thermal state we want to take an ensemble average over all possible infinite-temperature states to describe the observed features of ensembles. This point will now be explored further.

Suppose we denote by $\boldsymbol{\Phi}$ the system of random phases labelling the $\beta = 0$ states for a given quantum ensemble. The thermal state $\psi^a_{\boldsymbol{\Phi}}$ is, thus, itself a random state, since $\boldsymbol{\Phi}$ is random. Then, if $X_{ab}$ is a typical observable, satisfying the Hermitian property,

we can write $E_{\boldsymbol{\Phi}}[X] = X_{ab}\psi_{\boldsymbol{\Phi}}^a\psi_{\boldsymbol{\Phi}}^b$ for the conditional expectation of $X_{ab}$, given $\boldsymbol{\Phi}$. This is in keeping with the line of argument introduced earlier for classical thermal systems. The ensemble average $E[X]$ is given by the unconditional expectation

$$E[E_{\boldsymbol{\Phi}}[X]] = X_{ab}\rho^{ab}, \tag{4.13}$$

where

$$\rho^{ab}(\beta) = E[\psi_{\boldsymbol{\Phi}}^a\psi_{\boldsymbol{\Phi}}^b] \tag{4.14}$$

is the Hermitian density matrix characterizing the state of the ensemble. Here the unconditional expectation $E[-]$ averages over the phase variables appearing in random state $\psi_{\boldsymbol{\Phi}}^a(0,\beta)$. In this connection it is worth noting that the infinite-temperature states have the property that the ensemble average in formula (4.14) is proportional to the metric. That is,

$$\rho^{ab}(0) = \frac{1}{2n+2}g^{ab}, \tag{4.15}$$

where $2n+2$ is the dimension of the real Hilbert space.

It is interesting to note the topological characteristics of the infinite-temperature state manifold. As we already remarked, this manifold can be viewed as an $n$-torus $T^n$ in the state manifold $\mathrm{CP}^n$. This can be seen as follows. First, recall the fact that the $\beta = 0$ states are equidistant from the fixed points $u_k^\alpha$ of the Hamiltonian flow in $\mathrm{CP}^n$, and that these fixed points form a regular simplex. Thus, we would like to find the locus of points $Z^\alpha$, such that the distances between $Z^\alpha$ and $u_k^\alpha$ are equal for all $k$. Since the points $u_k^\alpha$ are vertices of a regular simplex, we can choose coordinates, without loss of generality, such that $u_0^\alpha = (1,0,0,\dots)$, $u_1^\alpha = (0,1,0,\dots)$, and so on. Thus, by writing the equidistance condition in terms of the cross ratio explicitly, we find that the locus of $Z^\alpha = (Z^0, Z^1, \dots, Z^n)$ is given by the simultaneous solution of the following system of equations:

$$Z^0\bar{Z}_0 = Z^1\bar{Z}_1, Z^1\bar{Z}_1 = Z^2\bar{Z}_2, \dots, Z^n\bar{Z}_n = Z^0\bar{Z}_0. \tag{4.16}$$

We thus deduce that the solution takes the form

$$Z^\alpha = (\mathrm{e}^{\mathrm{i}\phi_0}, \mathrm{e}^{\mathrm{i}\phi_1}, \dots, \mathrm{e}^{\mathrm{i}\phi_n}), \tag{4.17}$$

for some phase variables $\phi_j$. This implies that the topology of the infinite-temperature state manifold has the structure of an $n$-torus $T^n = (S^1)^n$ (one of the $n+1$ phase variables $\phi_j$ is redundant, and can be scaled away). In summary, the fundamental torus $T^n \subset \mathrm{CP}^n$ associated with a given Hamiltonian function consists of all those states that are equally distant to the eigenstates of the specified Hamiltonian with respect to the Fubini–Study metric on $\mathrm{CP}^n$. The significance of $T^n$ is that it embodies a geometrical representation of the extraneous phases over which one has to integrate for many purposes in thermal physics.

### (c) The equilibrium condition

The property that the density matrix $\rho^{ab}(0)$ at infinite temperature is proportional to the metric is closely related to the so-called KMS condition (Kubo 1957; Martin & Schwinger 1959; Haag *et al.* 1967; Bratteli & Robinson 1979, 1981), which provides an alternative characterization of thermal equilibrium states. It is, therefore, interesting to see how the KMS construction fits into the present description of thermal

phenomena. First, recall that a general state, or density matrix, can be regarded as a semidefinite map $\rho(A)$ from Hermitian operators $A$ to the real numbers given by $\rho(A) = \text{Tr}(\rho A)$, satisfying $\rho(X\bar{X}) \geqslant 0$ for all $X$, and $\rho(g) = 1$, where $g$ is the identity operator. The trace operation has the commutative property $\text{Tr}(AB) = \text{Tr}(BA)$. For a general state $\rho$, on the other hand, clearly we do not have $\rho(AB) = \rho(BA)$. However, the thermal equilibrium states $\rho_\beta(-)$ are characterized by a slightly weakened form of commutation relation, which is the KMS condition.

Let $A$ be a Hermitian operator and write $\tau_t A$ for the unitary action of the one-parameter group of time translations given by $\tau_t A = e^{itH} A e^{-itH}$. We can extend this definition by analytic continuation to complex time by writing

$$\tau_{t+i\beta} A = e^{i(t+i\beta)H} A e^{-i(t+i\beta)H}. \tag{4.18}$$

The KMS condition on a state $\rho$ is that, for all Hermitian operators $A$ and $B$, it should satisfy

$$\rho(\tau_t AB) = \rho(B\tau_{t+i\beta}A). \tag{4.19}$$

It is not difficult to see that in finite dimensions this implies that $\rho$ is the thermal state $\rho_\beta$ discussed above, and that commutativity holds for infinite temperature.

Now let us look at these relations from a real point of view. First, we need to consider the action of the one-parameter group of time translations. Suppose we have a real Hilbert space of dimension $2n + 2$, with typical elements $\xi^a, \eta^a \in \mathcal{H}$. On $\mathcal{H}$ we have the metric $g_{ab}$ and the complex structure $J_a^b$, satisfying $J_a^c J_c^b = -\delta_a^b$ and $J_a^c J_b^d g_{cd} = g_{ab}$. The tensor $\Omega_{ab} = g_{ac}J_b^c$ is antisymmetric, and defines a symplectic structure on $\mathcal{H}$. With these ingredients at hand, we can define the action of the orthogonal and unitary groups on $\mathcal{H}$.

For a typical element $\xi^a \in \mathcal{H}$, the orthogonal group $O(2n+2)$ consists of transformations $\xi^a \to O_b^a \xi^b$, satisfying $O_a^c O_b^d g_{cd} = g_{ab}$. The unitary group $U(n+1)$ is then a subgroup of $O(2n+2)$, given by transformations $\xi^a \to U_b^a \xi^b$ satisfying

$$U_a^c U_b^d g_{cd} = g_{ab}, \qquad U_a^c U_b^d \Omega_{cd} = \Omega_{ab}. \tag{4.20}$$

In other words, transformations that preserve both the metric and the symplectic structure on $\mathcal{H}$. For typical real elements $\xi^a, \eta^a \in \mathcal{H}$, their ordinary real Hilbert-space inner product is given by $\eta^a g_{ab} \xi^b$, which is invariant if $\xi^a$ and $\eta^a$ are subject to orthogonal transformations. On the other hand, the standard Dirac product $\langle \eta \mid \xi \rangle$ between two real vectors $\xi^a$ and $\eta^a$, given by

$$\langle \eta \mid \xi \rangle = \tfrac{1}{2}\eta^a (g_{ab} - i\Omega_{ab})\xi^b, \tag{4.21}$$

is invariant only under unitary transformations. This is clear from the need to preserve the real and imaginary parts of (4.21) separately, which implies (4.20).

If $O_a^b$ is an orthogonal transformation, the corresponding inverse element $\bar{O}_a^b$ is given by the transpose $\bar{O}_a^b = g^{bc}g_{ad}O_c^d$, which satisfies $O_b^a \bar{O}_c^b = \delta_c^a$. In the case of a unitary transformation we can, without ambiguity, use the notation $\bar{U}_a^b = g^{bc}g_{ad}U_c^d$ for the conjugate transformation, from which it follows that the conditions

$$U_b^a \bar{U}_c^b = \delta_c^a, \qquad U_b^a J_c^b \bar{U}_d^c = \delta_d^a, \tag{4.22}$$

are equivalent to (4.20) in characterizing unitary transformations. In particular, (4.22) can be viewed as saying that the unitary transformations on $\mathcal{H}$ are orthogonal transformations that preserve the complex structure.

In equation (4.22) we see the action of the unitary group on the complex structure tensor. More generally, for a typical multi-index tensor $A_{bc}^a$ we define the action of the unitary group by

$$A_{bc}^a \to U_{a'}^a A_{b'c'}^{a'} \bar{U}_b^{b'} \bar{U}_c^{c'}. \tag{4.23}$$

Here, the primed indices merely serve to increase the size of the standard alphabet.

Now suppose we consider one-parameter subgroups of the orthogonal group, continuous with the identity. Such transformations are of the form

$$O_a^b = \exp[tM_{ac}g^{bc}], \tag{4.24}$$

where the tensor $M_{ab}$ is antisymmetric. For a one-parameter family of unitary transformations we have to specialize further, and require $M_{ab}$ to be of the form $M_{ab} = \Omega_{ac}H_b^c$, where $H_{ab}$ is symmetric and Hermitian. Indeed, if $H_{ab}$ is symmetric, then a necessary and sufficient condition that it should be Hermitian is that $\Omega_{ac}H_b^c$ is antisymmetric. It follows that the general one-parameter group of unitary transformations, continuous with the identity, can be written

$$\xi^a \to \xi_t^a = \exp[tJ_b^a H_c^b]\xi^c. \tag{4.25}$$

If $H_{ab}$ is the Hamiltonian, then $\xi_t^a$ is the Schrödinger evolution generated from the given initial state. The point of view here continues to be purely 'real' in the sense that the complex structure tensor $J_b^a$ is playing the role of the factor of 'i' in the conventional expression of unitary evolution given by $|\xi\rangle \to e^{itH}|\xi\rangle$. Our goal is to formulate a similar geometrization for the KMS condition.

For this purpose, we need to consider the changes implied for the picture noted above when we go to the Heisenberg representation. Suppose $A_{ab}$ is an observable (i.e. a symmetric, Hermitian operator) and we consider the evolution of its expectation $A_{ab}\xi^a\xi^b$ under the action of the unitary transformation $\xi^a \to U_b^a \xi^b$ with $U_b^a = \exp[tJ_c^a H_b^c]$. The Heisenberg picture is obtained if we let $\xi^a$ be stationary, and let $A_{ab}$ evolve according to the scheme

$$A_{ab} \to U_a^c U_b^d A_{cd}. \tag{4.26}$$

It will be appreciated that the evolution of $A_{ab}$ in the Heisenberg representation is contragradient to the 'natural' action of the unitary group $A_{ab} \to \bar{U}_a^{a'} \bar{U}_b^{b'} A_{a'b'}$ discussed earlier, since the natural action of the unitary group is defined to be the action on $A_{ab}$ that preserves $A_{ab}\xi^a\xi^b$ when $\xi^a$ is evolved in the Schrödinger representation. By the same token, in the Heisenberg representation the action of the time evolution operator on an observable, represented in 'operator' form by $A_a^b$ (rather than $A_{ab}$), is

$$A_a^b \to U_a^{a'} A_{a'}^{b'} \bar{U}_{b'}^b, \tag{4.27}$$

where

$$U_a^{a'} = \exp[tJ_b^{a'} H_a^b] \quad \text{and} \quad \bar{U}_a^{a'} = \exp[-tJ_b^{a'} H_a^b].$$

Now, finally, we are in a position to address the KMS condition. First, we note that for a pair of observables $A_{ab}$ and $B_{ab}$, their quantum-mechanical 'Dirac' product, $C = AB$, is given by

$$C_{ab} = A_{c(a}\Delta^{cd}B_{b)d}, \tag{4.28}$$

where $\Delta^{ab} = \frac{1}{2}(g^{ab} - \mathrm{i}\Omega^{ab})$. The real and imaginary parts of the complex symmetric tensor $C_{ab}$ are given, respectively, by the Jordan product $\frac{1}{2}(AB + BA)$ and the commutator $\frac{1}{2}\mathrm{i}(AB - BA)$. In particular, we have

$$C_{ab} = \tfrac{1}{2}A^c_{(a}B^d_{b)}g_{cd} - \tfrac{1}{2}\mathrm{i}A^c_{(a}B^d_{b)}\Omega_{cd}. \tag{4.29}$$

It follows that if $\rho_{ab}$ is a density matrix, the expectation $\rho(AB)$ in the corresponding state is given by $\rho(AB) = \rho^a_b A^b_c \Delta^c_d B^d_a$, or, equivalently,

$$\rho(AB) = \tfrac{1}{2}\rho_{ab}A^a_c B^{bc} - \tfrac{1}{2}\mathrm{i}\rho_{ab}A^a_c B^b_d \Omega^{cd}. \tag{4.30}$$

As we noted earlier, the action of the one-parameter group $\tau_t$ on $A$ is given by

$$A^b_a \to \tau_t A^b_a := \exp[tJ^c_a H^{a'}_c]A^b_{a'}\exp[-tJ^d_{b'}H^b_d]. \tag{4.31}$$

The KMS complexification of this action, corresponding to replacing $t$ with $t + \mathrm{i}\beta$, replaces $tJ^b_a$ with $tJ^b_a - \beta\delta^b_a$, and we have

$$\tau_{t+\mathrm{i}\beta}A^b_a := \exp[tJ^c_a H^{a'}_c - \beta H^{a'}_a]A^b_{a'}\exp[-tJ^d_{b'}H^b_d + \beta H^b_{b'}]. \tag{4.32}$$

It is then straightforward to verify that in the case of a thermal state, for which

$$\rho^b_a = \frac{\exp[-\beta H^b_a]}{\delta^c_d \exp[-\beta H^d_c]}, \tag{4.33}$$

that the KMS condition, $\rho_\beta(\tau_t AB) = \rho_\beta(B\tau_{t+\mathrm{i}\beta}A)$, is indeed satisfied.

### (*d*) *Phase space dynamics and temperature*

In the foregoing formulation, we have adopted the viewpoint that, given a system in a heat bath with inverse temperature $\beta$, we let the equilibrium thermal states evolve quantum mechanically under the influence of the Hamiltonian. We have also considered the case of an ensemble of particles in a thermalized box, whereby while each particle takes a definite energy value, the probability law of the energy for a random element of the ensemble is given by the Boltzmann distribution. In either case, the value of $\beta$ is taken as an 'input' variable that specifies the thermal state.

This point of view can be inverted by shifting from the canonical description to the microcanonical ensemble. In this case, the essential equivalence of the two approaches is typically recognized only for large systems (see, for example, Ellis 1985). The dynamical formulation we offer here may be useful, in particular, for small systems such as those on a quantum or mesoscopic scale.

First, we regard as given, a Hamiltonian function $H$ defined on the quantum phase space $\Gamma$, which is distinguished by the fact that it possesses both a symplectic structure and a metric structure. Thus, suppose we foliate the real state manifold with constant energy surfaces, given by $H(x) = E$. Then the metrical volume $\mathcal{V}(E)$ of such a surface $\mathcal{E}^{2n-1}_E$ is given by the integral

$$\mathcal{V}(E) = \int_{\mathcal{E}_E} \frac{\nabla_a H \, \mathrm{d}\sigma^a}{(\nabla_b H \nabla^b H)^{1/2}}. \tag{4.34}$$

Here, we have written $\mathrm{d}\sigma^a = g^{ab}\epsilon_{bc\cdots d}\,\mathrm{d}x^c \cdots \mathrm{d}x^d$ for the natural vector $(2n-1)$-form on $\Gamma$. Note that the denominator of the integrand in (4.34) is the energy uncertainty

$(\Delta H^2)^{1/2}$. The measure for the number of microscopic configurations, on the other hand, is given by the state density

$$\begin{aligned} \Omega(E) &= \int_{\Gamma} \delta(H(x) - E)\,\mathrm{d}V \\ &= \int_{\mathcal{E}_E} \frac{\nabla_a H\,\mathrm{d}\sigma^a}{\nabla_b H \nabla^b H}, \end{aligned} \tag{4.35}$$

that is, the weighted 'canonical' volume of the energy surface. The number of microscopic quantum states with expected energy in the small range between $E$ and $E + \Delta E$ is $\Omega(E)\Delta E$. Therefore, using the Boltzmann relation, the entropy $S(E)$ associated with this energy band is $S(E) = \ln(\Omega(E)\Delta E)$. As a result, the inverse temperature can be calculated by use of the familiar prescription

$$\beta = \frac{\mathrm{d}S(E)}{\mathrm{d}E}. \tag{4.36}$$

In this way, we can calculate the temperature of the system directly from the given quantum dynamics of the system, or, equivalently, given the interrelationship of the energy surfaces. The natural equation of state associated with the microcanonical ensemble is given by the relation $\beta P = \partial S/\partial \mathcal{V}$, where $P(E)$ is the generalized pressure conjugate to the metrical volume $\mathcal{V}(E)$.

The construction just noted for obtaining temperature from the underlying phase-space dynamics is well known in classical statistical mechanics (see, for example, Thompson 1972; Rugh 1997). The novelty here, however, is to regard the quantum state space $\mathrm{CP}^n$ as a real $2n$-dimensional manifold $\Gamma$, playing the role of a quantum phase space, hence allowing us to obtain the temperature for quantum-dynamical systems in the case of the microcanonical ensemble. Based upon this formulation, we can also investigate the possibility of various generalizations of the temperature, when the underlying quantum-mechanical dynamics is modified. There are a number of distinct approaches that can be pursued in this context.

First, we can replace the Hamiltonian function defined on the state manifold by a more general observable. In this case, the underlying theory would correspond to the Kibble–Weinberg-type theories of nonlinear quantum mechanics, for which the above prescription would provide the temperature in the case of a nonlinear dynamical system of this sort.

Alternatively, a more radical extension is obtained by replacing the state-space manifold itself with a general Kähler manifold. In such generalizations, the notion of a particle state may no longer survive. This might appear to be an undesirable feature. However, our view is that, at high energies, as in the instance of particle collisions, the notion of a particle state as such may be lost in any case, and recovered only at an asymptotic level, i.e. on the tangent space. For such generalizations, the notion of temperature we have outlined above would nonetheless survive, indicating a strong link between the geometrical structure of the state manifold and the thermodynamics of associated statistical ensembles.

The above formulation for the quantum microcanonical ensemble can naturally be extended to the regime of a quantum canonical ensemble (Brody & Hughston 1998$d$). In this case, we have in mind a system in contact with a heat reservoir having inverse temperature $\beta$. The probability density for the energy distribution of such a system

is given by

$$\rho(E) = \frac{\Omega(E)\exp(-\beta E)}{Q(\beta)}, \tag{4.37}$$

where

$$Q(\beta) = \int \Omega(E)\exp(-\beta E)\,\mathrm{d}E \tag{4.38}$$

is the partition function. Therefore, the conventional thermodynamic functions can be calculated in terms of derivatives of the function $Q(\beta)$. Note that expression (4.37) for energy distribution does not agree, as such, with the conventional treatment in quantum statistical mechanics, for which the state density $\Omega(E)$ is replaced by a sum of delta-function distributions $\delta(E - E_i)$ centred on energy eigenvalues. Presumably, at least for a bulk object, the state of the system spontaneously reduces to one of the eigenstates, due to strong internal interactions, with transition probability given by the corresponding Boltzmann weights. For a microscopic system maintaining quantum coherence, on the other hand, we might in some circumstances expect quantum effects to exhibit a different behaviour, as given by (4.37).

### (e) *Density matrix approach*

The formulation based on pure thermal states presented above can be re-expressed in a natural way by using density operators. Such a formulation paves the way between the standard approach used in operator algebra and the dynamical approach that has been introduced in this paper.

We write, as usual, $Z^\alpha$ for the homogeneous coordinates of a point in $\mathrm{CP}^n$, and let $\boldsymbol{Z}^\alpha$ denote a random point in $\mathrm{CP}^n$. In terms of the coordinates $Z^\alpha$, the natural volume element on $\mathrm{CP}^n$ is written

$$\mathrm{d}V = \frac{D^n Z D^n \bar{Z}}{(Z^\alpha \bar{Z}_\alpha)^n}, \tag{4.39}$$

where $D^n Z = \epsilon_{\alpha\beta\cdots\gamma} Z^\alpha \,\mathrm{d}Z^\beta \cdots \mathrm{d}Z^\gamma$ is the natural twisted holomorphic $n$-form on $\mathrm{CP}^n$, and $D^n \bar{Z}$ is its complex conjugate. Now, let $p(Z^\alpha, \bar{Z}_\alpha)$ be the probability density function on $\mathrm{CP}^n$ for the random variable $\boldsymbol{Z}^\alpha$. Then, if $\phi(\boldsymbol{Z}^\alpha, \bar{\boldsymbol{Z}}_\alpha)$ is any real-valued function of the random state $\boldsymbol{Z}^\alpha$, we have

$$E[\phi(\boldsymbol{Z}^\alpha, \bar{\boldsymbol{Z}}_\alpha)] = \int_{\mathrm{CP}^n} \phi(Z^\alpha, \bar{Z}_\alpha) p(Z^\alpha, \bar{Z}_\alpha)\,\mathrm{d}V, \tag{4.40}$$

where $E[-]$ is the unconditional expectation. In particular, suppose that $\boldsymbol{Z}^\alpha$ has a uniform distribution over the energy surface $\mathcal{E}_E$, defined by the constraint

$$\frac{H_\beta^\alpha Z^\beta \bar{Z}_\alpha}{Z^\gamma \bar{Z}_\gamma} = E, \tag{4.41}$$

and has vanishing probability outside the surface $\mathcal{E}_E$. Then, for an arbitrary observable $F_\beta^\alpha$, its unconditional expectation is given by

$$\begin{aligned}
\langle F \rangle_{\mathrm{MC}} &= E\left[\frac{F_\beta^\alpha Z^\beta \bar{Z}_\alpha}{Z^\gamma \bar{Z}_\gamma}\right] \\
&= \frac{1}{\Omega(E)} \int_{\mathrm{CP}^n} \delta\left[\frac{H_\beta^\alpha Z^\beta \bar{Z}_\alpha}{Z^\gamma \bar{Z}_\gamma} - E\right] \frac{F_\beta^\alpha Z^\beta \bar{Z}_\alpha}{Z^\gamma \bar{Z}_\gamma}\,\mathrm{d}V,
\end{aligned} \tag{4.42}$$

where $\Omega(E)$ is the state density defined in (4.35), and $\langle - \rangle_{\mathrm{MC}}$ indicates the microcanonical ensemble average.

Alternatively, we can write the expectation of $F_\beta^\alpha$ in the form $\langle F \rangle_{\mathrm{MC}} = F_\beta^\alpha \mu_\alpha^\beta$. Here, the relevant density matrix for the microcanonical $\Gamma$ ensemble is

$$\mu_\alpha^\beta = \frac{1}{\Omega(E)} \int_{\mathrm{CP}^n} \delta_E(Z^\alpha, \bar{Z}_\alpha) \frac{Z^\beta \bar{Z}_\alpha}{Z^\gamma \bar{Z}_\gamma} \, \mathrm{d}V, \tag{4.43}$$

where we have written $\delta_E(Z^\alpha, \bar{Z}_\alpha)$ for the delta function $\delta(H - E)$ that restricts the distribution to the hypersurface $\mathcal{E}_E$. Note that the microcanonical density operator $\mu_\alpha^\beta$ is the uniform average of the projection operator

$$\Pi_\alpha^\beta = Z^\beta \bar{Z}_\alpha / Z^\gamma \bar{Z}_\gamma, \tag{4.44}$$

over the energy surface $\mathcal{E}_E$.

Now we consider the canonical $\Gamma$ ensemble, for which the random state $\boldsymbol{Z}^\alpha$ is uniformly distributed over each fixed energy surface, while the density function $\rho(E)$ for the energy level $E$ is given as in (4.37). The unconditional expectation for an observable $F_\beta^\alpha$ in the canonical $\Gamma$ ensemble is then given by

$$\langle F \rangle_C = E\left[ \frac{F_\beta^\alpha Z^\beta \bar{Z}_\alpha}{Z^\gamma \bar{Z}_\gamma} \right] = E[E[F_\beta^\alpha \boldsymbol{\Pi}_\alpha^\beta | \mathcal{E}_E]], \tag{4.45}$$

where $E[- \mid \mathcal{E}_E]$ is the expectation *conditioned* on the energy surface $\mathcal{E}_E$, and $\boldsymbol{\Pi}_\alpha^\beta$ is the random projection operator, given by

$$\boldsymbol{\Pi}_\alpha^\beta = \boldsymbol{Z}^\beta \bar{\boldsymbol{Z}}_\alpha / \boldsymbol{Z}^\gamma \bar{\boldsymbol{Z}}_\gamma. \tag{4.46}$$

It follows that $E[F_\beta^\alpha \boldsymbol{\Pi}_\alpha^\beta \mid \mathcal{E}_E]$ is the expectation in the microcanonical distribution with energy $E$, so we can write

$$\langle F \rangle_C = F_\beta^\alpha \int_E \frac{\Omega(E) \mathrm{e}^{-\beta E}}{Q(\beta)} \mu_\alpha^\beta(E) \, \mathrm{d}E, \tag{4.47}$$

where $\mu_\alpha^\beta(E)$ is the microcanonical density operator for energy $E$. In particular, it follows that the quantum canonical ensemble is characterized by the density matrix

$$\rho_\alpha^\beta = \int_E \int_{\mathrm{CP}^n} \frac{\mathrm{e}^{-\beta E} \delta_E(Z^\alpha, \bar{Z}_\alpha) Z^\beta \bar{Z}_\alpha}{Z^\gamma \bar{Z}_\gamma} \, \mathrm{d}V \mathrm{d}E. \tag{4.48}$$

This expression differs from the standard canonical ensemble, for which it is, in effect, assumed that the density matrix is given by a superposition of projection operators onto energy eigenstates, with the familiar Boltzmann weights.

### (f) Spin-1/2 particle

Here, we study the thermal dynamics of systems having two energy levels. For such systems, most physical quantities of interest depend only on the energy difference but not on the actual values of the energies. Therefore, for practical purposes, any two-level system can be viewed as equivalent to a system consisting of a spin-1/2 particle interacting with an external field.

We study the classical situation first. In this case, we have an Ising spin in a constant magnetic field whose strength is $h$. The associated real Hilbert space $\mathcal{H}^2$ is

two dimensional, with orthogonal axes given by the spin-up state $u_\uparrow^a$ and the spin-down state $u_\downarrow^a$. If we view $\mathcal{H}^2$ as an $x$–$y$ plane, then the infinite-temperature ($\beta = 0$) state is the intersection of the segment $S_+^1$ of the unit circle with the line $y = x$. As the temperature decreases, the equilibrium state moves towards $u_\uparrow^a$ or $u_\downarrow^a$ depending on the sign of the magnetic field $h$.

The unnormalized thermal state $\psi^a$ in $\mathcal{H}^2$ can be obtained by solving the equation $\mathrm{d}\psi^a/\mathrm{d}\beta = -\frac{1}{2}H_b^a\psi^b$, with the result

$$\psi^a(\beta) = \mathrm{e}^{\beta h}u_\uparrow^a + \mathrm{e}^{-\beta h}u_\downarrow^a. \tag{4.49}$$

This can be normalized by dividing the right-hand side by the partition function $Q(\beta) = 2\cosh(\beta h)$. We now project this space down to $\mathrm{RP}^1$, which is effectively a circle. Since a circle in $\mathcal{H}^2$ is a double covering of $\mathrm{RP}^1 \sim S^1$, the two eigenstates are mapped to opposite points in $\mathrm{RP}^1$.

Our geometric 'quantization' procedure for this system is as follows. For the given thermal state $\psi^a(\beta)$, we assign a phase factor, and thus obtain a quantum-mechanical state space $\mathrm{CP}^1$, which is viewed as a sphere $S^2$. The north and south poles of the sphere correspond to the two energy eigenstates, and the great circles passing through these two points correspond to unitary equivalent thermal state-space trajectories parametrized by the phase factor. The equator of the sphere, in particular, corresponds to the infinite-temperature states. This circle is of course a 1-torus, in accordance with the general description of the infinite-temperature state manifold given earlier. For the Schrödinger dynamics, we have a rigid rotation of the sphere about the axis that passes through the north and south poles, the two stationary points. This rotation gives rise to a Killing vector field on the sphere, where the angular velocity is given by the strength of the external field $h$. Thus, the temperature determines the latitude on the sphere and the Schrödinger evolution corresponds to a latitudinal circle.

To pursue this in more detail, we introduce a complex Hilbert space with coordinates $Z^\alpha$ ($\alpha = 0, 1$), which we regard as homogeneous coordinates for $\mathrm{CP}^1$. The complex conjugate of $Z^\alpha$ is the 'plane' $\bar{Z}_\alpha$ in $\mathrm{CP}^1$, which in this dimension is a point. The point corresponding to $\bar{Z}_\alpha$ is given by $\bar{Z}^\alpha = \epsilon^{\alpha\beta}\bar{Z}_\beta$, where $\epsilon^{\alpha\beta}$ is the natural symplectic form on the two-dimensional Hilbert space. The formalism in this case is equivalent to the algebra of two-component spinors. In particular, by use of the spinor identity $2X^{[\alpha}Y^{\beta]} = \epsilon^{\alpha\beta}X_\gamma Y^\gamma$, where $X^\alpha\epsilon_{\alpha\beta} = X_\beta$, we obtain

$$\mathrm{d}s^2 = \frac{4Z_\alpha\,\mathrm{d}Z^\alpha\,\bar{Z}_\beta\,\mathrm{d}\bar{Z}^\beta}{(\bar{Z}_\gamma Z^\gamma)^2} \tag{4.50}$$

for the Fubini–Study metric in this situation, and

$$Z_\gamma\,\mathrm{d}Z^\gamma = \mathrm{i}H_{\alpha\beta}Z^\alpha Z^\beta\,\mathrm{d}t \tag{4.51}$$

for the projective Schrödinger equation. The Hamiltonian $H_\alpha^\beta$ has a symmetric representation of the form

$$H_{\alpha\beta} = \frac{2hP_{(\alpha}\bar{P}_{\beta)}}{\bar{P}_\gamma P^\gamma}, \tag{4.52}$$

where $P^\alpha$ and $\bar{P}^\alpha$ correspond to the stationary points. In particular, we have $H_\beta^\alpha P^\beta = hP^\alpha$ and $H_\beta^\alpha \bar{P}^\beta = -h\bar{P}^\alpha$, which follows as a consequence of the identity $\bar{P}_\gamma P^\gamma = -P_\gamma \bar{P}^\gamma$.

By insertion of (4.52) into (4.51) and then (4.51) into (4.50), we deduce that

$$\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 = 16h^2 \operatorname{Prob}[Z^\alpha \to P^\alpha] \operatorname{Prob}[Z^\alpha \to \bar{P}^\alpha], \tag{4.53}$$

where $\operatorname{Prob}[Z^\alpha \to P^\alpha]$ is the transition probability from $Z^\alpha$ to the north pole $P^\alpha$, given by the cross ratio

$$\frac{(\epsilon_{\alpha\beta}Z^\alpha \bar{P}^\beta)(\epsilon_{\gamma\delta}\bar{Z}^\gamma P^\delta)}{(\epsilon_{\alpha\beta}Z^\alpha \bar{Z}^\beta)(\epsilon_{\gamma\delta}P^\gamma \bar{P}^\delta)} = \tfrac{1}{2}(1 + \cos\theta). \tag{4.54}$$

Here, $\theta$ is the distance from $Z^\alpha$ to $P^\alpha$, given by the angular coordinate measured down from the north pole. The complementary probability $\operatorname{Prob}[Z^\alpha \to \bar{P}^\alpha]$ is $\tfrac{1}{2}(1 - \cos\theta)$, and it follows that

$$\frac{\mathrm{d}s}{\mathrm{d}t} = 2h \sin\theta, \tag{4.55}$$

a special case of the Anandan–Aharonov relation noted above. For the evolutionary trajectory we obtain

$$Z^\alpha = \cos(\tfrac{1}{2}\theta)\mathrm{e}^{\mathrm{i}(ht+\phi)}P^\alpha + \sin(\tfrac{1}{2}\theta)\mathrm{e}^{-\mathrm{i}(ht+\phi)}\bar{P}^\alpha, \tag{4.56}$$

where $\theta$ and $\phi$ are the initial coordinates on the sphere for $t = 0$. If appropriate, $Z^\alpha$ can be normalized by setting $P^\alpha \bar{P}_\alpha = 1$. A short calculation then shows that the expectation $E$ of the energy is given by $E = h \cos\theta$, and that the variance of the energy is given by $h^2 \sin^2 \theta$.

Now we are in a position to examine the primitive thermal trajectories associated with the spin-1/2 case. In this case, the thermal equation is

$$\frac{\mathrm{d}Z^\alpha}{\mathrm{d}\beta} = -\tfrac{1}{2}\tilde{H}^\alpha_\beta Z^\beta, \tag{4.57}$$

where

$$\tilde{H}^\alpha_\beta = H^\alpha_\beta - \delta^\alpha_\beta H^\mu_\nu Z^\nu \bar{Z}_\mu / Z^\gamma \bar{Z}_\gamma. \tag{4.58}$$

From (4.52) we then obtain

$$Z^\alpha = \frac{\mathrm{e}^{-(1/2)\beta h + \mathrm{i}\phi}P^\alpha + \mathrm{e}^{(1/2)\beta h - \mathrm{i}\phi}\bar{P}^\alpha}{(\mathrm{e}^{-\beta h} + \mathrm{e}^{\beta h})^{1/2}}, \tag{4.59}$$

where we assume the normalization $P^\alpha \bar{P}_\alpha = 1$. This shows that at infinite temperature ($\beta = 0$), the state lies on the equator, given by $Z^\alpha = \mathrm{e}^{\mathrm{i}\phi}P^\alpha + \mathrm{e}^{-\mathrm{i}\phi}\bar{P}^\alpha$, where $\phi$ lies in the interval from 0 to $2\pi$. The zero-temperature state is obtained by taking the limit $\beta \to \infty$, and we find that $Z^\alpha$ approaches $\bar{P}^\alpha$, the south pole, provided $h > 0$. For the expectation of the energy $E = H^\alpha_\beta Z^\beta \bar{Z}_\alpha$ we have

$$E = h\frac{\mathrm{e}^{-\beta h} - \mathrm{e}^{\beta h}}{\mathrm{e}^{-\beta h} + \mathrm{e}^{\beta h}}, \tag{4.60}$$

which ranges from 0 to $-h$, as $\beta$ ranges from 0 to $\infty$. We note, in particular, that $E$ is independent of the phase angle $\phi$, and that the relation $E = h \tanh(\beta h)$ agrees with the result for a classical spin.

For other observables, this is not necessarily the case, and we have to average over the random state $\boldsymbol{Z}^\alpha$ obtained by replacing $\phi$ with a random variable $\boldsymbol{\Phi}$, having a uniform distribution over the interval $(0, 2\pi)$. Then, for the density matrix $\rho^\alpha_\beta := E[\boldsymbol{Z}^\alpha \bar{\boldsymbol{Z}}_\beta]$, where $E[-]$ is the unconditional expectation, we obtain

$$\rho^\alpha_\beta = \frac{\mathrm{e}^{-\beta h} P^\alpha \bar{P}_\beta - \mathrm{e}^{\beta h} \bar{P}^\alpha P_\beta}{\mathrm{e}^{-\beta h} + \mathrm{e}^{\beta h}}. \tag{4.61}$$

The fact that $\rho^\alpha_\beta$ has trace unity follows from the normalization $Z^\alpha \bar{Z}_\alpha = 1$. For the energy expectation $H^\alpha_\beta \rho^\beta_\alpha$ we then recover (4.60).

Alternatively, we can consider the phase-space volume approach based on the $\Gamma$ ensembles outlined earlier. In the case of a spin-1/2 particle, such as a single electron, it is feasible to consider the canonical $\Gamma$ ensemble, rather than microcanonical distribution. The phase-space volume element in the case of $\mathrm{CP}^1$ is $\mathrm{d}V = 2\pi \sin\theta \, \mathrm{d}\theta$, where $\theta$ is the angle measured from the pole of $\mathrm{CP}^1 \sim S^2$. The state density $\Omega(E)$ here is a constant. Hence, by use of (4.37), along with the energy expectation $E = h\cos\theta$, we deduce that the energy distribution is given by the probability density:

$$\rho(E) \, \mathrm{d}E = \frac{2\pi \sin\theta \exp(-\beta h \cos\theta)}{Q(\beta)} \, \mathrm{d}\theta. \tag{4.62}$$

Since $E \leqslant 0$ and $E^2 \leqslant h^2$, the inverse temperature $\beta$ is positive. Furthermore, we see that $E = 0$ implies $\theta = \pi/2$, the equator of the sphere, which gives infinite temperature ($\beta = 0$), and $E^2 = h^2$ corresponds to $\theta = \pi$, which gives the zero temperature ($\beta = \infty$) state. The partition function is given by

$$Q(\beta) = 4\pi(\beta h)^{-1} \sinh(\beta h), \tag{4.63}$$

from which the energy expectation $E = -\partial \ln Q/\partial\beta$ is

$$E = \beta^{-1} - h\coth(\beta h). \tag{4.64}$$

We note that for positive energies $E > 0$, the temperature takes negative values. If we take $E < 0$ and then flip the direction of the external field $h$, this situation can be achieved in practice. Although the concept of such a negative temperature is used frequently in the study of Laser phenomena, it is a transient phenomenon (Kubo 1951), and is thus not, as such, subject to the principles of equilibrium thermodynamics. We note, incidentally, that the distinct energy–temperature relationships obtained here in equation (4.60) for the standard canonical ensemble, and in (4.64) for the canonical $\Gamma$ ensemble, have qualitatively similar behaviour.

## 5. Discussion

The principal results of this paper are the following. Firstly, we have formulated a projective geometric representation for classical probability states. By specializing on the canonical ensemble of statistical mechanics, we have been able to develop a natural geometric theory of thermal trajectories, which are expressed in terms of a Hamiltonian gradient flow. This flow is then shown to be a special case of a projective automorphism on the state space when it is endowed with the natural $\mathrm{RP}^n$ metric. The same formalism, and essentially the same results, apply also to the grand canonical and the pressure–temperature distributions. We have also been able to clarify certain aspects of the measurement problem for thermal states by introducing the concept of a random state, an idea that links up directly with quantum mechanics.

The quantum-mechanical dynamics of equilibrium thermal states can be studied by consideration of the Hopf-type map $RP^{2n+1} \to CP^n$, which, in the present context, allows one to regard the quantum-mechanical state space as the base space in a fibre manifold that has the structure of an essentially classical thermal state space. The fact that a projective automorphism on a space of constant curvature can be decomposed into two distinct terms suggests the identification of the Killing term with the Schrödinger evolution of the Hamiltonian gradient flow with respect to the symplectic structure, and the other term with thermal evolution of the Hamiltonian gradient flow with respect to the metric; the former gives rise to a linear transformation, while the latter is nonlinear.

A number of open questions remain. First, much of our formulation has been based on the consideration of finite-dimensional examples. The study of phase transitions, however, will require a more extensive treatment of the infinite-dimensional case. Our analysis of the projective automorphism group suggests that the infinite-dimensional case can also be handled in the geometric framework. For most of the paper we have adopted the Schrödinger picture, which has the disadvantage of being inappropriate for relativistic covariance. It would be desirable to reformulate the theory in a manifestly covariant manner to study the case of relativistic fields.

Nevertheless, the present formulation is sufficiently rich in allowing us to speculate on a scenario for the spontaneous symmetry breaking of, say, a pure gauge group, in the infinite-dimensional situation. In such cases, the hypersurface of the parameter space (a curve in the one-parameter case considered here) proliferates into possibly infinite thermally inequivalent hypersurfaces, corresponding to the multiplicity of the ground-state degeneracy, at which the symmetry is spontaneously broken. The hyperline characterizing the proliferation should presumably be called the spinodal line (cf. Brody & Rivier 1995), along which the Riemann curvature of the parameter-space manifold is expected to diverge. Furthermore, a pure thermal state in the 'high-temperature' region should evolve into a mixed state, obtained by averaging over all possible surfaces, by passing through the geometrical singularity (the spinodal boundary) characterizing phase transitions. By a suitable measurement that determines which one of the ground states the system is in, this mixed state will reduce back to a pure state.

In any case, the remarkable advantages of the use of projective space should be stressed. As we have observed, the structure of projective space allows us to identify probabilistic operations with precise geometric relations. One of the problems involved in developing nonlinear (possibly relativistic) generalizations of quantum mechanics concerns their probabilistic interpretation. Formulated in a projective space, such generalizations can be obtained, for example, by replacing the Hamiltonian function by a more general function, or by introducing a more general manifold structure. In this way, the assignment of a suitable probability theory can be approached in an appropriate way. In particular, as we have observed, the canonical ensemble of statistical mechanics has an elegant characterization in projective space—but this is an example of a theory that is highly nonlinear and yet purely probabilistic.

It is also interesting to observe that the nonlinear generalization of quantum mechanics considered by Kibble (1979) and others can be applied in the thermal situation, in the sense that the Hamiltonian function defined on the state manifold can be replaced by a general observable. For such generalizations, it is not yet clear

what physical interpretation can be assigned. Naively, one might expect that by a suitable choice of an observable, the resulting trajectory characterizes some kind of non-equilibrium process. One of the goals of this paper has been to formulate quantum theory at finite temperature in such a way as to allow for the possibility of various natural generalizations. These might include, for example, the stochastic approach to measurement theory, or nonlinear relativistic extensions of standard quantum theory. These generalizations will be pursued further elsewhere.

## References

Anandan, J. & Aharonov, Y. 1990 *Phys. Rev. Lett.* **65**, 1697.

Ashtekar, A. & Schilling, T. A. 1998 In *On Einstein's path* (ed. A. Harvey). Berlin: Springer.

Bratteli, O. & Robinson, D. W. 1979 *Operator algebras and quantum statistical mechanics*, vol. I. New York: Springer.

Bratteli, O. & Robinson, D. W. 1981 *Operator algebras and quantum statistical mechanics*, vol. II. New York: Springer.

Brody, D. C. & Hughston, L. P. 1996 *Phys. Rev. Lett.* **77**, 2851.

Brody, D. C. & Hughston, L. P. 1997 *Phys. Lett.* A **236**, 257.

Brody, D. C. & Hughston, L. P. 1998*a* In *The geometric universe: science, geometry, and the work of Roger Penrose* (ed. S. A. Huggett, L. J. Mason, K. P. Tod, S. T. Tsou & N. M. J. Woodhouse). Oxford University Press.

Brody, D. C. & Hughston, L. P. 1998*b* *Phys. Lett.* A **245**, 73.

Brody, D. C. & Hughston, L. P. 1998*c* *Proc. R. Soc. Lond.* A **454**, 2445.

Brody, D. C. & Hughston, L. P. 1998*d* *J. Math. Phys.* **39**, 6502.

Brody, D. C. & Hughston, L. P. 1999 *J. Math. Phys.* **40**, 12.

Brody, D. & Rivier, N. 1995 *Phys. Rev.* E **51**, 1006.

Ellis, R. S. 1985 *Entropy, large deviations, and statistical mechanics*. New York and Berlin: Springer.

Feshbach, H. 1987 *Phys. Today* (November), p. 10.

Gibbons, G. W. 1992 *J. Geom. Phys.* **8**, 147.

Gibbons, G. W. 1997 *Class. Quant. Grav.* **14**, A155.

Gibbons, G. W. & Pohle, H. J. 1993 *Nucl. Phys.* B **410**, 117.

Gisin, N. 1989 *Helv. Phys. Acta.* **62**, 363.

Haag, R. & Kastler, D. 1964 *J. Math. Phys.* **5**, 848.

Haag, R., Hugenholtz, N. M. & Winnink, M. 1967 *Commun. Math. Phys.* **5**, 215.

Hughston, L. P. 1995 In *Twistor theory* (ed. S. Huggett). New York: Marcel Dekker.

Hughston, L. P. 1996 *Proc. R. Soc. Lond.* A **452**, 953.

Hughston, L. P. & Hurd, T. R. 1983 *Phys. Rep.* **100**, 273.

Hughston, L. P. & Tod, K. P. 1990 *An introduction to general relativity*. Cambridge University Press.

Kibble, T. W. B. 1978 *Commun. Math. Phys.* **64**, 239.

Kibble, T. W. B. 1979 *Commun. Math. Phys.* **65**, 189.

Kittel, C. 1988 *Phys. Today* (May), p. 93.

Kobayashi, S. & Nomizu, K. 1963 *Foundations of differential geometry*, vol. 1. New York: Wiley.

Kobayashi, S. & Nomizu, K. 1969 *Foundations of differential geometry*, vol. 2. New York: Wiley.

Kubo, R. 1951 *Statistical mechanics*. Tokyo: Kyoritsu.

Kubo, R. 1957 *Phys. Soc. Japan* **12**, 570.

Landsman, N. P. 1998 *Mathematical topics between classical and quantum mechanics* New York: Springer.

Mandelbrot, B. 1962 *Ann. Math. Stat.* **33**, 1021.

Mandelbrot, B. 1989 *Phys. Today* (January), p. 71.

Martin, P. C. & Schwinger, J. 1959 *Phys. Rev.* **115**, 1342.

Pearle, P. 1993 *Phys. Rev.* A **48**, 913.

Percival, I. 1994 *Proc. R. Soc. Lond.* A **447**, 189.

Ross, S. 1996 *Stochastic process.* New York: Wiley.

Rugh, H. H. 1997 *Phys. Rev. Lett.* **78**, 772.

Thompson, C. J. 1972 *Mathematical statistical mechanics.* Princeton University Press.

Tomonaga, Y. 1970 *Riemannian geometry.* Tokyo: Kyoritsu.

Weinberg, S. 1989a *Phys. Rev. Lett.* **62**, 485.

Weinberg, S. 1989b *Ann. Phys.* **194**, 336.

Xia, D. X. 1972 *Measure and integration theory on infinite-dimensional spaces.* New York: Academic.