THERMODYNAMIC INTERPRETATION OF INFORMATION GEOMETRIC CURVATURE

RAMIL NIGMATULLIN, MIKHAIL PROKOPENKO

The aim of statistical thermodynamics is to explain the thermodynamic behaviour of large systems using microscopic models and probability theory. A statistical mechanical model for an equilibrium ensemble specifies a probability distribution $p(x|[\theta^i])$ for a microscopic state of the system, x, given microscopic parameters $[\theta^i]$, e.g. interaction strength between neighbouring atomic spins. The macroscopic observables are the expectations of collective variables, $X(x)$, for example, total magnetization of material. The traditional thermodynamic quantities such as entropy, S, heat Q, work W, internal energy U, free energy F can be shown to arise as expectations of certain collective functions [1].

A probability distribution $p(x|[\theta^i])$ possesses a natural differential geometric structure, since the distribution can be regarded as a point on a differential manifold with coordinates $(\theta^1, ..., \theta^n)$. The analysis of probability distribution from the point of view of differential geometry is known as information geometry [2]. Central to information geometry is the interpretation of Fisher information as the natural metric of the manifold and the concept of α-connection which specifies the curvature of the manifold. The rich analytical framework provided by information geometry has been applied in statistical inference, machine learning, signal processing, and optimization [3]. Naturally, information geometry is also useful for the analysis and design of statistical mechanical thermodynamic systems and processes. For example, Fisher metric provides a measure of distance between thermodynamic states [4] and provides a way of detecting proximity to thermodynamic critical points [5, 6, 7]. So far, the role of information geometric α-connection in thermodynamics remains largely unexplored. The aim of this work is to clarify the role of information geometry in thermodynamics and provide an thermodynamic interpretation of α-connection and related concepts, such as parallel transport, covariant derivatives and geodesics.

We start by recalling that the probability of a statistical mechanical system in thermal equilibrium to be in state x is given by the Gibbs measure

$$p(x|[\lambda^i]) = \frac{e^{-\beta H(\lambda^i, x)}}{Z(\lambda^i)} = \frac{e^{-\beta\lambda^i X_i(x)}}{Z(\lambda^i)}, \tag{1}$$

where x denotes the state of the system, $Z(\lambda^i) \equiv \sum_x e^{-\beta\lambda^i X_i(x)}$ is the partition function that normalizes the distribution, β is inverse temperature, $H = \lambda^i X_i$ is the Hamiltonian of the system, λ^i are the generalized forces conjugate to the collective variables X_i. The probability distribution given by equation (1) belongs to the well studied exponential family of distributions. In the notation commonly used in information geometry equation (1) is written as

$$p(x|[\theta^i]) = e^{\theta^i X_i - \psi(\theta_i)}, \tag{2}$$

where $\theta^i \equiv -\beta\lambda^i$ and $\psi(\theta_i) \equiv \log Z(\theta_i) = \log \sum_x e^{\theta^i X_i(x)}$ is the so-called Massieu potential, which in thermodynamics is also known as free entropy. From the point of view of probability theory, $\psi(\theta)$ is the moment generating function.

A natural metric on the manifold is the Fisher information given by

$$g_{ij} = \partial_i \partial_j \psi, \tag{3}$$

where $\partial_i \equiv \partial/\partial\theta^i$.

In information geometry, of central importance is one parameter family of affine connections called the α-connections. This is because there exists an important duality structure. Given a coordinate system $[\theta^i]$ which is α-flat, one can always find a dual coordinate system $[\eta_j]$ which is $(-\alpha)$-flat. Dual coordinate systems are connected via a Legendre transformation. For a distribution (1), the Legendre transformation is given by

$$\varphi(\theta) = -\psi(\theta) + \theta^i \eta_i(\theta), \tag{4}$$

where φ and ψ are dual potential of the Legendre transformation and θ^i and η_i are dual coordinate systems. From the point of view of statistical mechanics φ is the negative of the configuration entropy, i.e. $\varphi = -S$, and η_i are the expectation of the collective variables X_i

$$\eta_i = \partial_i \psi = \langle X_i \rangle \, . \tag{5}$$

In differential geometry a standard way of specifying a connection is through Christoffel symbol Γ^k_{ij}, or through $\Gamma_{ij,k} \equiv \Gamma^h_{ij} g_{hk}$, where g_{hk} is the Riemannian metric. $\{\Gamma^k_{ij}\}$ are the connection coefficients with respect to the coordinate system $[\theta^i]$. The α-connections are denoted by $\Gamma^{(\alpha)}_{ij,k}$. It can be shown [2] that

$$\Gamma^{(-1)}_{ij,k} = \partial_i \partial_j \partial_k \psi \tag{6}$$

$$\Gamma^{(1)ij,k} = \partial^i \partial^j \partial^k \varphi \tag{7}$$

$$\Gamma^{(1)}_{ij,k} = \Gamma^{(-1)ij,k} = 0 \tag{8}$$

Thus in the $[\theta^i]$ coordinate systems the Gibbs distribution is 1-flat and in the $[\eta_i]$ coordinate system it is -1-flat. Since both ψ and φ are moment generating functions, equation (6) and (7) imply that both -1-connection and 1-connection are the third moments of the collective variables occurring in the $p(x|[\theta^i])$ and $p(x|[\eta_i])$ distributions.

In order to begin interpreting the connection thermodynamically, we must obtain an equation linking the connection coefficients to quantities with thermodynamic meaning. Such equation is obtained by differentiating equation (4) twice with respect to the control parameter $[\theta^i]$ giving

$$\theta^i \partial_i \partial_j \partial_k \psi = -\left(\partial_i \partial_j S + \partial_i \partial_j \psi\right). \tag{9}$$

The connection coefficients appear on the left hand side of (9). The right hand side of (9) is the difference between curvatures (Hessians) of two functions - the configurational entropy, S, and the free entropy, ψ. The difference between these two curvatures is the quantity $\mathbb{S}^+$ that was defined in a recent work discussing thermodynamics of natural computation in swarms [8], but without an information geometric interpretation.

If we consider a quasi-static thermodynamic driving protocol in the space $[\theta^i]$, the system always remains in thermal equilibrium and it can be shown [8] that the changes in ψ correspond to work done, W. Hence, for a quasi-static protocol, we can interpret equation (9) as

$$\theta^i \partial_i \partial_j \partial_k \psi = \partial_i \partial_j W - \partial_i \partial_j S. \tag{10}$$

Thus the curvature of the information geometric manifold is related to the functional curvatures of work done and configuration change in quasi-static process. We discuss the notions of covariant derivatives, geodesics and parallel transport in thermodynamic terms using an example of trapped colloidal particle model system.

References

[1] U. Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics*, 75(12):126001, 2012.

[2] H. Nagaoka and S.-I. Amari. *Methods of Information Geometry*. AMS, OUP, 2007.

[3] S.-Ichi Amari. *Information Geometry and Its Applications (Applied Mathematical Sciences)*. Springer, 2016.

[4] G. E. Crooks. Measuring thermodynamic length. *Phys. Rev. Lett.*, 99:100602, 2007.

[5] W. Janke, D.A. Johnston, and R. Kenna. Information geometry and phase transitions. *Physica A: Statistical Mechanics and its Applications*, 336(1):181 – 186, 2004.

[6] M. Prokopenko, J. T. Lizier, O. Obst, and X. R. Wang. Relating Fisher information to order parameters. *Phys. Rev. E*, 84:041116, 2011.

[7] D. Brody and N. Rivier. Geometrical aspects of statistical mechanics. *Phys. Rev. E*, 51:1006–1011, 1995.

[8] E. Crosato, R. E. Spinney, R. Nigmatullin, J. T. Lizier, and M. Prokopenko. Thermodynamics and computation during collective motion near criticality. *arXiv:1708.04004*, 2017.