

The Information Geometry of Space-Time [†]

Ariel Caticha

Physics Department, University at Albany-SUNY, Albany, NY 12222, USA; ariel@albany.edu

[†] Presented at the 39th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Garching, Germany, 30 June–5 July 2019.

Published: 28 November 2019



Abstract: The method of maximum entropy is used to model curved physical space in terms of points defined with a finite resolution. Such a blurred space is automatically endowed with a metric given by information geometry. The corresponding space-time is such that the geometry of any embedded spacelike surface is given by its information geometry. The dynamics of blurred space, its *geometrodynamics*, is constructed by requiring that as space undergoes the deformations associated with evolution in local time, it sweeps a four-dimensional space-time. This reproduces Einstein's equations for vacuum gravity. We conclude with brief comments on some of the peculiar properties of blurred space: There is a minimum length and blurred points have a finite volume. There is a relativistic “blur dilation”. The volume of space is a measure of its entropy.

Keywords: information geometry; general relativity; geometrodynamics

1. Introduction

The problem of reconciling quantum theory (QT) and general relativity (GR) has most commonly been addressed by preserving the framework of QT essentially unchanged while modifying the structure and dynamics of space-time. This is not unreasonable. Einstein's equation, $G_{\mu\nu} = 8\pi G T_{\mu\nu}$, relates geometry on the left to matter on the right. Since our best theories for the matter right hand side are QTs it is natural to try to construct a theory in which the geometrical left hand side is also of quantum mechanical origin [1–3].

Further thought however shows that this move carries a considerable risk, particularly because the old process of quantization involves ad hoc rules which, however successful in the past, have led to conceptual difficulties that would immediately spread and also infect the gravitational field. One example is the old quantum measurement problem and its closely related cousin the problem of macroscopic superpositions. Do quantum superpositions of space-times even make sense? In what direction would the future be? Another example is the cosmological constant problem. Does the zero point energy of quantum fields gravitate? Why does it not give rise to unacceptably large space-time curvatures? Considerations such as these suggest that the issue of whether and how to quantize gravity hinges on a deeper understanding of the foundations of QT and also on a deeper understanding of GR and of geometry itself—what, after all, is *distance*? Why are QT and GR framed in such different languages? Recent developments indicate that they might be closer than previously thought—the link is entropy. Indeed, in the entropic dynamics approach [4–6] QT is derived as an application of entropic methods of inference [7] with a central role assigned to concepts of information geometry [8–16]. And, on the GR side, the link between gravity and entropy has been recognized from the early work of Bekenstein and Hawking and further reaffirmed in more recent thermodynamic approaches to GR [17–22].

In a previous paper [23] we used the method of maximum entropy to construct a model of physical space in which points are blurred; they are defined with a finite resolution. Such a blurred

space is a statistical manifold and therefore it is automatically endowed with a Riemannian metric given by information geometry. Our goal here is to further close the gap between QT and GR by formulating the corresponding Lorentzian geometry of space-time.

The extension from space to space-time is not just a simple matter of applying information geometry to four dimensions rather than three. The problem is that information geometry leads to metrics that are positive—statistical manifolds are inevitably Riemannian—which cannot reproduce the light-cone structure of space-time. Some additional ingredient is needed. We do not model space-time as a statistical manifold. Instead, space-time is modelled as a four-dimensional manifold such that the geometry of all space-like embedded surfaces is given by information geometry. We find that in the limit of a flat space-time our model coincides with a stochastic model of space-time proposed long ago by Ingraham by following a very different line of argument [24].

Blurred space is a curious hybrid: some features are typical of discrete spaces while other features are typical of continuous manifolds [25,26]. For example, there is a minimum length and blurred points have a finite volume. The volume of a region of space is a measure of the number of blurred points within it, and it is also a measure of its bulk entropy. Under Lorentz transformations the minimum length suffers a dilation which is more analogous to the relativistic time dilation than to the familiar length contraction.

The dynamics of blurred space, its *geometrodynamics*, is constructed by requiring that as three-dimensional space undergoes the deformations associated with time evolution it sweeps a four-dimensional space-time. As shown in a remarkable paper by Hojman, Kuchař, and Teitelboim [27] in the context of the familiar sharp space-time this requirement is sufficient to determine the dynamics. Exactly the same argument can be deployed here. The result is that in the absence of matter the geometrodynamics of four-dimensional blurred space-time is given by Einstein's equations. The coupling of gravity to matter will not be addressed in this work.

2. The Information Geometry of Blurred Space

To set the stage we recall the model of blurred space as a smooth three-dimensional manifold \mathbf{X} the points of which are defined with a finite resolution [23]. It is noteworthy that, unlike the very rough space-time foams expected in some models of quantum gravity, one expects blurred space to be very smooth because irregularities at scales smaller than the local uncertainty are suppressed. Blurriness is implemented as follows: when we say that a test particle is located at $x \in \mathbf{X}$ (with coordinates x^a , $a = 1, 2, 3$) it turns out that it is actually located at some unknown neighboring x' . The probability that x' lies within d^3x' is $p(x'|x)d^3x'$. Since to each point $x \in \mathbf{X}$ one associates a distribution $p(x'|x)$ the space \mathbf{X} is a statistical manifold automatically endowed with a metric. Indeed, when points are blurred one cannot fully distinguish the point at x described by the distribution $p(x'|x)$ from another point at $x + dx$ described by $p(x'|x + dx)$. The quantitative measure of distinguishability [7,11] is the *information distance*,

$$d\ell^2 = g_{ab}(x)dx^a dx^b, \quad (1)$$

where the metric tensor g_{ab} —the *information metric*—is given by,

$$g_{ab}(x) = \int dx' p(x'|x) \partial_a \log p(x'|x) \partial_b \log p(x'|x). \quad (2)$$

(We adopt the standard notation $\partial_a = \partial/\partial x^a$ and $dx' = d^3x'$.) Thus, in a blurred space *distance* is *distinguishability*.

In Section 4 we will briefly address the physical/geometrical interpretation of $d\ell$. For now we merely state [23] that $d\ell$ measures the distance between two neighboring points in units of the local uncertainty defined by the distribution $p(x'|x)$, that is, information length is measured in units of the local blur.

In order to completely define the information geometry of \mathbf{X} which will allow us to introduce notions of parallel transport, curvature, and so on, one must specify a connection or covariant derivative ∇ . The natural choice is the Levi-Civita connection, defined so that $\nabla_a g_{bc} = 0$. Indeed, as argued in [28], the Levi-Civita connection is to be preferred because, unlike the other α -connections [11], it does not require imposing any additional structure on the Hilbert space of functions $(p)^{1/2}$.

The next step is to use the method of maximum entropy to assign the blur distribution $p(x'|x)$. The challenge is to identify the constraints that capture the physically relevant information. One might be tempted to consider imposing constraints on the expected values of $\langle x'^a - x^a \rangle$ and $\langle (x'^a - x^a)(x'^b - x^b) \rangle$ but this does not work because in a curved space neither of these constraints is covariant. This technical difficulty is evaded by maximizing entropy on the flat space \mathbf{T}_P that is tangent to \mathbf{X} at P and then using the *exponential map* (see [23]) to “project” the distribution from the flat \mathbf{T}_P to the curved space \mathbf{X} . It is important to emphasize that the validity of this construction rests on the assumption that the normal neighborhood of every point x —the region about x where the exponential map is 1-1—is sufficiently large. The assumption is justified provided the scale of the blur is much smaller than the scale over which curvature effects are appreciable.

Consider a point $P \in \mathbf{X}$ with generic coordinates x^a and a positive definite tensor field $\gamma^{ab}(x)$. The components of $y \in \mathbf{T}_P$ are y^a . The distribution $\hat{p}(y|P)$ on \mathbf{T}_P is assigned on the basis of information about the expectation $\langle y^a \rangle_P$ and the variance-covariance matrix $\langle y^a y^b \rangle_P$,

$$\langle y^a \rangle_P = 0 \quad \text{and} \quad \langle y^a y^b \rangle_P = \gamma^{ab}(P). \quad (3)$$

On \mathbf{X} it is always possible to transform to new coordinates

$$x^i = X^i(x^a), \quad (4)$$

such that

$$\gamma^{ij}(P) = \delta^{ij} \quad \text{and} \quad \partial_k \gamma^{ij}(P) = 0, \quad (5)$$

where $i, j, \dots = 1, 2, 3$. If γ^{ab} were a metric tensor the new coordinates would be called Riemann Normal Coordinates at P (RNC $_P$). The new components of y are

$$y^i = X_a^i y^a \quad \text{where} \quad X_a^i = \frac{\partial x^i}{\partial x^a}, \quad (6)$$

and the constraints (3) take the simpler form,

$$\langle y^i \rangle_P = 0 \quad \text{and} \quad \langle y^i y^j \rangle_P = \delta^{ij}. \quad (7)$$

We can now maximize the entropy

$$S[\hat{p}, q] = - \int d^3 y \hat{p}(y|P) \log \frac{\hat{p}(y|P)}{\hat{q}(y)} \quad (8)$$

relative to the measure $\hat{q}(y)$ subject to (7) and normalization. Since \mathbf{T}_P is flat we can take $\hat{q}(y)$ to be constant and we may ignore it. The result in RNC $_P$ is

$$\hat{p}(y^i|P) = \frac{1}{(2\pi)^{3/2}} \exp \left[-\frac{1}{2} \delta_{ij} y^i y^j \right]. \quad (9)$$

Using the inverse of Equation (6) we can transform back to the original coordinates y^a ,

$$y^a = X_i^a y^i \quad \text{and} \quad \gamma_{ab} = X_a^i X_b^j \delta_{ij}. \quad (10)$$

The resulting distribution is also Gaussian,

$$\hat{p}(y^a|P) = \frac{(\det \gamma_{ab})^{1/2}}{(2\pi)^{3/2}} \exp \left[-\frac{1}{2} \gamma_{ab} y^a y^b \right], \quad (11)$$

and the matrix γ_{ab} of Lagrange multipliers turns out to be the inverse of the correlation matrix γ^{ab} , $\gamma_{ab} \gamma^{bc} = \delta_a^c$.

Next we use the exponential map to project y^i coordinates on the flat T_P to the RNC_P coordinates on the curved \mathbf{X} ,

$$x'^i = x^i(P) + y^i. \quad (12)$$

The corresponding distribution $p(x'^i|P)$ induced on \mathbf{X} by $\hat{p}(y^i|P)$ on T_P is

$$p(x'^i|P) d^3 x' = \hat{p}(y^i|P) d^3 y, \quad (13)$$

or

$$p(x'^i|x^i) = \frac{1}{(2\pi)^{3/2}} \exp \left[-\frac{1}{2} \delta_{ij} (x'^i - x^i)(x'^j - x^j) \right]. \quad (14)$$

Thus, in RNC_P the distribution $p(x'^i|x^i)$ retains the Gaussian form. We can now invert (4) and transform back to the original generic frame of coordinates x^a and define $p(x'^a|x^a)$ by

$$p(x'^a|x^a) d^3 x'^a = p(x'^i|x^i) d^3 x'^i, \quad (15)$$

which is an identity between scalars and holds in all coordinate systems. In the original x^a coordinates the distribution $p(x'^a|x^a)$ will not, in general, be Gaussian,

$$p(x'^a|x^a) = \frac{(\det \gamma_{ab})^{1/2}}{(2\pi)^{3/2}} \exp \left[-\frac{1}{2} \delta_{ij} \left(X^i(x'^a) - X^i(x^a) \right) \left(X^j(x'^a) - X^j(x^a) \right) \right]. \quad (16)$$

Finally we substitute (16) into (2) to calculate the information metric g_{ab} . (The integral is easily handled in RNC_P .) The result is deceptively simple,

$$g_{ab} = X_a^i X_b^j \delta_{ij} = \gamma_{ab}. \quad (17)$$

The main result of [23] was to show that the metric g_{ab} of a blurred space is a statistical concept that measures the “degree of distinguishability” between neighboring points. The metric is given by the Lagrange multipliers γ_{ab} associated to the covariance tensor γ^{ab} that describes the blurriness of space.

3. Space-Time and the Geometrodynamics of Pure Gravity

The constraint that determines the dynamics is the requirement that blurred space be a three-dimensional spacelike “surface” embedded in four-dimensional space-time. As shown in [27] the reason this condition is so constraining is that when evolving from an initial to a final surface every intermediate surface must also be embeddable in the same space-time and, furthermore, the sequence of intermediate surfaces—the path or foliation—is not unique. Such a “foliation invariance”, which amounts to the local relativity of simultaneity, is a requirement of consistency: if there are two alternative paths to evolve from an initial to a final state, then the two paths must lead to the same result.

Space-time is foliated by a sequence of space-like surfaces $\{\Sigma\}$. Points on the surface Σ are labeled by coordinates x^a ($a = 1, 2, 3$) and space-time events are labeled by space-time coordinates X^μ ($\mu = 0, 1, 2, 3$). The embedding of Σ within space-time is defined by four functions $X^\mu = X^\mu(x)$.

An infinitesimal deformation of Σ to a neighboring Σ' is specified by $X^\mu(x) \rightarrow X^\mu(x) + \delta X^\mu(x)$. The deformation vector $\delta X^\mu(x)$ is decomposed into normal and tangential components,

$$\delta X^\mu = \delta X^\perp n^\mu + \delta X^a X_a^\mu, \quad (18)$$

where n^μ is the unit normal to the surface and the three vectors $X_a^\mu = \partial X^\mu / \partial x^a$ are tangent to the coordinate lines x^a ($n_\mu n^\mu = -1$, $n_\mu X_a^\mu = 0$).

We assume a phase space endowed with a symplectic structure: the basic dynamical variables are the surface metric $g_{ab}(x)$ and its canonically conjugate momentum $\pi^{ab}(x)$. This leads to a Hamiltonian dynamics where the super-Hamiltonian $H_\perp(x)[g, \pi]$ and the super-momentum $H_a(x)[g, \pi]$ generate normal and tangential deformations respectively. In order for the dynamics to be consistent with the kinematics of deformations the Poisson brackets of H_\perp and H_a must obey two sets of conditions [29,30]. First, they must close in the same way as the “group” of deformations, that is, they must provide a representation of the “algebra” of deformations [31],

$$[H_\perp(x), H_\perp(x')] = \left(g^{ab}(x) H_b(x) + g^{ab}(x') H_b(x') \right) \partial_{ax} \delta(x, x'), \quad (19)$$

$$[H_a(x), H_\perp(x')] = H_\perp(x) \partial_{ax} \delta(x, x'), \quad (20)$$

$$[H_a(x), H_b(x')] = H_a(x') \partial_b \delta(x, x') + H_b(x) \partial_a \delta(x, x'). \quad (21)$$

And second, the initial values of the variables g_{ab} and π^{ab} must be restricted to obey the weak constraints

$$H_\perp(x) \approx 0 \quad \text{and} \quad H_a(x) \approx 0. \quad (22)$$

A remarkable feature of the resulting dynamics is that once the constraints (22) are imposed on one initial surface Σ they will be satisfied automatically on all subsequent surfaces. As shown in [27] the generators that satisfy (19)–(21) are

$$H_a = -2 \nabla_b \pi_a^b, \quad (23)$$

$$H_\perp = 2\kappa G_{abcd} \pi^{ab} \pi^{cd} - \frac{1}{2\kappa} g^{1/2} (R - 2\Lambda), \quad (24)$$

$$G_{abcd} = \frac{1}{2} g^{-1/2} (g_{ac} g_{bd} + g_{ad} g_{bc} - g_{ab} g_{cd}), \quad (25)$$

where κ and Λ are constants which, once the coupling to matter is introduced, can be related to Newton’s constant $G = c^4 \kappa / 8\pi$ and to the cosmological constant Λ . Equations (22)–(25) are known to be equivalent to Einstein’s equations in vacuum.

To summarize: (a) Space-time is constructed so that the geometry of any embedded spacelike surface is given by information geometry. (b) The geometrodynamics of blurred space is given by Einstein’s equations. These are the main conclusions of this paper.

4. Discussion

Dimensionless Distance?—As with any information geometry the distance $d\ell$ given in Equations (1) and (2) turns out to be dimensionless. The interpretation [23] is that an information distance is measured distances in units of the local uncertainty—the blur. To make this explicit we write the distribution (14) that describes a blurred point in RNC_P in the form

$$p(x'^i | x^i) = \frac{1}{(2\pi\ell_0^2)^{3/2}} \exp \left[-\frac{1}{2\ell_0^2} \delta_{ij} (x'^i - x^i)(x'^j - x^j) \right], \quad (26)$$

so that the information distance between two neighboring points is

$$d\ell^2 = \frac{1}{\ell_0^2} \delta_{ij} dx^i dx^j . \quad (27)$$

Since the blur ℓ_0 is the only unit of length available to us (there are no external rulers) it follows that $\ell_0 = 1$ but it is nevertheless useful to write our equations showing ℓ_0 explicitly. In (26) the two points x and x' are meant to be simultaneous.

Minimum Length—To explore the geometry of blurred space it helps to distinguish the abstract “mathematical” points that are sharply defined by the coordinates x from the more “physical” blurred points. We shall call them c-points and b-points respectively. In RNC_P the distance between two c-points located at x and at $x + \Delta x$ is given by (27). To find the corresponding distance $\Delta\lambda$ between two b-points located at x and at $x + \Delta x$ we recall that when we say a test particle is at x it is actually located at $x' = x + y$ so that

$$\Delta\lambda^2 = \frac{1}{\ell_0^2} \delta_{ij} (\Delta x^i + \Delta y^i) (\Delta x^j + \Delta y^j) . \quad (28)$$

Taking the expectation over y with the probability (26)—use $\langle y^i \rangle = 0$ and $\langle y^i y^j \rangle = \ell_0^2 \delta^{ij}$ —we find

$$\langle \Delta\lambda^2 \rangle = \frac{1}{\ell_0^2} \delta_{ij} \langle (\Delta x^i + \Delta y^i) (\Delta x^j + \Delta y^j) \rangle = \Delta\ell^2 + 6 . \quad (29)$$

We see that even as $\Delta x \rightarrow 0$ and the two b-points coincide we still expect a minimum *rms* distance of $\sqrt{6}\ell_0$.

Blur Dilation—The size of the blur of space is a length but it does not behave as the length of a rod. When referred to a moving frame it does not undergo a Lorentz contraction. It is more analogous to time dilation: just as a clock marks time by ticking along the time axis, so are lengths measured by ticking ℓ_0 s along them. By the principle of relativity all inertial observers measure the same blur in their own rest frames — the proper blur ℓ_0 . Relative to another inertial frame the blur is dilated to $\gamma\ell_0$ where γ is the usual relativistic factor. This implies the proper blur ℓ_0 is indeed the minimum attainable.

The Volume of a Blurred Point: Is Space Continuous or Discrete?—A b-point is smeared over the whole of space but we can still define a useful measure of its volume by adding all volume elements $g^{1/2}(x') d^3x'$ weighed by the scalar density $p(x'|x)/g^{1/2}(x')$. Therefore in ℓ_0 units a blurred point has unit volume. This means that we can measure the volume of a finite region of space by *counting* the number of b-points it contains. It also means that the number of distinguishable b-points within a region of finite volume is finite which is a property one would normally associate to discrete spaces. In this sense blurred space is both continuous and discrete. (See also [26].)

The Entropy of Space—The statistical state of blurred space is the joint distribution of all the y_x variables associated to every b-point x . We assume that the y_x variables at different x s are independent, and therefore their joint distribution is a product,

$$\hat{P}[y] = \prod_x \hat{p}(y_x|x) . \quad (30)$$

From (11) and (17) the distribution $\hat{p}(y_x^a|x)$ in the tangent space \mathbf{T}_x is Gaussian,

$$\hat{p}(y_x|x) = \frac{(\det g_x)^{1/2}}{(2\pi)^{3/2}} \exp \left[-\frac{1}{2} g_{ab}(x) y_x^a y_x^b \right] , \quad (31)$$

which shows explicitly how the information metric g_{ab} determines the statistical state of space.

Next we calculate the total entropy of space,

$$S[\hat{P}, \hat{Q}] = - \int Dy \hat{P}[y] \log \frac{\hat{P}[y]}{\hat{Q}[y]} \stackrel{\text{def}}{=} S[g] \quad (32)$$

relative to the uniform distribution

$$\hat{Q}[y|g] = \prod_x g^{1/2}(x), \quad (33)$$

which is independent of y — a constant. Since the y 's in Equation (30) are independent variables the entropy is additive, $S[g] = \sum_x S(x)$, and we only need to calculate the entropy $S(x)$ associated to a b-point at a generic location x ,

$$S(x) = - \int d^3y \hat{p}(y|x) \log \frac{\hat{p}(y|x)}{g^{1/2}(x)} = \frac{3}{2} \log 2\pi e = s_0. \quad (34)$$

Thus, the entropy per b-point is a numerical constant s_0 and the entropy of any region R of space, $S_R[g]$, is just its volume,

$$S_R[g] = \sum_{x \in R} S(x) = s_0 \int_R d^3x g^{1/2}(x). \quad (35)$$

Thus, the entropy of a region of space is proportional to the number of b-points within it and is proportional to its volume.

Canonical Quantization of Gravity?—The picture of space as a smooth blurred statistical manifold stands in sharp contrast to ideas inspired from various models of quantized gravity in which the short distance structure of space is dominated by extreme fluctuations. From our perspective it is not surprising that attempts to quantize gravity by imposing commutation relations on the metric tensor g_{ab} have not been successful. The information geometry approach suggests a reason why: quantizing the Lagrange multipliers $g_{ab} = \gamma_{ab}$ would be just as misguided as formulating a quantum theory of fluids by imposing commutation relations on those Lagrange multipliers like temperature, pressure, or chemical potential, that define the thermodynamic macrostate.

Physical Consequences of a Minimum Length?—A minimum length will eliminate the short wavelength divergences in QFT. This in turn will most likely illuminate our understanding of the cosmological constant and affect the scale dependence of running coupling constants. One also expects that QFT effects that are mediated by short wavelength excitations should be suppressed. For example, the lifetime of the proton ought to be longer than predicted by grand-unified theories formulated in Minkowski space-time. The nonlocality implicit in a minimum length might lead to possible violations of CPT symmetry with new insights into matter-antimatter asymmetry. Of particular interest would be the early universe cosmology where inflation might amplify minimum-length effects possibly making them observable.

Acknowledgments: I would like to thank N. Carrara, N. Caticha, S. Ipek, and P. Pessoa, for valuable discussions.

References and Notes

1. For an introduction to the extensive literature on canonical quantization of gravity, loop quantum gravity, string theory, and causal sets see e.g., [2,3].
2. Kiefer, C. *Quantum Gravity*; Oxford U.P.: Oxford, UK, 2007.
3. Ashtekar, A.; Berger, B.; Isenberg, J.; MacCallum, M. (Eds.) *General Relativity and Gravitation*; Cambridge U.P.: Cambridge, UK, 2015.
4. Caticha, A. The Entropic Dynamics approach to Quantum Mechanics. *Entropy* **2019**, *21*, 943; doi:10.3390/e21100943, *arXiv* **2019** arXiv:1908.04693.
5. Ipek, S.; Abedi, M.; Caticha, A. Entropic Dynamics: Reconstructing Quantum Field Theory in Curved Spacetime. *Class. Quantum Grav.* **2019**, *36*, 205013, *arXiv* **2018**, arXiv:1803.07493.
6. Ipek, S.; Caticha, A. An Entropic Dynamics approach to Geometroynamics. *arXiv* **2019**, arXiv:1910.01188.

7. Caticha, A. *Entropic Inference and the Foundations of Physics*; International Society for Bayesian Analysis-ISBrA: Sao Paulo, Brazil, 2012. Available online: <http://www.albany.edu/physics/ACaticha-EIFP-book.pdf> (accessed on 20 September 2019).
8. The subject of information geometry was introduced in statistics by Fisher [9] and Rao [10] with important later contributions by other authors [11–14]. Important aspects were also independently discovered in thermodynamics [15,16].
9. Fisher, R.A. Theory of statistical estimation. *Math. Proc. Camb. Philos. Soc.* **1925**, *22*, 700–725.
10. Rao, C.R. Information and the accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.* **1945**, *37*, 81.
11. Amari, S. *Differential-Geometrical Methods in Statistics*; Springer: Berlin, Germany, 1985.
12. Čencov, N.N. *Statistical Decision Rules and Optimal Inference*; American Mathematical Soc.: Providence, RI, USA, 1981; Volume 53.
13. Rodríguez, C.C. The metrics generated by the Kullback number. In *Maximum Entropy and Bayesian Methods*; Skilling, J., Ed.; Kluwer: Dordrecht, The Netherlands, 1989.
14. Ay, N.; Jost, J.; Văn Lê, H.; Schwachhöfer, L. *Information Geometry*; Springer: Berlin, Germany, 2017.
15. Weinhold, F. Metric geometry of equilibrium thermodynamics. *J. Chem. Phys.* **1975**, *63*, 2479.
16. Ruppeiner, G. Thermodynamics: A Riemannian geometric model. *Phys. Rev. A* **1979**, *20*, 1608.
17. Bekenstein, J.D. Black holes and entropy. *Phys. Rev. D* **1973**, *7*, 2333.
18. Hawking, S. Black Holes and Thermodynamics. *Phys. Rev. D* **1976**, *13*, 191.
19. Jacobson, T. Thermodynamics of space-time: the Einstein equation of state. *Phys. Rev. Lett.* **1995**, *75*, 1260.
20. Padmanabhan, T. Thermodynamical aspects of gravity: New insights. *Rep. Prog. Phys.* **2010**, *73*, 046901.
21. Verlinde, E.P. On the origin of gravity and the laws of Newton. *J. High Energy Phys.* **2011**, *2011*, 29.
22. Jacobson, T. Entanglement equilibrium and the Einstein equation. *Phys. Rev. Lett.* **2016**, *116*, 201101.
23. Caticha, A. Geometry from Information Geometry. In *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*; Giffin, A., Knuth, K., Eds.; AIP American Institute of Physics: College Park, MD, USA, 2016; Volume 1757, p. 030001, *arXiv* **2015**, arXiv:1512.09076.
24. Ingraham, R.L. Stochastic Space-time. *Nuovo Cimento* **1964**, *34*, 182.
25. It is possible that there is some connection with ideas proposed by Kempf [26] expressed in the language of spectral geometry. This is a topic for future research.
26. Kempf, A. Information-theoretic natural ultraviolet cutoff for spacetime. *Phys. Rev. Lett.* **2009**, *103*, 231301..
27. Hojman, S.A.; Kuchař, K.; Teitelboim, C. Geometrodynamics Regained. *Ann. Phys.* **1976**, *96*, 88.
28. Brodie, D.J.; Hughston, L.P. Statistical Geometry in Quantum Mechanics. *Proc. R. Soc. Lond. Ser. A* **1998**, *454*, 2445–2475.
29. Teitelboim, C. How Commutators of Constraints Reflect the Spacetime Structure. *Ann. Phys.* **1973**, *79*, 542.
30. Kuchař, K. Canonical Quantization of Gravity. In *Relativity, Astrophysics, and Cosmology*; Israel, W., Ed.; Reidel: Dordrecht, The Netherlands, 1973; pp. 237–288.
31. The quotes in “group” and “algebra” are a reminder that the set of deformations do not form a group. The composition of two successive deformations is itself a deformation but it depends on the surface to which the first deformation is applied.

