General Relativity and Gravitation, Vol. 35, No. 12, December 2003 (© 2003)

ESSAY

Why Gravity Has No Choice: Bulk Spacetime Dynamics Is Dictated by Information Entanglement across Horizons[1]

T. Padmanabhan[2]

Received May 21, 2003

The principle of equivalence implies that gravity affects the light cone (causal) structure of the space-time. It follows that there will exist observers (in any space-time) who do not have access to regions of space-time bounded by horizons. Since physical theories in a given coordinate system must be formulated entirely in terms of variables which an observer using that coordinate system can access, gravitational action functional *must* contain a *foliation dependent* surface term which encodes the information inaccessible to the particular observer. I show that: (i) It is possible to determine the nature of this surface term from general symmetry considerations and *prove* that the entropy of any horizon is proportional to its area. (ii) The gravitational action can be determined using a differential geometric identity related to this surface term. The dynamics of spacetime is dictated by the nature of quantum entanglements across the horizons and the flow of information, making *gravity inherently quantum mechanical at all scales*. (iii) In static space-times, the action for gravity can be given a purely thermodynamic interpretation and the Einstein equations have a formal similarity to laws of thermodynamics. (iv) The horizon area must be quantized with $A_{\rm horizon} = (8\pi G\hbar/c^3)m$ with $m = 1, 2, \cdots$ in the semi-classical limit.

KEY WORDS: Gravitational action; horizon.

[1] Fifth Award in the 2003 Essay Competition of the Gravity Research Foundation. - *Ed.*

[2] IUCAA, Post Bag 4, Ganeshkhind, Pune - 411 007; e-mail: nabhan@iucaa.ernet.in

0001-7701/03/1200-2097/0 © 2003 Plenum Publishing Corporation

The principle of equivalence makes it is possible to define a local inertial frame around any event $\mathcal{P}$ in which the laws of special relativity are valid. This allows one to determine the interaction of gravity with other fields by expressing the laws in a generally covariant manner in the local inertial frame and then extending them to curved space-time. An immediate consequence is the effect of gravity on light rays which determine the causal structure of spacetime and restrict the flow of information. A region of spacetime, described in some coordinate system with a non-trivial metric tensor $g_{ab}(x^k)$, could have a light cone structure such that information about one sub-region is not accessible to observers in another region. It should be stressed that such a limitation is *always* observer/coordinate dependent. To appreciate this fact, let us begin by noting that the freedom of choice of the coordinates allows 4 out of 10 components of the metric tensor to be pre-specified, which we shall take to be $g_{00} = N^2$, $g_{0\alpha} = N_\alpha$. (We use the signature $(+,-,-,-)$ and units with $G = \hbar = c = 1$; the Latin indices vary over 0–3, while the Greek indices cover 1–3.) These four variables allow us to characterize the observer-dependent information. For example, with the choice $N = 1$, $N_\alpha = 0$, $g_{\alpha\beta} = -\delta_{\alpha\beta}$, the $\mathbf{x} =$ constant trajectories correspond to a class of inertial observers in flat spacetime while with $N = (ax)^2$, $N_\alpha = 0$, $g_{\alpha\beta} = -\delta_{\alpha\beta}$ the $\mathbf{x} =$ constant trajectories represent a class of accelerated observers with a horizon at $x = 0$. *We only need to change the form of N to make this transition in which a class of time-like trajectories, $\mathbf{x} =$ constant, acquire a horizon.* Similarly observers plunging into a black hole will find it natural to describe the Schwarzschild metric in the synchronous gauge with $N = 1$, $N_\alpha = 0$ (see e.g., ref. [1]) in which they can indeed access the information contained inside the horizon. The less masochistic observers will use a more standard foliation which has $N^2 = (1 - 2M/r)$ and the surface $N = 0$ will act as the horizon which restricts the flow of information from $r < 2M$ to the observers at $r > 2M$.

This aspect, viz. that different observers [defined as different families of time-like curves] may have access to different regions of space-time and hence differing amount of information, introduces a very *new* feature into physics. It is now necessary to ensure that *physical theories in a given coordinate system are formulated entirely in terms of the variables that an observer using that coordinate system can access*. This "principle of effective theory" is analogous to the renormalization group arguments used in high energy physics which "protects" the low energy theories from the unknown complications of the high energy sector. For example, one can use QED to predict results at, say, 10 GeV without worrying about the structure of the theory at 10^{19} GeV, as long as one uses coupling constants and variables defined around 10 GeV and determined observationally. In this case, one invokes the effective field theory approach in the momentum space. We can introduce the same reasoning in coordinate space and demand—for example—that the observed physics outside a black hole horizon must not depend on the unobservable processes beyond the horizon [4].

In fact, this is a natural extension of a more conventional procedure used in flat spacetime physics. Let us recall that, in standard description of flat spacetime physics, one often divides the spacetime by a space-like surface $t = t_1 =$ constant. Given the necessary information on this surface, one can predict the evolution for $t > t_1$ *without knowing the details at* $t < t_1$. In the case of curved spacetime with horizon, similar considerations apply. For example, if the spacetime contains a Schwarzschild black hole, say, then the light cone structure guarantees that the processes inside the black hole horizon cannot affect the outside events *classically*. What makes our demand non trivial is the fact that the situation in *quantum theory* is quite different. Quantum fluctuations of fields [especially gravity, treated as spin-2 modes propagating in the classical metric] will have nontrivial correlations across the horizon and will lead to entanglement of modes across the horizon. Our principle of effective theory states that it must be possible to "protect" the physical processes outside the horizon from such effects influencing it across the horizon. Since the horizon surface is the only common element to inside and outside regions, the effect of these entanglements across a horizon can only appear as a surface term in the action. Hence it is an inevitable consequence of principle of equivalence that the action functional describing gravity must contain certain boundary terms which are capable of encoding the information equivalent to that present beyond the horizon. This relic of quantum entanglements will survive in the classical limit but—being a surface term—will not affect the equations of motion.

In order to provide a local, Lagrangian, description of physics this boundary term must be expressible as an integral of a four-divergence, allowing us to write the action functional for gravity formally as

$$A_{\rm grav} = \int d^4x\sqrt{-g}L_{\rm grav} = \int d^4x\sqrt{-g}(L_{\rm bulk} + \nabla_i V^i) = A_{\rm bulk} + A_{\rm surface} \tag{1}$$

where $\nabla_i V^i \equiv (-g)^{-1/2}\partial_i[(-g)^{1/2}V^i]$ irrespective of whether V^i is a genuine four vector or not. In fact, since different observers will have different levels of access to information, we do expect $A_{\rm surface}$ to depend on the foliation of spacetime. On the other hand, since the overall dynamics should be the same for all observers, $A_{\rm grav}$ should be a scalar. It follows that neither $A_{\rm bulk}$ nor $A_{\rm surface}$ are covariant but their sum should be a covariant scalar. As we shall see, the fact that such a relic of quantum microstructure, $A_{\rm surface}$, must exist, encoding the entanglements across the horizon, is powerful enough to determine the the form of action functional $A_{\rm grav}$ and the bulk dynamics of spacetime in classical limit! (In fact, we will see that the concept of classical limit of quantum gravity is very nontrivial and *cannot* be obtained by a naive $\hbar \to 0$ rule). The dynamics of spacetime is dictated by the nature of quantum entanglements across the horizons and the microscopic flow

of information, making *gravity inherently quantum mechanical at all scales* in a precise manner.

Let us now determine the form of $A_{\rm surface}$. The horizon for a class of observers arises in a specific gauge and resultant $A_{\rm surface}$ will in general depend on the gauge variables $N,\ N_\alpha$. Among these, the lapse function N plays a more important role than N_α. To see this explicitly, let us start with a spacetime described in the synchronous gauge (see [1]; section 97) in which $N = 1$, $N_\alpha = 0$. Consider now the *infinitesimal* transformations $t \to t + \phi(t, x^\alpha)$; $x^\alpha \to x^\alpha + \xi^\alpha(t, x^\alpha)$ with the condition $g_{\alpha\beta}\dot{\xi}^\beta = -(\partial\phi/\partial x^\alpha)$. Such transformations maintain $N_\alpha = 0$ but change N from $N = 1$ to $N \to (1 + \dot{\phi})$, [as well as the form of $g_{\alpha\beta}$]; this, in turn, should change the value of $A_{\rm surface}$. In what follows, we shall set $N_\alpha = 0$ without loss of generality and our results are independent of this assumption. We next introduce a $(1 + 3)$ foliation with the standard notation for the metric components $(g_{00} = N^2,\ g_{0\alpha} = N_\alpha)$. Let $u^i = (N^{-1}, 0, 0, 0)$ be the four-velocity of observers corresponding to this foliation, i.e. the normal to the foliation, and let $a^i = u^j\nabla_j u^i$ be the related acceleration. Let $K_{ab} = -\nabla_a u_b + u_a a_b$ be the extrinsic curvature of the foliation, with $K \equiv K^i_i = -\nabla_i u^i$.

Given this structure, we can list all possible vector fields V^i which can be used in (1). This vector has to be built out of u^i, g_{ab} and the covariant derivative operator ∇_j acting only once. The last restriction arises because the equations of motion should be of no order higher than two. Given these conditions, (i) there is only one vector field — viz., the u^i itself — which has no derivatives and (ii) only three vectors $(u^j\nabla_j u^i, u^j\nabla^i u_j, u^i\nabla^j u_j)$ which are linear in covariant derivative operator. The first one is the acceleration $a^i = u^j\nabla_j u^i$; the second identically vanishes since u^j has unit norm; the third can be written as $-u^i K$. Thus V^i in the surface term must be a linear combination of u^i, $u^i K$ and a^i. The corresponding term in the action *must* have the form

$$A_{\rm surface} = \int d^4x\,\sqrt{-g}\nabla_i V^i = \int d^4x\,\sqrt{-g}\nabla_i[\lambda_0 u^i + \lambda_1 K u^i + \lambda_2 a^i] \qquad (2)$$

where λ's are numerical constants to be determined.

Let the region of integration be a four volume $\mathcal{V}$ bounded by two space-like surfaces Σ_1 and Σ_2 and two time-like surfaces $\mathcal{S}$ and $\mathcal{S}_1$. The space-like surfaces are constant time slices with normals u^i, and the time-like surfaces have normals n^i and we shall choose $n_i u^i = 0$. The induced metric on the space-like surface Σ is $h_{ab} = g_{ab} - u_a u_b$, while the induced metric on the time-like surface $\mathcal{S}$ is $\gamma_{ab} = g_{ab} + n_a n_b$. These two surfaces intersect on a two-dimensional surface $\mathcal{Q}$, with the induced metric $\sigma_{ab} = h_{ab} + n_a n_b = g_{ab} - u_a u_b + n_a n_b$. In this foliation, the first two terms of (2) contribute only on the $t =$ constant hyper-surfaces (Σ_1 and Σ_2) while the third term is the one which contributes on a horizon (which we shall treat as the null limit of a time-like surface $\mathcal{S}$, like the limit $r \to 2M+$ in the

black hole spacetime). Hence we get, on the horizon,

$$A_{\rm surface} = \lambda_2 \int d^4x\, \sqrt{-g}\, \nabla_i a^i = \lambda_2 \int dt \int_{\mathcal{S}} d^2x\, N\sqrt{|\sigma|}(n_\alpha a^\alpha) \tag{3}$$

Further, in any static spacetime with a horizon: (i) The integration over t becomes multiplication by $\beta \equiv 2\pi/\kappa$ where κ is the surface gravity of the horizon, since there is a natural periodicity in the Euclidean sector. (ii) As the surface $\mathcal{S}$ approaches the horizon, the quantity $N(a_i n^i)$ tends to κ which is constant over the horizon [2]. Using $\beta\kappa = 2\pi$, the surface term gives, on the horizon, the contribution

$$A_{\rm surface} = \lambda_2 \kappa \int_0^\beta dt \int d^2x\, \sqrt{\sigma} = 2\pi\lambda_2 \mathcal{A}_H \tag{4}$$

where $\mathcal{A}_H$ is the area of the horizon. We thus arrive at the conclusion that *the information blocked by a horizon, and encoded in the surface term, must be proportional to the area of the horizon*. Taking into consideration the non compact horizons, like the Rindler horizon, we may state that the entropy [or the information content] per unit area of the horizon is a constant related to λ_2. Writing $\lambda_2 \equiv -(1/8\pi\mathcal{A}_P)$, where $\mathcal{A}_P$ is a fundamental constant with the dimensions of area, the entropy associated with the horizon will be $S_H = (1/4)(\mathcal{A}_H/\mathcal{A}_P)$.

Having determined the form of $A_{\rm surface}$ we now turn to the nature of $A_{\rm grav}$ and $A_{\rm bulk}$. We need to express the Lagrangian $\nabla_i V^i$ as a difference between two Lagrangians L_{grav} and L_{bulk} such that: (a) L_{grav} is a generally covariant scalar. (b) L_{bulk} is utmost quadratic in the time derivatives of the metric tensor. (c) Neither $L_{\rm grav}$ nor $L_{\rm bulk}$ should contain four divergences since such terms are already taken into account in $L_{\rm surface}$. This is in fact just an exercise in differential geometry. To do this formally, we shall first write the sum $[\lambda_1 K u^i + \lambda_2 a^i + \lambda_0 u^i]$ as $[-(Ku^i + a^i)/8\pi\mathcal{A}_P + \lambda_3 K u^i + \lambda_0 u^i]$ where $\lambda_3 = \lambda_1 + (8\pi\mathcal{A}_P)^{-1}$ is another constant. We next note that there is a differential geometric identity (see e.g., [3])

$$2\nabla_i(Ku^i + a^i) = R - \left[{}^3\mathcal{R} - K_{ab}K^{ab} + K^a_a K^b_b\right] \tag{5}$$

where R and ${}^3\mathcal{R}$ are the scalar curvatures of the spacetime and the $t =$ constant surfaces respectively. We thus find that

$$\begin{aligned} L_{\rm surface} &= -\frac{1}{8\pi\mathcal{A}_P}\nabla_i[Ku^i + a^i] + \nabla_i(\lambda_3 K u^i + \lambda_0 u^i) \\ &= -\frac{R}{16\pi\mathcal{A}_P} - \frac{1}{16\pi\mathcal{A}_P}\left[{}^3\mathcal{R} - K_{ab}K^{ab} + K^a_a K^b_b\right] + \nabla_i(\lambda_3 K u^i + \lambda_0 u^i) \end{aligned}$$

It follows that the Lagrangian $\nabla_i V^i$ can be expressed as a difference between two Lagrangians $L_{grav} \equiv (R/16\pi\mathcal{A}_P) = L_{\rm EH}$ and

$$L_{bulk} \equiv \frac{1}{16\pi\mathcal{A}_P}\left[{}^3\mathcal{R} - K_{ab}K^{ab} + K^a_a K^b_b\right] = L_{\rm ADM} \tag{6}$$

with the necessary properties (a), (b), (c) listed above, if and only if $\lambda_3 = \lambda_0 = 0$. The gravitational action $A_{\rm grav}$ is just the Einstein-Hilbert action while $A_{\rm bulk}$ is the standard ADM action. No other possibilities exist (except for a trivial addition of a cosmological constant). We see that the structure of gravitational action can be determined *uniquely* using the form of $A_{\rm surface}$ which—on using the known form of λ's—turns out to be the integral of $\nabla_i(Ku^i + a^i)$.

When unobserved degrees of freedom inside the horizon are integrated out, the resulting effective theory will have the surviving term $A_{\rm surface}$ of (3). This will contribute a phase factor $\exp(iA_{\rm surface})$ to the path integral amplitude outside the horizon. Though such a term is innocuous classically, it changes the quantum amplitude for processes. The principle of effective theory demands that this should *not* happen which, in turn, requires the quantization condition $A_{\rm surface} = 2\pi j$. From (4) we find that the areas of all horizons should be quantized in terms of a fundamental area element in the WKB limit [4] leading to

$$\mathcal{A}_H = (8\pi\mathcal{A}_P)\, j; \qquad j = 1, 2, 3, \ldots. \tag{7}$$

The boundary term—which is not generally covariant—may be different for different observers, but the corresponding operators will not commute thereby eliminating any possible contradiction. (This is analogous to the fact that, in quantum mechanics, the component of angular momentum J_z measured along any axis is quantized irrespective of the orientation of the axis.). In fact, detailed analysis shows that the horizon area is analogous to J_z (and quantized in integer units, j) while the area operator itself is like J^2 and has the spectrum $j(j+1)$.

The fact that the information content entangled across a horizon is proportional to the area of the horizon arises very naturally in the above derivation. This, in turn, shows that the fundamental constant characterizing gravity is the quantum of area $\mathcal{A}_P$ which can hold approximately one bit of information. (It is the introduction of a quantity with dimensions of area, which frees us from having to worry about $\hbar$; the only quantum mechanical input we used is the periodicity in Euclidean time.) What is more, the conventional gravitational constant is given by $G = \mathcal{A}_P c^3/\hbar$ and will *diverge* when $\hbar \to 0$! This is strikingly reminiscent of the structure of bulk matter made of atoms. Though one can describe bulk matter using various elastic constants etc., such a description cannot be strictly considered as the $\hbar \to 0$ limit of quantum mechanics — since no atomic system can exist in this limit. Similarly, spacetime and gravity are inherently quantum mechanical just as bulk solids are.

This implies that spacetime dynamics is like the thermodynamic limit in solid state physics. In fact, this paradigm arises very naturally for any static spacetime with a horizon. Such a spacetime has a metric

$$ds^2 = N^2(\mathbf{x})\, dt^2 - \gamma_{\alpha\beta}(\mathbf{x})\, dx^\alpha dx^\beta\ , \tag{8}$$

with the horizon occurring at the surface $N = 0$. In this case, we have $R = {}^3\mathcal{R} + 2\nabla_i a^i$, where $a_i = (0, \partial_\alpha N/N)$ is the acceleration of $\mathbf{x} =$ constant world lines. Then, limiting the time integration to $[0, \beta]$, the gravitational action has the explicitly thermodynamic form

$$A_{\rm grav} = -\frac{\beta}{16\pi\mathcal{A}_P}\int_{\mathcal{V}} d^3x\ N\sqrt{\gamma}\,{}^3\mathcal{R} - \frac{\beta}{8\pi\mathcal{A}_P}\int_{\partial\mathcal{V}} d^2\mathcal{S}\ N(n_\alpha a^\alpha) \equiv \beta E - S, \tag{9}$$

In the Euclidean sector, the integral in the first term is proportional to energy (in the sense of spatial integral of the ADM Hamiltonian), and the second term is proportional to entropy in the presence of a horizon. (The signs are correct in the Euclidean sector with the signature we are using). $A_{\rm grav}$ thus indeed represents the free energy of the space-time, and various thermodynamic identities follow from its variation [5].

In summary, the basic fact that gravity can make regions inaccessible implies a loss of information, which—in turn—requires a surface term in the action describing the entropy. The physics of spacetime, like that of any other system with non-zero entropy, now needs to be obtained by extremising the free energy (and not the energy). The action has two naturally arising terms, neither of which can be covariant, since different foliations will lead to different levels of information loss. But the sum of the two terms is indeed covariant. The dynamics of spacetime is governed by the variation of the information–energy content and quantum entanglements across the horizon under small variations of the metric.

REFERENCES

[1] Landau, L. D. and Lifshitz, E. M. (1975). *Classical Theory of Fields*, Pergamon, New York, p. 310, eq. (102.3).
[2] Brown, J. D. (1995). *Phys. Rev. D* **52**, 7011 (Appendix B).
[3] Padmanabhan, T. (2003). *Astrophys. Space Sci.* **285**, (gr-qc/0209088).
[4] Padmanabhan, T. and Patel, A. (2003). (hep-th/0305165).
[5] Padmanabhan, T. (2002). *Class. Quant. Grav.* **19**, 5387 (gr-qc/0204019).