Information geometry in vapour–liquid equilibrium

2009 J. Phys. A: Math. Theor. 42 023001

(http://iopscience.iop.org/1751-8121/42/2/023001)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **42** (2009) 023001 (33pp) doi:10.1088/1751-8113/42/2/023001

TOPICAL REVIEW

Information geometry in vapour–liquid equilibrium

Dorje C Brody[1] **and Daniel W Hook**[2]

[1] Department of Mathematics, Imperial College London, London SW7 2AZ, UK
[2] Blackett Laboratory, Imperial College London, London SW7 2AZ, UK

Received 9 September 2008, in final form 4 November 2008
Published 4 December 2008
Online at stacks.iop.org/JPhysA/42/023001

Abstract

Using the square-root map $p \to \sqrt{p}$ a probability density function p can be represented as a point of the unit sphere $\mathcal{S}$ in the Hilbert space of square-integrable functions. If the density function depends smoothly on a set of parameters, the image of the map forms a Riemannian submanifold $\mathfrak{M} \subset \mathcal{S}$. The metric on $\mathfrak{M}$ induced by the ambient spherical geometry of $\mathcal{S}$ is the Fisher information matrix. Statistical properties of the system modelled by a parametric density function p can then be expressed in terms of information geometry. An elementary introduction to information geometry is presented, followed by a precise geometric characterization of the family of Gaussian density functions. When the parametric density function describes the equilibrium state of a physical system, certain physical characteristics can be identified with geometric features of the associated information manifold $\mathfrak{M}$. Applying this idea, the properties of vapour–liquid phase transitions are elucidated in geometrical terms. For an ideal gas, phase transitions are absent and the geometry of $\mathfrak{M}$ is flat. In this case, the solutions to the geodesic equations yield the adiabatic equations of state. For a van der Waals gas, the associated geometry of $\mathfrak{M}$ is highly nontrivial. The scalar curvature of $\mathfrak{M}$ diverges along the spinodal boundary which envelopes the unphysical region in the phase diagram. The curvature is thus closely related to the stability of the system.

PACS numbers: 02.40.Ky, 02.50.Tt, 05.20.−y, 05.70.Fh, 64.60.A, 64.70.F

1. Statistical geometry

This paper is an overview of the information-geometric description of vapour–liquid phase transitions in equilibrium statistical mechanics. The present section begins with a reasonably self-contained account of the relevant background material on information geometry. As an illustrative example we shall examine in some detail the geometry of the space of Gaussian density functions. The relation between the information measure of Fisher and that of Shannon

and Wiener is also briefly discussed. In later sections these ideas are applied to the information-geometric characterization of the equilibrium properties of noninteracting and interacting gas molecules. The relevant references are provided in the bibliographical notes in section 4, where we also provide a brief and perhaps incomplete history of information geometry.

1.1. From probability to geometry

The concept of 'information geometry' is a simple one which emerged from an attempt to discriminate among different probabilities in statistical analysis. The idea can be sketched as follows. Let $\{p_i\}_{i=1,2,\ldots,N}$ denote a set of probabilities satisfying

$$0 \leqslant p_i \leqslant 1 \qquad \text{and} \qquad \sum_{i=1}^{N} p_i = 1. \tag{1}$$

We introduce the following *square-root map*:

$$p_i \to \xi_i = \sqrt{p_i}. \tag{2}$$

By construction, the square-root probabilities $\{\xi_i\}$ satisfy the normalization condition

$$\sum_{i=1}^{N} \xi_i^2 = 1. \tag{3}$$

If we regard the variables $\{\xi_i\}_{i=1,2,\ldots,N}$ as the coordinates of a vector in an N-dimensional Euclidean space $\mathbb{R}^N$, then the normalization condition (3) implies that the endpoint of the vector $\{\xi_i\}$ lies on the unit sphere $\mathcal{S}$ in $\mathbb{R}^N$. Now suppose that $\{\eta_i\}_{i=1,2,\ldots,N}$ corresponds to a second set of square-root probabilities. Then the vector $\{\eta_i\}$ also lies on the unit sphere in $\mathbb{R}^N$. Hence, we can measure the *relative separation* or *overlap* of two sets of probabilities in terms of the angle

$$\phi = \cos^{-1} \sum_{i=1}^{N} \xi_i \eta_i \tag{4}$$

between the associated square-root probability vectors. The angular separation ϕ clearly vanishes if $\{\xi_i\}$ and $\{\eta_i\}$ are equal. Conversely, if $\{\xi_i\}$ and $\{\eta_i\}$ are orthogonal then ϕ achieves its maximum value $\frac{1}{2}\pi$. The angular separation ϕ defined in (4) is known as the *Bhattacharyya spherical distance*. Note that the cosine square of the spherical distance resembles the transition probability in quantum mechanics modelled on a finite-dimensional Hilbert space.

We turn to the notion of the so-called *statistical geometry*, which arises from the embedding of probability density functions in the Hilbert space of square-integrable functions. In probability theory one typically deals with a probability density function $p(x)$ on, say, the real line $\mathbb{R}$. For the function $p : x \to p(x)$ to represent the density of some random variable X we require that $p(x) \geqslant 0$ for all $x \in \mathbb{R}$ and that $\int_{\mathbb{R}} p(x)\,\mathrm{d}x = 1$. If we consider the *square-root map*

$$p(x) \to \xi(x) = \sqrt{p(x)}, \tag{5}$$

then the function $\xi(x)$ defined in this way belongs to the space $\mathcal{H} = L^2(\mathbb{R})$ of square-integrable functions on the real line. In other words, we *embed* the density functions in Hilbert space via the square-root map (5). In particular, since the square-root density functions satisfy the normalization condition

$$\int_{\mathbb{R}} \xi(x)^2\,\mathrm{d}x = 1, \tag{6}$$

the images of the map lie on the unit sphere $\mathcal{S} \subset \mathcal{H}$.

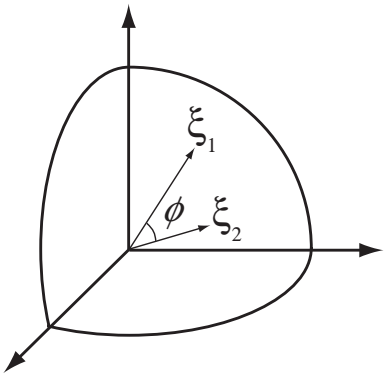

Figure 1. *Bhattacharyya's spherical distance*. Two vectors $\xi_1(x)$ and $\xi_2(x)$, corresponding to a pair of probability density functions $p_1(x)$ and $p_2(x)$, lie on the surface of the positive orthant of the unit sphere $\mathcal{S}$ in Hilbert space. The spherical distance between two unit vectors is given by the angle ϕ defined in equation (7).

The advantage of working in Hilbert space $\mathcal{H}$ rather than the space of density functions is that $\mathcal{H}$ is a vector space endowed with various geometric features that are familiar from other branches of physics, such as quantum mechanics or general relativity.

Suppose we have a pair of density functions $p_1(x)$ and $p_2(x)$, and wish to compare the overlap or separation of these two density functions. If the associated Hilbert space vectors are given respectively by $\xi_1(x)$ and $\xi_2(x)$, then the overlap is measured in terms of the inner product $\int_{\mathbb{R}} \xi_1(x)\xi_2(x)\,\mathrm{d}x$. Since $\|\xi_1\| = \|\xi_2\| = 1$, i.e. both vectors have unit norm, this overlap is given by the cosine of the angular separation. It follows that the Bhattacharyya spherical distance between two square-root density functions is

$$\phi = \cos^{-1} \int_{\mathbb{R}} \xi_1(x)\xi_2(x)\,\mathrm{d}x. \tag{7}$$

This idea is illustrated schematically in figure 1.

1.2. Parametric density and Fisher–Rao geometry

In theoretical statistics one typically deals with a parametric family of probability density functions $p_\theta(x) = p(x|\theta)$. Here θ denotes one or more real parameters. For example, a Gaussian density function is characterized by two parameters, i.e. the mean μ and the variance σ^2. For each value, or set of values, of θ we have the normalization condition $\int_{\mathbb{R}} p_\theta(x)\,\mathrm{d}x = 1$.

In problems of statistical inference, it is often convenient to consider the *log-likelihood function*

$$l_\theta(x) = \ln p(x|\theta). \tag{8}$$

However, in physics it is more natural to work with the square-root density function

$$\xi_\theta(x) = \sqrt{p(x|\theta)}, \tag{9}$$

since, as indicated above, this permits formulation of the problem in a real Hilbert-space context. As before, for each given θ the density function is mapped to a point on the unit sphere $\mathcal{S} \subset \mathcal{H}$ by the prescription (9). If the value of θ is changed, the image under the map in general also varies on $\mathcal{S}$. We assume that the density function is at least twice differentiable

with respect to the parameters. Then as the parameters change continuously, the image point on $\mathcal{S}$ will vary smoothly over a parametric subspace $\mathfrak{M}$ of the sphere $\mathcal{S}$.

Given a parametric subspace $\mathfrak{M} \subset \mathcal{S}$ the metric of the ambient sphere $\mathcal{S}$ induces a Riemannian metric on the subspace in the usual way. This can be seen as follows. Recall that the Hilbert space inner product is defined by

$$\langle \xi, \eta \rangle = \int_{\mathbb{R}} \xi(x)\eta(x)\,\mathrm{d}x. \tag{10}$$

Therefore, if we set

$$\xi(x) = \xi_\theta(x) \qquad \text{and} \qquad \eta(x) = \xi_\theta(x) + \partial_i \xi_\theta(x)\,\mathrm{d}\theta^i, \tag{11}$$

where $\partial_i = \partial/\partial\theta^i$, then the squared distance $\mathrm{d}s^2$ of the difference vector $\xi(x) - \eta(x)$ is given by

$$\mathrm{d}s^2 = \left(\int_{\mathbb{R}} \partial_i \xi_\theta(x) \partial_j \xi_\theta(x)\,\mathrm{d}x \right) \mathrm{d}\theta^i\,\mathrm{d}\theta^j. \tag{12}$$

Before we proceed further with the derivation of the metric, let us introduce the statistical notion of the *Fisher information matrix*, which is usually defined by

$$G_{ij} = \int_{\mathbb{R}} p_\theta(x) \partial_i l_\theta(x) \partial_j l_\theta(x)\,\mathrm{d}x, \tag{13}$$

where $l_\theta(x)$ is the log-likelihood density (8). The Fisher information matrix is important in statistics because it provides a lower bound for the variance of a parameter estimate. Consider, for example, the case of a one-parameter family of density functions. That is, we have a density function $p_\theta(x)$ that depends upon a single unknown parameter θ. The objective is thus to estimate the parameter by performing observations. If $T(x)$ is an unbiased estimator for θ, i.e. if the expectation of $T(x)$ with respect to $p_\theta(x)$ yields θ, then we have

$$\int_{\mathbb{R}} (T(x) - \theta)\xi_\theta(x)^2\,\mathrm{d}x = 0. \tag{14}$$

Differentiating with respect to θ we obtain

$$\int_{\mathbb{R}} (T(x) - \theta)\xi_\theta(x)\partial_\theta \xi_\theta(x)\,\mathrm{d}x = \frac{1}{2}, \tag{15}$$

where $\partial_\theta = \partial/\partial\theta$. By the Schwarz inequality

$$\left(\int_{\mathbb{R}} (T(x) - \theta)\xi_\theta(x)\partial_\theta \xi_\theta(x)\,\mathrm{d}x \right)^2 \leqslant \left(\int_{\mathbb{R}} (T(x) - \theta)^2 \xi_\theta(x)^2\,\mathrm{d}x \right) \left(\int_{\mathbb{R}} (\partial_\theta \xi_\theta(x))^2\,\mathrm{d}x \right) \tag{16}$$

we thus find

$$\int_{\mathbb{R}} (T(x) - \theta)^2 \xi_\theta(x)^2\,\mathrm{d}x \geqslant \frac{1}{4\int_{\mathbb{R}} (\partial_\theta \xi_\theta(x))^2\,\mathrm{d}x}. \tag{17}$$

Note that the left-hand side is the variance of the estimator, whereas the denominator of the right-hand side is the one-parameter form of the Fisher information matrix. The relation (17) provides a lower bound for the variance, and is known as the *Cramér–Rao inequality*. The inequality in (16) is attained only when the two vectors are proportional, that is, $\partial_\theta \xi_\theta(x) = c(T(x) - \theta)\xi_\theta(x)$ for some constant c. By scaling θ we can set $c = \frac{1}{2}$ without loss of generality. Hence, the lower bound of the variance is attained only if the square-root density assumes an exponential form:

$$\xi_\theta(x) = \frac{\exp\left(\frac{1}{2}\theta T(x)\right)}{\left(\int_{\mathbb{R}} \exp(\theta T(x))\,\mathrm{d}x\right)^{1/2}}. \tag{18}$$

The exponential family (18) plays an important role in the applications to statistical mechanics considered below.

In a multi-parameter context the reciprocal of the Fisher information matrix determines lower bounds for the variance in an analogous manner. From the geometrical viewpoint, the significance of the Fisher information matrix is that it defines the induced Riemannian metric on the parametric subspace $\mathfrak{M}$ of the unit sphere $\mathcal{S}$ in $\mathcal{H}$. Specifically, comparing (12) and (13) we see that

$$\mathrm{d}s^2 = \tfrac{1}{4} G_{ij}\, \mathrm{d}\theta^i\, \mathrm{d}\theta^j. \tag{19}$$

The metric $\frac{1}{4}G_{ij}$ on $\mathfrak{M}$ will be referred to as the *Fisher–Rao metric*. The factor of a quarter is purely a matter of convention, and the Fisher–Rao metric is thus given by a quarter of the Fisher information matrix.

1.3. Riemannian structure of the exponential family

We introduce here some elementary concepts in Riemannian geometry that are relevant to the ensuring discussion. We note first that all equilibrium distributions that we consider here are represented in the exponential form

$$p_\theta(x) = q(x) \exp\left(-\sum_i \theta^i H_i(x) - \psi(\theta) \right), \tag{20}$$

where $\{\theta^i\}_{i=1,2,\ldots}$ are parameters, $q(x)$ represents the prescribed equilibrium state at $\theta^i = 0$ for all i, and the functions $\{H_i(x)\}_{i=1,2,\ldots}$ determine the form of the energy. In other words, we shall only consider equilibrium states represented in the *canonical* form. The parameters $\{\theta^i\}$ may include inverse temperature, chemical potential, pressure, magnetic field and so on, whereas the functions $H_i(x)$ may represent system energy, particle number, system volume, magnetization and so on. The variable x ranges over the phase space $\mathbf{\Gamma}$ of the system. The function

$$\psi(\theta) = \ln \int_{\mathbf{\Gamma}} \exp\left(-\sum_i \theta^i H_i(x) \right) q(x)\, \mathrm{d}x \tag{21}$$

determines the overall normalization. We refer to $\psi(\theta)$ as the *thermodynamic potential* of the system. It should be evident by inspection that

$$\begin{aligned} -\frac{\partial \psi}{\partial \theta^i} &= \frac{\int_{\mathbf{\Gamma}} H_i(x) \exp\left(-\sum_i \theta^i H_i(x)\right) q(x)\, \mathrm{d}x}{\int_{\mathbf{\Gamma}} \exp\left(-\sum_i \theta^i H_i(x)\right) q(x)\, \mathrm{d}x} \\ &= \int_{\mathbf{\Gamma}} H_i(x)\, p_\theta(x)\, \mathrm{d}x. \end{aligned} \tag{22}$$

In other words, the first derivative of $\psi(\theta)$ with respect to θ^i determines the expectation value of $H_i(x)$ in the equilibrium state (20). As we shall indicate below, analogous calculations show that higher derivatives of the thermodynamic potential $\psi(\theta)$ determine higher moments of the functions $\{H_i(x)\}$.

If the equilibrium density function assumes the form (20), then the expressions for the corresponding Fisher–Rao metric and the coefficients of the associated metric connection on the statistical manifold $\mathfrak{M}$ are simplified. Let us state the results first.

Proposition 1. *For a density function of the exponential form (20) the Fisher–Rao metric G_{ij} and the Christoffel symbols $\Gamma_{ijk} = G_{il}\Gamma^l_{\ jk}$ are given, respectively, by*

$$G_{ij} = \partial_i \partial_j \psi(\theta) \tag{23}$$

and

$$\Gamma_{ijk} = \tfrac{1}{2}\partial_i\partial_j\partial_k\psi(\theta), \tag{24}$$

in terms of the canonical parameterization $\{\theta^i\}$ *on* $\mathfrak{M}$*, where* $\partial_i = \partial/\partial\theta_i$.

The Christoffel symbols characterize the geodesics on $\mathfrak{M}$. Specifically, to find the shortest path from a to b on $\mathfrak{M}$ we consider the variational problem:

$$\delta\int_a^b \mathrm{d}s = 0. \tag{25}$$

From (19) we find

$$4\delta\,\mathrm{d}s^2 = \mathrm{d}\theta^i\,\mathrm{d}\theta^k\frac{\partial G_{ik}}{\partial\theta^l}\delta\theta^l + 2G_{ik}\,\mathrm{d}\theta^i\,\mathrm{d}\delta\theta^k. \tag{26}$$

Bearing in mind that the right-hand side of (26) equals $8\mathrm{d}s\,\delta\,\mathrm{d}s$, we obtain

$$\int_a^b\left[\frac{1}{2}\frac{\mathrm{d}\theta^i}{\mathrm{d}s}\frac{\mathrm{d}\theta^k}{\mathrm{d}s}\frac{\partial G_{ik}}{\partial\theta^k}\delta\theta^l + G_{ik}\frac{\mathrm{d}\theta^i}{\mathrm{d}s}\frac{\mathrm{d}\delta\theta^k}{\mathrm{d}s}\right]\mathrm{d}s = 0. \tag{27}$$

Integrating the second term in the integrand by parts and writing $u^i = \mathrm{d}\theta^i/\mathrm{d}s$ we see that (27) reduces to

$$\int_a^b\left[\frac{1}{2}u^iu^k\frac{\partial G_{ik}}{\partial\theta^l} - \frac{\mathrm{d}}{\mathrm{d}s}(G_{il}u^i)\right]\delta\theta^l\,\mathrm{d}s = 0. \tag{28}$$

Since this must hold for arbitrary $\delta\theta^l$ we have

$$\frac{1}{2}u^iu^k\frac{\partial G_{ik}}{\partial\theta^l} - \frac{\mathrm{d}}{\mathrm{d}s}(G_{il}u^i) = 0. \tag{29}$$

Writing

$$\Gamma^i_{kl} = \frac{1}{2}G^{im}\left(\frac{\partial G_{mk}}{\partial\theta^l} + \frac{\partial G_{ml}}{\partial\theta^k} - \frac{\partial G_{kl}}{\partial\theta^m}\right) \tag{30}$$

for the Christoffel symbol we find that (29) can be expressed in the form

$$\frac{\mathrm{d}^2\theta^i}{\mathrm{d}s^2} + \Gamma^i_{kl}\frac{\mathrm{d}\theta^k}{\mathrm{d}s}\frac{\mathrm{d}\theta^l}{\mathrm{d}s} = 0. \tag{31}$$

This is the *geodesic equation* that determines the shortest paths on $\mathfrak{M}$. Owing to their nonlinearity, geodesic equations do not generally admit elementary analytic solutions, although in some cases one can solve (31) in the closed form, as in the Gaussian example discussed below.

Proof of proposition 1. From (20) we have

$$\partial_i l_\theta(x) = -(H_i(x) + \partial_i\psi(\theta)). \tag{32}$$

On the other hand, differentiating the normalization condition

$$\int_\Gamma p_\theta(x)\,\mathrm{d}x = 1 \tag{33}$$

once with respect to θ^i and using $\partial_i p_\theta(x) = p_\theta(x)\partial_i l_\theta(x)$ we obtain

$$\int_\Gamma p_\theta(x)(H_i(x) + \partial_i\psi(\theta))\,\mathrm{d}x = 0, \tag{34}$$

whence it follows that the expectation of $H_i(x)$ with respect to $p_\theta(x)$ is given by $-\partial_i\psi(\theta)$, as shown in (22). Differentiating (34) with respect to θ^j, we find that

$$-\int_\Gamma p_\theta(x)(H_i(x) + \partial_i\psi(\theta))(H_j(x) + \partial_j\psi(\theta))\,\mathrm{d}x + \partial_i\partial_j\psi(\theta) = 0. \tag{35}$$

In view of (32) and (13), this implies that the Fisher–Rao metric G_{ij} is given by (23). From (30) we have

$$\Gamma_{ikl} = \tfrac{1}{2}(\partial_l G_{ik} + \partial_k G_{il} - \partial_i G_{kl}), \tag{36}$$

but since the metric is given by (23) we immediately deduce expression (24) for the Christoffel symbols. □

It is important to note that if we choose an alternative parameterization for $\mathfrak{M}$, then the components of the metric tensor and the Christoffel symbol cannot be calculated using the simple expressions given in proposition 1, and we must use the defining equations (13) and (30) to determine these quantities. Also note that the metric tensor is the covariance matrix of the functions $\{H_i(x)\}$, whereas the components of Γ_{ijk} are third-order cross-moments of $\{H_i(x)\}$.

In terms of the Christoffel symbols Γ^i_{jk} the Riemann curvature tensor R^i_{ljk} can be expressed as

$$R^i_{ljk} = \partial_k \Gamma^i_{lj} - \partial_j \Gamma^i_{lk} - \Gamma^i_{jh}\Gamma^h_{lk} + \Gamma^i_{kh}\Gamma^h_{lj}. \tag{37}$$

The Riemann tensor encodes the information concerning the local geometry of $\mathfrak{M}$, and is related to the parallel transport of vectors on $\mathfrak{M}$. In particular, if we define the covariant derivative by

$$\nabla_j A_i = \frac{\partial A_i}{\partial \theta^j} - \Gamma^k_{ij} A_k, \tag{38}$$

then the commutator of the covariant derivatives defines the Riemann tensor:

$$\nabla_j \nabla_k A_i - \nabla_k \nabla_j A_i = A_l R^l_{ijk}. \tag{39}$$

The symmetry properties of the Riemann tensor can be derived by lowering the index with the metric and writing $R_{ijkl} = G_{im} R^m_{jkl}$. Specifically, this is given by

$$R_{ijkl} = \frac{1}{2}\left(\frac{\partial^2 G_{il}}{\partial\theta^j \partial\theta^k} + \frac{\partial^2 G_{jk}}{\partial\theta^i \partial\theta^l} - \frac{\partial^2 G_{ik}}{\partial\theta^j \partial\theta^l} - \frac{\partial^2 G_{jl}}{\partial\theta^i \partial\theta^k} \right) + G_{nm}\big(\Gamma^n_{jk}\Gamma^m_{il} - \Gamma^n_{jl}\Gamma^m_{ik}\big), \tag{40}$$

whence it follows that $R_{ijkl} = R_{klij}$. Along with the relation $R^i_{jkl} = -R^i_{jlk}$ that follows from (37) we find that $R_{ijkl} = -R_{ijlk} = -R_{jikl}$. Therefore, the components of R_{ijkl} vanish when $i = j$ or $k = l$. In the case of a two-dimensional manifold $\mathfrak{M}$ the only nonvanishing components of the Riemann tensor are given by

$$R_{1212} = -R_{1221} = -R_{2112} = R_{2121}. \tag{41}$$

In other words, R_{1212} is the sole independent component in two dimensions.

Given a Riemann tensor R^i_{jkl} we define the associated Ricci tensor by the symmetric expression

$$R_{jl} = R^k_{jkl}. \tag{42}$$

Equivalently, we can write $R_{jl} = G^{ik} R_{ijkl}$. A further contraction with the metric defines the scalar curvature:

$$R = G^{jl} R_{jl}. \tag{43}$$

We call (43) the *Ricci scalar curvature*. If the Ricci tensor R_{jl} is proportional to the metric tensor G_{jl} then the manifold $\mathfrak{M}$ is called an *Einstein space*. This is because the metric of such a space satisfies the vacuum Einstein equation

$$R^i_k - \tfrac{1}{2}\delta^i_k R = 0, \tag{44}$$

where $R^i_k = G^{il}R_{lk}$. The significance of the Einstein equation in statistics or statistical mechanics can be seen if we relax the normalization condition on the square-root density function $\xi(x)$ and thus eliminate the physically irrelevant degree of freedom associated with the norm $\|\xi(x)\|$. Specifically, if $T(x)$ represents an observable function on the phase space $\mathbf{\Gamma}$ then its expectation value in the generic 'state' $\xi(x)$ is

$$\langle T(x)\rangle = \frac{\int_\Gamma \xi(x)T(x)\xi(x)\,\mathrm{d}x}{\int_\Gamma \xi(x)^2\,\mathrm{d}x}. \tag{45}$$

Evidently, the expectation so defined is invariant under the scale transformation $\xi(x) \to \lambda\xi(x)$, where λ is an arbitrary nonzero number. Thus, all the relevant statistical information is encoded in the direction of the vector $\xi(x) \in \mathcal{H}$, irrespective of its length. Via the identification $\xi(x) \sim \lambda\xi(x)$ we obtain a space of rays in $\mathcal{H}$, which is known as the real projective Hilbert space. Suppose now that we consider the Einstein equation (44) for the metric on the projective Hilbert space. Then there is a unique solution which is induced by the infinitesimal form of the Bhattacharyya spherical distance (7). Specifically, this is obtained by setting $\xi_1(x) = \xi(x)$ and $\xi_2(x) = \xi(x) + \mathrm{d}\xi(x)$ in

$$\cos^2\phi = \frac{\left(\int_{\mathbb{R}} \xi_1(x)\xi_2(x)\,\mathrm{d}x\right)^2}{\left(\int_{\mathbb{R}} \xi_1(x)\xi_1(x)\,\mathrm{d}x\right)\left(\int_{\mathbb{R}} \xi_2(x)\xi_2(x)\,\mathrm{d}x\right)}, \tag{46}$$

Taylor expanding each side, and retaining terms of quadratic order. Then with the notation of (10) we can write

$$\phi^2 = \frac{\langle \mathrm{d}\xi, \mathrm{d}\xi\rangle}{\langle \xi, \xi\rangle} - \frac{\langle \xi, \mathrm{d}\xi\rangle^2}{\langle \xi, \xi\rangle^2}, \tag{47}$$

which defines a metric on the projective Hilbert space. It follows that the Einstein equation uniquely determines the probabilistic properties of the space of densities.

For a statistical manifold $\mathfrak{M}$ associated with a distribution of the exponential type (20), the Riemann tensor assumes a simple form because the first two terms in (37) cancel, and only the contractions of the Christoffel symbols remain.

The examples of statistical mechanical systems considered here are parameterized by a pair of external variables, so that the statistical manifold $\mathfrak{M}$ is two dimensional. In this case, the expression for the scalar curvature admits further simplifications. Specifically, we have the following:

Proposition 2. *In terms of the canonical parameterization (θ^1, θ^2), the scalar curvature of a two-dimensional statistical model corresponding to the density function (20) is given by the determinant*

$$R = -\frac{1}{2G^2}\begin{vmatrix} \psi_{11}(\theta) & \psi_{12}(\theta) & \psi_{,22}(\theta) \\ \psi_{111}(\theta) & \psi_{112}(\theta) & \psi_{122}(\theta) \\ \psi_{112}(\theta) & \psi_{122}(\theta) & \psi_{222}(\theta) \end{vmatrix}, \tag{48}$$

where $G = \det(G_{ij})$ is the determinant of the Fisher–Rao metric, and where $\psi_{12}(\theta) = \partial^2\psi(\theta)/\partial\theta^1\partial\theta^2$, $\psi_{112}(\theta) = \partial^3\psi(\theta)/\partial\theta^1\partial\theta^1\partial\theta^2$, and so on.

Proof. Recall that the scalar curvature is determined by the contraction

$$R_{ijkl} = (\Gamma_{kmi}\Gamma_{jln} - \Gamma_{kmj}\Gamma_{iln})G^{mn}. \tag{49}$$

Substituting (24) and the inverse of (23) into (49), we obtain (48) after some rearrangement of terms. □

1.4. Geometry of Gaussian distributions

As an elementary illustrative example, consider the Gaussian (normal) distribution $N(\mu, \sigma)$ on the real line $\mathbb{R}$ with mean μ and standard deviation $\sigma > 0$. For the parameterized density function we have

$$p(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right). \tag{50}$$

The normal density function can be rewritten in the canonical form

$$p(x|\theta) = \exp[-\theta^1 x^2 - \theta^2 x - \psi(\theta)], \tag{51}$$

where

$$\theta^1 = \frac{1}{2\sigma}, \qquad \theta^2 = -\frac{\mu}{\sigma^2} \qquad \text{and} \qquad \psi(\theta) = \frac{1}{8}\left(\frac{\theta^2}{\theta^1}\right)^2 - \ln\left(\frac{\sqrt{2\pi}}{2\theta^1}\right). \tag{52}$$

By differentiating $\psi(\theta)$ with respect to the parameters $\{\theta^i\}$ we can determine the components of the metric tensor $G_{ij}(\theta)$ in the coordinate system specified by the canonical parameterization $\{\theta^i\}$.

Alternatively, we may regard the mean μ and standard deviation σ as coordinates on the statistical manifold $\mathfrak{M}$. In terms of the parameters (μ, σ) the metric does not admit a simple representation (23), and we must perform the Gaussian integration in the defining relation (13). The line element then becomes

$$\mathrm{d}s^2 = \frac{1}{\sigma^2}(\mathrm{d}\mu^2 + 2\mathrm{d}\sigma^2), \tag{53}$$

which is defined on the upper-half plane $-\infty < \mu < \infty$ and $0 < \sigma < \infty$. Since the metric G_{ij} is diagonal in these coordinates, it is easily inverted and we obtain

$$G^{ij} = \sigma^2 \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}. \tag{54}$$

A short calculation shows that the Christoffel symbols are given by

$$\Gamma^1_{12} = \Gamma^1_{21} = -2\Gamma^2_{11} = \Gamma^2_{22} = -\frac{1}{\sigma}, \qquad \Gamma^1_{11} = \Gamma^1_{22} = \Gamma^2_{12} = \Gamma^2_{21} = 0. \tag{55}$$

Since the inverse of the metric tensor (54) is diagonal, we need to only determine the diagonal components of the Ricci tensor in order to calculate the scalar curvature (that is, the off-diagonal elements of the Ricci tensor vanish). These are

$$R_{11} = -\frac{1}{2\sigma^2} \qquad \text{and} \qquad R_{22} = -\frac{1}{\sigma^2}, \tag{56}$$

respectively. Hence, the resulting geometry is that of a hyperbolic space, which is a homogeneous manifold of constant negative curvature:

$$R = -1. \tag{57}$$

This space has many interesting properties. For example, the geodesic equations (31) characterizing trajectories of shortest paths on $\mathfrak{M}$ are determined by the equations

$$\frac{\mathrm{d}^2\mu(s)}{\mathrm{d}s^2} - 2\frac{1}{\sigma(s)}\frac{\mathrm{d}\mu(s)}{\mathrm{d}s}\frac{\mathrm{d}\sigma(s)}{\mathrm{d}s} = 0 \tag{58}$$

and

$$\frac{\mathrm{d}^2\sigma(s)}{\mathrm{d}s^2} + \frac{1}{\sigma(s)}\left\{\frac{1}{2}\left(\frac{\mathrm{d}\mu(s)}{\mathrm{d}s}\right)^2 - \left(\frac{\mathrm{d}\sigma(s)}{\mathrm{d}s}\right)^2\right\} = 0. \tag{59}$$

Since $\sigma > 0$ we can divide (58) and (59) by σ, obtaining

$$\left(\frac{\mu'}{\sigma}\right)' - \frac{\mu'}{\sigma}\frac{\sigma'}{\sigma} = 0 \qquad \text{and} \qquad \left(\frac{\sigma'}{\sigma}\right)' + \frac{1}{2}\frac{\mu'}{\sigma}\frac{\mu'}{\sigma} = 0, \tag{60}$$

where the prime indicates $\mathrm{d}/\mathrm{d}s$. It follows that

$$\left(\left(\frac{\mu'}{\sigma}\right)^2 + 2\left(\frac{\sigma'}{\sigma}\right)^2\right)' = 0, \tag{61}$$

and hence that

$$\left(\frac{\mu'}{\sigma}\right)^2 + 2\left(\frac{\sigma'}{\sigma}\right)^2 = v^2, \tag{62}$$

where $v \geqslant 0$ is a constant. On the other hand, if we define $X = \mu'/\sigma$ then from the first equation of (60) we have $X\sigma' - X'\sigma = 0$, or equivalently $(X/\sigma)' = 0$, and thus $\mu'/\sigma = c\sigma$, where c is a constant. Substituting this into (62) we deduce that

$$c^2\sigma^2 + \left(\frac{\sigma'}{\sigma}\right)^2 = v^2, \tag{63}$$

where we have rescaled the integration constants, i.e. $c \to \sqrt{2}c$ and $v \to \sqrt{2}v$. There are now two cases to consider, depending on whether c is zero or nonzero.

If $c = 0$, then μ is constant, and $\sigma' = v\sigma$, so that $\sigma(s) = a\,\mathrm{e}^{vs}$ for some constants a and v such that $a > 0$ and $v > 0$. This represents a straight line parallel to the σ-axis in the μ–σ plane. If $c \neq 0$, then $c\sigma \leqslant v$ from (63), hence $\sigma(s) = \frac{v}{c}\sin\gamma(s)$ for some $\gamma(s)$. Substituting this into (63) we obtain

$$\left(\frac{\mathrm{d}\gamma}{\mathrm{d}s}\right)^2 = v^2\sin^2\gamma(s). \tag{64}$$

Since $\gamma' \neq 0$, $\gamma(s)$ is monotonic and thus invertible. We may assume, without loss of generality, that $\gamma' > 0$ so that (64) implies $\mathrm{d}\gamma/\mathrm{d}s = v\sin\gamma$. Using $\sigma(s) = \frac{v}{c}\sin\gamma(s)$ we find that

$$\frac{\sigma'}{\sigma} = \frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\gamma}\frac{\mathrm{d}\gamma}{\mathrm{d}s} = v\cos\gamma, \tag{65}$$

and further, using (62) with rescaled c and v we obtain

$$\mu'(s) = v\sigma(s)\sin\gamma(s). \tag{66}$$

If we regard $s = s(\gamma)$ as parameterized by γ we can write

$$\mu(s) = \int \frac{\mathrm{d}\mu}{\mathrm{d}\gamma}\,\mathrm{d}\gamma = \int \mu'\frac{\mathrm{d}s}{\mathrm{d}\gamma}\,\mathrm{d}\gamma = \frac{v}{c}\int \sin\gamma\,\mathrm{d}\gamma = -\frac{v}{c}\cos\gamma + b, \tag{67}$$

where b is an integration constant.

To summarize, if we regard $\gamma(s)$ as the independent parameter, then the solutions to the geodesic equations for the Gaussian family of densities (50) are

$$\mu(s) = -\frac{v}{c}\cos\gamma(s) + b \qquad \text{and} \qquad \sigma(s) = \frac{v}{c}\sin\gamma(s). \tag{68}$$

These equations represent half-circles on the μ–σ plane centred on the μ axis ($\sigma = 0$) with radius v/c. In figure 2 we sketch examples of geodesic curves for the Gaussian family, in the case $c \neq 0$.

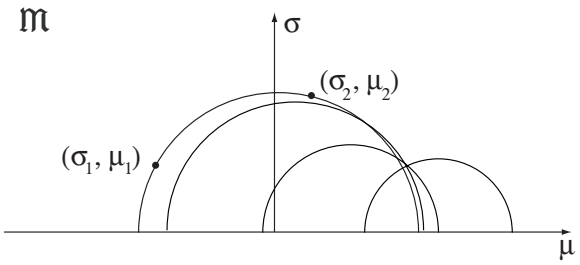

Figure 2. *Geodesic curves for Gaussian distributions*. The statistical manifold $\mathfrak{M}$ in this case is the upper half plane parameterized by μ and σ. We have $-\infty < \mu < \infty$ and $0 < \sigma < \infty$. The shortest path joining the two normal distributions $N(\mu_1, \sigma_1)$ and $N(\mu_2, \sigma_2)$ is given by the unique semi-circular arc through the given two points and centred on the boundary line $\sigma = 0$.

If follows from the solutions to the geodesic equations that the separation of a pair of normal distributions $N(\mu_1, \sigma_1)$ and $N(\mu_2, \sigma_2)$ is given by

$$D(\rho_1, \rho_2) = \frac{1}{\sqrt{2}} \log \frac{1 + \delta_{1,2}}{1 - \delta_{1,2}}, \tag{69}$$

where the function $\delta_{1,2}$ defined by

$$\delta_{1,2} = \sqrt{\frac{(\mu_2 - \mu_1)^2 + 2(\sigma_2 - \sigma_1)^2}{(\mu_2 - \mu_1)^2 + 2(\sigma_2 + \sigma_1)^2}} \tag{70}$$

lies between 0 and 1. These results follow directly from the fact that the geodesics are semi-circular arcs centred on the boundary line $\sigma = 0$ (this line itself is not part of the manifold $\mathfrak{M}$ because $\sigma > 0$). In the exceptional case when $\mu_1 = \mu_2$, the geodesic is a straight line $\mu =$ constant, and

$$D(\rho_1, \rho_2) = \frac{1}{\sqrt{2}} \left| \log \frac{\sigma_1}{\sigma_2} \right|. \tag{71}$$

The above example illustrates how various geometric aspects of a statistical manifold $\mathfrak{M}$ can be investigated in a systematic manner. It is interesting to note, in particular, that the Gaussian distributions define an elementary hyperbolic geometry with constant negative curvature. Before examining various geometric characterizations of ideal and interacting gases in thermal equilibrium, let us discuss the relation between the Fisher–Rao distance measure and various measures of entropy, a topic of some interest.

1.5. From entropy to Fisher information

We have observed how the notion of information geometry arises from the Fisher information matrix commonly used in statistical analysis. On the other hand, the term 'information' often suggests the concept of entropy, rather than the Fisher matrix. Indeed, the entropy concept is essential in thermodynamics and statistical mechanics. Therefore, it would be appropriate to clarify the interrelation between these two notions of information.

We shall discuss entropy in a fairly general context, and consider as before a parametric family of probability density functions $p(x|\theta)$ which we assume to be defined, say, on the real line $\mathbb{R}$, where $\theta = \{\theta^i\}$. Then with respect to any twice-differentiable concave function $\varphi(p)$ we define the associated *entropy functional* by the expression

$$H_\varphi(p) = \int_{\mathbb{R}} \varphi[p(x|\theta)]\, \mathrm{d}x. \tag{72}$$

Now, if f represents a vector in the tangent space (at p) of the manifold of density functions, then the derivative of the entropy H_φ at p in the direction f is defined by

$$\begin{aligned} \mathrm{d}H_\varphi(p;\, f) &= \frac{\mathrm{d}}{\mathrm{d}s} H_\varphi(p+sf)\bigg|_{s=0} \\ &= \int_{\mathbb{R}} \varphi'[p(x)] f(x)\, \mathrm{d}x, \end{aligned} \tag{73}$$

where $\varphi'(p) = \mathrm{d}\varphi(p)/\mathrm{d}p$. Similarly, if f and g are two vectors in the tangent space at p, we define the Hessian of H_φ by

$$\mathrm{d}^2 H_\varphi(p;\, f, g) = \int_{\mathbb{R}} \varphi''[p(x)] f(x) g(x)\, \mathrm{d}x. \tag{74}$$

The corresponding quadratic form is

$$\triangle_f H_\varphi(p) = 4\, \mathrm{d}^2 H_\varphi(p;\, f,\, f), \tag{75}$$

or equivalently,

$$\triangle_f H_\varphi(p) = 4 \int_{\mathbb{R}} \varphi''[p(x)] f^2(x)\, \mathrm{d}x, \tag{76}$$

where the factor of 4 here is purely conventional. The concavity of φ then implies that

$$-\triangle_f H_\varphi(p) \geqslant 0. \tag{77}$$

In particular, if we chose f to be $\partial_i p$, where $\partial_i = \partial/\partial\theta^i$, then we have

$$\triangle_{\partial p} H_\varphi(p) = 4 \int_{\mathbb{R}} \varphi''[p(x|\theta)]\, (\partial_i p(x|\theta))^2\, \mathrm{d}x. \tag{78}$$

Thus far, we have not specified the form of the function φ, except for the requirements of concavity and twice differentiability. As a special case, let us consider the one-parameter family of concave functions

$$\varphi_\alpha(z) = (\alpha - 1)^{-1}(z - z^\alpha), \tag{79}$$

where $\alpha > 0$. This determines a one-parameter family of entropies $H_\alpha(p)$ given by

$$H_\alpha(p) = \frac{1}{\alpha - 1}\left(1 - \int p^\alpha(x)\, \mathrm{d}x\right). \tag{80}$$

Note that when $\alpha = 1$ we have $\varphi_1(z) = -z \ln z$ and hence

$$H_1(p) = -\int p(x) \ln p(x)\, \mathrm{d}x. \tag{81}$$

In other words, we recover the expression for the familiar Shannon–Wiener entropy in the limit $\alpha \to 1$. In the general case, expression (80) defines the *Havrda–Charvát entropy* (also known as the α-order entropy), which is related to the well-known *Rényi entropy* $R_\alpha(p)$ as follows:

$$R_\alpha(p) = \frac{1}{1 - \alpha} \ln(1 + (1 - \alpha) H_\alpha(p)). \tag{82}$$

In particular, R_α is monotonic in H_α.

Also, choosing φ_α as in (79) we find that

$$\mathrm{d}s_\alpha^2 = -\frac{1}{4\alpha} \triangle_{\partial \vec{p}} H_\varphi(p) = -\frac{1}{\alpha}\, \mathrm{d}^2 H_\varphi(p;\, \partial_i p,\, \partial_j p)\, \mathrm{d}\theta^i\, \mathrm{d}\theta^j \tag{83}$$

is positive definite and defines a Riemannian metric. We can summarize this as follows.

Proposition 3 (Burbea–Rao metric)**.** *The coefficients of the differential metric*

$$\mathrm{d}s_\alpha^2 = G_{ij}^{(\alpha)}\,\mathrm{d}\theta^i\,\mathrm{d}\theta^j \tag{84}$$

associated with the Hessian of the α-order entropy (80) are

$$G_{ij}^{(\alpha)} = \int p^\alpha(\partial_i \ln p)(\partial_j \ln p)\,\mathrm{d}x. \tag{85}$$

In particular, when $\alpha = 1$, $G_{ij}^{(\alpha)}$ reduces to the Fisher–Rao metric.

Proof. The expression in (85) follows at once from (83) by virtue of the relation $\varphi_\alpha''(p) = -\alpha p^{\alpha-2}$. □

We find, therefore, that the so-called α-order entropy metric is closely related to the Fisher–Rao geometry of the statistical manifold. In addition, there is another significant relationship between the derivative of the entropy and the α-order Fisher information matrix. This can be established as follows. From the defining equation (80) we have

$$\partial_i\partial_j H_\alpha(p) = -\alpha\int_{\mathbb{R}} p^{\alpha-2}(\partial_i p)(\partial_j p)\,\mathrm{d}x - \frac{\alpha}{\alpha-1}\int_{\mathbb{R}} p^{\alpha-1}\partial_i\partial_j p\,\mathrm{d}x, \tag{86}$$

and therefore we deduce that

$$G_{ij}^{(\alpha)} = -\frac{1}{\alpha}\partial_i\partial_j H_\alpha(p) + \frac{1}{\alpha-1}\int_{\mathbb{R}} p^{\alpha-1}\partial_i\partial_j p\,\mathrm{d}x. \tag{87}$$

For the canonical distribution (20), the limit $\alpha \to 1$ of this relation yields the Shannon–Wiener entropy

$$H_1(p) = \sum_i \theta^i \int p(x|\theta)H_i(x)\,\mathrm{d}x + \psi(\theta). \tag{88}$$

In other words, the thermodynamic potential and the entropy are related by a Legendre transformation. Consequently, the Fisher–Rao geometry and the geometry arising from the Hessian of the Shannon–Wiener entropy are related by the general theory of Legendre transforms.

2. Classical ideal gas

We shall now characterize the geometry of the statistical manifold that arises from the equilibrium distribution of a gas of noninteracting particles. Although this system displays no phase transition, the analysis presented here will provide an enlightening contrast with the results of section 3 where we shall examine the geometry of the van der Waals gas, which does exhibit a liquid–vapour transition.

2.1. Partition function in P–T distribution

To elucidate the geometrical representation of gaseous systems in statistical mechanics, we begin our analysis with a system of noninteracting identical particles in the absence of potential energy. Physically, this system corresponds to a classical ideal gas immersed in a heat bath. As we shall show below, not only the Riemann curvature but also the geodesic equations for this system can be solved exactly. We consider, in particular, a pressure–temperature (P– T) distribution (also known as the Boguslavski distribution) of the form

$$p(H, V|\alpha, \beta) = \frac{\exp(-\beta H - \alpha V)}{Z(\alpha, \beta)} \tag{89}$$

defined on the phase space $\mathbf{\Gamma}$ of a system of noninteracting particles. Here the partition function $Z(\alpha, \beta)$ is determined by the phase-space and volume integral

$$Z(\alpha, \beta) = \frac{1}{N!h^{3N}} \int_0^\infty \left(\int_{\mathbf{\Gamma}} \exp(-\beta H)\, \mathrm{dq}\, \mathrm{dp} \right) \exp(-\alpha V)\, \mathrm{d}V. \tag{90}$$

As usual, we have $\beta = 1/k_B T, \alpha = P/k_B T$, where P denotes the pressure, h the Planck constant and N the number of particles. The Hamiltonian H is just the free particle kinetic energy

$$H = \sum_{i=1}^{N} \frac{\mathrm{p}_i^2}{2m}. \tag{91}$$

Thus, we consider a closed system of noninteracting gas molecules immersed in a heat bath at inverse temperature β and effective pressure α. Since the system is in contact with a bath at fixed temperature and pressure, the system energy and volume fluctuate. In thermal equilibrium, the distribution of these variables is determined by (89). For a real gas, the constituent particles inevitably interact. Nevertheless, the ideal gas represented by the distribution (89) adequately characterizes the properties of a real gas at high temperature or low density, where the effects of inter-particle interactions can be neglected.

Comparing (89) and (20) we observe that the thermodynamic potential is given by $\psi(\alpha, \beta) = \ln Z(\alpha, \beta)$. Therefore, to determine the Fisher–Rao metric (23) we must perform the integration (90). Noting the fact that each q-integration in (90) gives the volume V of the system, one obtains the partition function

$$Z(\alpha, \beta) = \left(\frac{2\pi m}{h^2 \beta} \right)^{3N/2} \alpha^{-(N+1)}. \tag{92}$$

This follows from the fact that

$$\int_{\mathbf{\Gamma}} \exp(-\beta H)\, \mathrm{dq}\, \mathrm{dp} = \int_{\mathrm{q}} \left(\int_{\mathrm{p}} \mathrm{e}^{-\beta \sum_{i=1}^{N} \frac{\mathrm{p}_i^2}{2m}} \prod_{i=1}^{N} \mathrm{dp}_i \right) \prod_{i=1}^{N} \mathrm{dq}_i \tag{93}$$

is just a product of Gaussian integrals, and the identity

$$\frac{1}{N!} \int_0^\infty V^N \mathrm{e}^{-\alpha V} \mathrm{d}V = \alpha^{-(N+1)} \tag{94}$$

that holds for $\alpha > 0$.

Note that the partition function $z(\beta, V)$ in the canonical ensemble is

$$\begin{aligned} z(\beta, V) &= \frac{1}{N!h^{3N}} \int_{\mathbf{\Gamma}} \exp(-\beta H)\, \mathrm{dq}\, \mathrm{dp} \\ &= \frac{1}{N!} \left(\frac{2\pi m}{h^2 \beta} \right)^{3N/2} V^N, \end{aligned} \tag{95}$$

from which one can calculate the *Helmholtz free energy*

$$F(\beta, V) = -k_B T \ln z(\beta, V) \tag{96}$$

and thus obtain the equation of state

$$P = -\left(\frac{\partial F}{\partial V} \right)_\beta = N k_B T / V \tag{97}$$

satisfied by a classical ideal gas.

2.2. *Curvature and geodesics for ideal gas*

Expression (92) for the partition function clearly shows that the Riemannian geometry of the statistical model $\mathfrak{M}$ associated with the classical ideal gas depends upon the number N of particles. Although finite size effects in small systems are sometimes of interest, here we are primarily concerned with the geometry that arises in the so-called thermodynamic limit $N \to \infty$. Thus, we consider the thermodynamic potential $\psi(\alpha, \beta)$ per particle in the thermodynamic limit, given by

$$\psi(\alpha, \beta) = \lim_{N\to\infty} N^{-1} \ln Z(\alpha, \beta) = \frac{3}{2} \ln \frac{2\pi m}{h^2 \beta} - \ln \alpha. \tag{98}$$

The components of the Fisher–Rao metric, with respect to the parameterization (α, β), can then be calculated by differentiation, with the result

$$G_{ij} = \begin{pmatrix} \alpha^{-2} & 0 \\ 0 & \frac{3}{2}\beta^{-2} \end{pmatrix}. \tag{99}$$

From this expression we deduce the following.

Proposition 4. *All components of the Riemann tensor, and consequently also the scalar curvature, of the statistical manifold $\mathfrak{M}$ associated with the classical ideal gas vanish, and thus the manifold is flat.*

Proof. From the components of the metric (99) one can calculate the components of the Christoffel symbol Γ^i_{jk} and the Riemann tensor R^i_{jkl} using the definitions (24) and (37). Alternatively, to show that this manifold is flat, it suffices to display a change of coordinates which transforms the metric (99) into a Euclidean metric. Here, we adopt the latter approach because this also permits more expeditious solution of the geodesic equations. We recall that under a coordinate transformation $x^i \to \bar{x}^i$ the metric of a Riemannian manifold transforms in the usual tensorial manner, so that the components of the inverse metric in the new coordinate system are determined by the contraction

$$\bar{G}^{ij} = G^{kl} \frac{\mathrm{d}\bar{x}^i}{\mathrm{d}x^k} \frac{\mathrm{d}\bar{x}^j}{\mathrm{d}x^l}. \tag{100}$$

Consider the following coordinate transformation:

$$\alpha \to \alpha' = \ln \alpha \qquad \text{and} \qquad \beta \to \beta' = \sqrt{\tfrac{3}{2}} \ln \beta. \tag{101}$$

A straightforward calculation then shows that the components of the inverse metric in the (α', β') coordinate system are

$$\bar{G}^{ij} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{102}$$

and thus the manifold is indeed flat. □

Since the statistical manifold associated with the ideal gas is flat, solution of the geodesic equations is straightforward. The result can be summarized as follows.

Proposition 5. *The geodesic curves on the statistical manifold $\mathfrak{M}$ associated with the classical ideal gas are given by*

$$\frac{P}{P_0} = \left(\frac{k_B T}{k_B T_0}\right)^{1+c}, \tag{103}$$

where P_0, T_0 and c are integration constants. In particular, the geodesics include the adiabatic equation of state for the ideal gas, corresponding to the choice $c = -C_V/Nk_B$, where C_V is the constant-volume heat capacity.

Proof. The geodesic equations for the variables α and β assume identical forms, i.e.

$$\frac{\mathrm{d}^2x}{\mathrm{d}s^2} - \frac{1}{x}\left(\frac{\mathrm{d}x}{\mathrm{d}s}\right)^2 = 0 \tag{104}$$

for $x = \alpha, \beta$. This can be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}s}\ln\left(\frac{\mathrm{d}x}{\mathrm{d}s}\right) = \frac{\mathrm{d}}{\mathrm{d}s}\ln x, \tag{105}$$

from which we see that the general solution is $x(s) = c_1\,\mathrm{e}^{c_2 s}$. Thus, we obtain

$$\frac{P}{k_BT} = c_1\,\mathrm{e}^{c_0 s} \qquad \text{and} \qquad \frac{1}{k_BT} = c_3\,\mathrm{e}^{c_2 s} \tag{106}$$

as the general solution to the geodesic equations. Combining these two equations, we have

$$P = \frac{c_1}{c_3^{c_0/c_2}}(k_BT)^{1-c_0/c_2}. \tag{107}$$

Setting $s = 0$ we find $c_1 = P_0/k_BT_0$ and $c_3 = 1/k_BT_0$, which yields at once the expression in (103). □

3. van der Waals gas

The geometry of the statistical manifold changes considerably if the gas molecules interact. In particular, if the system exhibits a phase transition, then the curvature tends to become singular at the transition point. This property seems to be universal and appears in many systems exhibiting critical phenomena. The van der Waals gas model is not only of physical interest, but also illustrates many of the universal geometrical features of the associated manifold of equilibrium states.

3.1. Equation of state

The idealized system of noninteracting particles considered above is inadequate for the description of phase transition phenomena, that is, the condensation of gas molecules. Here we shall extend the model to include inter-particle interactions, which leads to the *van der Waals equation of state*:

$$\left(P + a\frac{N^2}{V^2}\right)(V - bN) = Nk_BT, \tag{108}$$

where N is the total number of molecules and a, b are constants determined by the properties of the molecule. The liquid–vapour transition occurs at the critical point where the temperature T, pressure P and volume V simultaneously assume the values

$$P_\mathrm{c} = \frac{a}{27b^2}, \qquad V_\mathrm{c} = 3bN \qquad \text{and} \qquad T_\mathrm{c} = \frac{8a}{27k_Bb}. \tag{109}$$

The critical point is determined by the simultaneous solution of the equations

$$\frac{\partial P}{\partial V} = 0 \qquad \text{and} \qquad \frac{\partial^2 P}{\partial V^2} = 0. \tag{110}$$

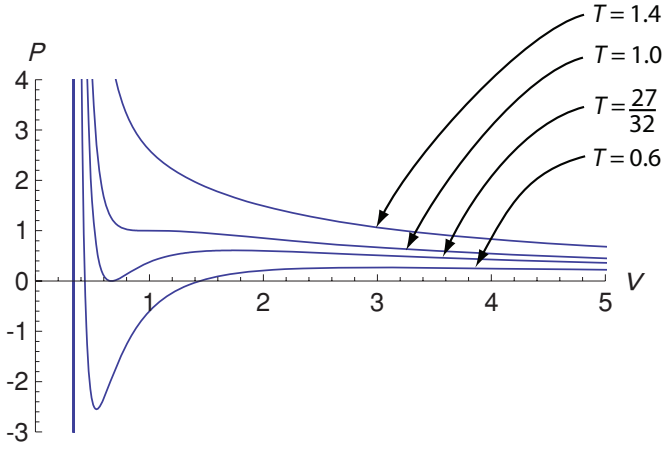

Figure 3. Equations of state for the van der Waals gas in terms of dimensionless variables. The isothermal curves correspond to $\hat{T} = 1.4, \hat{T} = 1, \hat{T} = 27/32$ (maximum superheating temperature) and $\hat{T} = 0.6$. Note that the isothermal curves associated with temperatures below $T_{\rm m}$ allow metastable regions for which $\hat{P} < 0$.

Using the dimensionless variables

$$\hat{P} = \frac{P}{P_{\rm c}}, \qquad \hat{V} = \frac{V}{V_{\rm c}}, \qquad \text{and} \qquad \hat{T} = \frac{T}{T_{\rm c}}, \tag{111}$$

the equation of state can be rewritten in the universal form

$$(\hat{P} + 3\hat{V}^{-2})\left(\hat{V} - \tfrac{1}{3}\right) = \tfrac{8}{3}\hat{T}, \tag{112}$$

independent of the parameters a and b. In figure 3 we plot the pressure $\hat{P}$ as a function of the volume $\hat{V}$ for $\hat{T} > 1$, $\hat{T} = 1$ and $\hat{T} < 1$.

Note that the positivity of the pressure implies a bound on the temperature. In particular, from (112) we deduce that the condition $\hat{P} \geqslant 0$ is equivalent to the bound:

$$\hat{T} \geqslant \frac{9\hat{V} - 3}{8\hat{V}^2}, \tag{113}$$

in terms of the dimensionless volume $\hat{V}$. If we demand the positivity of $\hat{P}$ for all volumes $\hat{V} \geqslant \frac{1}{3}$, then we must require $\hat{T} \geqslant \frac{27}{32}$, or equivalently

$$T \geqslant \frac{27}{32} T_{\rm c}. \tag{114}$$

The temperature

$$T_{\rm m} = \frac{27}{32} T_{\rm c} \approx 0.85 T_{\rm c} \tag{115}$$

is known as the *temperature of maximum superheating* and is related to the nucleation of bubbles when the liquid is heated very abruptly. In particular, if $T < T_{\rm m}$ then the liquid can be contained at low external pressure, whereas if $T > T_{\rm m}$ the liquid cannot exist under low external pressure and thus evaporates. Experimental data show, for example, that $T_{\rm m} = 0.89 T_{\rm c}$ for ether, $T_{\rm m} = 0.92 T_{\rm c}$ for alcohol and $T_{\rm m} = 0.84 T_{\rm c}$ for water, indicating the fairly accurate predictive power of the van der Waals equations of state.

Turning to the equations of state, the pressure P as a function of the volume V has three distinct roots when the temperature is below its critical value $T_{\rm c}$. Amongst these three roots,

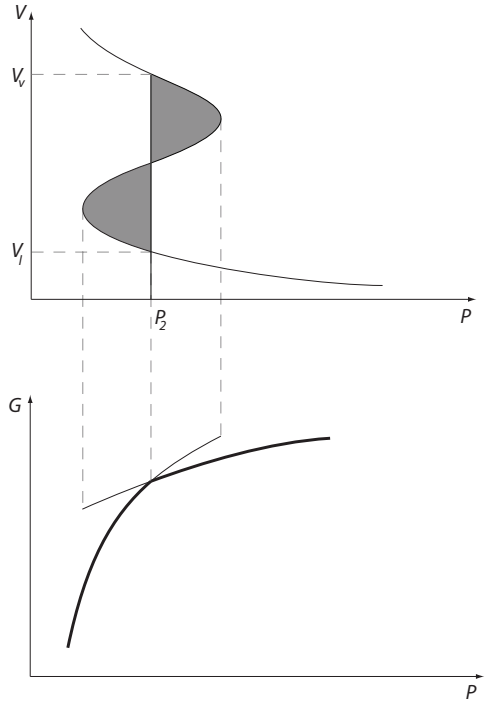

Figure 4. Isothermal curve and equal area law in the pressure–volume plane (above), and pressure dependence of the Gibbs free energy (below). As the pressure P of the gas is slowly increased, the Gibbs free energy $G(P)$ increases along the thick solid line in the lower diagram, until P reaches the coexisting pressure P_2. The gas then undergoes a phase transition and condenses. During this transition the volume changes from $V_{\rm v}$ to $V_{\rm l}$, at which point the entire system enters the liquid phase. The value of the coexisting pressure P_2 is determined by Maxwell's equal area law.

the intermediate root corresponds to a point at which $(\partial P/\partial V)_T > 0$. Hence, this root is unstable, since the pressure increases with volume for fixed temperature. It follows that one of the remaining two roots should correspond to thermal equilibrium. To ascertain which of the two roots is stable, we recall that the condition for stability is determined by the minimization of the free energy. If we let $G(T, P)$ denote the Gibbs free energy, then Maxwell's relation $V = (\partial G/\partial P)_T$ implies that

$$G(T, P) = G(T, P_0) + \int_{P_0}^{P} V(u, T)\,\mathrm{d}u. \tag{116}$$

Therefore, when viewed as a function of pressure P for a fixed temperature $T < T_{\rm c}$ below the critical point, the free energy $G(T, P)$ describes one of the two distinct curves, depending on whether P is reduced from high values or increased from low values. This is shown schematically in figure 4.

If the free energy assumes its minimum value, then as the value of P changes, $G(T, P)$ must describe one of the two curves in figure 4 until its intersection with the other curve at pressure $P = P_2$, whereafter $G(T, P)$ follows the other curve. At the point $P = P_2$ the liquid and vapour phases coexist. Therefore, if we, say, reduce the pressure from high values, then after reaching the value P_2 the pressure remains constant until the liquid is entirely converted into vapour. During this transition the volume changes from $V_{\rm l}$ to $V_{\rm v}$ as indicated in figure 4.

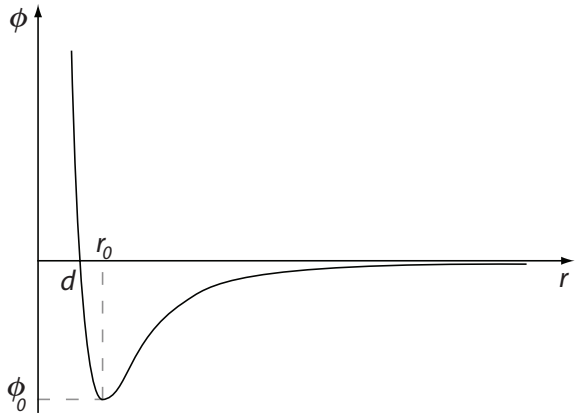

Figure 5. Lennard–Jones potential. There is a strong repulsive force at short distances $r \leqslant r_0$ and a weak attractive force at long distances $r > r_0$, where $r_0 = \sqrt[6]{2}d$ and d represents the radius of the gas molecule.

The value of the coexisting pressure P_2 is determined, for each fixed $T < T_c$, by Maxwell's equal area principle. That is, the vertical line in figure 4 is chosen so that the volumes of the two shaded regions are exactly equal.

3.2. Canonical partition function

The equation of state (108) was first deduced empirically by van der Waals, directly from experimental observations. However, it can also be derived analytically from the canonical partition function associated with an empirically postulated intermolecular potential. Assume that the interaction energy between a pair of molecules separated by a distance r is given by the Lennard–Jones potential

$$\phi(r) = 4\phi_0 \left[\left(\frac{d}{r}\right)^{12} - \left(\frac{d}{r}\right)^{6} \right] = \phi_0 \left(\frac{r_0}{r}\right)^{12} - 2\phi_0 \left(\frac{r_0}{r}\right)^{6}, \tag{117}$$

where $r_0 = 2^{1/6}d$ and d is a parameter which can be regarded as the radius of the gas molecule. Clearly, $\phi(d) = 0$ and $\phi(r)$ assumes its minimum value at $r = r_0$. As we see from figure 5, this inter-molecular potential energy gives rise to a weak long-range attractive force and a strong short-range repulsive force between each pair of molecules.

The canonical partition function can thus be written as

$$z(\beta, V) = \frac{1}{N!h^{3N}} \prod_{i=1}^{N} \int \mathrm{d}^3\mathrm{p}_i \int \mathrm{d}^3 r_i \exp\left(-\beta \sum_{i=1}^{N} \frac{\mathrm{p}_i^2}{2m} - \beta \sum_{(ij)} \phi_{ij} \right), \tag{118}$$

where $\phi_{ij} = \phi(r_{ij})$, with r_{ij} denoting the distance between the ith and jth molecules. Thus, the canonical partition function can be expressed as a product

$$z(\beta, V) = z_0(\beta, V) Q(\beta, V), \tag{119}$$

where $z_0(\beta, V)$ is the canonical partition function for the ideal gas (95) and

$$Q(\beta, V) = \frac{1}{V^N} \int \mathrm{d}^3 r_1 \cdots \int \mathrm{d}^3 r_N \exp\left(-\beta \sum_{(ij)} \phi_{ij} \right) \tag{120}$$

is the contribution from the interaction energy.

Now, as an approximation to the Lennard–Jones potential we assume that $\exp(-\beta\phi_{ij}) = 0$ for $r_{ij} < d$. In other words, we regard the molecules as hard spheres of radius d, which cannot overlap. As a consequence, the overlapping region can be removed from the range of the volume integration (120). Defining the so-called *Mayer function* $f_{ij} = f(r_{ij})$ by

$$f_{ij} = \exp(-\beta\phi_{ij}) - 1, \tag{121}$$

we rewrite the integral (120) as

$$\begin{aligned} Q(\beta, V) &= \frac{1}{V^N} \int_{r_1>d} \mathrm{d}^3 r_1 \cdots \int_{r_N>d} \mathrm{d}^3 r_N \prod_{(ij)} (1 + f_{ij}) \\ &= \frac{1}{V^N} \int_{r_1>d} \mathrm{d}^3 r_1 \cdots \int_{r_N>d} \mathrm{d}^3 r_N \left(1 + \sum_{(ij)} f_{ij} + \sum_{(ij)} \sum_{(kl)} f_{ij} f_{kl} + \cdots \right). \end{aligned} \tag{122}$$

Assuming that the parameter ϕ_0 in (117) is sufficiently small, the contribution arising from f_{ij} in the specified integration range can be regarded as an infinitesimal. The first term on the right-hand side of (122), i.e. the integral of unity, can, on the other hand, be approximated by

$$V(V - v_0) \cdots [V - (N-2)v_0][V - (N-1)v_0] \approx (V - bN)^N, \tag{123}$$

where we put $v_0 = 2b = \frac{4}{3}\pi d^3$. The integrations are performed consecutively, so that the first particle can occupy volume V without constraints, the second particle can occupy volume V less the volume v_0 occupied by the first particle, the third particle can occupy volume V less the volume $2v_0$ occupied by the first two particles, and so on. Similarly, the second term on the right-hand side of (122) can be approximated by

$$\begin{aligned} \int_{r_1>d} \mathrm{d}^3 r_1 \cdots \int_{r_N>d} \mathrm{d}^3 r_N f_{ij} &= (V - bN)^{N-1} \int_{r_j>d} \mathrm{d}^3 r_j f_{ij} \\ &\approx -(V - bN)^{N-1} \beta\pi \int_d^\infty \phi(r) r^2 \, \mathrm{d}r. \end{aligned} \tag{124}$$

Assembling these results, we can approximate $Q(\beta, V)$ in the following closed form:

$$\begin{aligned} Q(\beta, V) &\cong \left(1 - b\frac{N}{V}\right)^N \left(1 + \beta\frac{aN^2}{V} + \cdots\right) \\ &\cong \left(1 - b\frac{N}{V}\right)^N \exp\left(\beta\frac{aN^2}{V}\right), \end{aligned} \tag{125}$$

where we have defined

$$a = -\pi \int_d^\infty r^2 \phi(r) \, \mathrm{d}r. \tag{126}$$

Using the above expression for $Q(\beta, V)$ we finally obtain the canonical partition function

$$z(\beta, V) = \frac{1}{N!} \left(\frac{2\pi m}{\beta h^2}\right)^{3N/2} (V - bN)^N \exp(a\beta N^2/V). \tag{127}$$

3.3. Critical behaviour of the van der Waals gas

From the expression for the partition function of the canonical distribution we deduce the equation of state

$$P = \frac{1}{\beta} \left(\frac{\partial \ln z(\beta, V)}{\partial V}\right)_\beta = \frac{N k_B T}{V - bN} - a\frac{N^2}{V^2}, \tag{128}$$

where for clarity we have substituted $\beta = 1/k_B T$. Observe that this is precisely the van der Waals equation introduced in (108). If we had not applied various approximations in the derivation of (127), then additional terms of order $(N/V)^3$ and higher would have appeared on the right-hand side of (128).

To analyse the behaviour of the system near a critical point we introduce the deviation parameters

$$p = \hat{P} - 1, \qquad v = \hat{V} - 1 \qquad \text{and} \qquad t = \hat{T} - 1. \tag{129}$$

In terms of these shifted variables the equation of state (128) becomes

$$p = \frac{8(t+1)}{3v+2} - \frac{3}{(v+1)^2} - 1. \tag{130}$$

We can then expand the equation of state (130) for small v and t, obtaining

$$p = t(4 - 6v) - \tfrac{3}{2}v^3 + \cdots. \tag{131}$$

Similarly the Gibbs free energy

$$G = PV - k_B T \ln z(T, V) \tag{132}$$

can be expanded as follows:

$$G = P_c V_c \left[(p - 4t)v + 3tv^2 + \frac{3}{8}v^4 + 1 + p \right] - \frac{3}{2} N k_B T \ln\left(\frac{2\pi m k_B T}{h^2}\right). \tag{133}$$

For fixed pressure p and temperature t, the volume v in thermal equilibrium is that which minimizes the Gibbs free energy G. The equation of state (131) is a necessary but not sufficient condition for the Gibbs free energy to assume its minimum. Therefore, at the coexisting pressure $p = 4t$ below the critical temperature ($t < 0$) we have the three roots

$$v = \pm 2\sqrt{-t}, \quad 0 \tag{134}$$

for the volume determined by the equation of state (131). Differentiating (133) with respect to v, we find that the first two roots $\pm 2\sqrt{-t}$ minimize the free energy G and thus correspond to stable states, while the root $v = 0$ maximizes G and thus corresponds to an unstable state. Stable states represent the coexisting phase of liquid and vapour, with pressure given by

$$P_2 = P_c c(1 + 4t) = P_c \left(1 - 4\frac{T_c - T}{T_c}\right). \tag{135}$$

The liquid phase is more stable when $P_2 < P < P_c$, and the vapour phase more stable when $P < P_2$.

3.4. The thermodynamic limit

The existence of the instability in the van der Waals system studied above is related to the fact that in the canonical distribution the volume V of the system is held fixed, whereas in a real gas *volume fluctuations are significant in the vicinity of the critical point*. In other words, the canonical distribution does not provide a completely accurate physical description of the vapour–liquid equilibrium. Therefore, as in the case of an ideal gas, we consider the pressure–temperature distribution, with the corresponding partition function

$$Z(\alpha, \beta) = \frac{1}{b} \int_{bN}^{\infty} z(\beta, V) \exp(-\alpha V)\, \mathrm{d}V \tag{136}$$

wherein the volume fluctuation is integrated out. Recall that $b = \frac{2}{3}\pi d^3$ represents the smallest volume each molecule can occupy. Hence, the random variable V representing

the total volume ranges from bN to infinity. When the canonical partition function (127) is substituted into (136), the resulting integral does not admit an elementary analytical expression. Nevertheless, in the thermodynamic limit $N \to \infty$ we can implicitly determine the potential $\psi(\alpha, \beta) = N^{-1} \ln Z(\alpha, \beta)$ by the method of steepest descent.

We proceed as follows. First we write the integrand in (136) as

$$\exp(-\alpha V) z(\beta, V) = \exp[N g(\alpha, \beta, \hat{v})], \tag{137}$$

where $\hat{v} \equiv V/N$ and

$$g(\alpha, \beta, \hat{v}) = 1 - \alpha \hat{v} + \ln \left[\left(\frac{2\pi m}{\beta h^2} \right)^{\frac{3}{2}} (\hat{v} - b) \right] + \frac{\beta a}{\hat{v}}. \tag{138}$$

In deriving (138) we have used the Stirling formula $\ln N! \simeq N \ln N - N$. It should be evident from (132) that

$$G = -\beta^{-1} g(\alpha, \beta, \hat{v}) \tag{139}$$

is the Gibbs free energy. Also, note that $g(\alpha, \beta, \hat{v})$ must have at least one maximum in the range $\hat{v} \in [b, \infty)$ corresponding to the minimum Gibbs free energy, because $g(\alpha, \beta, \hat{v}) \to -\infty$ in the limits $\hat{v} \to b$ and $\hat{v} \to \infty$. The value of $\hat{v}$ at which $g(\alpha, \beta, \hat{v})$ is maximized therefore determines the equation of state (128) for the canonical distribution. However, in the P–T distribution the volume is a random variable, hence we must take its expectation to obtain the equation of state:

$$\langle \hat{v} \rangle = -\frac{1}{N} \frac{\partial \ln Z(\alpha, \beta)}{\partial \alpha}, \tag{140}$$

where $\langle \hat{v} \rangle$ denotes the expected volume per particle in the P–T distribution characterized by the density function $Z^{-1}(\alpha, \beta) \exp[N g(\alpha, \beta, \hat{v})]$.

Applying the change of variable $V \to \hat{v}$, the partition function (136) can be written in the form

$$Z(\alpha, \beta) = \frac{N}{b} \int_b^\infty \exp[N g(\alpha, \beta, \hat{v})] \, \mathrm{d}\hat{v}. \tag{141}$$

Recall that we are interested in the thermodynamic potential per particle in the thermodynamic limit:

$$\psi(\alpha, \beta) = \lim_{N \to \infty} N^{-1} \ln Z(\alpha, \beta). \tag{142}$$

Using the method of steepest descent, we find that $\psi(\alpha, \beta)$ in this limit is given by

$$\psi(\alpha, \beta) = g(\alpha, \beta, \bar{v}) = -\alpha \bar{v} + \ln z(\beta, \bar{v}), \tag{143}$$

where $\bar{v} = \bar{v}(\alpha, \beta)$ is the function of α and β which maximizes $g(\alpha, \beta, \hat{v})$. Since $\bar{v}$ minimizes the Gibbs free energy, it is the solution of the van der Waals equation of state. Although the exact functional form of $\bar{v}(\alpha, \beta)$ is not at our disposal owing to the cubic nature of the equation of state, we can nonetheless determine the exact expression for the scalar curvature in terms of the variables β and $\bar{v}$. Before we proceed, however, we first establish the following result.

Proposition 6. *The thermal expectation value of the volume $\hat{v}$ per particle in the P–T distribution is given by $\bar{v}$, that is, $\langle \hat{v} \rangle = \bar{v}$.*

Proof. Differentiating (143) and using the chain rule we find

$$\frac{\partial \psi}{\partial \alpha} = -\bar{v} + \left(-\alpha + \frac{\partial \ln z}{\partial \bar{v}} \right) \frac{\partial \bar{v}}{\partial \alpha}. \tag{144}$$

However, by definition $\bar{v}$ maximizes $g(\alpha, \beta, \hat{v})$ so that

$$\frac{\partial g}{\partial \bar{v}} = -\alpha + \frac{\partial \ln z}{\partial \bar{v}} = 0, \tag{145}$$

and hence

$$\frac{\partial \psi}{\partial \alpha} = -\bar{v}. \tag{146}$$

On the other hand, from (140) we have $\langle \hat{v} \rangle = -\partial\psi/\partial\alpha$, and thus $\langle \hat{v} \rangle = \bar{v}$. □

3.5. Geometry of the van der Waals manifold

As we have just indicated, the functional form of $\bar{v}(\alpha, \beta)$ is unknown. Nevertheless, we can implicitly determine expressions for $\partial\bar{v}/\partial\alpha$, $\partial\bar{v}/\partial\beta$, and so on, in the following manner. First, define Ω by

$$\Omega \equiv -\frac{\partial g}{\partial \bar{v}} = \alpha - \frac{1}{\bar{v} - b} + \beta \frac{a}{\bar{v}^2}. \tag{147}$$

Since $\hat{v} = \bar{v}$ maximizes $g(\alpha, \beta, \hat{v})$ we have by definition the relation $\Omega = 0$. This, however, is just the equation of state for the van der Waals gas. Now, consider the total derivative of Ω:

$$\mathrm{d}\Omega = \frac{\partial \Omega}{\partial \alpha}\,\mathrm{d}\alpha + \frac{\partial \Omega}{\partial \beta}\,\mathrm{d}\beta + \frac{\partial \Omega}{\partial \bar{v}}\,\mathrm{d}\bar{v}. \tag{148}$$

Since $\mathrm{d}\Omega = 0$ it follows that

$$\begin{aligned} \mathrm{d}\bar{v} &= -\left[\frac{\partial \Omega}{\partial \alpha} \Big/ \frac{\partial \Omega}{\partial \bar{v}}\right] \mathrm{d}\alpha - \left[\frac{\partial \Omega}{\partial \beta} \Big/ \frac{\partial \Omega}{\partial \bar{v}}\right] \mathrm{d}\beta \\ &= \left(\frac{\partial \bar{v}}{\partial \alpha}\right)_{\beta} \mathrm{d}\alpha + \left(\frac{\partial \bar{v}}{\partial \beta}\right)_{\alpha} \mathrm{d}\beta, \end{aligned} \tag{149}$$

where we have used the general identity:

$$\left(\frac{\partial \alpha}{\partial \beta}\right)_{\gamma} \left(\frac{\partial \beta}{\partial \gamma}\right)_{\alpha} \left(\frac{\partial \gamma}{\partial \alpha}\right)_{\beta} = -1. \tag{150}$$

On the other hand, from (147) we have the relations

$$\frac{\partial \Omega}{\partial \alpha} = 1, \qquad \frac{\partial \Omega}{\partial \beta} = \frac{a}{\bar{v}^2} \qquad \text{and} \qquad \frac{\partial \Omega}{\partial \bar{v}} = \frac{1}{(\bar{v} - b)^2} - \frac{2a\beta}{\bar{v}^3}. \tag{151}$$

Therefore, substituting these into (149) we deduce that

$$\frac{\partial \bar{v}}{\partial \alpha} = \frac{1}{D} \qquad \text{and} \qquad \frac{\partial \bar{v}}{\partial \beta} = \frac{1}{D}\frac{a}{\bar{v}^2}, \tag{152}$$

where D is defined by

$$D(\alpha, \beta) = \frac{2a\beta}{\bar{v}^3} - \frac{1}{(\bar{v} - b)^2}. \tag{153}$$

The derivatives of $\bar{v}$ with respect to the parameters α and β are required in order to determine the components of the Fisher–Rao metric on the van der Waals manifold. Specifically we obtain the following result:

Proposition 7. *In terms of the pressure–temperature coordinates* (α, β) *the Fisher–Rao metric on the van der Waals manifold is given by*

$$G_{ij} = \frac{1}{D} \begin{pmatrix} -1 & -a/\bar{v}^2 \\ -a/\bar{v}^2 & \frac{3}{2}\beta^{-2} - (a/\bar{v}^2)^2 \end{pmatrix}. \tag{154}$$

In particular, in the ideal gas limit $a \to 0$ and $b \to 0$, the metric (154) reduces to the metric (99) for the ideal gas.

Proof. The components of the metric are determined by the matrix $\partial_i \partial_j \psi(\alpha, \beta)$. We have, in proposition 6, established that $\partial\psi/\partial\alpha = -\bar{v}$, and, using (145), we have

$$\frac{\partial \psi}{\partial \beta} = -\alpha \frac{\partial \bar{v}}{\partial \beta} + \frac{\partial \ln z}{\partial \bar{v}} \frac{\partial \bar{v}}{\partial \beta} + \frac{\partial \ln z}{\partial \beta} = \frac{\partial \ln z}{\partial \beta}. \tag{155}$$

Therefore, we obtain

$$\frac{\partial^2 \psi}{\partial \alpha^2} = -\frac{\partial \bar{v}}{\partial \alpha}, \qquad \frac{\partial^2 \psi}{\partial \alpha \partial \beta} = -\frac{\partial \bar{v}}{\partial \beta} \qquad \text{and} \qquad \frac{\partial^2 \psi}{\partial \beta^2} = \frac{\partial^2 \ln z}{\partial \beta^2} + \frac{\partial^2 \ln z}{\partial \beta \partial \bar{v}} \frac{\partial \bar{v}}{\partial \beta}, \tag{156}$$

whence the desired expression for the metric follows from the formula (127) for the canonical partition function. In the ideal gas limit $a \to 0$ and $b \to 0$, we have $D \to \bar{v}$. However, from the ideal gas equation of state we have $\bar{v} = \alpha^{-1}$, hence we recover (99) at once in this limit. □

To describe the geometry of the van der Waals gas it will be convenient to introduce the concept of a *spinodal curve*. In general a spinodal curve consists of the points in the thermodynamic phase space at which the second derivative of the free energy with respect to an order parameter vanishes. For a gas of interacting molecules, the mean volume $\bar{v}$ constitutes the order parameter of the system, and the vanishing of the second derivative of $\ln z(\beta, \bar{v})$ with respect to $\bar{v}$ thus determines the spinodal curve in the phase diagram. For the van der Waals gas, by (145), we have the relation

$$\frac{\partial \alpha}{\partial \bar{v}} = -\frac{\partial^2 \ln z}{\partial \bar{v}^2} = 0 \tag{157}$$

that determines the spinodal curve. In other words, the locus of points at which the derivative of pressure with respect to volume vanishes for some temperature determines the spinodal curve. This is schematically illustrated for the van der Waals equation in figure 6.

It is evident from figure 6 that the region in the phase diagram enveloped by the spinodal curve is unstable because $(\partial P/\partial V)_T > 0$ in this region, i.e. the pressure increases with increasing volume. The spinodal curve thus forms the boundary of a semi-stable region in the phase diagram. In view of the first relation in (152), the spinodal curve is determined by the condition $D = 0$. On the other hand, from expression (154) for the Fisher–Rao metric on the van der Waals manifold we see that each component of the metric G_{ij} as well as its determinant $-3/(2\beta^2 D)$ is singular along the spinodal curve. Is this singular behaviour merely due to the specific choice of coordinates or is it an intrinsic feature of the van der Waals manifold? We can answer this question by calculating the scalar curvature of the manifold. The exact expression for the curvature is as follows.

Proposition 8. *The scalar curvature of the van der Waals manifold is given by*

$$R = \frac{4}{3D^2} \left(\frac{a\beta}{\bar{v}} \right) \left(\frac{a\beta}{\bar{v}^3} - D \right). \tag{158}$$

In particular, R diverges along the entire spinodal curve specified by $D = 0$, which includes the critical point (P_c, V_c, T_c). The scalar curvature vanishes in the ideal gas limit for which $a \to 0$ and $b \to 0$.

Proof. Since we have chosen the canonical parameterization (α, β), we can use the determinant in (48) to calculate the curvature. To compute the entries in the determinant we differentiate

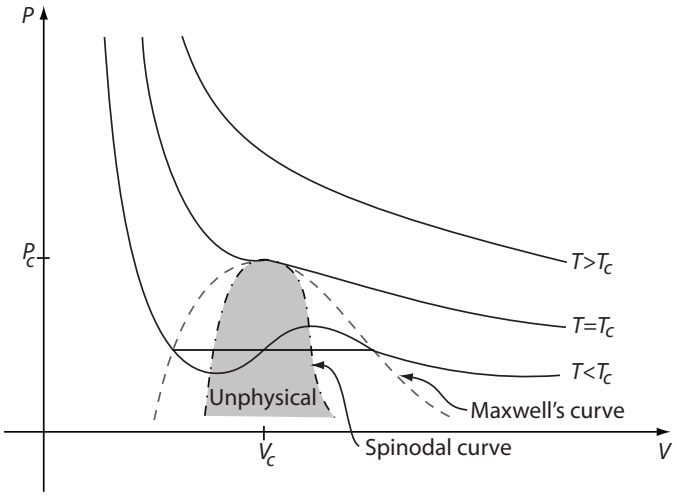

Figure 6. Schematic illustration of equations of state for van der Waals gas. The scalar curvature on the parameter space diverges along the spinodal boundary which envelopes the unphysical region. The critical point is that where the spinodal curve is tangent to the Maxwell equal area boundary. The divergence of the curvature along the spinodal boundary may be interpreted as 'preventing', in some sense, entry into the unphysical domain in the phase diagram.

(156) and use the chain rule, together with (152), thereby obtaining

$$\begin{aligned} \psi_{111} &= \frac{2}{D^3}\left(\frac{1}{(\bar{v}-b)^3} - 3\frac{a\beta}{\bar{v}^4}\right) \equiv X, \qquad \psi_{112} = \frac{a}{\bar{v}^2}X + \frac{2a}{\bar{v}D^2}, \\ \psi_{122} &= \frac{a^2}{\bar{v}^4}\left(X + \frac{4}{\bar{v}D^2}\right), \qquad \psi_{222} = \frac{a^3}{\bar{v}^6}\left(X + \frac{6}{\bar{v}D^2}\right) - \frac{3}{\beta^3}. \end{aligned} \tag{159}$$

Substituting these results into (48) we obtain, after some algebra, the desired expression in (158). The fact that the curvature diverges along the spinodal curve $D = 0$ is evident from expression (158). Also, from the definition (157) of the spinodal curve and the condition (110) for the critical point, it is clear that the critical point lies on the spinodal curve. The vanishing of the curvature in the ideal gas limit $a, b \to 0$ is also evident from the expression in (158). □

The van der Waals manifold possesses the structure of a Riemann surface over a planar base space with coordinates (α, β), branched around the singularities specified by the spinodal curve. Now, suppose we slowly change the variables (α, β) along a closed contour C in the planar base space. Then, the lifted curve in the van der Waals manifold does not, in general, return to the same sheet (i.e. to the same thermodynamic state) if C encloses the critical point or crosses the spinodal curve, while Maxwell's relation ensures that an infinitesimal closed contour enclosing no singularities is thermodynamically trivial. Thus, the presence of singularities may give rise to changes in the thermodynamic state $\bar{v}$ of the system upon following a closed contour in the parameter base space which encloses a point of divergency. Conversely, if a closed contour in the parameter space does not enclose the critical point or cross the spinodal curve, then the corresponding curve in $\mathfrak{M}$ is closed and thus gives rise to a well-defined holonomy. This leads naturally to the following open problem: what is the physical interpretation or relevance of the holonomy (analogue of the geometric phase in quantum mechanics) in classical statistical mechanics?

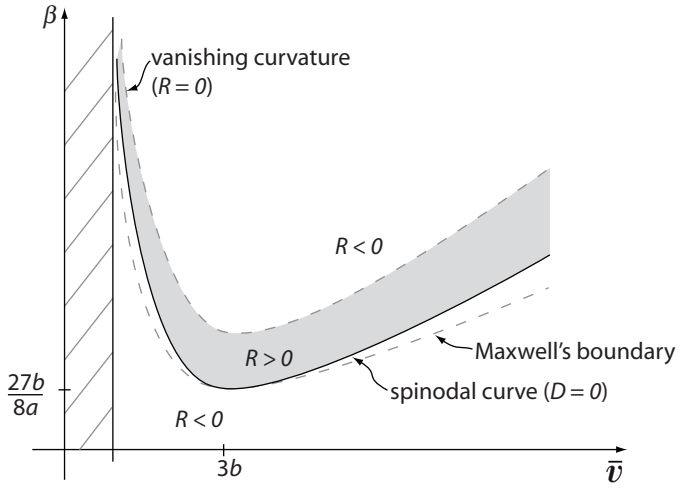

Figure 7. Geometric phase diagram for the van der Waals gas. The gas is divided into positive and negative curvature phases by the vanishing curvature $R = 0$ curve. The change of phase along $R = 0$ is analytic, while the curvature exhibits singular behaviour along the spinodal curve $D = 0$.

Finally we note that the scalar curvature R on the van der Waals manifold vanishes along the curve specified by

$$\beta = \frac{\bar{v}^3}{a(\bar{v} - b)^2}, \tag{160}$$

and its sign changes smoothly from positive to negative as one decreases the temperature in the $(\bar{v}, \beta)$ plane. The sign of the scalar curvature in the $(\bar{v}, \beta)$ plane and its relation to the spinodal and Maxwell boundaries are schematically illustrated in figure 7. One might refer to the smooth change in the sign of the scalar curvature in the phase diagram as a *geometric phase transition*. Unlike the conventional phase transitions associated with singular behaviour, however, such geometric phase transitions are not associated with any divergence. There are indications that the scalar curvature can be viewed as a measure of stability of the system. However, the precise correspondence between physical characteristics of the system and properties of the curvature, apart from the singular behaviour along the spinodal boundary, remains an open research problem.

4. Bibliographical notes

When the first author was invited to write a survey article on a topic of interest, it seemed appropriate to utilize this opportunity by briefly reviewing some applications of information geometry in physics. This reflects the fact that interest in this area has continued to grow: two international conferences on applications of information geometry have been held in recent years, while diverse new applications continue to emerge—for example, in the area of shape recognition in computer science (Maybank 2005, Peter and Rangarajan 2006), in connection with out of equilibrium measures (Crooks 2007), in the characterization of quantum phase transitions (Zanardi *et al* 2007), in various mathematical extensions (e.g., Cena and Pistone 2007), in anyon statistics (Mirza and Mohammadzadeh 2008) or in black hole thermodynamics (Ruppeiner 2008). The application of information geometry to statistical physics, however, has

generated a vast amount of literature, and it seemed neither feasible nor helpful to cover all the results which have emerged in this area. What seemed more appropriate was to focus attention upon one specific topic that nevertheless incorporates all the essential ingredients. It also appeared desirable to explain the basic ideas of information geometry and its emergence from probabilistic and statistical considerations in a language accessible to a graduate student in theoretical physics. For these reasons we begin the paper with a rather elementary background material, and then consider its application to the theory of vapour–liquid equilibrium. This leads to the analysis of the van der Waals gas model, which, in our opinion, not only embodies all the essential features of a phase transition in statistical mechanics but also admits an elegant geometric characterization. To keep the exposition at a fairly elementary level, we have excluded *ad hoc* citations from the main text so as to avoid impeding the continuity of the exposition. Instead, references to the literature are presented more informally in these bibliographical notes.

To the authors' best knowledge, the use of geometric methods in statistical analysis was first introduced by P C Mahalanobis, the founder of the Indian Statistical Institute and also the founding editor of Sankhyā (*The Indian Journal of Statistics*), in the early 1930s (Mahalanobis 1930, 1936). Mahalanobis was a physicist and statistician who taught relativity and wrote an introduction to the translation by M Saha of Minkowski's work on relativity. His articles on relativity, written jointly with S N Bose, were published by the Calcutta University. He introduced a measure of mutual separation in the study of statistical data arising from anthropometric measurements. An alternative measure of separation was subsequently introduced by Bhattacharyya (1943, 1946), and was defined in section 1 as the Bhattacharyya spherical distance between two probability densities.

The geometric description of the parameter-space manifold was initiated at around the same time by Rao (1945, 1947, 1954). The seminal paper by Rao (1945) is significant in two respects: on the one hand, it introduced the so-called Cramér–Rao inequality as a lower bound for the variance, while on the other hand it pointed out that the information measure introduced previously by Fisher (1925) defines a Riemannian metric on the parameter-space manifold of a statistical model. Rao then proposed the associated geodesic distance as a measure of dissimilarity between probability distributions. The formulation considered by Rao was based on the Hilbert space embedding $p_\theta(x) \to \xi_\theta(x) = \sqrt{p_\theta(x)}$. As a consequence, many of the constructions in Rao (1945) bear a formal resemblance to the geometric formulation of quantum mechanics, developed by physicists some time later in the 1980s and 1990s (see Brody and Hughston 2001 and references cited therein). Concurrently with Rao's work on the application of geometry in statistics, Jeffreys (1946) also introduced the concept of uninformative priors based on the use of a Riemannian metric.

It is worth noting incidentally that the inner product of probability measures via the square-root embedding was introduced earlier by Hellinger (1909) in connection with unitary invariants of self-adjoint operators in Hilbert space. The distance measure of Hellinger was subsequently extended by von Neumann and Kakutani (see Kakutani 1946), who introduced an inner product of probability measures in an abstract measure-theoretic context and applied this to investigate equivalence and orthogonality relations between product measures. The Kakutani inner product was used by Brody (1971) to provide a simple proof of the Gaussian dichotomy theorem, and has also been extended by Bures (1969) in the context of operator algebras.

Interest in the application of geometrical techniques to statistical inference appear to have somewhat diminished subsequently, but reemerged with the appearance of Efron's seminal paper (1975) on information loss and statistical curvature. Efron considered the logarithmic embedding $p_\theta(x) \to l_\theta(x) = \ln p_\theta(x)$ and demonstrated that in the space of log-likelihood

density functions the curvature of the curve l_θ measures the deviation of the density function from the exponential family. Furthermore, the squared statistical curvature was shown to determine the loss of information resulting from the use of the maximum likelihood estimator for the unknown parameter θ.

The work of Efron (1975, 1978)—and to some extent that of Čencov (1982) who showed that the Fisher information metric is unique under certain assumptions including invariance—evoked considerable interest in the geometrical approach to asymptotic inference and related topics in statistics. Numerous research papers (for example, Atkinson and Mitchell 1981), as well as review papers (for example, Kass 1989) and monographs (for example, Amari 1985, Amari *et al* 1987, Murray and Rice 1993, and Amari and Nagaoka 2000, Arwini and Dodson 2008), were subsequently published.

In parallel with these developments in statistics, the application of information geometry to the description of the equilibrium properties of thermodynamic systems was considered by a number of authors. One of the initiators was Ingarden (1981), who considered various Banach space embeddings of probability density functions and their relation to the concepts of entropy and Fisher information, and suggested their application to statistical physics. The works of Ingarden and his collaborators led to the establishment of a Polish School of researchers systematically investigating various geometric aspects of physical systems described by equilibrium distributions. Janyszek and Mrugała (1989a), for instance, clarified the relation of the Fisher–Rao geometry to contact geometry (the latter characterizes the geometry of Legendre transformations), and investigated the physical interpretation of the metric tensor for a system of gas molecules characterized by the pressure–temperature distribution. Janyszek and Mrugała (1989b) calculated the scalar curvatures of the parameter space manifolds of the one-dimensional Ising model in the thermodynamic limit and of the mean-field model. Janyszek and Mrugała (1990) also extended information geometric analysis to the investigation of the stability of ideal quantum gases. Some of these ideas were further extended by others in the context of various spin models in statistical mechanics (see, for example, Janke *et al* 2002, Brody and Ritz 2003, Johnston *et al* 2003).

An independent line of investigation on the various geometric properties of thermodynamic state spaces was initiated by Weinhold (1975) and also by Ruppeiner (1979). Weinhold proposed the existence of a metric structure for the thermodynamic state space arising from empirical laws of thermodynamics. This line of thinking, which led to the notion of the so-called *thermodynamic length*, was extended in a variety of ways by various authors (see, for example, Salamon and Berry 1983, Schlögl 1985, Mrugała *et al* 1990).

Ruppeiner went a step further and considered the second derivative of the entropy, discussed briefly here in section 1.5, as a Riemannian metric on the thermodynamic state space. The metric considered by Ruppeiner, based upon the Shannon–Wiener entropy, agrees with Rao's entropy derivative metric (cf Rao 1984), and is also related to the Fisher–Rao metric through a Legendre transformation. The key idea is that the convexity of entropy implies the positive-definiteness of the associated Hessian matrix, which can therefore be used to define a Riemannian metric on the thermodynamic state space. The simplest system to consider in this respect is naturally that of a noninteracting gas of classical particles. This was investigated by Ruppeiner (1979), and also by Mijatović *et al* (1987). Geometric aspects of various other physical systems have also been investigated along these lines (Ruppeiner 1990, 1991). For a comprehensive bibliography on this topic, see the reference list in the review article by Ruppeiner (1995).

Ingarden and Tamassy (1993) also applied an entropic measure of divergence to explain the thermodynamic arrow of time. While the infinitesimal form of the entropic measure of divergence (relative entropy) gives rise to a Riemannian structure characterized in general by

the Burbea–Rao metric (85), the metrics arising from entropies generally possess Finslerian structure. In particular, the Finsler metric arising from relative entropy is not symmetric. The idea of Ingarden and Tamassy (1993) consists in exploiting the lack of symmetry in such metrics to explain the thermodynamic arrow of time, without the introduction of dissipation.

An important application of information geometry to the properties of renormalization group flow in statistical physics—motivated in part by the observation of Dawid (1975) that Efron's results might be represented more concisely in terms of Hilbert space geometry—was proposed by Brody (1987). Closely related ideas were developed further by O'Connor and Stephens (1993), Dolan (1998), Brody and Ritz (1998), and Brody (2000). (See also Diósi *et al* (1984) for an alternative approach to analysis of renormalization group flow using Rao's entropy derivative metric.)

The ideas of information geometry have also been extended to the quantum domain by replacing the density functions of classical probability theory by density matrices. Substantial work has been done on quantum information geometry (for example, Petz and Sudar 1996, Uhlmann 1996, Grasselli and Streater 2001, Petz 2002, Streater 2004, Jenčová and Petz 2006, Gibilisco and Isola 2007, and Gibilisco *et al* 2007), as well as its application to quantum statistical inference (Brody and Hughston 1998, Barndorff-Nielsen and Gill 2000).

Turning more specifically to the subject matter of the present review, as indicated above, the spherical distance (7) representing the dissimilarity of probability densities was introduced by Bhattacharyya (1943), and the concept of a statistical manifold represented by the metric (23) was introduced by Rao (1945). The uniqueness of the Einstein metric on a complex projective space was conjectured by Calabi, and later proven by Yau (1977). The implication of this result in quantum mechanics—that the solution to the vacuum Einstein equation in the space of pure quantum states determines transition probabilities—was pointed out to the first author by G W Gibbons in the late 1990s. The relevance of the projective space and the associated metric (47) to statistical mechanics was demonstrated by Brody and Hughston (1999).

The expression in proposition 2 for the scalar curvature in terms of the determinant of a 3×3 matrix, valid for the two-dimensional statistical manifold associated with a canonical density function, is given in Janyszek and Mrugała (1989b). The fact that the statistical manifold associated with the normal density function possesses constant negative curvature, shown in equation (57), was observed by Amari (1982). However, the significance of the scalar curvature (or in fact the Riemann tensor itself) in statistical analysis remains somewhat obscure.

A survey article by Burbea (1986) deals systematically with expressions for geodesic curves associated with a number of standard density functions used in statistics (see also the article by Rao in Amari *et al* 1987). This includes, in particular, the distance between Gaussian density functions as given in (69) and (71). Analogous results for gamma-distributed densities have been calculated in some detail by Burbea *et al* (2002).

The fact that the leading-order term in the Taylor expansion of the relative entropy of a neighbouring pair of parametric density functions gives rise to the Fisher–Rao metric was observed by Ingarden (1981). A more detailed and thorough analysis was given by Burbea and Rao (1984), and constitutes the basis for the discussion in section 1.5. The specific form of entropy defined in (80), sometimes referred to as the α-order entropy, was introduced by Havrad and Charvát (1967) in the context of quantifying classification schemes. See also Burbea and Rao (1982a, 1982b) for details concerning various properties of this entropy. The use of the α-order entropy in statistical mechanics has been proposed by Tsallis (1988).

In section 2 we considered the information geometry of the pressure–temperature distribution representing the equilibrium state of a gas of noninteracting particles (classical

ideal gas). The flatness of the information manifold of a classical ideal gas was pointed out by Ruppeiner (1979) using the entropy derivative metric. The solution to the associated geodesic equations, in the form expressed in proposition 5, does not seem to appear elsewhere. (An alternative representation of the geodesics appears in Mijatović *et al* (1987).)

In sections 3.1–3.4 we have provided a brief account of the classical theory of the van der Waals gas, as a background for the subsequent geometric description. Our exposition follows closely the classic treatise of Mayer and Mayer (1940). There is a series of inspiring papers by Kac *et al* (1963), Uhlenbeck *et al* (1963), Hemmer *et al* (1964), and also Hemmer (1964), analysing the vapour–liquid equilibrium of the van der Waals gas in great detail. These papers extend the earlier work of Kac (1959), which provides a method for determining the partition function of interacting gas molecules.

Other related work on systems of interacting gas molecules includes the following: Tonks (1936) determined the equations of state for gases composed of hard elastic spheres with finite radius. van Hove (1950) calculated the free energy of a system of molecules with nonvanishing incompressible radii, interacting according to a finite range force. He showed that in one dimension the system exhibits no phase transition. Lebowitz and Percus (1963) studied the properties of the correlation functions. Van Kampen (1964) showed that a gas of molecules with hard sphere repulsive forces and long-range attractive interactions exhibits condensation, and calculated the density fluctuations. Rigorous bounds for the free energy of the van der Waals gas were derived by Lebowitz and Penrose (1966).

Detailed analyses of the curvature of some of these classical systems of interacting molecules were presented by Ruppeiner and Chance (1990). The geometry of the van der Waals gas associated with the entropy derivative metric is considered in papers by Diósi and Lukács (1986) and in Diósi *et al* (1989), wherein the authors determine the scalar curvature using the density and temperature as coordinates. Using these coordinates, they have also shown that on this statistical manifold there exists no solution to the Killing equations (i.e. no vector field such that the associated flow preserves geodesic distances).

The description of the geometry of the van der Waals manifold presented in section 3.5 follows closely the analysis outlined by Janyszek (1990) and also by Brody and Rivier (1995). In particular, the expressions for the metric tensor in proposition 7 and for the scalar curvature in proposition 8 were derived by Brody and Rivier (1995), who also suggested that the curvature of the statistical manifold might play a role in statistical mechanics analogous to that of geometric phases in quantum mechanics. This remains an open issue, although recent work on quantum phase transitions indicate that there is indeed a close analogy between these two concepts.

Finally, the present authors regret that owing to the huge volume of literature on this subject there are many other valuable contributions which have not been mentioned in these brief biographical notes.

References

Amari S 1982 Differential geometry of curved exponential families—curvatures and information loss *Ann. Stat.* **10** 357–85

Amari S 1985 *Differential-Geometrical Methods in Statistics (Lecture Notes in Statistics* vol 28*)* (New York: Springer)

Amari S, Barndorff-Nielsen O E, Kass R E, Lauritzen S L and Rao C R 1987 *Differential Geometry in Statistical Inference (Institute of Mathematical Statistics Lecture Notes. Monograph Series* vol 10*)* (Hayward, CA: Institute of Mathematical Statistics)

Amari S and Nagaoka H 2000 *Methods of Information Geometry (AMS Translations of Mathematical Monograph* vol 191*)* (Oxford: Oxford University Press)

Arwini K A and T Dodson C J 2008 *Information Geometry (Lecture Notes in Mathematics* vol 1953*)* (Berlin: Springer)

Atkinson C and S Mitchell A F 1981 Rao's distance measure *Sankhyā* **43** 345–65
Barndorff-Nielsen O E and Gill R D 2000 Fisher information in quantum statistics *J. Phys. A: Math. Gen.* **33** 4481–90
Bhattacharyya A 1943 On a measure of divergence between two statistical populations defined by their probability distributions *Bull. Calcutta Math. Soc.* **35** 99–109
Bhattacharyya A 1946 On a measure of divergence between two multinomial populations *Sankhyā* **7** 401–6
Brody D C 2000 Differential renormalisation flow in random lattice gauge theories *Phys. Lett.* B **485** 422–8
Brody D C and Hughston L P 1998 Statistical geometry in quantum mechanics *Proc. R. Soc. Lond.* A **454** 2445–75
Brody D C and Hughston L P 1999 Geometrisation of statistical mechanics *Proc. R. Soc. Lond.* A **455** 1683–715
Brody D C and Hughston L P 2001 Geometric quantum mechanics *J. Geom. Phys.* **677** 1–35
Brody D C and Ritz A 1998 On the symmetry of real-space renormalisation *Nucl. Phys.* B **522** 588–604
Brody D C and Ritz A 2003 Information geometry of finite Ising models *J. Geom. Phys.* **47** 207–20
Brody D C and Rivier N 1995 Geometrical aspects of statistical mechanics *Phys. Rev.* E **51** 1006–11
Brody E J 1971 An elementary proof of the Gaussian dichotomy theorem *Z. Wahrscheinlichkeitstheor. Verwandte Geb.* **20** 217–26
Brody E J 1987 Applications of the Kakutani metric to real-space renormalization *Phys. Rev. Lett.* **58** 179–82
Burbea J 1986 Informative geometry of probability spaces *Expositiones Math.* **4** 347–78
Burbea J, Oller J M and Reverter F 2002 Some remarks on the information geometry of the Gamma distribution *Commun. Stat. Theory Methods* **31** 1959–75
Burbea J and Rao C R 1982a On the convexity of some divergence measures based on entropy functions *IEEE Trans. Inf. Theory* **IT-28** 489–95
Burbea J and Rao C R 1982b On the convexity of higher order Jensen differences based on entropy functions *IEEE Trans. Inf. Theory* **IT-28** 961–3
Burbea J and Rao C R 1984 Differential metrics in probability spaces *Probab. Math. Stat.* **3** 241–58
Bures D 1969 An extension of Kakutani's theorem on infinite product measures to the tensor product of semifinite w^*-algebras *Trans. Am. Math. Soc.* **135** 199–212
Cena A and Pistone G 2007 Exponential statistical manifold *Ann. Inst. Stat. Math.* **59** 27–56
Čencov N N 1982 *Statistical Decision Rules and Optimal Inference (Translations of Mathematical Monographs* vol 53*)* (Providence, RI: American Mathematical Society) (Originally published as Статистические Решающие Правила и Оптималбные Выводы Москва: Наука 1972.)
Crooks G E 2007 Measuring thermodynamic length *Phys. Rev. Lett.* **99** 100602
Dawid A P 1975 Discussion on Professor Efron's paper *Ann. Stat.* **3** 1231–4
Diósi L, Forgács G, Lukács B and Frisch H L 1984 Metricization of thermodynamic-state space and the renormalization group *Phys. Rev.* A **29** 3343–5
Diósi L and Lukács B 1986 Spatial correlations in diluted gases from the viewpoint of the metric of the thermodynamic state space *J. Chem. Phys.* **84** 5081–4
Diósi L, Lukács B and Rácz A 1989 Mapping the van der Waals state space *J. Chem. Phys.* **91** 3061–7
Dolan B P 1998 Geometry and thermodynamic fluctuations of the Ising model on a Bethe lattice *Proc. R. Soc. Lond.* A **454** 2655–65
Efron B 1975 Defining the curvature of a statistical problem (with applications to second order efficiency). With a discussion by C R Rao, D A Pierce, D R Cox, D V Lindley, L LeCam, J K Ghosh, J Pfanzagl, N Keiding, A P Dawid, J Reeds and with a reply by the author *Ann. Stat.* **3** 1189–242
Efron B 1978 The geometry of exponential families *Ann. Stat.* **6** 362–76
Fisher R A 1925 Theory of statistical estimation *Proc. Camb. Phil. Soc.* **122** 700–25
Gibilisco P and Isola T 2007 Uncertainty principle and quantum Fisher information *Ann. Inst. Stat. Math.* **59** 147–59
Gibilisco P, Imparato D and Isola T 2007 Uncertainty principle and quantum Fisher information: II *J. Math. Phys.* **48** 072109
Grasselli M R and Streater R F 2001 On the uniqueness of the Chentsov metric in quantum information geometry *Infinite Dimens. Anal. Quantum Probab. Relat. Top.* **4** 173–82
Havrda J and Charvát F 1967 Quantification method of classification processes *Kybernetika* **3** 30–5
Hellinger E 1909 Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen *J. Reine Angew. Math.* **136** 210–71
Hemmer P C 1964 On the van der Waals theory of the vapour–liquid equilibrium: IV. The pair correlation function and equation of state for long-range forces *J. Math. Phys.* **5** 75–84
Hemmer P C, Kac M and Uhlenbeck G E 1964 On the van der Waals theory of the vapour–liquid equilibrium: III. Discussion of the critical region *J. Math. Phys.* **5** 60–74
Ingarden R S 1981 Information geometry in functional spaces of classical and quantum finite statistical systems *Int. J. Eng. Sci.* **19** 1609–33

Ingarden R S and Tamassy L 1993 On parabolic geometry and irreversible macroscopic time *Rep. Math. Phys.* **32** 11–33
Janke W, Johnston D A and C. Malmini R P K 2002 Information geometry of the Ising model on planar random graphs *Phys. Rev.* E **66** 056119
Janyszek H 1990 Riemannian geometry and stability of thermodynamical equilibrium systems *J. Phys. A: Math. Gen.* **23** 477–90
Janyszek H and Mrugała R 1989a Geometrical structure of the state space in classical stastical and phenomological thermodynamics *Rep. Math. Phys.* **27** 145–59
Janyszek H and Mrugała R 1989b Riemannian geometry and the thermodynamics of model magnetic systems *Phys. Rev.* A **39** 6515–23
Janyszek H and Mrugała R 1990 Riemannian geometry and the stability of ideal quantum gases *J. Phys. A: Math. Gen.* **23** 467–76
Jeffreys H 1946 An invariant form for the prior probability in estimation problems *Proc. R. Soc. Lond.* A **186** 453–61
Jenčová A and Petz D 2006 Sufficiency in quantum statistical inference: a survey with examples *Infinite Dimens. Anal. Quantum Probab. Relat. Top.* **9** 331–51
Johnston D A, Janke W and Kenna R 2003 Information geometry, one, two, three (and four) *Acta Phys. Pol.* B **34** 4923–37
Kac M 1959 On the partition function of a one-dimensional gas *Phys. Fluids* **2** 8–12
Kac M, Uhlenbeck G E and Hemmer P C 1963 On the van der Waals theory of the vapour–liquid equilibrium: I. Discussion of a one-dimensional model *J. Math. Phys.* **4** 216–28
Kakutani S 1948 On equivalence of infinite product measures *Ann. Math.* **49** 214–24
Kass R 1989 The geometry of asymptotic inference. With comments and a rejoinder by the author *Stat. Sci.* **4** 188–234
Lebowitz J L and Penrose O 1966 Rigorous treatment of the van der Waals–Maxwell theory of liquid–vapour transition *J. Math. Phys.* **7** 98–113
Lebowitz J L and Percus J K 1963 Asymptotic behaviour of the radial distribution function *J. Math. Phys.* **4** 248–54
Mahalanobis P C 1930 On tests and meassures of groups divergence *J. Asiatic Soc. Bengal* **26** 541–88
Mahalanobis P C 1936 On the generalised distance in statistics *Proc. Natl. Inst. Sci. India* A **2** 49–55
Maybank S 2005 The Fisher–Rao metric for projective transformations of the line *Int. J. Comput. Vis.* **63** 191–206
Mayer J E and Mayer M G 1940 *Statistical Mechanics* (New York: Wiley)
Mijatović M, Veselinović V and Trenčevski K 1987 Differential geometry of equilibrium thermodynamics *Phys. Rev.* A **35** 1863–7
Mirza B and Mohammadzadeh H 2008 Ruppeiner geometry of anyon gas *Phys. Rev.* E **78** 021127
Mrugała R, Nulton J D, Schön J C and Salamon P 1990 Statistical approach to the geometric structure of thermodynamics *Phys. Rev.* A **41** 3156–60
Murray M K and Rice J W 1993 *Differential Geometry and Statistics* (London: Chapman and Hall)
O'Connor D and Stephens C R 1993 Geometry, the renormalisation group and gravity *Directions in General Relativity Proc. 1993 Int. Symp., MA* vol 1, ed Hu B. L., M P Ryan (Jr) and C V Vishveshwava (Cambridge: Cambridge University Press)
Peter A and Rangarajan A 2006 Shape analysis using the Fisher–Rao Riemannian metric: unifying shape representation and deformation *Proc. 3rd IEEE Int. Symp. on Biomedical Imaging: Nano to Macro* pp 1164–7
Petz D 2002 Covariance and Fisher information in quantum mechanics *J. Phys. A: Math. Gen.* **35** 929–39
Petz D and Sudar C 1996 Geometries of quantum states *J. Math. Phys.* **37** 2662–73
Rao C R 1945 Information and the accuracy attainable in the estimation of statistical parameters *Bull. Calcutta Math. Soc.* **37** 81–91
Rao C R 1947 The problem of classification and distance between two populations *Nature* **159** 30–1
Rao C R 1954 On the use and interpretation of distance functions in statistics *Bull. Inst. Int. Stat.* **34** 90–7
Rao C R 1984 Convexity properties of entropy functions and analysis of diversity *Inequalities in Statistics and Probability Proc. Symp. on Inequalities in Statistics and Probability (Lincoln, Nebraska 1982) (Institute of Mathematical Statistics Lecture Notes Monograph Series* vol 5*)* ed Y L Tong (Hayward, CA: Institute of Mathematical Statistics)
Ruppeiner G 1979 Thermodynamics: a Riemannian geometric model *Phys. Rev.* A **20** 1608–13
Ruppeiner G 1990 Thermodynamic curvature of a one-dimensional fluid *J. Chem. Phys.* **92** 3700–9
Ruppeiner G 1991 Riemannian geometric theory of critical phenomena *Phys. Rev.* A **44** 3583–95
Ruppeiner G 1995 Reimannian geometry in thermodynamic fluctuation theory *Rev. Mod. Phys.* **67** 605–59
Ruppeiner G 2008 Thermodynamic curvature and phase transitions in Kerr–Newman black holes *Phys. Rev.* D **78** 024016
Ruppeiner G and Chance J 1990 Reimannian geometry in thermodynamic fluctuation theory *J. Chem. Phys.* **92** 3700–9
Salamon P and Berry R S 1983 Thermodynamic length and dissipated availability *Phys. Rev. Lett.* **51** 1127–30

Schlögl F 1985 Thermodynamic metric and stochastic measures *Z. Phys.* B **59** 449–54
Streater R F 2004 Duality in quantum information geometry *Open Syst. Inf. Dyn.* **11** 71–7
Tonks L 1936 The complete equation of state of one, two, and three-dimensional gases of hard elastic spheres *Phys. Rev.* **50** 955–63
Tsallis C 1988 Possible generalization of Boltzmann–Gibbs statistics *J. Stat. Phys.* **52** 479–87
Uhlenbeck G E, Hemmer P C and Kac M 1963 On the van der Waals theory of the vapour–liquid equilibrium: II. Discussion of the distribution functions *J. Math. Phys.* **4** 229–47
Uhlmann A 1996 Spheres and hemispheres as quantum state space *J. Geom. Phys.* **18** 76–92
van Hove L 1950 Sur l'intégral de configuration pour les systèmes de particules à une dimension *Physica* **16** 137–43
van Kampen N G 1964 Condensation of a classical gas with long-range attraction *Phys. Rev.* **135** A362–9
Weinhold F 1975 Metric geometry of equilibrium thermodynamics *J. Chem. Phys.* **63** 2479–83
Yau S T 1977 Calabi's conjecture and some new results in algebraic geometry *Proc. Natl. Acad. Sci. USA* **74** 1798–9
Zanardi P, Giorda P and Cozzini M 2007 Information-theoretic differential geometry of quantum phase transitions *Phys. Rev. Lett.* **99** 100603