Riemannian geometry of fluctuation theory: 
An introduction

Luisberis Velazquez
Departamento de Física, Universidad Católica del Norte, Av. Angamos 0610,
Antofagasta, Chile
E-mail: lvelazquez@ucn.cl

Abstract. Fluctuation geometry was recently proposed as a counterpart approach of
Riemannian geometry of inference theory (information geometry), which describes the
geometric features of the statistical manifold $\mathcal{M}$ of random events that are described
by a family of continuous distributions $dp_{\zeta}(x|\theta)$. This theory states a connection
among geometry notions and statistical properties: separation distance as a measure
of relative probabilities, curvature as a measure about the existence of irreducible
statistical correlations, among others. In statistical mechanics, fluctuation geometry
arises as the mathematical apparatus of a Riemannian extension of Einstein fluctuation
theory, which is also closely related to Ruppeiner geometry of thermodynamics.
Moreover, the curvature tensor allows to express some asymptotic formulae that
account for the system fluctuating behavior beyond the gaussian approximation, while
curvature scalar appears as a second-order correction of Legendre transformation
between thermodynamic potentials.

1. Introduction
Riemannian geometries defined on statistical manifolds establish a direct correspondence
among statistical properties of a parametric family of continuous distributions:

$$dp_{\zeta}(x|\theta) = \rho_{\zeta}(x|\theta)dx \quad (1)$$

and geometrical notions of certain statistical manifolds $\mathcal{M}$ and $\mathcal{P}$ associated to them.
The advantage of these formalisms is that they enable a direct application of powerful
tools of Riemannian geometry for statistical analysis. There exist two possible
Riemannian geometries in the framework of continuous distribution (1). The first one
is Riemannian geometry of inference theory, which is widely known as information
geometry [1]. Distance notion of this geometry:

$$ds^2 = g_{\alpha\beta}(\theta)d\theta^\alpha d\theta^\beta \quad (2)$$
Figure 1. Continuous distributions $dp_x(x|\theta)$ and $dp_{\zeta}(\tilde{x}|\theta)$ are diffeomorphic distributions, that is, a same abstract distribution $dp(\epsilon|\mathcal{E})$ expressed into two different coordinate representations of the abstract statistical manifold $\mathcal{M}$.

establishes a statistical separation between two close distributions of parametric family (1), which characterizes distinguishing probability of these distribution during an statistical inference of control parameters ($\theta, \theta + d\theta$). The second one is Riemannian geometry of fluctuation theory, or more briefly, fluctuation geometry [2, 3, 4]. Its distance notion:

$$ ds^2 = g_{ij}(x|\theta) dx^i dx^j $$

establishes a statistical separation between two close values $(x, x + dx)$ of a random quantity $\xi$ for a given member of parametric family (1).

A great advantage of differential geometry is the possibility to perform a coordinate-free treatment. An important concept here is the notion of diffeomorphic distributions [4]. There are those distributions whose random quantities $\xi$ and $\zeta$ are related by a diffeomorphism $\phi: \xi \to \zeta$, that is, a bijective and differentiable map that leaves invariant their respective probability distributions (see scheme in Fig.1):

$$ \phi: dp_\xi(x|\theta) = dp_\zeta(\tilde{x}|\theta) \Rightarrow \rho_\zeta(\tilde{x}|\theta) = \rho_\xi(x|\theta) \left| \frac{\partial \tilde{x}}{\partial x} \right|^{-1}. $$

All these distributions are regarded as different representations of a same abstract distribution defined on the manifolds $\mathcal{M}$ and $\mathcal{P}$.

A simple example of transformation among random quantities is the one associated with Box-Muller transformation [5]:

$$ \zeta_1 = \sqrt{-2 \ln(\xi_1)} \cos(2\pi \xi_2) \quad \text{and} \quad \zeta_2 = \sqrt{-2 \ln(\xi_1)} \sin(2\pi \xi_2) $$

which is employed to generate Gaussian random numbers $\zeta_1$ and $\zeta_2$ from uniform random numbers $\xi_1$ and $\xi_2$. Continuous distributions whose associated manifolds $\mathcal{M}$ are
diffeomorphic to the real one-dimensional space $\mathbb{R}$ are always diffeomorphic distributions because of the only possible Riemannian geometry for these manifolds is the Euclidean one. In particular, Gaussian distribution:

$$dp_\xi(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-(x - \mu)^2/2\sigma^2\right] dx, -\infty < x < +\infty. \quad (6)$$

Cauchy distribution:

$$dp_\xi(\tilde{x}|\nu, \gamma) = \frac{1}{\pi \gamma} \frac{\gamma d\tilde{x}}{\tilde{x}^2 + (\tilde{x} - \nu)^2}, -\infty < \tilde{x} < +\infty. \quad (7)$$

Bimodal Gaussian distribution:

$$dp_\xi(\tilde{x}|\mu, \sigma) = \frac{1}{2\sqrt{2\pi\sigma}} \{ \exp\left[-(\tilde{x} - \mu)^2/2\sigma^2\right] + \exp\left[-(\tilde{x} + \mu)^2/2\sigma^2\right] \} d\tilde{x}, -\infty < \tilde{x} < +\infty. \quad (8)$$

are fully equivalent from this geometric perspective, namely, all they can be regarded as different representations of a same abstract distribution. Of course, not all distributions can be regarded as diffeomorphic distributions. For random quantities $\xi$ whose abstract statistical manifold $\mathcal{M}$ has a dimension $n \geq 2$ are possible the notions of curvature and statistical correlations. In particular, distributions family [4]:

$$dp_\xi(x,y|\theta) = \frac{1}{\mathcal{Z}(\theta)} \exp\left[-\frac{1}{2}(x^2 + y^2)\right] \frac{\theta dx dy}{2\pi \sqrt{x^2 + y^2 + \theta^2}} \quad (9)$$

with normalization constant:

$$\mathcal{Z}(\theta) = \sqrt{\pi e^{2\sigma^2}} \theta \frac{e^{\frac{\theta^2}{2}}}{\sqrt{2}} \text{erfc}\left(\frac{\theta}{\sqrt{2}}\right) \quad (10)$$

can be associated with curved geometry of surface of revolution represented in Fig.2. This family cannot be map to the product of two Gaussian distributions:

$$dp_\xi(x,y|\sigma) = \frac{1}{2\pi\sigma} \exp\left[-\frac{(x^2 + y^2)}{2\sigma^2}\right] dx dy \quad (11)$$

because of this last has Euclidean geometry of two-dimensional real space $\mathbb{R}^2$. Geometrical non-equivalence means that distributions (9) cannot be decomposed into the product of two independent distributions.

2. Fundamental equations and results of fluctuation geometry

For the sake of simplicity in notations, let us hereinafter omit the subindex of random quantity $\xi$ in all mathematical expressions. Riemannian structure of the statistical manifold $\mathcal{M}$ allows us to introduce the invariant volume element $d\mu(x|\theta)$:

$$d\mu(x|\theta) = \sqrt{|g_{ij}(x|\theta)|/2\pi} dx, \quad (12)$$
Figure 2. The geometry of the statistical manifold $\mathcal{M}$ associated with the distributions family (9) is fully equivalent to curved geometry defined on the revolution surface represented here.

which replaces the ordinary volume element $dx$ (Lebesgue measure) that is employed in equation (1). The notation $|T_{ij}|$ represents the determinant of a given tensor $T_{ij}$ of second-rank, while the factor $2\pi$ has been introduced for convenience. Additionally, one can define the probabilistic weight [3]:

$$\omega(x|\theta) = \rho(x|\theta) \sqrt{|2\pi g^{ij}(x|\theta)|},$$

(13)

which is a scalar function that arises as a local invariant measure of the probability. Although the mathematical form of the probabilistic weight $\omega(x|\theta)$ depends on the coordinates representations of the statistical manifolds $\mathcal{M}$ and $\mathcal{P}$; the values of this function are the same in all coordinate representations. Using the above notions, the family of continuous distributions (1) can be rewritten as follows:

$$dp(x|\theta) = \omega(x|\theta) d\mu(x|\theta),$$

(14)

which is a form that explicitly exhibits the invariance of this family of distributions.

The notion of probability weight $\omega(x|\theta)$ can be employed to redefine the notion of information entropy for continuous distributions [6]:

$$S_d[\omega|g, \mathcal{M}] = -\int_{\mathcal{M}} \omega(x|\theta) \log \omega(x|\theta) d\mu(x|\theta).$$

(15)

as a global invariant measure that depends on the metric tensor $g_{ij}(x|\theta)$ of the manifold $\mathcal{M}$. The quantity $\mathcal{I}(x|\theta)$:

$$\mathcal{I}(x|\theta) = -\log \omega(x|\theta)$$

(16)
represents a local invariant measurement of the information content, where differential entropy (15) exhibits the same value for all diffeomorphic distributions. Introducing the information potential $S(x|\theta)$ as the negative of the information content (16):

$$S(x|\theta) = \log \omega(x|\theta) \equiv -\mathcal{I}(x|\theta),$$  

(17)

the metric tensor can be rewritten as follows [3]:

$$g_{ij}(x|\theta) = -D_t D_j S(x|\theta) = \frac{\partial^2 S(x|\theta)}{\partial x^i \partial x^j} + \Gamma^k_{ij}(x|\theta) \frac{\partial S(x|\theta)}{\partial x^k},$$  

(18)

where $D_t$ is the covariant derivative associated with the Levi-Civita affine connections $\Gamma^k_{ij}(x|\theta)$ [10]:

$$\Gamma^k_{ij}(x|\theta) = g^{km}(x|\theta) \frac{1}{2} \left[ \frac{\partial g_{jm}(x|\theta)}{\partial x^i} + \frac{\partial g_{jm}(x|\theta)}{\partial x^i} - \frac{\partial g_{ij}(x|\theta)}{\partial x^m} \right].$$  

(19)

Covariant set of differential equations (18) can be rewritten into the alternative form:

$$g_{ij}(x|\theta) = \frac{\partial^2 \log \rho(x|\theta)}{\partial x^i \partial x^j} + \Gamma^k_{ij}(x|\theta) \frac{\partial \log \rho(x|\theta)}{\partial x^k} + \frac{\partial \Gamma^k_{ij}(x|\theta)}{\partial x^i} - \Gamma^k_{ij}(x|\theta) \Gamma^l_{kj}(x|\theta)$$  

(20)

in terms of probability density. According to expression (18), the metric tensor $g_{ij}(x|\theta)$ defines a positive definite distance notion (3), while the information potential $S(x|\theta)$ is locally concave everywhere. This last behavior guarantees the uniqueness of the point $\bar{x}$ where the information potential reaches a global maximum, that is, the uniqueness of the point of global maximum $\bar{x}$ of the probabilistic weight $\omega(x|\theta)$.

The main consequence derived from equation (18) is the possibility to rewrite the distributions family (14) into the following Riemannian gaussian representation [2, 3]:

$$dp(x|\theta) = \frac{1}{\mathcal{Z}(\theta)} \exp \left[ -\frac{1}{2} \ell_\theta(x, \bar{x}) \right] d\mu(x|\theta),$$  

(21)

where $\ell_\theta(x, \bar{x})$ denotes the separation distance between the arbitrary point $x$ and the point $\bar{x}$ with maximum information potential $S(x|\theta)$ (the arc-length $\Delta s$ of the geodesics that connects these points). Moreover, the negative of the logarithm of gaussian partition function $\mathcal{Z}(\theta)$ defines the so-called gaussian potential:

$$\mathcal{P}(\theta) = -\log \mathcal{Z}(\theta),$$  

(22)

which appears as the first integral of the problem (18):

$$\mathcal{P}(\theta) = S(x|\theta) + \frac{1}{2} \psi^2(x|\theta).$$  

(23)

Here, $\psi^2(x|\theta) = \psi^i(x|\theta) \psi_i(x|\theta) = g^{ij}(x|\theta) \psi_i(x|\theta) \psi_j(x|\theta)$ is the square norm of covariant vector field $\psi_i(x|\theta)$ defined by the gradient of the information potential $S(x|\theta)$:

$$\psi_i(x|\theta) = -D_i S(x|\theta) \equiv -\partial S(x|\theta) / \partial x^i.$$  

(24)
The factor $2\pi$ of definition (12) guarantees that the gaussian partition function $Z(\theta)$ drops the unity when the Riemannian structure of the manifold $M$ is the same of Euclidean real space $\mathbb{R}^n$.

Riemannian gaussian representation (46) rephrases the distributions family (1) in term of geometric notions of the manifold $M$. According to this result, the distance $\ell_\theta(x, \bar{x})$ is a measure of the occurrence probability of a deviation from the state $\bar{x}$ with maximum information potential. This result can be obtained combining equations (14) and (23) with the following the identity:

$$\psi^2(x|\theta) \equiv \ell_\theta^2(x, \bar{x}).$$

This last relation is a consequence of the geodesic character of the curves $x_g(s) \in M$ derived from the following set of ordinary differential equations [3]:

$$\frac{dx^i_g(s)}{ds} = \upsilon^i(x|\theta).$$

Here, $\upsilon^i(x|\theta) = g^{ij}(x|\theta)\upsilon_j(x|\theta)$ is the contravariant form of the unitary vector field $v_i(x|\theta)$ associated with the vector field (24):

$$v_i(x|\theta) = \psi_i(x|\theta) / \psi(x|\theta),$$

while the parameter $s$ is the arc-length of the curve $x_g(s)$. It is easy to check that this unitary vector field obeys the geodesic differential equation:

$$\upsilon^j(x|\theta)D_j v_i(x|\theta) = \upsilon^j(x|\theta) [g_{ij}(x|\theta) - v_i(x|\theta)v_j(x|\theta)] \equiv 0.$$  

Identity (25) follows from the directional derivatives:

$$\frac{dS(x_g(s)|\theta)}{ds} = \psi(x_g(s)|\theta) \text{ and } \frac{d^2S(x_g(s)|\theta)}{ds^2} \equiv -1,$$

which can be obtained from equation (26).

Let us now talk about the notion of curvature of fluctuation geometry. The affine connections $\Gamma^k_{ij} = \Gamma^k_{ij}(x|\theta)$ are employed to introduce of the curvature tensor $R^l_{ijk} = R^l_{ijk}(x|\theta)$ of the manifold $M$:

$$R^l_{ijk} = \frac{\partial}{\partial X^l} \Gamma^l_{jk} - \frac{\partial}{\partial X^l} \Gamma^l_{ik} + \Gamma^l_{im} \Gamma^m_{jk} - \Gamma^l_{jm} \Gamma^m_{ik}.$$  

Generally, the affine connections $\Gamma^k_{ij}(x|\theta)$ and the metric tensor $g_{ij}(x|\theta)$ are independent entities of Riemannian geometry. However, the knowledge of the metric tensor allows to introduce natural affine connections: the Levi-Civita connections (19). These affine connections are also referred to in the literature as the metric connections or the Christoffel symbols. The same ones follow from the consideration of a torsion-free covariant differentiation $D_i$ that obeys the condition of Levi-Civita parallelism [10]:

$$D_k g_{ij}(x|\theta) = 0.$$
Using the Levi-Civita connections, the curvature tensor can be expressed in terms of the metric tensor $g_{ij}(x|\theta)$ and its first and second partial derivatives. Additionally, one can introduce the Ricci curvature tensor $R_{ij}(x|\theta)$:

$$R_{ij}(x|\theta) = R^{k}_{kiij}(x|\theta)$$

(32)

as well as the curvature scalar $R(x|\theta)$:

$$R(x|\theta) = g^{ij}(x|\theta)R^{k}_{kiij}(x|\theta) = g^{ij}(x|\theta)g^{kl}(x|\theta)R_{kijl}(x|\theta).$$

(33)

According to Riemannian geometry [10], the curvature scalar $R(x|\theta)$ is the only invariant derived from the first and second partial derivatives of the metric tensor $g_{ij}(x|\theta)$.

The curvature tensor characterizes the deviation of local geometric properties of a manifold $\mathcal{M}$ from the properties of the Euclidean geometry (see scheme of Fig.3). For example, the volume of a small sphere about a point $x$ has smaller (larger) volume (area) than a sphere of the same radius defined on the $n$-dimensional real space $\mathbb{R}^n$ when the scalar curvature $R(x|\theta)$ is positive (negative) at that point. Quantitatively, this behavior is described by the following approximation formulae:

$$\frac{\text{Vol} \left[ S^{(n-1)}(x|\ell) \subset \mathcal{M} \right]}{\text{Vol} \left[ S^{(n-1)}(x|\ell) \subset \mathbb{R}^n \right]} = 1 - \frac{R(x|\theta)}{6(n+2)}\ell^2 + O(\ell^4),$$

(34)

$$\frac{\text{Area} \left[ S^{(n-1)}(x|\ell) \subset \mathcal{M} \right]}{\text{Area} \left[ S^{(n-1)}(x|\ell) \subset \mathbb{R}^n \right]} = 1 - \frac{R(x|\theta)}{6n}\ell^2 + O(\ell^4),$$

(35)

where the notation $S^{(m)}(x|\ell)$ represents a $m$-dimensional sphere with small radius $\ell$ centered at the point $x$. Accordingly, the local effects associated with the curvature of the manifold $\mathcal{M}$ appears as second-order (and higher) corrections of the Euclidean formulæ. The corresponding asymptotic formula for distribution (46) using spherical coordinates $(\ell, q)$ for radius $\ell$ sufficiently small:

$$dp(\ell, q|\theta) = \frac{1}{Z(\theta)} \left[ 1 - \frac{1}{24} \ell^2 F(q|\theta) + O(\ell^4) \right] dp_G(\ell, q|\theta).$$

(36)
Here, $dp_G(\ell, q|\theta)$ denotes the spherical coordinate representation of a gaussian distribution associated with the local Euclidean properties of the manifold $\mathcal{M}$ at the point $\bar{x}$ with maximum information potential:

$$
dp_G(\ell, q|\theta) = \exp \left( -\frac{1}{2} \ell^2 \right) \left( \sqrt{2} \pi \right)^{n-1} \frac{1}{\sqrt{\kappa(q)}} \, dq.
$$

(37)

where $\kappa(q) = \bar{g}_{ij} \xi_i(q) \xi_j(q)$. The \((n-1)\) vector fields $\xi_\alpha(q) = \{\xi^i_\alpha(q)\}$ are obtained from the unitary vector field $e(q) = \{e^i(q)\}$ associated with the spherical coordinates at the point $\bar{x}$ as follows:

$$
\xi_\alpha(q) = \frac{\partial e^i(q)}{\partial q^\alpha}.
$$

(38)

$\mathcal{F}(q|\theta)$ is a function on the spherical coordinates $q$ defined as follows:

$$
\mathcal{F}(q|\theta) = \bar{R}_{ijkl} \kappa^{\alpha\beta}(q) S^i_\alpha(q) S^j_\beta(q),
$$

(39)

which is referred to as the spherical function. Moreover, $\bar{R}_{ijkl} = R_{ijkl}(\bar{x}|\theta)$ is the curvature tensor evaluated at the point $\bar{x}$, while the quantities $S^i_\alpha(q)$ are defined as:

$$
S^i_\alpha(q) = e^i(q) \xi^j_\alpha(q) - e^j(q) \xi^i_\alpha(q).
$$

(40)

Curvature of statistical manifold $\mathcal{M}$ is directly related to the notion of irreducible statistical correlations. Specifically, it is said that a continuous distribution $dp(x|\theta)$ exhibits a reducible statistical dependence if it possesses a diffeomorphic distribution $dp(\hat{x}|\theta)$ that admits to be decomposed into independent distribution functions $dp^{(i)}(\hat{x}^i|\theta)$ for each coordinate as follows:

$$
dp(\hat{x}|\theta) = \prod_{i=1}^{n} dp^{(i)}(\hat{x}^i|\theta).
$$

(41)

Otherwise, the distribution function $dp(x|\theta)$ exhibits an irreducible statistical dependence. The existence (or nonexistence) of a reducible statistical dependence for a given distributions family (1) is fully equivalent to the existence (or nonexistence) of a Cartesian decomposition of its associated statistical manifold $\mathcal{M}$ into two (or more) independent statistical manifolds $\{\mathcal{A}_{\theta}^{(i)}\}$:

$$
\mathcal{M} = \mathcal{A}^{(1)} \otimes \mathcal{A}^{(2)} \ldots \otimes \mathcal{A}^{(l)}.
$$

(42)
A given manifold $\mathcal{A}$ is said to be an irreducible manifold when the same one does not admit the Cartesian decomposition (42). Moreover, a given Cartesian decomposition (42) is said to be an irreducible Cartesian decomposition if each independent manifold $\mathcal{A}^{(k)}$ is an irreducible manifold. In general, the question about the Cartesian decomposition of a Riemannian manifold into independent manifolds with arbitrary dimensions is better phrased and understood in the language of holonomy groups. The relation of holonomy of a connection with the curvature tensor is the main content of Ambrose-Singer theorem, while de Rham theorem states the conditions for a global Cartesian decomposition [10].

The flat character of the statistical manifold $\mathcal{M}$ implies the existence of a reducible statistical dependence for the family of distributions (1), while its curved character implies the existence of an irreducible statistical dependence.

3. Relevance in statistical mechanics

Redefining information potential (17) in units of Boltzmann constant $k$, the probability distribution (14) can be rewritten as follow:

$$ dp(x|\theta) = \exp[S(x|\theta)/k] \, d\mu(x|\theta). $$

Formally, this expression represents a sort of covariant extension of Einstein postulate of classical fluctuation theory [11], where the information potential $S(x|\theta)$ is identified with the thermodynamic entropy of closed system (up to the precision of an additive constant). Hereinafter, the coordinates $x = (x^1, x^2, \ldots, x^n)$ are the relevant macroscopic observables of the closed system, e.g.: the internal energy $U$, the volume $V$, the total angular momentum $M$, the magnetization $\mathcal{M}$, etc. Moreover, $\theta$ represents the set of control parameters of the given situation of thermodynamic equilibrium. The metric tensor $g_{ij}(x|\theta)$ of fluctuation geometry:

$$ g_{ij}(x|\theta) = -D_iD_jS(x|\theta) = -\partial_i\partial_jS(x|\theta) + \Gamma^k_{ij}(x|\theta)\partial_kS(x|\theta) $$

establishes a constraint between the entropy $S(x|\theta)$ and the metric tensor $g_{ij}(x|\theta)$ of the statistical manifold $\mathcal{M}$ of macroscopic observables $x$. Expression (44) provides a generalization for the thermodynamic metric tensor of Ruppeiner geometry [7, 8]:

$$ g_{ij}(\bar{x}) = -\frac{\partial^2S(\bar{x}|\theta)}{\partial x^i \partial x^j}, $$

while the Riemannian Gaussian representation:

$$ dp(x|\theta) = \frac{1}{Z(\theta)} \exp\left[ -\frac{1}{2k} F^2(x, \bar{x}) \right] \, d\mu(x|\theta), $$

is an exact improvement of Gaussian approximation of classical fluctuation theory [11]:

$$ dp(x|\theta) \simeq \exp \left[ -g_{ij}(\bar{x})(x - \bar{x})^i(x - \bar{x})^j / 2k \right] \sqrt{|g_{ij}(\bar{x})|/2\pi k} \, d^n x. $$

According to the asymptotic formula (36), curvature characterizes deviation exact distribution beyond Gaussian approximation for thermodynamical fluctuations.
Applicability of Gaussian approximation is breakdown during the occurrence of phase transitions and critical phenomena, and therefore, curvature can be a useful notion to study these situations. It is noteworthy that similar connections have been established for curvature notion of information geometry [12, 13], as well as Ruppeiner geometry [9]. Direct application of the asymptotic formula (36) for the case of Boltzmann-Gibbs distributions [11]:

$$\frac{dp(x|\theta)}{Z(\theta)} = \exp\left[-\theta_i x^i/k\right] \Omega(x) dx$$

(48)
can be employed to obtain the following result [4]:

$$P(\theta) \simeq \theta_i \bar{x}^i - s(\bar{x}|\theta) + \frac{k^2 R(\bar{x}|\theta)}{6}. \quad (49)$$

Here, $P(\theta)$ is the Planck thermodynamic potential [11]:

$$P(\theta) = -k \log Z(\theta), \quad (50)$$

while $s(x|\theta)$ is referred to as the entropy of the open system. This function is directly associated with the density of states $\Omega(x)$ via the metric tensor $g_{ij}(x|\theta)$:

$$\exp\left[s(x|\theta)/k\right] \sqrt{\frac{g_{ij}(x|\theta)}{2\pi k}} \equiv \Omega(x). \quad (51)$$

The entropy $s(x|\theta)$ is not an intrinsic property of the open system. Certainly, this entropy also depends on the metric tensor $g_{ij}(x|\theta)$, which accounts for the underlying environmental influence. Result (49) exhibits a very simple interpretation. Gaussian or zeroth-order approximation:

$$P(\theta) \simeq \bar{P}(\theta) = \theta_i \bar{x}^i - s(\bar{x}|\theta). \quad (52)$$

is just the known Legendre transformation that estimates the Planck thermodynamic potential $P(\theta)$ from the entropy of the open system $s(x|\theta)$. The curvature scalar $R(\bar{x}|\theta)$ introduces a correction of second-order of this transformation.

References
[1] Amari Sh 1990 Differential-Geometrical Methods in Statistics: Lecture notes in Statistics Vol. 28 (Berlin: Springer).
[2] Velazquez L 2011 J. Stat. Mech. P11007.
[3] Velazquez L 2012 J. Phys. A: Math and Theo 45 175002.
[4] Velazquez L 2013 J. Phys. A: Math and Theo 46 345003.
[5] Box G E P and Muller M E 1958 Annals Math. Stat. 29 610.
[6] Jaynes E T 1963 Information Theory and Statistical Mechanics, in Statistical Physics, Ford K ed (New York: Benjamin).
[7] Ruppeiner G 1979 Phys. Rev. A 20 1608.
[8] Ruppeiner G 1995 Rev. Mod. Phys. 67 605.
[9] Ruppeiner G 2010 Am. J. Phys. 78 1170.
[10] Berger A 2002 A panoramic view of Riemannian geometry (Berlin: Springer).
[11] Reichl L E 1980 A modern course in Statistical Mechanics, (Austin, TX: University of Texas Press).
[12] Janke W et al 2004 Physica A 336 181-6.
[13] Brody D C and Hook D W 2009 J. Phys. A: Math. Theor. 42 023001.