Chapter 8

Jaynes' Maximum Entropy Principle, Riemannian Metrics and Generalised Least Action Bound

Robert K. Niven[1] and Bjarne Andresen[2]

[1]*School of Aerospace, Civil and Mechanical Engineering, The University of New South Wales at ADFA, Canberra, ACT, 2600, Australia.*
r.niven@adfa.edu.au

[2]*Niels Bohr Institute, University of Copenhagen, Copenhagen Ø, Denmark. andresen@fys.ku.dk*

The set of solutions inferred by the generic maximum entropy (MaxEnt) or maximum relative entropy (MaxREnt) principles of Jaynes – considered as a function of the moment constraints or their conjugate Lagrangian multipliers – is endowed with a Riemannian geometric description, based on the second differential tensor of the entropy or its Legendre transform (negative Massieu function). The analysis provides a generalised *least action bound* applicable to all Jaynesian systems, which provides a lower bound to the cost (in generic entropy units) of a transition between inferred positions along a specified path, at specified rates of change of the control parameters. The analysis therefore extends the concepts of "finite time thermodynamics" to the generic Jaynes domain, providing a link between purely static (stationary) inferred positions of a system, and dynamic transitions between these positions (as a function of time or some other coordinate). If the path is unspecified, the analysis gives an absolute lower bound for the cost of the transition, corresponding to the geodesic of the Riemannian hypersurface. The analysis is applied to (i) an equilibrium thermodynamic system subject to mean internal energy and volume constraints, and (ii) a flow system at steady state, subject to constraints on the mean heat, mass and momentum fluxes and chemical reaction rates. The first example recovers the *minimum entropy cost* of a transition between equilibrium positions, a widely used result of finite-time thermodynamics. The second example leads to a new *minimum entropy production principle*, for the cost of a transition

284   R.K. Niven and B. Andresen

between steady state positions of a flow system.

Contents

8.1. Introduction

Jaynes' *maximum entropy principle* (MaxEnt) and its extension, the *maximum relative entropy principle* (MaxREnt), based on the principles of inductive (probabilistic) rather than deductive reasoning, arguably constitutes one of the most important tools for the solution of indeterminate problems of all kinds.[1–7] In this method, one maximises the entropy function of a system – a measure of its statistical spread over its parameter space – subject to the set of constraints on the system, to determine its "least informative" or "most probable" probability distribution.[1,2,7] By a series of generic "Jaynes relations", this can then be used to calculate the macroscopic properties of the system, providing the best (inferred) description of the system, subject to all that is known about the system. Since its inception half a century ago, the MaxEnt and MaxREnt principles have been successfully applied to the analysis of a diverse range of systems, including in thermodynamics (its first and foremost application), solid and fluid mechanics, mathematical biology, transport systems, networks, economic, social and human systems.[1–9]

The aim of this study is to examine a valuable extension to Jaynes' generic approach, by endowing the set of solutions inferred by Jaynes' method – considered as a function of the set of moment constraints and/or their conjugate Lagrangian multipliers – with a Riemannian geometric interpretation, using a metric tensor furnished directly by Jaynes' method. The analysis leads to a generalised *least action bound* applicable to all

Jaynesian systems, which provides a lower bound for the cost (in generic entropy units) of a transition between different inferred positions of the system. The analysis therefore extends the concepts of "finite time thermodynamics", developed over the past three decades,[10–43] to the generic Jaynes domain. The analysis reveals a deep, underlying connection between the essentially static manifold of stationary positions predicted by Jaynes' method, and lower bounds for the cost of dynamic transitions between these positions.

The manuscript proceeds as follows. In §8.2, the theoretical principles of Jaynes' MaxEnt and MaxREnt methods are discussed, followed by an appraisal of a generalised free energy (generalised potential) concept associated with Jaynes' method. In §8.3, the concepts of a Riemannian metric, arc length and action sums and integrals are developed in a generic Jaynesian context, leading to a generic least action bound for transitions on the manifold of Jaynes solutions. Considerations of minimum path lengths, involving calculation of the geodesic in Riemannian space, are also discussed. In §8.4, the foregoing principles are applied to (i) an equilibrium thermodynamic system subject to mean internal energy and volume constraints, and (ii) a flow system at steady state, subject to constraints on the mean heat, mass and momentum fluxes and chemical reaction rates. The first example (§8.4.1) recovers the *minimum entropy cost* of a transition between equilibrium positions, a widely used result of finite-time thermodynamics. The second example (§8.4.2) leads to a new *minimum entropy production principle*, for the cost of a transition between steady state positions of a flow system. The analyses reveal the tremendous utility of Jaynes' MaxEnt and MinXEnt methods augmented by the least action bound, for the analysis of probabilistic systems of all kinds.

8.2. Jaynes' Generic Formulation (MaxREnt)

8.2.1. *Theoretical Principles*

The usefulness of Jaynes' method for statistical inference arises from its *generic* formulation, first expounded by Jaynes and other workers in the context of information theory,[1–7] but which can be reinterpreted using a combinatorial framework (the "Boltzmann principle").[44–51] In consequence, the method can be applied to *any* probabilistic system involving the allocation of entities to categories; this includes – but is not restricted to – thermodynamic systems. For maximum generality, it is useful to include

286   *R.K. Niven and B. Andresen*

source or prior probabilities q_i associated with each category $i = 1, ..., s$, to give the maximum relative entropy (MaxREnt) or minimum cross-entropy (MinXEnt) principle. In the event of equal q_i, this reduces to the special case of Jaynes' maximum entropy (MaxEnt) principle.[1–7]

The MaxREnt method proceeds as follows. To infer the "least informative" or "most probable" distribution of a probabilistic system, we wish to identify its observable *realization* or *macrostate* of maximum probability $\mathbb{P}$. This is equivalent to maximising the following dimensionless function, chosen for several "nice" mathematical properties:[44,45]

$$\mathfrak{H} = \frac{1}{N} \ln \mathbb{P}, \tag{8.1}$$

For a system of N distinguishable entities allocated to s distinguishable categories, it can be shown that the distribution is governed by the multinomial distribution $\mathbb{P} = N! \prod_{i=1}^{s} q_i^{n_i}/n_i!$, where n_i is the occupancy of the ith category and $N = \sum_{i=1}^{s} n_i$. In the asymptotic limit $N \to \infty$, (8.1) reduces to the relative entropy function[2] (the negative of the Kullback-Leibler function[52,53]):

$$\mathfrak{H} = -\sum_{i=1}^{s} p_i \ln \frac{p_i}{q_i} \tag{8.2}$$

where $p_i = n_i/N$ is the frequency or probability of occupancy of the ith category. Maximisation of (8.2) is subject to the normalisation constraint and any moment constraints on the system:

$$\sum_{i=1}^{s} p_i = 1, \tag{8.3}$$

$$\sum_{i=1}^{s} p_i f_{ri} = \langle f_r \rangle, \quad r = 1, ..., R, \tag{8.4}$$

where f_{ri} is the value of the property f_r in the ith category and $\langle f_r \rangle$ is the mathematical expectation of f_{ri}. Applying Lagrange's method of undetermined multipliers to (8.2)-(8.4) gives the stationary or "most probable" distribution of the system (denoted *):

$$p_i^* = q_i \exp\left(-\lambda_0 - \sum_{r=1}^{R} \lambda_r f_{ri}\right) = \frac{1}{Z} q_i \exp\left(-\sum_{r=1}^{R} \lambda_r f_{ri}\right),$$

$$Z = e^{\lambda_0} = \sum_{i=1}^{s} q_i \exp\left(-\sum_{r=1}^{R} \lambda_r f_{ri}\right) \tag{8.5}$$

where λ_r is the Lagrangian multiplier associated with the rth constraint, Z is the partition function and $\lambda_0 = \ln Z$ is the Massieu function.[4] In thermodynamics, the constraints $\langle f_r \rangle$ are usually taken to represent conserved quantities, and thus correspond to extensive variables (e.g. internal energy, volume and numbers of particles), whilst the multipliers λ_r emerge as functions of the intensive variables of the system (e.g. temperature, pressure and chemical potentials). It is useful to preserve this distinction between extensive and intensive variables, even beyond a thermodynamic context.

By subsequent analyses,[1–7,54] one can derive the maximum relative entropy $\mathfrak{H}^*$ and the derivatives of $\mathfrak{H}^*$ and λ_0 for the system:

$$\mathfrak{H}^* = \lambda_0 + \sum_{r=1}^{R} \lambda_r \langle f_r \rangle \tag{8.6}$$

$$\frac{\partial \mathfrak{H}^*}{\partial \langle f_r \rangle} = \lambda_r \tag{8.7}$$

$$\frac{\partial^2 \mathfrak{H}^*}{\partial \langle f_m \rangle \partial \langle f_r \rangle} = \frac{\partial \lambda_r}{\partial \langle f_m \rangle} = g_{mr} \in \mathbf{g} \tag{8.8}$$

$$\frac{\partial \lambda_0}{\partial \lambda_r} = -\langle f_r \rangle \tag{8.9}$$

$$\frac{\partial^2 \lambda_0}{\partial \lambda_m \partial \lambda_r} = \langle f_r f_m \rangle - \langle f_r \rangle \langle f_m \rangle = -\frac{\partial \langle f_r \rangle}{\partial \lambda_m} = -\gamma_{mr} \in -\boldsymbol{\gamma} \tag{8.10}$$

The second derivatives of λ_0 in (8.10) express the dependence of each constraint on each multiplier, and therefore give the "capacities" or "susceptibilities" of the system (e.g. in thermodynamics, they define the heat capacity, compressibility, coefficient of thermal expansion and other material properties[12,55,56]). Their matrix $\boldsymbol{\gamma}$, the variance-covariance matrix of the constraints (with change of sign), is equal to the inverse of the matrix $\mathbf{g}$ of second derivatives of $\mathfrak{H}^*$ in (8.8), yielding the generic Legendre transformation between the $\mathfrak{H}^*(\langle f_1 \rangle, \langle f_2 \rangle, ...)$ and $\lambda_0(\lambda_1, \lambda_2, ...)$ descriptions of the system:[2]

$$\mathbf{g}\boldsymbol{\gamma} = \mathbf{I}, \tag{8.11}$$

where $\mathbf{I}$ is the identity matrix.[2] From (8.8) or (8.10) and the equality of mixed derivatives, we also obtain the generic reciprocal relations $\partial \langle f_r \rangle / \partial \lambda_m = \partial \langle f_m \rangle / \partial \lambda_r$ for the system.

Jaynes also showed that the incremental change in the relative entropy

can be expressed as:[1]

$$d\mathfrak{H}^* = \sum_{r=1}^{R} \lambda_r \Big(d\langle f_r \rangle - \langle df_r \rangle \Big) = \sum_{r=1}^{R} \lambda_r \delta Q_r \tag{8.12}$$

where $\delta W_r = \langle df_r \rangle = \sum_{i=1}^{s} p_i^* df_{ri}$ and $\delta Q_r = \sum_{i=1}^{s} dp_i^* f_{ri}$ can be identified, respectively, as the increments of "generalised work" and "generalised heat" associated with a change in the rth constraint, and $\delta(\cdot)$ indicates a path-dependent differential. Eq. (8.12) gives a "generalised Clausius equality",[57] applicable to all multinomial systems in the asymptotic limit.

It is again emphasised that the above relations (8.5)-(8.12) apply to *any* probabilistic system of multinomial form, in the asymptotic limit. Although originally derived in thermodynamics, the above-mentioned quantities need not be interpreted as thermodynamic constructs, but have far broader application. Furthermore, the relations (8.5)-(8.12) apply to the stationary position of any multinomial probabilistic system. The derivatives (8.7)-(8.10) therefore relate to transitions of the system between different stationary positions, or in other words, to paths on the manifold of stationary positions. Whilst the lack of inclusion of non-stationary positions may seem unnecessarily restrictive, such geometry provides a sufficient foundation for most of engineering and chemical equilibrium thermodynamics. As will be shown, it is also useful for the analysis of many other systems of similar probabilistic structure.

8.2.2. *The Generalised Free Energy Concept*

It is instructive to insert (8.12) into the differential of (8.6) and rearrange in the form:

$$
\begin{aligned}
d\phi = -d\lambda_0 = -d\ln Z &= \sum_{r=1}^{R} \lambda_r \delta W_r + \sum_{r=1}^{R} d\lambda_r \langle f_r \rangle \\
&= -d\mathfrak{H}^* + \sum_{r=1}^{R} \lambda_r d\langle f_r \rangle + \sum_{r=1}^{R} d\lambda_r \langle f_r \rangle
\end{aligned}
\tag{8.13}
$$

The negative Massieu function $-\lambda_0$ is therefore equivalent to a potential function ϕ which captures all possible changes in the system, whether they be in the entropy, constraints or multipliers. For constant multipliers, it simplifies to the weighted sum of generalised work on the system. It thus provides a dimensionless analogue of the free energy concept used in thermodynamics. For constant multipliers, $\phi|_{\{\lambda_r\}}$ also provides a measure of the

dimensionless "availability", or the available "weighted generalised work", which can be extracted from a system. By extension of the principles of equilibrium thermodynamics, we can thus adopt the potential ϕ as a measure of distance from the stationary state. The system will converge towards a position of minimum ϕ, representing the balance between maximisation of entropy within the system $\mathfrak{H}^*$, and maximisation of the entropy generated and exported to the rest of the universe by the transfer of generalised heats δQ_r (see[63] for further discussion). The advantage of Jaynes' generic formulation is that ϕ can be defined for *any* multinomial probabilistic system, and is not restricted to thermodynamic systems.[1–4,7]

Returning to the second derivatives in the last section, we see that λ_0 can be replaced by $-\phi$ in (8.9)-(8.10). The latter provides a clean (nonnegative) Legendre transformation between matrices $\mathbf{g}$ and $\boldsymbol{\gamma}$, and thus between the $\mathfrak{H}^*(\langle f_1 \rangle, \langle f_2 \rangle, ...)$ and $\phi(\lambda_1, \lambda_2, ...)$ representations of a system.

8.3.   Riemannian Geometric Concepts

8.3.1.   *Generalised Riemannian Metrics and Arc Lengths*

Since the time of Gibbs,[58,59] examination of the geometry of the manifold of stationary positions has been of tremendous interest to scientists and engineers. In thermodynamics, this has typically involved analysis of the concave hypersurface defined by the Euler relation $S(\tilde{X}_1, \tilde{X}_2, ...)$, where S is the thermodynamic entropy and $\tilde{X}_r$ are the extensive variables, or alternatively of its Legendre transform, the convex hypersurface $\psi(Y_1, Y_2, ...)$ or $F(Y_1, Y_2, ...)$, where $\psi = F/T$ is a Planck potential function, F is a free energy and Y_r are the intensive variables.[55,56] Such interpretations have led to major advances in the understanding and analysis of thermodynamic processes and cycles.[58,59] However, adoption of the Jaynes MaxEnt framework (§8.2) permits a rather different insight, based on a Riemannian geometric interpretation. As will be evident from the previous discussion, this interpretation extends well beyond "mere" thermodynamics, forming a natural adjunct of Jaynes' generic formulation (§8.2).

Consider the R-dimensional hypersurface parameterised by the constraints $\{\langle f_r \rangle\}$ or their conjugate Lagrangian multipliers $\{\lambda_r\}$, representing the hypersurface of stationary states within the $(R+1)$-dimensional space given by $(\mathfrak{H}^*, \{\langle f_r \rangle\})$ or $(\phi, \{\lambda_r\})$. If the R parameters are linearly independent, the matrices of the second derivatives $\mathbf{g}$ (8.8) or $\boldsymbol{\gamma}$ (8.10) are positive definite (i.e. $\boldsymbol{x}^\top \mathbf{g} \boldsymbol{x} > 0$ or $\boldsymbol{x}^\top \boldsymbol{\gamma} \boldsymbol{x} > 0$ for any non-zero vector $\boldsymbol{x}$[60–62]). The

R.K. Niven and B. Andresen

matrices $\mathbf{g}$ or $\boldsymbol{\gamma}$ can therefore be adopted as Riemannian metric tensors associated with the stationary state hypersurface defined by $\{\langle f_r \rangle\}$ or $\{\lambda_r\}$, and used to interpret its geometric properties. Indeed, even if the R parameters are not always independent, whence $\mathbf{g}$ or $\boldsymbol{\gamma}$ are positive semidefinite (i.e. $\boldsymbol{x}^\top \mathbf{g} \boldsymbol{x} \geq 0$ or $\boldsymbol{x}^\top \boldsymbol{\gamma} \boldsymbol{x} \geq 0$ for $\boldsymbol{x} \neq \mathbf{0}$), the latter can still be adopted as pseudo-Riemannian metric tensors on the stationary hypersurface. This representation was first proposed by Weinhold,[10–14] and its implications in terms of a least action bound were subsequently developed, largely within a thermodynamic context, by Salamon, Berry, Andresen, Nulton and co-workers[15–36] and also by Beretta,[37–39] Diósi and co-workers,[40] Crooks and Feng[41,42] and Brody and Hook.[43] Some theoretical aspects of the adopted Riemannian formulation are discussed in Appendix A. It must be noted that the Riemannian formulation replaces – it cannot be used in conjunction with – the traditional convex or concave hypersurface interpretation normally used in thermodynamics and information theory.[30]

Firstly, the Riemannian geometric interpretation provides an *intrinsic differential* or *line element* (its square, a *metric*) with which to measure distances along a specified path on the manifold[60,62a]:

$$ds_{\mathfrak{H}^*} = \sqrt{d^2 \mathfrak{H}^*} = \sqrt{\sum_{m,r=1}^{R} d\langle f_m \rangle \, g_{mr} \, d\langle f_r \rangle} = \sqrt{d\boldsymbol{f}^\top \mathbf{g} \, d\boldsymbol{f}}, \qquad (8.14)$$

$$ds_{\phi} = \sqrt{d^2 \phi} = \sqrt{\sum_{m,r=1}^{R} d\lambda_m \, \gamma_{mr} \, d\lambda_r} = \sqrt{d\boldsymbol{\Lambda}^\top \boldsymbol{\gamma} \, d\boldsymbol{\Lambda}}. \qquad (8.15)$$

where $\boldsymbol{f} = [\langle f_1 \rangle, \langle f_2 \rangle, ..., \langle f_s \rangle]^\top$ and $\boldsymbol{\Lambda} = [\lambda_1, \lambda_2, ..., \lambda_s]^\top$. Integration between points a and b along a path on the manifold, defined by the set of increments $d\boldsymbol{f}$ or $d\boldsymbol{\Lambda}$, gives the *arc length* along that path between those points:[17,41,60,62]

$$L_{\mathfrak{H}^*} = \int_a^b ds_{\mathfrak{H}^*} = \int_a^b \sqrt{\sum_{m,r=1}^{R} d\langle f_m \rangle \, g_{mr} \, d\langle f_r \rangle} = \int_a^b \sqrt{d\boldsymbol{f}^\top \mathbf{g} \, d\boldsymbol{f}}, \qquad (8.16)$$

$$L_{\phi} = \int_a^b ds_{\phi} = \int_a^b \sqrt{\sum_{m,r=1}^{R} d\lambda_m \, \gamma_{mr} \, d\lambda_r} = \int_a^b \sqrt{d\boldsymbol{\Lambda}^\top \boldsymbol{\gamma} \, d\boldsymbol{\Lambda}}. \qquad (8.17)$$

The shortest such path is known as the *geodesic*. An infinite number of other paths on the manifold are also possible, of longer arc length, as also given

---

[a]Strictly, this line element is not a first fundamental form in Riemannian geometry;[60,62] its use as a distance measure is discussed in Appendix A.

by (8.16) or (8.17). If the manifold is parameterised by some parameter ξ – which can, but need not, correspond to time t – the arc lengths can be expressed in continuous form as:

$$L_{\mathfrak{H}^*} = \int\limits_0^{\xi_{max}} \sqrt{\sum_{m,r=1}^R \frac{d\langle f_m \rangle}{d\xi} g_{mr} \frac{d\langle f_r \rangle}{d\xi}} \, d\xi = \int\limits_0^{\xi_{max}} \sqrt{\dot{\boldsymbol{f}}^\top \mathbf{g} \, \dot{\boldsymbol{f}}} \, d\xi, \qquad (8.18)$$

$$L_\phi = \int\limits_0^{\xi_{max}} \sqrt{\sum_{m,r=1}^R \frac{d\lambda_m}{d\xi} \gamma_{mr} \frac{d\lambda_r}{d\xi}} \, d\xi = \int\limits_0^{\xi_{max}} \sqrt{\dot{\boldsymbol{\Lambda}}^\top \boldsymbol{\gamma} \, \dot{\boldsymbol{\Lambda}}} \, d\xi \qquad (8.19)$$

where the overdot indicates differentiation with respect to ξ.

The symmetry of the Legendre transformation (8.11) also permits a further insight. From (8.8) and (8.10), the metrics g_{mr} or γ_{mr} within the intrinsic differentials (8.14)-(8.15) can be substituted respectively by $\partial \lambda_r / \partial \langle f_m \rangle$ or $\partial \langle f_r \rangle / \partial \lambda_m$, to give:

$$ds_{\mathfrak{H}^*} = \sqrt{\sum_{m,r=1}^R d\langle f_m \rangle \frac{\partial \lambda_r}{\partial \langle f_m \rangle} d\langle f_r \rangle} = \sqrt{\sum_{r=1}^R d\lambda_r \, d\langle f_r \rangle} = \sqrt{d\boldsymbol{\Lambda} \cdot d\boldsymbol{f}} \quad (8.20)$$

$$ds_\phi = \sqrt{\sum_{m,r=1}^R d\lambda_m \frac{\partial \langle f_r \rangle}{\partial \lambda_m} d\lambda_r} = \sqrt{\sum_{r=1}^R d\lambda_r \, d\langle f_r \rangle} = \sqrt{d\boldsymbol{\Lambda} \cdot d\boldsymbol{f}} \qquad (8.21)$$

In consequence, the intrinsic differentials are equal, $ds = ds_{\mathfrak{H}^*} = ds_\phi$, and so too are the arc lengths:

$$L = L_{\mathfrak{H}^*} = L_\phi = \int\limits_0^{\xi_{max}} \sqrt{\dot{\boldsymbol{\Lambda}} \cdot \dot{\boldsymbol{f}}} \, d\xi, \qquad (8.22)$$

From a Riemannian geometric perspective, it therefore does not matter whether one examines a system using its $\mathfrak{H}^*(\langle f_1 \rangle, \langle f_2 \rangle, ...)$ or $\phi(\lambda_1, \lambda_2, ...)$ representation. The above identities – touched on by several workers[27,34,40,41] – are not surprising, since the Legendre transforms $\mathfrak{H}^*$ and ϕ both have the character of entropy-related quantities, respectively indicating the (generic) entropy of a system and the capacity of a system to generate (generic) entropy.[63] The quantity $d\boldsymbol{\Lambda} \cdot d\boldsymbol{f}$ therefore expresses the second differential of generic entropy produced due to incremental changes in $\boldsymbol{\Lambda}$ and $\boldsymbol{f}$ (a generalised force-displacement or fluctuation-response relation). For all changes, $d\boldsymbol{\Lambda} \cdot d\boldsymbol{f} \geq 0$ must be valid, to preserve a positive

R.K. Niven and B. Andresen

definite metric (whence $\dot{\boldsymbol{\Lambda}} \cdot \dot{\boldsymbol{f}} \geq 0$);[12] this is in sympathy with a generalised form of the second law of thermodynamics, namely "*each net mean increment of (generic) entropy produced along a path must be positive*".

One further consideration arises from the recognition that most probabilistic systems involve *quantised* phenomena, which can only be approximated by the above continuous representation. For a system capable only of discrete increments in line elements $\Delta s_{\mathfrak{H}^*}$ or Δs_ϕ associated with a minimum dissipation parameter $\Delta\xi$ (e.g. a minimum dissipation time if $\xi = t$), the arc lengths are more appropriately given as:[25]

$$L_{\mathfrak{H}^*} = \sum_{v=1}^{M} \Delta s_{\mathfrak{H}^*, v} = \sum_{v=1}^{M} \sqrt{\Delta \boldsymbol{f}_v^\top \, \mathbf{g}_v \, \Delta \boldsymbol{f}_v} = \sum_{v=1}^{M} \sqrt{\dot{\boldsymbol{f}}_v^\top \, \mathbf{g}_v \, \dot{\boldsymbol{f}}_v} \, \Delta\xi_v, \quad (8.23)$$

$$L_\phi = \sum_{v=1}^{M} \Delta s_{\phi, v} = \sum_{v=1}^{M} \sqrt{\Delta \boldsymbol{\Lambda}_v^\top \, \boldsymbol{\gamma}_v \, \Delta \boldsymbol{\Lambda}_v} = \sum_{v=1}^{M} \sqrt{\dot{\boldsymbol{\Lambda}}_v^\top \, \boldsymbol{\gamma}_v \, \dot{\boldsymbol{\Lambda}}_v} \, \Delta\xi_v \quad (8.24)$$

where v is the index of each increment. The last terms in (8.23)-(8.24) invoke the finite difference forms $\dot{\boldsymbol{f}}_v = \Delta\boldsymbol{f}_v / \Delta\xi_v$ or $\dot{\boldsymbol{\Lambda}}_v = \Delta\boldsymbol{\Lambda}_v / \Delta\xi_v$, strictly valid only in the limits $\Delta\xi_v \to 0$. The two discrete length scales (8.23)-(8.24) are again equivalent, but there will most likely be some discrepancy between their values due to their finite difference formulation.

8.3.2. *Generalised Action Concepts and Least Action Bound*

A Riemannian geometry can also be examined from a different perspective,[15,17,25,33,41] discussed with reference to Figure 8.1; the following analysis largely follows that given by Nulton and coworkers,[25] converted into generic form. Although applied to $\mathfrak{H}^*$, an analogous derivation can be given for the ϕ representation. Consider a system on the manifold of stationary positions, subject to displacements $\{\Delta\langle f_r \rangle\}$ in its stationary position. The modified (generic) entropy $\mathfrak{H}^*(\{\langle f_r \rangle + \Delta\langle f_r \rangle\})$ of the system can be ex-

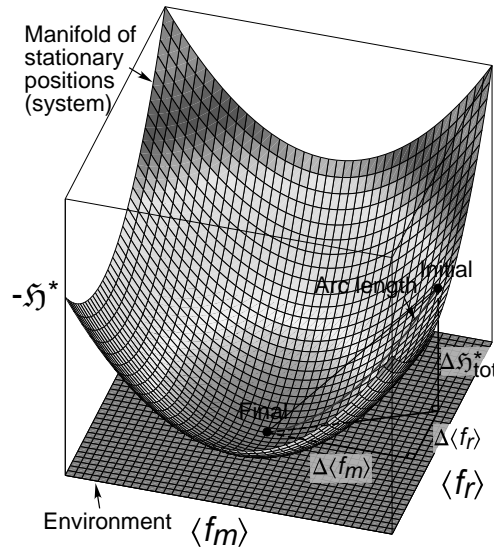*Jaynes' MaxEnt, Riemannian Metrics and Least Action Bound*   293

Fig. 8.1.   Illustration of Riemannian geometry concepts, for a two-constraint system represented by $\mathfrak{H}^*(\langle f_m \rangle, \langle f_r \rangle)$ (for convenience the environment is shown as horizontal).

panded in a Taylor series about $\mathfrak{H}^*(\{\langle f_r \rangle\})$:

$$\mathfrak{H}^*(\{\langle f_r \rangle + \Delta\langle f_r \rangle\}) = \mathfrak{H}^*(\{\langle f_r \rangle\}) + \sum_{r=1}^{R} \lambda_r |_{\{\langle f_r \rangle\}} \Delta\langle f_r \rangle$$

$$+ \frac{1}{2!} \sum_{m,r=1}^{R} \left. \frac{\partial^2 \mathfrak{H}^*}{\partial\langle f_m \rangle \partial\langle f_r \rangle} \right|_{\{\langle f_r \rangle\}} \Delta\langle f_m \rangle \Delta\langle f_r \rangle$$

$$+ \frac{1}{3!} \sum_{m,r,\ell=1}^{R} \left. \frac{\partial^3 \mathfrak{H}^*}{\partial\langle f_\ell \rangle \partial\langle f_m \rangle \partial\langle f_r \rangle} \right|_{\{\langle f_r \rangle\}} \Delta\langle f_\ell \rangle \Delta\langle f_m \rangle \Delta\langle f_r \rangle + ...$$

$$(8.25)$$

where use is made of (8.7). The corresponding change in entropy of the "reservoir" or "environment" of constant $\{\lambda_r^{env}\}$, by which this change is effected, is given (exactly) by:[17,25]

$$\mathfrak{H}^{env}(\{\langle f_r \rangle + \Delta\langle f_r \rangle\}) = \mathfrak{H}^{env}(\{\langle f_r \rangle\}) + \sum_{r=1}^{R} \lambda_r^{env} |_{\{\langle f_r \rangle\}} \Delta\langle f_r \rangle^{env} \qquad (8.26)$$

At the stationary state, $\lambda_r = \lambda_r^{env}$, whilst from the constraints (conservation laws), $\Delta\langle f_r \rangle = -\Delta\langle f_r \rangle^{env}$.[25] Addition of (8.25)-(8.26) thus yields

the total change in the entropy of the system and environment for the step process:

$$
\begin{aligned}
\Delta\mathfrak{H}^* = &\frac{1}{2!} \sum_{m,r=1}^{R} \left.\frac{\partial^2\mathfrak{H}^*}{\partial\langle f_m\rangle\partial\langle f_r\rangle}\right|_{\{\langle f_r\rangle\}} \Delta\langle f_m\rangle\Delta\langle f_r\rangle \\
&+ \frac{1}{3!} \sum_{m,r,\ell=1}^{R} \left.\frac{\partial^3\mathfrak{H}^*}{\partial\langle f_\ell\rangle\partial\langle f_m\rangle\partial\langle f_r\rangle}\right|_{\{\langle f_r\rangle\}} \Delta\langle f_\ell\rangle\Delta\langle f_m\rangle\Delta\langle f_r\rangle + ...
\end{aligned}
\tag{8.27}
$$

Provided the manifold is smooth, continuous, continuously differentiable (i.e., there are no phase changes in the neighbourhood) and the step sizes $\{\Delta\langle f_r\rangle\}$ are small, we can neglect the higher order terms in (8.27), giving:

$$
\Delta\mathfrak{H}_v^* \approx \frac{1}{2} \sum_{m,r=1}^{R} \Delta\langle f_m\rangle_v \, g_{mr,v}|_{\{\langle f_r\rangle\}} \, \Delta\langle f_r\rangle_v = \frac{1}{2}\Delta\boldsymbol{f}_v{}^\top \mathbf{g}_v \, \Delta\boldsymbol{f}_v
\tag{8.28}
$$

where the subscript denotes the vth equilibration step. In the ϕ representation, the analogous form is obtained (in this case, giving the loss in ϕ):

$$
-\Delta\phi_v \approx \frac{1}{2} \sum_{m,r=1}^{R} \Delta\lambda_{m,v} \, \gamma_{mr,v}|_{\{\lambda_r\}} \, \Delta\lambda_{r,v} = \frac{1}{2}\Delta\boldsymbol{\Lambda}_v{}^\top \boldsymbol{\gamma}_v \, \Delta\boldsymbol{\Lambda}_v
\tag{8.29}
$$

The total increase in entropy or decrease in potential of the system and environment subject to a M-step process is therefore:

$$
\Delta\mathfrak{H}_{tot}^* = \sum_{v=1}^{M} \Delta\mathfrak{H}_v^* \approx \sum_{v=1}^{M} \frac{1}{2}\Delta\boldsymbol{f}_v{}^\top \mathbf{g}_v \, \Delta\boldsymbol{f}_v = \sum_{v=1}^{M} \frac{1}{2}\dot{\boldsymbol{f}}_v{}^\top \mathbf{g}_v \, \dot{\boldsymbol{f}}_v \, \Delta\xi_v\Delta\xi_v,
\tag{8.30}
$$

$$
-\Delta\phi_{tot} = -\sum_{v=1}^{M} \Delta\phi_v \approx \sum_{v=1}^{M} \frac{1}{2}\Delta\boldsymbol{\Lambda}_v{}^\top \boldsymbol{\gamma}_v \, \Delta\boldsymbol{\Lambda}_v = \sum_{v=1}^{M} \frac{1}{2}\dot{\boldsymbol{\Lambda}}_v{}^\top \boldsymbol{\gamma}_v \, \dot{\boldsymbol{\Lambda}}_v \, \Delta\xi_v\Delta\xi_v
\tag{8.31}
$$

Recognising $\Delta\xi_v$ as the minimum dissipation parameter for the vth step (e.g. the minimum dissipation time if $\xi = t$), one such term may be factored out of each sum, to give the *mean minimum dissipation parameter* $\bar{\epsilon}_n$ for

$n \in \{\mathfrak{H}^*, \phi\}.^{17,25}$ This gives $\Delta \mathfrak{H}^*_{tot} = \bar{\epsilon}_{\mathfrak{H}^*} \mathcal{J}_{\mathfrak{H}^*}$ or $-\Delta \phi_{tot} = \bar{\epsilon}_\phi \mathcal{J}_\phi$, with:

$$\mathcal{J}_{\mathfrak{H}^*} = \sum_{\upsilon=1}^{M} \frac{1}{2} \dot{\boldsymbol{f}}_\upsilon^\top \, \mathbf{g}_\upsilon \, \dot{\boldsymbol{f}}_\upsilon \, \Delta \xi_\upsilon, \qquad (8.32)$$

$$\mathcal{J}_\phi = \sum_{\upsilon=1}^{M} \frac{1}{2} \dot{\boldsymbol{\Lambda}}_\upsilon^\top \, \boldsymbol{\gamma}_\upsilon \, \dot{\boldsymbol{\Lambda}}_\upsilon \, \Delta \xi_\upsilon \qquad (8.33)$$

The summands $\frac{1}{2} \dot{\boldsymbol{f}}_\upsilon^\top \mathbf{g}_\upsilon \dot{\boldsymbol{f}}_\upsilon$ or $\frac{1}{2} \dot{\boldsymbol{\Lambda}}_\upsilon^\top \boldsymbol{\gamma}_\upsilon \dot{\boldsymbol{\Lambda}}_\upsilon$ in (8.32)-(8.33) can be viewed as *generalised energy* terms, akin to the kinetic energy in mechanics, with the metric $\mathbf{g}_\upsilon$ or $\boldsymbol{\gamma}_\upsilon$ representing the "mass" and $\dot{\boldsymbol{f}}_\upsilon$ or $\dot{\boldsymbol{\Lambda}}_\upsilon$ the "velocity".[64] The terms $\mathcal{J}_{\mathfrak{H}^*}$ or $\mathcal{J}_\phi$ can then be interpreted as the discrete *generalised action* of the specified process,[41] again by analogy with mechanics[b]. From the previous considerations (§8.3), the two action sums are equivalent, although once again, discrepancies may emerge from their finite difference formulation.

From the discrete form of the Cauchy-Schwarz inequality:

$$\left(\sum_{\upsilon=1}^{M} a_\upsilon^2 \right) \left(\sum_{\upsilon=1}^{M} b_\upsilon^2 \right) \geq \left(\sum_{\upsilon=1}^{M} a_\upsilon b_\upsilon \right)^2 \qquad (8.34)$$

with $a_\upsilon = \sqrt{\dot{\boldsymbol{f}}^\top \mathbf{g}_\upsilon \dot{\boldsymbol{f}}} \, \Delta \xi_\upsilon$ or $\sqrt{\dot{\boldsymbol{\Lambda}}^\top \boldsymbol{\gamma}_\upsilon \dot{\boldsymbol{\Lambda}}} \, \Delta \xi_\upsilon$ and $b_\upsilon = 1$, it can be shown that:[25]

$$\bar{\epsilon}_n \mathcal{J}_n \geq \frac{L_n^2}{2M} \qquad (8.35)$$

Physically, the number of steps is equal to $M = \xi_{max}/\bar{\epsilon}_n$, whence (8.35) reduces to:[17,25]

$$\bar{\epsilon}_n \mathcal{J}_n \geq \frac{\bar{\epsilon}_n L_n^2}{2 \, \xi_{max}} \qquad \text{or} \qquad \mathcal{J}_n \geq \frac{L_n^2}{2 \, \xi_{max}} \qquad (8.36)$$

Eqs. (8.35)-(8.36) can be considered a generalised *least action bound*,[41] applicable to all probabilistic systems amenable to analysis by Jaynes' method. Its physical interpretation is that it specifies the minimum cost or penalty, in units of dimensionless entropy per unit ξ, to move the system from one stationary position (at $\xi = 0$) to another (at $\xi = \xi_{max}$) along the given path at the specified rates $\dot{\boldsymbol{\Lambda}}$ and/or $\dot{\boldsymbol{f}}$. If the latter rates proceed infinitely slowly, the lower bound of the action is zero, indicating that

---

[b]Crooks[41] applies the terms "energy" and "action" interchangeably; we consider that the present definitions are more in keeping with those used in mechanics. Many authors include the $\bar{\epsilon}_n$ term within $\mathcal{J}_n$, but we here wish to preserve the mathematical structure of a generalised action principle.

the process can be conducted at zero cost; otherwise, it is necessary to "do generalised work" to move the system along the manifold of stationary positions within a finite parameter duration ξ_{max}.

The generalised least action bound thus provides a lower bound for the "transition cost" of a process (in entropy-related units). If the process is reversible, the cost would be zero, but no process can be reversible in practice. Identification of this minimum cost is of paramount importance: there is no point in undertaking expensive changes to the process, or initiating costly social or political changes, in the attempt to do better than the minimum predicted by (8.35)-(8.36). Taking a thermodynamic example, the method can be applied to determine the minimum cost of industrial processes such as work extraction from combustion, a question of fundamental importance to human society. Most thermodynamics and engineering textbooks give the Carnot limit as the theoretical limit of efficiency, but the limits imposed by finite time thermodynamics are more restrictive (see §8.4.1).

The generalised least action bound therefore emerges from the Riemannian geometry of the state space, and hence from somewhat different considerations than the principle of least action employed in mechanics.[64] We consider that the two principles are connected, but are unable to examine this topic further here. For further exploratory expositions, the reader is referred to the work of Crooks,[41] Caticha[65] and Wang.[67,68]

The above discrete sums (8.30)-(8.31) can also be presented in integral form. Consider a system represented by $\mathfrak{H}^*$, subjected to a finite change in the multipliers $\Delta\lambda_r$ due to movement of the reference environment. The incremental change in entropy is, again to first order (compare (8.27)):[17,25,33,41]

$$d\mathfrak{H}^* \approx \frac{1}{2!} \sum_{r=1}^{R} \Delta\lambda_r d\langle f_r \rangle \tag{8.37}$$

Substituting $\Delta\lambda_r = \sum_{m=1}^{R} g_{mr}\Delta\langle f_m \rangle$ from (8.8), and assuming a first order decay process:

$$\langle \dot{f}_m \rangle = \frac{\langle f_m \rangle - \langle f_m \rangle_{env}}{\epsilon_{\mathfrak{H}^*}} = \frac{\Delta\langle f_m \rangle}{\epsilon_{\mathfrak{H}^*}} \tag{8.38}$$

where $\epsilon_{\mathfrak{H}^*}$ is a minimum dissipation parameter (reciprocal rate constant),

(8.37) yields:

$$d\mathfrak{H}^* = \frac{1}{2} \sum_{m,r=1}^{R} \langle \dot{f}_m \rangle \, g_{mr} \, d\langle f_r \rangle \, \epsilon_{\mathfrak{H}^*} \qquad (8.39)$$

The total change in entropy $\Delta \mathfrak{H}^*_{tot} = \int_0^{\xi_{max}} d\mathfrak{H}^*$ is then obtained as:

$$\Delta \mathfrak{H}^*_{tot} = \int\limits_0^{\xi_{max}} \frac{1}{2} \dot{\boldsymbol{f}}^\top \mathbf{g} \, \dot{\boldsymbol{f}} \, \epsilon_{\mathfrak{H}^*} \, d\xi = \bar{\epsilon}_{\mathfrak{H}^*} \int\limits_0^{\xi_{max}} \frac{1}{2} \dot{\boldsymbol{f}}^\top \mathbf{g} \, \dot{\boldsymbol{f}} \, d\xi = \bar{\epsilon}_{\mathfrak{H}^*} \mathcal{J}_{\mathfrak{H}^*} \quad (8.40)$$

Similarly, in the ϕ representation, we obtain:

$$-\Delta \phi_{tot} = \int\limits_0^{\xi_{max}} \frac{1}{2} \dot{\boldsymbol{\Lambda}}^\top \boldsymbol{\gamma} \, \dot{\boldsymbol{\Lambda}} \, \epsilon_\phi \, d\xi = \bar{\epsilon}_\phi \int\limits_0^{\xi_{max}} \frac{1}{2} \dot{\boldsymbol{\Lambda}}^\top \boldsymbol{\gamma} \, \dot{\boldsymbol{\Lambda}} \, d\xi = \bar{\epsilon}_\phi \mathcal{J}_\phi \qquad (8.41)$$

In the continuous representation, the process does not proceed by a series of finite steps; instead, the reference variables continuously move ahead of those of the system.[17,25,33] However, we still see the influence of a finite decay parameter ϵ_n, which on integration yields the mean parameter $\bar{\epsilon}_n$. Each $\mathcal{J}_n$ term above can be regarded as the *action integral* corresponding respectively to (8.32)-(8.33). Based on the integral form of the Cauchy-Schwarz inequality,[62] it can be shown that the integral actions also satisfy the least action bound (8.35)-(8.36), with L_n in integral form.[17,25,33,41]

Finally, for the least action bound (8.36) to achieve equality, the summands or integrands of the arc length L_n and action J_n must be constant. This gives the simple result that for slow processes with constant dissipation parameter $\epsilon_n = \bar{\epsilon}_n$, the minimum action (whence minimum in $\bar{\epsilon}_n J_n$) is attained by a process which proceeds at a constant speed $\dot{s}_n = \sqrt{\dot{\boldsymbol{f}}^\top \mathbf{g} \, \dot{\boldsymbol{f}}} = \sqrt{\dot{\boldsymbol{\Lambda}}^\top \boldsymbol{\gamma} \, \dot{\boldsymbol{\Lambda}}}$.[28,29,33] For a constant metric, this is equivalent to constant rates of change of the parameter vector $\dot{\boldsymbol{f}}$ and/or $\dot{\boldsymbol{\Lambda}}$. For systems with a variable dissipation parameter $\epsilon(\xi)$, it was first considered that the minimum is attained at the constant speed $ds/d\eta$, expressed in the "natural" parameter units $\eta = \xi/\epsilon$.[28,29,33,40] This however oversimplifies the minimisation problem, which is better handled within a discrete (stepwise) framework.[25,35] As discussed in §8.4.1, such principles have been widely applied to thermodynamic systems.

8.3.3. *Minimum Path Length Principle*

The above discrete or continuous forms of the least action bound (8.35)-(8.36) are based on consideration of a specified path on the manifold of stationary positions, of arc length L_n. In many situations, we may wish to determine the path of minimum arc length $L_{n,min}$ – the *geodesic* – on the manifold of stationary positions. From the calculus of variations, this is given by the Euler-Lagrange equations:[66]

$$\frac{\partial \dot{s}_{\mathfrak{H}^*}}{\partial \boldsymbol{f}} - \frac{d}{d\xi} \frac{\partial \dot{s}_{\mathfrak{H}^*}}{\partial \dot{\boldsymbol{f}}} = \boldsymbol{0} \qquad (8.42)$$

$$\frac{\partial \dot{s}_{\phi}}{\partial \boldsymbol{\Lambda}} - \frac{d}{d\xi} \frac{\partial \dot{s}_{\phi}}{\partial \dot{\boldsymbol{\Lambda}}} = \boldsymbol{0} \qquad (8.43)$$

where $\dot{s}_{\mathfrak{H}^*} = \sqrt{\dot{\boldsymbol{f}}^{\top} \mathbf{g} \, \dot{\boldsymbol{f}}}$ and $\dot{s}_{\phi} = \sqrt{\dot{\boldsymbol{\Lambda}}^{\top} \boldsymbol{\gamma} \, \dot{\boldsymbol{\Lambda}}}$ are the integrands respectively of $L_{\mathfrak{H}^*}$ or L_{ϕ} (8.16)-(8.19). For two-dimensional parameters $\boldsymbol{f}, \boldsymbol{\Lambda} \in \mathbb{R}^2$, (8.42)-(8.43) can be reduced further in terms of the three unit normals to the surface, giving the curve(s) on the manifold for which the geodesic curvature vanishes.[43,60,62] Depending on the specified problem, a geodesic may not exist, or there may be multiple or il-defined solutions. Provided it does exist, a geodesic leads to the double minimisation principle:

$$\mathcal{J}_n \geq \frac{L_n^2}{2\,\xi_{max}} \geq \frac{L_{n,min}{}^2}{2\,\xi_{max}} \qquad (8.44)$$

where the right hand side indicates the absolute lower bound for the action, irrespective of path. This principle has been applied to thermodynamic systems, as will be discussed in §8.4.1.

8.4. Applications

As noted, the foregoing Riemannian geometric interpretation (§8.3) has mainly been presented within an equilibrium thermodynamics context,[15–43] although it has been applied to non-equilibrium thermodynamic and flow systems,[69–73] information coding[74] and in economics.[75] In the following sections, the utility of Riemannian geometric properties and the least action bound are demonstrated for two types of system: a thermodynamic system at equilibrium, and a flow system at steady state.

8.4.1. *Equilibrium Thermodynamic Systems*

The application of Riemannian geometric principles to equilibrium thermodynamic systems has constituted a major new development over the past three decades, forming an important plank of *finite-parameter* or (with $\xi = t$) *finite-time thermodynamics*.[21,34,35] Such analyses have progressed in four overlapping stages:

- The initial studies by Weinhold[10–14] and early work by Salamon, Andresen, Berry, Nulton and coworkers[15,17,18,20,25,28,29] all examined a manifold based on an internal energy representation $U(X_1, X_2, ...)$, as a function of extensive variables X_r, which include the thermodynamic entropy S. The resulting quantity $\bar{\epsilon}_U \mathcal{J}_U$ (in the present notation) was interpreted as an *availability* or *exergy* function, with (8.36) indicating the most efficient path (defined by the minimum amount of work or minimum loss of availability) required to move the equilibrium position of the system.[17,18] Such analyses complement the thermodynamic geometry used by Gibbs,[58,59] and fit well with the traditional heat-work framework of 19th century thermodynamics.

- Subsequently, following earlier pioneering works,[69,76] the entropy manifold $S(\tilde{X}_1, \tilde{X}_2, ...)$ was examined from a Riemannian perspective,[20,21,23,25,31–33,33,34,41,42] where $\tilde{X}_r$ are the new extensive variables, of course related to the $U(X_1, X_2, ...)$ representation by Jacobian transformation.[20] The quantity $\bar{\epsilon}_S \mathcal{J}_S$ was interpreted as a measure of *energy dissipation* or *entropy production*, again providing a measure of process efficiency. It was realised that the lower bound in (8.36) provides a formal, mathematical definition of the degree of irreversibility of a transition between equilibrium positions, with reversibility only for $\mathcal{J}_S = 0$ (a definition vastly preferable to the cumbersome word-play still used in thermodynamics references; see the scathing criticism by Truesdell[77]). However, the primacy of the entropy representation over that based on internal energy was not fully appreciated in these early studies. The applicability of Riemannian geometry in other contexts – based directly on the MaxEnt framework of Jaynes[1,2] – is hinted at by Levine,[27] but unfortunately was not developed further at the time, nor, to the authors' knowledge, in any subsequent studies.

- Several studies have considered an entropy representation based on a metric defined on a probability space $\{p_i\}$, either from the Boltzmann principle[76] or using a Shannon or relative entropy measure.[22–24,26,27,37–39,41–43]

Several authors[37–39,41–43]) have extended this analysis, to establish a connection with the Fisher information matrix[78] and an "entropy differential metric" of Rao.[79] The analysis is also intimately connected with paths in a space of square root probabilities, and thence to formulations of quantum mechanics.[37–39] These insights – not examined further here – demand further detailed attention; they may well furnish an explanation for the utility of extremisation methods based on the Fisher information function in many physical problems.[80]

- Finally, several workers realised that Riemannian geometric principles can be applied to Legendre-transformed representations, e.g. based on various forms of the free energy F (conjugate to U) or the negative Planck potential F/T (conjugate to S), as functions of the intensive variables (or functions thereof).[23,27,41,42] This approach offers particular advantages for the analysis of real thermodynamic systems, in which the control parameters tend to be intensive rather than extensive variables (the canonical ensemble), and for which the intensive variables do not exhibit sharp transitions or singularities associated with phase changes, as is the case for extensive variables.[41] Furthermore, the resulting metric is equivalent to the variance-covariance matrix of the constraints (8.10), and is therefore connected to fluctuation-dissipation processes within the system.

For completeness, we demonstrate – for a microcanonical thermodynamic system – how Riemannian geometric properties emerge as an inherent feature of Jaynes' MaxEnt formulation. Consider an isolated thermodynamic system, containing molecules of possible energy levels ϵ_i and volume elements V_j, subject to constraints on the mean energy $\langle U \rangle$ and mean volume $\langle V \rangle$. We consider the joint probability p_{ij} of a particle simultaneously occupying an energy level and volume element, giving the entropy function:

$$\mathfrak{H}_{eq} = -\sum_i \sum_j p_{ij} \ln p_{ij}, \tag{8.45}$$

where, without knowledge of any additional influences, we assume that each joint level ij is equally probable (hence the priors q_{ij} cancel out). Eq. (8.45) is maximised subject to the constraints:

$$\sum_i \sum_j p_{ij} = 1, \tag{8.46}$$

$$\sum_i \sum_j p_{ij} \epsilon_i = \langle U \rangle, \tag{8.47}$$

$$\sum_i \sum_j p_{ij} V_j = \langle V \rangle, \tag{8.48}$$

to give the equilibrium position:

$$p_{ij}^* = \frac{\exp\left(-\lambda_U U_i - \lambda_V V_j\right)}{\sum_i \sum_j \exp\left(-\lambda_U U_i - \lambda_V V_j\right)} = \frac{1}{Z} \exp\left(-\lambda_U U_i - \lambda_V V_j\right), \tag{8.49}$$

where Z is the partition function. From the existing body of thermodynamics, we can identify the Lagrangian multipliers as $\lambda_U = 1/kT$ and $\lambda_V = P/kT$, where k is the Boltzmann constant, T is absolute temperature and P is absolute pressure. Eq. (8.49) and Jaynes' relations (8.6)-(8.11) and (8.13) then reduce to:

$$p_{ij}^* = \frac{1}{Z} e^{-U_i/kT - PV_j/kT}, \tag{8.50}$$

$$S^* = k\mathfrak{H}_{eq}^* = k \ln Z + \frac{\langle U \rangle}{T} + \frac{P \langle V \rangle}{T} \tag{8.51}$$

$$k\,\mathbf{\Lambda}_{eq} = \left[\frac{\partial S^*}{\partial \langle U \rangle}, \frac{\partial S^*}{\partial \langle V \rangle} \right]^\top = \left[\frac{1}{T}, \frac{P}{T} \right]^\top \tag{8.52}$$

$$k\,\mathbf{g}_{eq} = \begin{bmatrix} \dfrac{\partial^2 S^*}{\partial \langle U \rangle^2}, & \dfrac{\partial^2 S^*}{\partial \langle U \rangle \partial \langle V \rangle} \\ \dfrac{\partial^2 S^*}{\partial \langle V \rangle \partial \langle U \rangle}, & \dfrac{\partial^2 S^*}{\partial \langle V \rangle^2} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial}{\partial \langle U \rangle}\left(\dfrac{1}{T}\right), & \dfrac{\partial}{\partial \langle U \rangle}\left(\dfrac{P}{T}\right) \\ \dfrac{\partial}{\partial \langle V \rangle}\left(\dfrac{1}{T}\right), & \dfrac{\partial}{\partial \langle V \rangle}\left(\dfrac{P}{T}\right) \end{bmatrix} \tag{8.53}$$

$$\psi = k\phi_{eq} = -k \ln Z = -S^* + \frac{\langle U \rangle}{T} + \frac{P \langle V \rangle}{T} = \frac{G}{T} \tag{8.54}$$

$$\boldsymbol{f}_{eq} = \left[\frac{\partial \psi}{\partial\left(\frac{1}{T}\right)}, \frac{\partial \psi}{\partial\left(\frac{P}{T}\right)} \right]^\top = \left[\langle U \rangle, \langle V \rangle \right]^\top \tag{8.55}$$

$$\frac{\boldsymbol{\gamma}_{eq}}{k} = \begin{bmatrix} \dfrac{\partial^2 \psi}{\partial\left(\frac{1}{T}\right)^2}, & \dfrac{\partial^2 \psi}{\partial\left(\frac{1}{T}\right)\partial\left(\frac{P}{T}\right)} \\ \dfrac{\partial^2 \psi}{\partial\left(\frac{P}{T}\right)\partial\left(\frac{1}{T}\right)}, & \dfrac{\partial^2 \psi}{\partial\left(\frac{P}{T}\right)^2} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial \langle U \rangle}{\partial\left(\frac{1}{T}\right)}, & \dfrac{\partial \langle V \rangle}{\partial\left(\frac{1}{T}\right)} \\ \dfrac{\partial \langle U \rangle}{\partial\left(\frac{P}{T}\right)}, & \dfrac{\partial \langle V \rangle}{\partial\left(\frac{P}{T}\right)} \end{bmatrix} \tag{8.56}$$

$$k\,\mathbf{g}_{eq}\,\frac{\boldsymbol{\gamma}_{eq}}{k} = \mathbf{I} \tag{8.57}$$

where S^* is the thermodynamic entropy at an equilibrium position, ψ is the negative Planck potential[81,82] (negative Massieu function[83]) and G is the Gibbs free energy. By Jacobian transformation of variables, using the

following material properties (susceptibilities):[12,55,56]

$$
\text{Heat capacity at constant pressure:} \quad C_P = \left(\frac{\partial \langle H \rangle}{\partial T} \right)_P \tag{8.58}
$$

$$
\text{Isothermal compressibility:} \quad \kappa_T = -\frac{1}{\langle V \rangle} \left(\frac{\partial \langle V \rangle}{\partial P} \right)_T \tag{8.59}
$$

$$
\text{Coefficient of thermal expansion:} \quad \alpha = \frac{1}{\langle V \rangle} \left(\frac{\partial \langle V \rangle}{\partial T} \right)_P \tag{8.60}
$$

where $\langle H \rangle = \langle U \rangle + P \langle V \rangle$ is the enthalpy, as well as the equality of cross-derivatives (Maxwell relation):

$$
\frac{\partial \langle V \rangle}{\partial (\frac{1}{T})} = \frac{\partial \langle U \rangle}{\partial (\frac{P}{T})} \tag{8.61}
$$

the ψ metric (8.56) reduces to[c]:

$$
\frac{\boldsymbol{\gamma}_{eq}}{k} = T \langle V \rangle \begin{bmatrix} -\kappa_T P^2 + 2\,\alpha PT - \frac{C_P T}{\langle V \rangle}, & \kappa_T P - \alpha T \\ \kappa_T P - \alpha T, & -\kappa_T \end{bmatrix} \tag{8.62}
$$

whence from (8.57):

$$
k\mathbf{g}_{eq} = \frac{1}{T^2(\alpha^2 T \langle V \rangle - \kappa_T C_P)} \begin{bmatrix} \kappa_T, & \kappa_T P - \alpha T \\ \kappa_T P - \alpha T, & \kappa_T P^2 - 2\alpha PT + \frac{C_P T}{\langle V \rangle} \end{bmatrix} \tag{8.63}
$$

Using (8.52), (8.55) and (8.62)-(8.63), the (dimensional) arc lengths (8.18)-(8.19) and action integrals (8.40)-(8.41) are obtained as:

$$
\breve{L}_{S^*} = \int_0^{\xi_{max}} \sqrt{\dot{\boldsymbol{f}}_{eq}^{\top} k\mathbf{g}_{eq} \dot{\boldsymbol{f}}_{eq}}\, d\xi = \int_0^{\xi_{max}} \sqrt{\frac{-C_P \dot{T}^2 + 2\alpha\langle V \rangle T \dot{T}\dot{P} - \kappa_T \langle V \rangle T \dot{P}^2}{k^2 T^2}}\, d\xi \tag{8.64}
$$

$$
\breve{\mathcal{J}}_{S^*} = \int_0^{\xi_{max}} \frac{1}{2}\, \dot{\boldsymbol{f}}_{eq}^{\top} k\mathbf{g}_{eq} \dot{\boldsymbol{f}}_{eq}\, d\xi = \int_0^{\xi_{max}} \frac{-C_P \dot{T}^2 + 2\alpha\langle V \rangle T \dot{T}\dot{P} - \kappa_T \langle V \rangle T \dot{P}^2}{2k^2 T^2}\, d\xi \tag{8.65}
$$

---

[c]The first variance is given erroneously by Callen.[55,56]

Jaynes' MaxEnt, Riemannian Metrics and Least Action Bound   303

$$\breve{L}_\psi = \int\limits_0^{\xi_{max}} \sqrt{\dot{\boldsymbol{\Lambda}}_{eq}{}^\top \frac{\boldsymbol{\gamma}_{eq}}{k} \dot{\boldsymbol{\Lambda}}_{eq}}\, d\xi$$

$$= \int\limits_0^{\xi_{max}} \sqrt{\frac{\langle V\rangle\langle\dot{U}\rangle\left[\kappa_T\langle\dot{U}\rangle + 2\langle\dot{V}\rangle(\kappa_T P - \alpha T)\right] + \langle\dot{V}\rangle^2\left[P\langle V\rangle(\kappa_T P - 2\alpha T) + C_P T\right]}{T^2\langle V\rangle(-C_P\kappa_T + \alpha^2 T\langle V\rangle)}}\, d\xi$$

$$(8.66)$$

$$\breve{\mathcal{J}}_\psi = \int\limits_0^{\xi_{max}} \frac{1}{2} \dot{\boldsymbol{\Lambda}}_{eq}{}^\top \frac{\boldsymbol{\gamma}_{eq}}{k} \dot{\boldsymbol{\Lambda}}_{eq}\, d\xi$$

$$= \int\limits_0^{\xi_{max}} \frac{\langle V\rangle\langle\dot{U}\rangle\left[\kappa_T\langle\dot{U}\rangle + 2\langle\dot{V}\rangle(\kappa_T P - \alpha T)\right] + \langle\dot{V}\rangle^2\left[P\langle V\rangle(\kappa_T P - 2\alpha T) + C_P T\right]}{2T^2\langle V\rangle(-C_P\kappa_T + \alpha^2 T\langle V\rangle)}\, d\xi$$

$$(8.67)$$

Using (8.58)-(8.60), these two sets of measures can be shown to be equivalent. The above equations must be integrated along the particular thermodynamic path followed by the process, as defined by the velocities $\{\dot{T}, \dot{P}\}$ or $\{\langle\dot{U}\rangle, \langle\dot{V}\rangle\}$. For a process which follows a pre-determined path, e.g. an adiabatic, isothermal, isovolumetric or isopiezometric curve, this can be simplified by expressing the velocities (e.g. $\dot{P}$) as functions of one independent velocity (e.g. $\dot{T}$).

To comment on units: if the above quantities were calculated using the "pure" metrics $\mathbf{g}_{eq}$ or $\boldsymbol{\gamma}_{eq}$, in either case the line element ds_n, arc length L_n and the term $\bar{\epsilon}_n\mathcal{J}_n$ would be dimensionless (whence the action is in reciprocal ξ units). Use of the "natural" metric $k\mathbf{g}_{eq}$, as conducted here, gives the line element and arc length in $\sqrt{JK^{-1}}$ and the action in $JK^{-1}\xi^{-1}$, consistent with $\Delta S_{tot}^* = \bar{\epsilon}_{S^*}\mathcal{J}_{S^*}$ being in entropy units. In contrast, use of the "natural" metric $\boldsymbol{\gamma}_{eq}/k$ gives the line element and arc length in $\sqrt{KJ^{-1}}$ and the action in $KJ^{-1}\xi^{-1}$. The latter case can be rescued by use of a modified line element $d\breve{s}'_\psi = \sqrt{k\dot{\boldsymbol{\Lambda}}_{eq}{}^\top \frac{\boldsymbol{\gamma}_{eq}}{k} k\dot{\boldsymbol{\Lambda}}_{eq}}\, d\xi$ – as suggested by (8.52) and (8.56) – giving the line element and arc length in $\sqrt{JK^{-1}}$ and the action in $JK^{-1}\xi^{-1}$. Thus in both the S^* and ψ representations, the least action bound (8.36) can be used to determine the *minimum entropy cost* of a transition from one equilibrium position to another, along a specified path on the manifold of equilibrium positions. As noted earlier, for slow processes and constant ϵ_n, this is attained by a process which proceeds at a constant thermodynamic speed $\dot{s}$[28,29,33,40] (a more general result is available

for rapid processes[84]). For variable ϵ_n and/or for stepwise phenomena, the process should be divided into individual steps placed at equal distances along the arc length traversed by the process, giving the so-called "equal thermodynamics distance" principle.[34–36] Such considerations have been applied to the optimisation of a wide variety of engineering and industrial batch and flow processes, including engine cycles, heat engines and pumps, chemical reactors, distillation towers and many other systems.

A final important point is that the minimum path length (double minimisation) principle (8.44) – involving calculation of the geodesic – has been applied to the analysis of equilibrium systems. In early work, this bound was established by applying the calculus of variations directly to particular thermodynamic problems, without use of a metric.[16,19] More recently, such lower bounds have been examined for particular thermodynamic systems.[34,43,100] In either case, for the entropy representation, this method yields an *absolute minimum entropy cost* $\Delta S^* \geq \Delta S^*_{min}$ for a transition between two equilibrium positions at particular rates of change, irrespective of the path. For cyclic or flow processes, this therefore gives a *minimum entropy production principle* $\dot{S} \geq \dot{S}_{min}$, providing one of the key concepts of finite-time (or finite-parameter) thermodynamics.

8.4.2. *Flow Systems*

We now consider a flow system consisting of a control volume, subject to continuous flows of heat, particles and momentum, and within which chemical reactions may take place. A few workers have examined such non-equilibrium systems previously within a Riemannian context, including for the Onsager linear regime[69,70] and for extended irreversible thermodynamics.[71–73] A different perspective is provided here, based on a recent analysis of a flow system from a Jaynesian perspective.[63] This involves a probabilistic analysis of each infinitesimal element of the control volume, which experiences instantaneous values of the heat flux $\boldsymbol{j}_{Q,\mathcal{I}}$, mass fluxes $\boldsymbol{j}_{\mathcal{N}_c}$ of each species c, stress tensor $\boldsymbol{\tau}_{\mathcal{J}}$ and molar rate per unit volume $\hat{\dot{\xi}}_{\mathcal{L}_d}$ of each chemical reaction d, where the indices $\mathcal{I}, \mathcal{J}, \mathcal{L}_d, \mathcal{N}_c \in \{0, \pm 1, \pm 2, ...\}$. We therefore consider the joint probability $\pi_{\boldsymbol{I}} = \pi_{\mathcal{I}, \mathcal{J}, \{\mathcal{L}_d\}, \{\mathcal{N}_c\}}$ of instantaneous fluxes through the element and instantaneous reactions within the element, giving the (dimensionless) "flux entropy" function:

$$\mathfrak{H}_{st} = -\sum_{\boldsymbol{I}} \pi_{\boldsymbol{I}} \ln \pi_{\boldsymbol{I}}, \tag{8.68}$$

Again assuming that each joint level $\boldsymbol{I}$ is equally probable, (8.68) is maximised subject to constraints on the mean values of the heat flux $\langle \boldsymbol{j}_Q \rangle$, mass fluxes $\langle \boldsymbol{j}_c \rangle$, stress tensor $\langle \boldsymbol{\tau} \rangle$ and molar reaction rates $\langle \hat{\dot{\xi}}_d \rangle$ through or within the element, as well as by the natural constraint (8.3). This gives the *steady state position* of the system:

$$\pi_{\boldsymbol{I}}^* = \frac{1}{\mathcal{Z}} \exp\left(-\boldsymbol{\zeta}_Q \cdot \boldsymbol{j}_{Q,\mathcal{I}} - \sum_c \boldsymbol{\zeta}_c \cdot \boldsymbol{j}_{\mathcal{N}_c} - \boldsymbol{\zeta}_{\boldsymbol{\tau}} : \boldsymbol{\tau}_{\mathcal{J}} - \sum_d \zeta_d \hat{\dot{\xi}}_{\mathcal{L}_d} \right) \quad (8.69)$$

where $\boldsymbol{\zeta}_Q, \boldsymbol{\zeta}_c, \boldsymbol{\zeta}_{\boldsymbol{\tau}}$ and ζ_d are the Lagrangian multipliers associated with the heat, particle, momentum and chemical reaction constraints, and $\mathcal{Z} = e^{\zeta_0}$ is the partition function. By a traditional control volume analysis,[85–88] the multipliers can be identified as:[63]

$$\boldsymbol{\zeta}_Q = -\frac{\theta \mathcal{V}}{k} \boldsymbol{\nabla}\left(\frac{1}{T} \right) \quad (8.70)$$

$$\boldsymbol{\zeta}_c = \frac{\theta \mathcal{V}}{k} \left[\boldsymbol{\nabla}\left(\frac{\mu_c}{M_c T} \right) - \frac{\boldsymbol{F}_c}{T} \right] \quad (8.71)$$

$$\boldsymbol{\zeta}_{\boldsymbol{\tau}} = \frac{\theta \mathcal{V}}{k} \boldsymbol{\nabla}\left(\frac{\boldsymbol{v}}{T} \right)^{\top} \quad (8.72)$$

$$\zeta_d = \frac{\theta \mathcal{V}}{k} \frac{A_d}{T} \quad (8.73)$$

where μ_c is the chemical potential of the cth constituent, M_c is the molar mass of the cth constituent, $\boldsymbol{F}_c$ is the specific body force on species c, $\boldsymbol{v}$ is the mass-average velocity, A_d is the chemical affinity of the dth reaction (< 0 for a spontaneous reaction), $\boldsymbol{\nabla}$ is the Cartesian gradient operator, and θ and $\mathcal{V}$ respectively are characteristic time and volume scales of the system. Generalising each component of the above multipliers as ζ_r and constraints as $\langle j_r \rangle$ with $r \in \{1, ..., R\}$, Jaynes' relations (8.6)-(8.11) and (8.13) reduce to:

$$\mathfrak{H}_{st}^* = \ln \mathcal{Z} + \sum_{r=1}^{R} \zeta_r \langle j_r \rangle = -\phi_{st} - \frac{\theta \mathcal{V}}{k} \hat{\sigma} \quad (8.74)$$

$$\boldsymbol{\Lambda}_{st} = \left[\frac{\partial \mathfrak{H}_{st}^*}{\partial \langle j_1 \rangle}, ..., \frac{\partial \mathfrak{H}_{st}^*}{\partial \langle j_R \rangle} \right]^{\top} = \left[\zeta_1, ..., \zeta_R \right]^{\top} \quad (8.75)$$

$$\mathbf{g}_{st} = \begin{bmatrix} \dfrac{\partial^2 \mathfrak{H}_{st}^*}{\partial \langle j_1 \rangle^2} & \cdots & \dfrac{\partial^2 \mathfrak{H}_{st}^*}{\partial \langle j_1 \rangle \partial \langle j_R \rangle} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 \mathfrak{H}_{st}^*}{\partial \langle j_R \rangle \partial \langle j_1 \rangle} & \cdots & \dfrac{\partial^2 \mathfrak{H}_{st}^*}{\partial \langle j_R \rangle^2} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial \zeta_1}{\partial \langle j_1 \rangle} & \cdots & \dfrac{\partial \zeta_R}{\partial \langle j_1 \rangle} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial \zeta_1}{\partial \langle j_R \rangle} & \cdots & \dfrac{\partial \zeta_R}{\partial \langle j_R \rangle} \end{bmatrix} \quad (8.76)$$

$$\phi_{st} = -\ln \mathcal{Z} = -\mathfrak{H}_{st}^* + \sum_{r=1}^{R} \zeta_r \langle j_r \rangle = -\mathfrak{H}_{st}^* - \frac{\theta \mathcal{V}}{k} \hat{\sigma} \qquad (8.77)$$

$$\boldsymbol{f}_{st} = \left[\frac{\partial \phi_{st}}{\partial \zeta_1}, ..., \frac{\partial \phi_{st}}{\partial \zeta_R} \right]^\top = \left[\langle j_1 \rangle, ..., \langle j_R \rangle \right]^\top \qquad (8.78)$$

$$\boldsymbol{\gamma}_{st} = \begin{bmatrix} \dfrac{\partial^2 \phi_{st}}{\partial \zeta_1^2} & \cdots & \dfrac{\partial^2 \phi_{st}}{\partial \zeta_1 \partial \zeta_R} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 \phi_{st}}{\partial \zeta_R \partial \zeta_1} & \cdots & \dfrac{\partial^2 \phi_{st}}{\partial \zeta_R^2} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial \langle j_1 \rangle}{\partial \zeta_1} & \cdots & \dfrac{\partial \langle j_R \rangle}{\partial \zeta_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial \langle j_1 \rangle}{\partial \zeta_R} & \cdots & \dfrac{\partial \langle j_R \rangle}{\partial \zeta_R} \end{bmatrix} \qquad (8.79)$$

$$\mathbf{g}_{st} \, \boldsymbol{\gamma}_{st} = \mathbf{I} \qquad (8.80)$$

where $\hat{\sigma}$ can be identified as the local entropy production per unit volume (units of $JK^{-1}m^{-3}s^{-1}$). A flow system subject to constant flux and reaction rate constraints will therefore converge to a steady state position defined by a maximum in the flux entropy $\mathfrak{H}_{st}^*$ and a minimum in the flux potential ϕ_{st}. If these effects occur simultaneously, the system will converge to a position of maximum $\hat{\sigma}$, therefore providing a conditional, local derivation of the *maximum entropy production* (MEP) principle,[63] which has been applied as a discriminator to determine the steady state of many non-linear flow systems.[89–97]

In Onsager's analysis of transport phenomena in the vicinity of equilibrium,[98,99] the fluxes and reaction rates are considered to be linear functions of the "forces" (the driving gradients and chemical affinities). In the present terminology, this would be written as:

$$\langle j_r \rangle = K \sum_m L_{rm}^0 \zeta_m \qquad (8.81)$$

where L_{rm}^0 are the (constant) phenomenological coefficients at the zero-gradient position (i.e., at equilibrium) and $K = k/\theta \mathcal{V}$. In the present analysis, we do not claim linearity between $\langle j_r \rangle$ and ζ_m, nor consider that the system is "close to equilibrium", but simply adopt the partial derivatives $\partial \langle j_r \rangle / \partial \zeta_m$ within the metric $\boldsymbol{\gamma}_{st}$ (8.79) as a set of parameters (functions of ζ_m) with which to analyse the system. The present analysis therefore encompasses, but is not restricted to, Onsager's linear regime. The diagonal and many off-diagonal terms can readily be identified as functions of the conductivities (transport coefficients) and chemical reaction rate coef-

ficents:[88]

Heat conductivity:
$$\widetilde{\kappa}_{\imath\jmath} = -\frac{\partial\langle j_{Q\imath}\rangle}{\partial\left(\dfrac{\partial T}{\partial \jmath}\right)} \tag{8.82}$$

Diffusion coefficient, species c:
$$\widetilde{D}^c_{\imath\jmath} = -\frac{\partial\langle j_{c\imath}\rangle}{\partial\left(\dfrac{\partial \hat{C}_c}{\partial \jmath}\right)} \tag{8.83}$$

Viscosity coefficient:
$$\widetilde{\mu}_{\imath\jmath\kappa\ell} = -\frac{\partial\langle \tau_{\imath\jmath}\rangle}{\partial\left(\dfrac{\partial v_\kappa}{\partial \ell}\right)} \tag{8.84}$$

Rate coefficient, reaction d:
$$\widetilde{k}_d = \frac{\partial\langle \hat{\dot{C}}_{cd}\rangle}{\partial \hat{C}_c} = \nu_{cd} M_c \frac{\partial\langle \hat{\dot{\xi}}_d\rangle}{\partial \hat{C}_c} \tag{8.85}$$

where $\hat{C}_c$ is the concentration of species c (units of kg m^{-3}; often used as a proxy for the chemical potential μ_c), $\langle\hat{\dot{C}}_{cd}\rangle$ is the mean rate of change of concentration of species c in the dth reaction (units of kg m^{-3} s^{-1}), ν_{cd} is the stoichiometric coefficient of species c in the dth reaction (positive if a product), and the indices $\imath,\jmath,\kappa,\ell \in \{x,y,z\}$. The remaining off-diagonal terms consist of the cross-process conductivity coupling coefficients and conductivity-reaction rate coefficients. The Riemannian metric $\boldsymbol{\gamma}_{st}$ can therefore be regarded as a function of the material properties or susceptibilities of a flow and chemical reactive system, in the same way that the Riemannian metric for an equilibrium system $\boldsymbol{\gamma}_{eq}$ is a function of its various susceptibilities, such as C_P, κ_T and α (§8.4.1). As with equilibrium systems, an abrupt change in a given component $\gamma_{st,rm}$ with ζ_m can be interpreted as the boundary of a phase change in the system. Notice also that symmetry of $\boldsymbol{\gamma}_{st}$ yields a set of Maxwell-like relations:[63]

$$\frac{\partial\langle j_r\rangle}{\partial \zeta_m} = \frac{\partial\langle j_m\rangle}{\partial \zeta_r} \tag{8.86}$$

These apply to all infinitesimal volume elements of a flow system, not merely those in the vicinity of equilibrium. Eqs. (8.86) considerably simplify the set of parameters needed for analysis, from R^2 to $\binom{R+1}{2}$ coefficients; further simplifications may be attainable in certain systems due to geometric and tensor symmetries.[88]

The above relations (8.74)-(8.79) can now be applied to develop a Riemannian description of a flow system on the manifold of steady state positions. In terms of the generalised derivatives, the (dimensionless) arc

lengths (8.18)-(8.19) and action integrals (8.40)-(8.41) are obtained as:

$$
\begin{aligned}
L_{st} &= \int_0^{\xi_{max}} \sqrt{\dot{\boldsymbol{f}}_{st}^{\top} \mathbf{g}_{st} \dot{\boldsymbol{f}}_{st}}\, d\xi = \int_0^{\xi_{max}} \sqrt{\dot{\boldsymbol{\Lambda}}_{st}^{\top} \boldsymbol{\gamma}_{st} \dot{\boldsymbol{\Lambda}}_{st}}\, d\xi \\
&= \int_0^{\xi_{max}} \sqrt{\dot{\boldsymbol{\Lambda}}_{st} \cdot \dot{\boldsymbol{f}}_{st}}\, d\xi = \int_0^{\xi_{max}} \sqrt{\sum_{r=1}^R \frac{\partial \zeta_r}{\partial \xi} \frac{\partial \langle j_r \rangle}{\partial \xi}}\, d\xi
\end{aligned}
\tag{8.87}
$$

$$
\begin{aligned}
\mathcal{J}_{st} &= \int_0^{\xi_{max}} \frac{1}{2} \dot{\boldsymbol{f}}_{st}^{\top} \mathbf{g}_{st} \dot{\boldsymbol{f}}_{st}\, d\xi = \int_0^{\xi_{max}} \frac{1}{2} \dot{\boldsymbol{\Lambda}}_{st}^{\top} \boldsymbol{\gamma}_{st} \dot{\boldsymbol{\Lambda}}_{st}\, d\xi \\
&= \int_0^{\xi_{max}} \frac{1}{2} \dot{\boldsymbol{\Lambda}}_{st} \cdot \dot{\boldsymbol{f}}_{st}\, d\xi = \frac{1}{2} \int_0^{\xi_{max}} \sum_{r=1}^R \frac{\partial \zeta_r}{\partial \xi} \frac{\partial \langle j_r \rangle}{\partial \xi}\, d\xi
\end{aligned}
\tag{8.88}
$$

where, as shown, the two alternative $\mathfrak{H}_{st}^*$ and ϕ_{st} measures are equivalent. Once again, these equations must be integrated along the particular path taken between the initial and final steady state positions.

To comment on units: since the above quantities are calculated using the "pure" metrics $\mathbf{g}_{st}$ or $\boldsymbol{\gamma}_{st}$, the resulting line element ds_{st}, arc length L_{st} and the term $\bar{\epsilon}_{st} \mathcal{J}_{st}$ are dimensionless. Use of the "natural" metric $K\mathbf{g}_{st}$, for $K = k/\theta\mathcal{V}$, therefore gives the line element and arc length in $\sqrt{JK^{-1}m^{-3}s^{-1}}$ and the action in $JK^{-1}m^{-3}s^{-1}\xi^{-1}$, thereby giving $\bar{\epsilon}_{st} \mathcal{J}_{st}$ in units of entropy production per unit volume. Similarly, use of the "natural" metric $\boldsymbol{\gamma}_{st}/K$ in conjunction with the dimensional constraint vector $K\dot{\boldsymbol{\Lambda}}_{st}$ gives the line element and arc length in $\sqrt{JK^{-1}m^{-3}s^{-1}}$ and action in $JK^{-1}m^{-3}s^{-1}\xi^{-1}$, again giving $\bar{\epsilon}_{st} \mathcal{J}_{st}$ in units of entropy production per unit volume. The least action bound (8.36) therefore yields a *minimum entropy production principle*, which sets a lower bound for the entropy production associated with movement of a flow system from one steady state position to another along a specified path. From the previous analysis, this involves two separate minimisation principles:

- If the path is specified, the process of minimum entropy production will be one which proceeds at constant speed $\dot{s}$, assuming a slow process and a constant dissipation parameter ϵ. Alternately, if the dissipation parameter ϵ is not constant, the minimum entropy production process will be given by a constant arc length speed, in accordance with a steady state analogue of the "equal thermodynamic distance" principle.[25,34–36]

- If the path is not specified or can be varied, an absolute lower bound for

the entropy production is given by the geodesic in steady state parameter space, in accordance with the methods of §8.3.3.

Although they share a similar name, the minimum entropy production principle derived herein is quite different to that of Prigogine,[86] which concerns the selection of a steady state position relative to possible non-steady state positions, and which only applies to the Onsager linear regime. Similarly, it differs from the minimum entropy production principle obtained by the application of Riemannian geodesic calculations to the manifold of equilibrium positions, discussed at the end of §8.4.1.[16,19,34,43,100] The minimum principle derived herein is more general than both these principles, being applicable beyond the set of equilibrium positions, and also well outside the linear regime of non-equilibrium thermodynamics. In turn, it is based on the even broader generic formulation of the least action bound given herein, applicable to any system which can be analysed by Jaynes' method.

8.5. Conclusions

In this study, the manifold of stationary positions inferred by Jaynes' MaxEnt and MaxREnt principles – considered as a function of the moment constraints or their conjugate Lagrangian multipliers – is endowed with a Riemannian geometric description, based on the second differential tensor of the entropy or its Legendre transform (negative Massieu function) obtained from Jaynes' method. The analysis provides a generalised *least action bound* applicable to all Jaynesian systems, which provides a lower bound to the cost (in generic entropy units) of a transition between inferred positions along a specified path, at specified rates of change of the control parameters. The analysis therefore extends the concepts of "finite time thermodynamics", developed over the past three decades,[10–43] to the generic Jaynes domain, providing a link between purely static (stationary) inferred positions of a system, and dynamic transitions between these positions (as a function of time or some other coordinate). If the path is unspecified, the analysis gives an absolute lower bound for the cost of the transition, corresponding to the geodesic of the Riemannian hypersurface.

   The analysis is then applied to (i) an equilibrium thermodynamic system subject to mean internal energy and volume constraints, and (ii) a flow system at steady state, subject to constraints on the mean heat, mass and momentum fluxes and chemical reaction rates. The first example recovers the *minimum entropy cost* of a transition between equilibrium positions, a widely used result of finite-time thermodynamics. The second example

leads to a new *minimum entropy production principle*, for the cost of a transition between steady state positions of a flow system. The analyses reveal the tremendous utility of Jaynes' MaxEnt and MinXEnt methods augmented by the generalised least action bound, for the analysis of probabilistic systems of all kinds.

Acknowledgments

Appendix 1: Riemannian Geometric Considerations

It is necessary to examine several salient features of the Riemannian geometric interpretation adopted herein.[60,62] Consider a hypersurface represented by the position vector $\boldsymbol{x} = [x_1, ..., x_n]^\top$, embedded within the n-dimensional space defined by the coordinates $(x_1, ..., x_n)$. For analysis, this hypersurface can be converted to the parametric representation $\boldsymbol{x}(\boldsymbol{u}) = [x_1(\boldsymbol{u}), ..., x_n(\boldsymbol{u})]^\top$, where $\boldsymbol{u} = [u_1, ..., u_{n-1}]^\top$ is the $(n-1)$-dimensional vector of parameters u_j, consisting of coordinates on the hypersurface. The *first fundamental form* of this geometry is defined by the metric:[60,62]

$$d\varsigma^2 = d\boldsymbol{x} \cdot d\boldsymbol{x} = \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} a_{ij} du_i du_j = d\boldsymbol{u}^\top \mathbf{a}\, d\boldsymbol{u} \tag{A.1}$$

in which, by elementary calculus, the components of the tensor $\mathbf{a}$ can be shown to be:

$$a_{ij} = \frac{\partial \boldsymbol{x}}{\partial u_i} \cdot \frac{\partial \boldsymbol{x}}{\partial u_j} \tag{A.2}$$

Accordingly, $\mathbf{a}$ is symmetric. By Euclidean geometry, (A.1) can be used to calculate distances between two points a and b on the hypersurface $\boldsymbol{x}$, on the path defined by $\boldsymbol{u}$:

$$L_{\boldsymbol{x}} = \int_a^b d\varsigma = \int_a^b \sqrt{d\boldsymbol{u}^\top \mathbf{a}\, d\boldsymbol{u}} = \int_{\xi_a}^{\xi_b} \sqrt{d\dot{\boldsymbol{u}}^\top \mathbf{a}\, d\dot{\boldsymbol{u}}}\, d\xi \tag{A.3}$$

where the overdot indicates the derivative with respect to the path parameter ξ. The *second fundamental form* of the hypersurface is then defined by:[60,62]

$$-d\boldsymbol{x} \cdot d\boldsymbol{n} = \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} b_{ij} du_i du_j = d\boldsymbol{u}^\top \mathbf{b}\, d\boldsymbol{u} \qquad (A.4)$$

where $\boldsymbol{n}$ is the unit normal vector to the hypersurface. By differential calculus, it can be shown that:

$$b_{ij} = \frac{\partial \boldsymbol{x}}{\partial u_i \partial u_j} \cdot \boldsymbol{n} \qquad (A.5)$$

The second fundamental form is not considered as a metric with which to calculate distances, but is used to examine the tangency and curvature properties of the manifold $\boldsymbol{x}$.[60,62]

In the present study, we wish to adopt the Jaynesian matrix $\mathbf{g}$ or $\boldsymbol{\gamma}$ as a Riemannian metric tensor for the calculation of arc lengths on the R-dimensional stationary state hypersurface, embedded in the $(R+1)$-dimensional space defined by $(\mathfrak{H}^*, \{\langle f_r \rangle\})$ or $(\phi, \{\lambda_r\})$. We therefore adopt the (somewhat peculiar) approach in which the coordinates $[x_2, ..., x_{R+1}]^\top$ are selected as the surface parameters $[u_1, ..., u_R]^\top$; i.e. with the hypersurface $\boldsymbol{x}_{\mathfrak{H}^*} = [\mathfrak{H}^*, \langle f_1 \rangle, ..., \langle f_R \rangle]^\top$ parameterised by $\boldsymbol{u}_{\mathfrak{H}^*} = \boldsymbol{f}$ and with $\boldsymbol{x}_\phi = [\phi, \lambda_1, ..., \lambda_R]^\top$ parameterised by $\boldsymbol{u}_\phi = \boldsymbol{\Lambda}$. Two necessary conditions for the use of $\mathbf{g}$ or $\boldsymbol{\gamma}$ as metric tensors is that they be symmetric and positive definite (or semi-definite); since they constitute Hessian matrices of the concave generic entropy $\mathfrak{H}^*$ or convex potential function ϕ, these conditions are satisfied, not only in thermodynamic applications but within the generic Jaynes formulation (with semi-definite behaviour only at singularities).[6,15] However, $\mathbf{g}$ and $\boldsymbol{\gamma}$ are related to a second, rather than a first, fundamental form.[15,31] For $\mathbf{g}$ or $\boldsymbol{\gamma}$ to be considered as metric tensors, they must be able to generate the first fundamental form of some position vector which describes the hypersurface. In mathematical terms, from (A.1):

$$ds^2_{\mathfrak{H}^*} = d\boldsymbol{f}^\top \mathbf{g}\, d\boldsymbol{f} = d\boldsymbol{u}_{\mathfrak{H}^*}{}^\top \mathbf{a}_{\mathfrak{H}^*}\, d\boldsymbol{u}_{\mathfrak{H}^*}, \qquad (A.6)$$

$$ds^2_\phi = d\boldsymbol{\Lambda}^\top \boldsymbol{\gamma}\, d\boldsymbol{\Lambda} = d\boldsymbol{u}_\phi{}^\top \mathbf{a}_\phi\, d\boldsymbol{u}_\phi \qquad (A.7)$$

From (8.8), (8.10), (8.13) and (A.2), taking advantage of tensor symmetries,

312   *R.K. Niven and B. Andresen*

the metric components must therefore satisfy:

$$g_{mr} = a_{\mathfrak{H}^*, mr} = \frac{\partial \boldsymbol{\omega}}{\partial \langle f_m \rangle} \cdot \frac{\partial \boldsymbol{\omega}}{\partial \langle f_r \rangle} = \frac{\partial^2 \mathfrak{H}^*}{\partial \langle f_m \rangle \partial \langle f_r \rangle} = \frac{\partial \lambda_r}{\partial \langle f_m \rangle} \tag{A.8}$$

$$\gamma_{mr} = a_{\phi, mr} = \frac{\partial \boldsymbol{\Omega}}{\partial \lambda_m} \cdot \frac{\partial \boldsymbol{\Omega}}{\partial \lambda_r} = \frac{\partial^2 \phi}{\partial \lambda_m \partial \lambda_r} = \frac{\partial \langle f_r \rangle}{\partial \lambda_m} \tag{A.9}$$

where $\boldsymbol{\omega}(\boldsymbol{f})$ and $\boldsymbol{\Omega}(\boldsymbol{\Lambda})$ are new R-dimensional position vectors, which from (A.10), are related by:

$$\mathbf{a}_{\mathfrak{H}^*} \, \mathbf{a}_\phi = \mathbf{I}, \tag{A.10}$$

In consequence, the metrics (8.14)-(8.15) and (8.20)-(8.21) and arc lengths (8.16)-(8.19) used herein are not measures of distance on the stationary state hypersurface defined by $\mathfrak{H}^*(\{\langle f_r \rangle\})$ or $\phi(\{\lambda_r\})$, but rather, on the transformed hypersurface given by $\boldsymbol{\omega}$ or $\boldsymbol{\Omega}$. In addition to the symmetry and positive definiteness conditions, it is therefore also necessary and sufficient that the hypersurface defined by $\boldsymbol{\omega}$ or $\boldsymbol{\Omega}$ exists within $\mathbb{R}^R$, is continuous and continuously differentiable – at least up to first order – except in the neighbourhood of singularities.

References

1. E.T. Jaynes, Information theory and statistical mechanics, Phys. Rev., **106**, 620-630 (1957).
2. E.T. Jaynes, Information theory and statistical mechanics, *in* Ford, K.W. (ed), Brandeis University Summer Institute, Lectures in Theoretical Physics, Vol. 3: Statistical Physics, Benjamin-Cummings Publ. Co., 1963, 181-218.
3. M. Tribus, Information theory as the basis for thermostatics and thermodynamics, J. Appl. Mech., Trans. ASME, **28**, 1-8 (1961).
4. M. Tribus, Thermostatics and Thermodynamics, D. Van Nostrand Co. Inc., Princeton, NJ, 1961.
5. J.E. Shore, R.W. Johnson, Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy, IEEE Trans. Information Theory **IT-26**(1), 26-37 (1980).
6. J.N. Kapur, H.K. Kesevan, Entropy Optimization Principles with Applications, Academic Press, Inc., Boston, MA, 1992.
7. E.T. Jaynes (G.L. Bretthorst, ed.) Probability Theory: The Logic of Science, Cambridge U.P., Cambridge, 2003.
8. J.N. Kapur, H.K. Kesevan, *The Generalized Maximum Entropy Principle (with Applications)*, Sandford Educational Press, Waterloo, Canada, 1987.
9. J.N. Kapur, *Maximum-Entropy Models in Science and Engineering*, John Wiley, NY, 1989.
10. F. Weinhold, Metric geometry of equilibrium thermodynamics, J. Chem. Phys. **63**(6) 2479-2483 (1975).

11. F. Weinhold, Metric geometry of equilibrium thermodynamics. II. Scaling, homogeneity and generalized Gibbs-Duhem relations, J. Chem. Phys. **63**(6) 2484-2487 (1975).

12. F. Weinhold, Metric geometry of equilibrium thermodynamics. III. Elementary formal structure of a vector-algebraic representation of equilibrium thermodynamics, J. Chem. Phys. **63**(6) 2488-2495 (1975).

13. F. Weinhold, Metric geometry of equilibrium thermodynamics. IV. Vector-algebraic evaluation of thermodynamic derivatives, J. Chem. Phys. **63**(6) 2496-2501 (1975).

14. F. Weinhold, Metric geometry of equilibrium thermodynamics. V. Aspects of heterogeneous equilibrium, J. Chem. Phys. **65**(2) 559-564 (1976).

15. P. Salamon, B. Andresen, P.D. Gait, R.S. Berry, The significance of Weinhold's length, J. Chem. Phys. **73**(2) 1001-1002 (1980), erratum **73**(10) 5407 (1980).

16. P. Salamon, A. Nitzan, B. Andresen, R.S. Berry, Minimum entropy production and the optimization of heat engines, Phys. Rev. A **21** 2115-2129 (1980).

17. P. Salamon, R.S. Berry, Thermodynamic length and dissipated availability, Phys. Rev. Lett. **51**(13) 1127-1130 (1983).

18. P. Salamon, E. Ihrig, R.S. Berry, A group of coordinate transformations which preserve the metric of Weinhold, J. Math. Phys. **24**(10) 2515-2520 (1983).

19. B. Andresen, Finite-Time Thermodynamics, Physics Laboratory II, University of Copenhagen, Denmark, 1983.

20. P. Salamon, J. Nulton, E. Ihrig, On the relation between entropy and energy versions of thermodynamic length, J. Chem. Phys. **80**(1) 436-437 (1984).

21. B. Andresen, R.S. Berry, M.J. Ondrechen, P. Salamon, Thermodynamics for processes in finite time, Acc. Chem. Res. **17** 266-271 (1984).

22. P. Salamon, J.D. Nulton, R.S. Berry, Length in statistical thermodynamics, J. Chem. Phys. **82**(5) 2433-2436 (1985).

23. F. Schlögl, Thermodynamic metric and stochastic measures, Z. Phys. B **59** 449-454 (1985).

24. T. Feldman, B. Andresen, A. Qi, P. Salamon, Thermodynamic lengths and intrinsic time scales in molecular relaxation, J. Chem. Phys. **83**(11) 5849-5853 (1985).

25. J. Nulton, P. Salamon, B. Andresen, Q. Anmin, Quasistatic processes as step equilibrations, J. Chem. Phys. **83**(1) 334-338 (1985).

26. T. Feldman, R.D. Levine, P. Salamon, A geometrical measure for entropy changes, J. Stat. Phys. **42**(5/6) 1127-1134 (1986).

27. R.D. Levine, Geometry in classical statistical thermodynamics, J. Chem. Phys. **84**(2) 910-916 (1986).

28. J.D. Nulton, P. Salamon, Statistical mechanics of combinatorial optimization, Phys. Rev. A **37**(4) 1351-1356.

29. P. Salamon, J.D. Nulton, J.R. Harland, J. Pedersen, G. Ruppeiner, L. Liao, Simulated annealing with constant thermodynamic speed, Comput. Phys. Comm. **49** 423-428 (1988).

30. B. Andresen, R.S. Berry, R. Gilmore, E. Ihrig, P. Salamon, Thermodynamic geometry and the metrics of Weinhold and Gilmore, Phys. Rev. A **37**(3) 845-848 (1988).

31. B. Andresen, R.S. Berry, E. Ihrig, P. Salamon, Inducing Weinhold's metric from Euclidian and Riemannin metrics, Phys. Rev. A **37**(3) 849-851 (1988).

32. K.H. Hoffmann, B. Andresen, P. Salamon, Measures of dissipation, Phys. Rev. A **39**(7) 3618-3621 (1989).

33. B. Andresen, J.M. Gordon, Constant thermodynamic speed for minimizing entropy production in thermodynamic processes and simulated annealing, Phys. Rev. E **50**(6) 4346-4351 (1994).

34. P. Salamon, J.D. Nulton, The geometry of separation processes: A horse-carrot theorem for steady flow systems, Europhysics Letters **42**(5) 571-576 (1998).

35. P. Salamon, J.D. Nulton, G. Siragusa, T.R. Andersen, A. Limon, Principles of control thermodynamics, Energy **26** 307-319 (2001).

36. M. Schaller, K.H. Hoffmanm, G. Siragusa, P. Salamon, B. Andresen, Numerically optimized performance of diabatic distillation columns, Comput. Chem. Eng. **25** 1537-1548 (2001).

37. G.P. Beretta, A new approach to constrained-maximization nonequilibrium problems, in R.A. Gaggioli (ed.) Computer-Aided Engineering of Energy Systems: Second Law Analysis and Modeling, ASME Book H0341C-AES, **3**, 129-134 (1986).

38. G.P. Beretta, Dynamics of smooth constrained approach to maximum entropy, in M.J. Moran, E. Sciubba (eds), Second Law Analysis of Thermal Systems, ASME Book I00236, 17-24 (1987).

39. G.P. Beretta, Modeling non-equilibrium dynamics of a discrete probability distribution: General rate equation for maximal entropy generation in a maximum-entropy landscape with time-dependent constraints, Entropy **10** 160-182 (2008).

40. L. Diósi, K. Kulacsy, B. Lukács, A. Rácz, Thermodynamic length, time, speed, and optimum path to minimize entropy production, J. Chem. Phys. **105**(24) 11220-11225 (1996).

41. G.E. Crooks, Measuring thermodynamic length, Phys. Rev. Lett. **99** 100602 (2007).

42. E.H. Feng, G.E. Crooks, Far-from-equilibrium measurements of thermodynamic length, Phys. Rev. E **79** 012104 (2009).

43. D.C. Brody, D.W. Hook, Information geometry in vapour-liquid equilibrium, J. Phys. A: Math. Theor. **42** 023001 (33pp) (2009).

44. L. Boltzmann, Über die Beziehung zwischen dem zweiten Hauptsatze dewr mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung, respective den Sätzen über das Wärmegleichgewicht, Wien. Ber. **76**, 373-435 (1877); English transl.: J. Le Roux (2002) 1-63 *http://www.essi.fr/~leroux/.*

45. M. Planck, Über das gesetz der Energieverteilung im Normalspektrum, Annalen der Physik **4**, 553-563 (1901).

46. I. Vincze, On the maximum probability principle in statistical physics, Progress in Statistics **2**, 869-895 (1974).

47. M. Grendar, M. Grendar, What is the question that MaxEnt answers? A probabilistic interpretation, *in* A. Mohammad-Djafari (ed.) MaxEnt 2000, Gif-sur-Yvette, France, 8-13 July 2000, AIP Conf. Proc. **568**, 83-94 (2001).
48. R.K. Niven, Exact Maxwell-Boltzmann, Bose-Einstein and Fermi-Dirac statistics, Phys. Lett. A **342**(4) 286-293 (2005).
49. R.K. Niven, Cost of s-fold decisions in exact Maxwell-Boltzmann, Bose-Einstein and Fermi-Dirac statistics, Physica A **365**(1) 142-149 (2006).
50. R.K. Niven, M. Grendar, Generalized classical, quantum and intermediate statistics and the Polya urn model, Physics Letters A **373** 621-626 (2009).
51. R.K. Niven, Combinatorial entropies and statistics, European Physics Journal B, in press.
52. S. Kullback, R.A. Leibler, On information and sufficiency, Annals Math. Stat. **22**, 79-86 (1951).
53. S. Kullback, Information Theory and Statistics, John Wiley, NY, 1959.
54. R.K. Niven, Combinatorial information theory: I. Philosophical basis of cross-entropy and entropy, *http://arxiv.org/abs/cond-mat/0512017 v5*, 2007.
55. H.B. Callen, Thermodynamics, John Wiley & Sons, NY, 1960.
56. H.B. Callen, Thermodynamics and an Introduction to Thermostatistics, 2nd ed., John Wiley & Sons, NY, 1985.
57. R. Clausius, Über verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wrmetheorie, Poggendorfs Annalen **125**, 335-400 (1865); English transl.: R.B. Lindsay *in* J. Kestin (ed.) The Second Law of Thermodynamics, Dowden, Hutchinson & Ross, PA (1976) 162-193.
58. J.W. Gibbs, A method of graphical representation of the thermodynamic properties of substances by means of surfaces, Trans. Connecticut Acad. **2**, 382-404 (1873).
59. J.W. Gibbs, On the equilibrium of heterogeneous substances, Trans. Connecticut Acad. **3**, 108-248 (1875-1876); 343-524 (1877-1878).
60. E. Kreysig, Differential Geometry, Dover Publ., NY, 1991.
61. R. Trasarti-Battistoni, Euclidean and Riemannian geometrical approaches to non-extensive thermo-statistical mechanics, arXiv:cond-mat/0203536v2 (2002).
62. D. Zwillinger, CRC Standard Mathematical Tables and Formulae, Chapman & Hall / CRC Press, Boca Raton, FL, 2003.
63. R.K. Niven, Derivation of the maximum entropy production principle for flow-controlled systems at steady state, Physical Review E, in press *http://arxiv.org/abs/0902.1568* (2009).
64. C. Lanczos, The Variational Principles of Mechanics, 3rd ed., University of Toronto Press, Toronto, 1966.
65. A. Caticha, C. Cafaro, CP954, in Knuth, K.H., Caticha, A., Center, J.L., Giffon, A., Rodrguez, C.C. (eds), AIP Conference Proceedings 954, 165-174 (2007).
66. R. Weinstock, Calculus of Variations, with Applications to Physics and Engineering, Dover Publ., NY, 1974.

R.K. Niven and B. Andresen

67. Q.A. Wang, Maximum path information and the principle of least action for chaotic system, Chaos, Solitons and Fractals **23** 1253-1258 (2005).

68. Q.A. Wang, Maximum entropy change and least action principle for nonequilibrium systems, Astrophys. Space Sci. **305** 273-281 (2006).

69. G. Nathanson, O. Sinanoğlu, The geometry of near equilibrium irreversible thermodynamics, J. Chem. Phys. **72**(5) 3127-3129 (1980).

70. R. Gilmore, Le Châtelier reciprocal relations, J. Chem. Phys. **76**(11) 5551-5553 (1982).

71. S. Sieniutycz, R.S. Berry, Field thermodynamic potentials and geometric thermodynamics with heat transfer and fluid flow, Phys. Rev. A **43**(6) 2807-2818 (1991).

72. M. Chen, Symmetry transformations in extended irreversible thermodynamics, J. Math. Phys. **42**(6) 2531-2539 (2001).

73. M. Chen, On the intrinsic geometric structure of extended irreversible thermodynamics, J. Phys. A: Math. Gen. **36**4717-4727 (2003).

74. J.D. Flick, P. Salamon, B. Andresen, Metric bounds on losses in adaptive coding, Information Sci. **42** 239-253 (1987).

75. P. Salamon, J. Komlos, B. Andresen, J.D. Nulton, A geometric view of welfare gains with non-instantaneous adjustment, Math. Social Sci. **13** 153-163 (1987).

76. G. Ruppeiner, Thermodynamics: a Reimannian geometric model, Phys. Rev. A **20**(4) 1608-1613 (1979).

77. C. Truesdell, Rational Thermodynamics, McGraw-Hill, NY, 1969, chap. 7.

78. R.A. Fisher, On the mathematical foundations of theoretical statistics, Phil. Trans. Royal Soc. London A **222** 309-368 (1922).

79. C.R. Rao, Information and the accuracy attainable in the estimation of statistical parameters, Bull. Calcutta Math. Soc. **37** 81-91 (1945).

80. B.R. Frieden, Science from Fisher Information, An Introduction, 2nd ed., Cambridge U.P., 2004.

81. M. Planck, Treatise on Thermodynamics, Engl. transl., 3rd ed., Dover Publications, NY, 1945.

82. M. Planck, Introduction to Theoretical Physics, Vol. V: Theory of Heat, Engl. transl. H.L. Brose, Macmillan & Co., Ltd, 1932.

83. M. Massieu, Thermodynamique - Sur les fonctions caractéristiques des divers fluides, Comptes Rendus **69** 858-862; 1057-1061 (1869).

84. W. Spirkl, H. Ries, Optimal finite-time endoreversible processes, Phys. Rev. E. **52**(4) 3485-3489 (1995).

85. S.R. de Groot, P. Mazur, Non-Equilibrium Thermodynamics, Dover Publications, NY, 1984.

86. I. Prigogine, Introduction to Thermodynamics of Irreversible Processes, 3rd ed., Interscience Publ., NY.

87. H.J. Kreuzer, Nonequilibrium Thermodynamics and its Statistical Foundations, Clarendon Press, Oxford, 1981.

88. R.B. Bird, W.E. Stewart, E.N. Lightfoot, Transport Phenomena, 2nd ed., John Wiley & Sons, NY, 2002.

89. G.W. Paltridge, Global dynamics and climate - a system of minimum en-

tropy exchange, Quart. J. Royal Meteorol. Soc. **101**, 475-484 (1975).

90. G.W. Paltridge, The steady-state format of global climate, Quart. J. Royal Meteorol. Soc. **104**, 927-945 (1978).

91. G.W. Paltridge, Thermodynamic dissipation and the global climate system, Quart. J. Royal Meteorol. Soc. **107**, 531-547 (1981).

92. H. Ozawa, A. Ohmura, R.D. Lorenz, T. Pujol, The second law of thermodynamics and the global climate system: A review of the maximum entropy production principle, Rev. Geophys. **41**, article 4 (2003).

93. R.C. Dewar, Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states, J. Phys. A: Math. Gen. **36**, 631-641 (2003).

94. R.C. Dewar, Maximum entropy production and the fluctuation theorem, J. Phys. A: Math. Gen. **38**, L371-L381 (2005).

95. A. Kleidon, R.D. Lorenz (eds.) Non-equilibrium Thermodynamics and the Production of Entropy: Life, Earth and Beyond, Springer Verlag, Heidelberg, 2005.

96. L.M. Martyushev, V.D. Seleznev, Maximum entropy production principle in physics, chemistry and biology, Physics Reports **426**, 1-45 (2006).

97. S. Bruers, Classification and discussion of macroscopic entropy production principles, *arXiv:cond-mat/0604482v3*, 2007.

98. L. Onsager, Reciprocal relations in irreversible processes I, Phys. Rev. **37**, 405-426 (1931).

99. L. Onsager, Reciprocal relations in irreversible processes II, Phys. Rev. **38**, 2265-2279 (1931).

100. J.C. Schön, A thermodynamic distance criterion of optimality for the calculations of free energy changes from computer simulations, J. Chem. Phys. **105**(22) 10072 (1996).

318   *R.K. Niven and B. Andresen*