Gravity: The Inside Story

T. Padmanabhan*
21 March 2008

Abstract

It is well known that one could determine the *kinematics* of gravity by using the Principle of Equivalence and local inertial frames. I describe how the *dynamics* of gravity can be similarly understood by suitable thought experiments in a local Rindler frame. This approach puts in proper context several unexplained features of gravity and describes the dynamics of spacetime in a broader setting than in Einstein's theory.

History has a habit of playing tricks on physicists. We thought of electrons as particles and photons as waves, time as absolute and gravity as a force. I argue, in this essay, that we have similarly misunderstood the true nature of gravity because of the way the ideas evolved historically. I will show that, when seen with the 'right side up', the description of gravity becomes remarkably simple, beautiful, and explains features which we never thought needed explanation! The essay, building on previous results, provides a key conceptual consolidation.

So what was wrong with the historical development of gravity ? Einstein started with the Principle of Equivalence and — with a few thought experiments — motivated why gravity should be described by a metric of spacetime. This approach gave the correct backdrop for the equality of inertial and gravitational masses. But then he needed to write down the field equations which govern the evolution of g_{ab} and here is where the trouble started. There is no good guiding principle which Einstein could use that leads in a natural fashion to $G_{ab} = \kappa T_{ab}$ or to the corresponding action principle explaining several false starts he had. Sure, one can obtain them from a series of postulates but they just do not have the same compelling force as, for example, the Principle of Equivalence. Nevertheless, we all accepted general relativity, Einstein's equations and their solutions; after all, they agreed with observations so well!

But — conceptually — strange things happen as soon as: (i) we let the metric to be dynamical and (ii) allow for arbitrary coordinate transformations or, equivalently, observers on any timelike curve examining physics. *Horizons are inevitable in such a theory and they are always observer dependent.* This

*IUCAA, Post Bag 4, Ganeshkhind, Pune - 411 007, India; email: paddy@iucaa.ernet.in

conclusion arises very simply: (i) Principle of equivalence implies that trajectories of light will be affected by gravity. So in any theory which links gravity to spacetime dynamics, we can have nontrivial null surfaces which block information from certain class of observers. (ii) Similarly, one can construct timelike congruences (e.g., uniformly accelerated trajectories) such that all the curves in such a congruence have a horizon. You can't avoid horizons.

What is more, the horizon is *always* an observer dependent concept, even when it can be given a purely geometrical definition. For example, the $r = 2M$ surface in Schwarzschild geometry acts operationally as a horizon *only* for the class of observers who choose to stay at $r > 2M$ and not for the observers falling into the black hole.

Once we have horizons — which are inevitable — we get into more trouble. It is an accepted dictum that *all observers have a right to describe physics using an effective theory based only on the variables (s)he can access.* (This was, of course, the lesson from renormalization group theory. To describe physics at 10 GeV you shouldn't need to know what happens at 10^{14} GeV in "good" theories.) This raises the famous question first posed by Wheeler to Bekenstein: What happens if you mix cold and hot tea and pour it down a horizon, erasing all traces of "crime" in increasing the entropy of the world? The answer to such thought experiments *demands* that horizons should have an entropy which should increase when energy flows across it.

With hindsight, this is obvious. The Schwarschild horizon — or for that matter any metric which behaves locally like Rindler metric — has a temperature which can be identified by the Euclidean continuation. If energy flows across a hot horizon $dE/T = dS$ leads to the entropy of the horizon. Again, historically, nobody — including Wheeler and Bekenstein — looked at the Euclidean periodicity in the Euclidean time (in Rindler or Schwarzschild metrics) *before* Hawking's result came! And the idea of Rindler temperature came *after* that of black hole temperature! So in summary, the history proceeded as follows:

Principle of equivalence ($\sim$ 1908)

$\Rightarrow$ **Gravity is described by the metric g_{ab} ($\sim$ 1908)**

? **Postulate Einstein's equations *without a real guiding principle!* (1915)**

$\Rightarrow$ **Black hole solutions with horizons (1916) allowing the entropy of hot tea to be hidden ($\sim$ 1971)**

$\Rightarrow$ **Entropy of black hole horizon (1972)**

$\Rightarrow$ **Temperature of black hole horizon (1975)**

$\Rightarrow$ **Temperature of the Rindler horizon (1975-76)**.

This historical sequence raises a some serious issues for which there is no satisfactory answer in the conventional approach:

1. How can horizons have temperature without the spacetime having a microstructure?

It simply cannot. Recall that the thermodynamic description of matter at finite temperature provides a crucial window into the existence of the corpuscular substructure of solids. As Boltzmann taught us, heat is a form of motion and we will not have the thermodynamic layer of description if matter is a continuum all the way to the finest scale and atoms did not exist! *The mere existence of a thermodynamic layer in the description is proof enough that there are microscopic degrees of freedom.* — in a solid or in a spacetime. In the conventional approach, we are completely at a loss to understand why horizons are hot or what kind of 'motion' is this 'heat'?

2. Why is it that Einstein's equations reduces to a thermodynamic identity for virtual displacements of a horizon?

Here is the first algebraic mystery — which has no explanation in conventional approach — suggesting a deep connection between the dynamical equations governing the metric and the thermodynamics of horizons. The first example was provided in ref. [1] in which it was shown that, in the case of spherically symmetric horizons, Einstein's equations can be interpreted as a thermodynamic relation $TdS = dE + PdV$ arising out of virtual radial displacements of the horizon. Further work showed that this result is valid in *all* the cases for which explicit computation can be carried out — as diverse as the Friedmann models as well as rotating and time dependent horizons in Einstein's theory [2]. Treating them as just some solutions to Einstein's field equations we cannot understand these results.

3. Why is Einstein-Hilbert action is holographic with a surface term that encodes same information as the bulk?

The Einstein-Hilbert Lagrangian has the structure $L_{EH} \propto R \sim (\partial g)^2 + \partial^2 g$. In the usual approach the surface term arising from $L_{sur} \propto \partial^2 g$ has to be ignored or canceled to get Einstein's equations from $L_{bulk} \propto (\partial g)^2$. But there is a peculiar (again unexplained) relationship between L_{bulk} and L_{sur}:

$$\sqrt{-g}L_{sur} = -\partial_a \left(g_{ij} \frac{\partial \sqrt{-g} L_{bulk}}{\partial(\partial_a g_{ij})} \right) \tag{1}$$

This shows that the gravitational action is 'holographic' with the same information being coded in both the bulk and surface terms making either one of them to be sufficient. It is well known that varying g_{ab} in $L_{\rm bulk}$ leads to the standard field equations. More remarkable is the fact that one can also obtain Einstein's equations from an action principle which uses only the surface term and the virtual displacements of horizons [3] *without* treating the metric as a dynamical variable.

4. Why does the surface term in Einstein-Hilbert action give the horizon entropy?

Yet another algebraic result which defies physical understanding! You first throw away the surface term in the action, vary the rest to get the field equations, find a solution with a horizon, compute its entropy — only to discover that the surface term you threw away is intimately related to the entropy.

5. Why do all these results hold for a much wider class of theories than Einstein gravity, like Lanczos-Lovelock models?

There are more serious 'algebraic accidents' in store. Recent work has shown that *all the thermodynamic features described above extend far beyond Einstein's* *theory.* The connection between field equations and the thermodynamic relation $TdS = dE + PdV$ is not restricted to Einstein's theory (GR) alone, but is in fact true for the case of the generalized, higher derivative Lanczos-Lovelock gravitational theory in D dimensions as well [4]. The same is true for the holographic structure of the action functional [5]: the Lanczos-Lovelock action has the same structure and — again — the entropy of the horizons is related to the surface term of the action.

I want to argue that these (and several related features) are not algebraic accidents but indicate that we have been looking at gravity the wrong way around. In the proper perspective, these features should emerge as naturally as the equivalence of inertial and gravitational masses emerges in the geometric description of the kinematics of gravity. *These results show that the thermodynamic description is far more general than just Einstein's theory* and occurs in a wide class of theories in which the metric determines the structure of the light cones and null surfaces exist blocking the information. So instead of the historical path, I will proceed as follows reversing most of the arrows:

Principle of equivalence

$\Rightarrow$ **Gravity is described by the metric** g_{ab}

$\Rightarrow$ **Existence of local Rindler frames (LRFs) with a horizon around any event**

$\Rightarrow$ **Temperature of the local Rindler horizon** $\mathcal{H}$ **from the Euclidean continuation**

$\Rightarrow$ **Virtual displacements of** $\mathcal{H}$ **allow for flow of energy across a hot horizon hiding an entropy** $dS = dE/T$ **as perceived by a given observer**

$\Rightarrow$ **The local horizon must have an entropy,** S_{grav}

$\Rightarrow$ **The dynamics should arise from maximizing the total entropy of horizon (**S_{grav}**) plus matter (**S_m**) for all LRF's leading to field equations!**

Let me elaborate. Take an event P and introduce a local inertial frame (LIF) around it with coordinates X^a. Go from the LIF to a local Rindler frame (LRF) coordinates x^a by accelerating along, say, x-axis with an acceleration κ. This LRF and its local horizon $\mathcal{H}(x=0)$ will exist within a region of size $L \ll \mathcal{R}^{-1/2}$ as long as $\kappa^{-1} \ll \mathcal{R}^{-1/2}$ where $\mathcal{R}$ is a typical component of curvature tensor. Now people can pour tea across $\mathcal{H}$ as suggested by Wheeler. Alternatively, one can consider a virtual displacement of the $\mathcal{H}$ normal to itself engulfing the tea. Either way, some entropy will be lost to the outside observers unless the horizon has an entropy. That is, displacing a piece of local Rindler horizon should cost some entropy S_{grav}, say. It is then natural to demand that the dynamics should follow from the prescription $\delta[S_{grav} + S_{matter}] = 0$.

All we need is the expressions for S_{matter} and S_{grav}. I will now write down the general expressions for both [6], such that they have the correct interpretation in the LRF. Assuming a $D(\geq 4)$ dimensional spacetime for the later convenience, I take:

$$S_{\rm matt} = \int_{\mathcal{V}} d^D x \sqrt{-g} T_{ab} n^a n^b \tag{2}$$

$$S_{grav} = -4 \int_{\mathcal{V}} d^D x \sqrt{-g} P_{ab}{}^{cd} \nabla_c n^a \nabla_d n^b \qquad (3)$$

where n^a is vector field which will reduce to the null normal $[\partial_a(T - X)]$ on the horizon $\mathcal{H}$, T_{ab} is the matter energy momentum tensor and $P_{ab}{}^{cd}$ is defined below. The $S_{\rm matt}$ is easy to understand. In the LRF, (with $-g_{tt} = 2\kappa x = g^{xx}, \sqrt{-g} = 1$) an infinitesimal spacetime region will contribute $T_{ab} n^b n^b d^3 x \, dt = \delta E \, dt$ which on integration over t in the range $(0, \beta)$ where $\beta^{-1} = T = (\kappa/2\pi)$ gives $\delta S_{matter} = \beta \delta E = \beta T_{ab} n^a n^b d^3 x$ when the energy flows across a surface with normal n^a. Integrating, we get $S_{\rm matt}$ to which Eq. (2) reduces to in LRF. (For example, if T_{ab} is due to an ideal fluid at rest in LIF, $T_{ab} n^a n^b$ will contribute $(\rho + P)$, which — by Gibbs-Duhem relation — is just Ts where s is the entropy density. Integrating over $\sqrt{-g} d^4 x = dt d^3 x$ with $0 < t < \beta$ gives $S_{\rm matt}$.)

The S_{grav}, on the other hand, is a general quadratic functional of the derivatives of n_a which is the form of entropy of an elastic solid, say, if n^a is the displacement field. Here we interpret it as the entropy cost for virtual displacement of horizon. The crucial requirement is that, dynamics for the *background spacetime* should emerge when we set $(\delta S_{tot}/\delta n_a)$=0 for *all* null vectors n^a. Incredibly enough, this can be achieved if (and only if) (i) the tensor P_{abcd} has the algebraic symmetries similar to the Riemann tensor R_{abcd} and (ii) we have $\nabla_a P^{abcd} = 0 = \nabla_a T^{ab}$. One can now show that [6] such a tensor can be constructed as a a series in the powers of the derivatives of the metric:

$$P^{abcd}(g_{ij}, R_{ijkl}) = c_1 \, P_{(1)}{}^{abcd}(g_{ij}) + c_2 \, P_{(2)}{}^{abcd}(g_{ij}, R_{ijkl}) + \cdots , \qquad (4)$$

where $c_1, c_2, \cdots$ are coupling constants with the *unique* m-th order term being $P_{ab}{}^{cd} \propto \partial \mathcal{L}_m^{(D)} / \partial R^{ab}{}_{cd}$ where $\mathcal{L}_m^{(D)}$ is the m-th order Lanczos-Lovelock Lagrangian [3, 7]. Then maximizing $(S_{\rm grav} + S_m)$ gives [6]:

$$16\pi \left[P_b{}^{ijk} R^a{}_{ijk} - \frac{1}{2} \delta^a_b \mathcal{L}_m^{(D)} \right] = 8\pi T^a_b + \Lambda \delta^a_b \qquad (5)$$

These are identical to the field equations for Lanczos-Lovelock gravity with a cosmological constant arising as an undetermined integration constant. The lowest order term $P^{ab}_{cd} = (1/32\pi)(\delta^a_c \delta^b_d - \delta^a_d \delta^b_c)$ leads to Einstein's theory while the first order term gives the Gauss-Bonnet correction. One can show, in the general case of Lanczos-Lovelock theory, Eq. (3) *does* give the correct gravitational entropy justifying our choice.

Let us get back to physics behind all these. The fact that matter crossing a hot horizon (or the horizon crossing the matter, in a virtual displacement) in the LRF should cost entropy is my starting point. One then writes down an expression for $(S_{matter} + S_{grav})$ and demands that it should be maximized with respect to all the null vectors — which are normals to local patches of null surfaces that can act locally as horizons for a suitable class of observers — in the spacetime. This puts a constraint on the *background* spacetime leading to our field equations. To the lowest order, this gives Einstein's equations with calculable corrections. These equations are no more fundamental than equations of elasticity but the thermodynamic identities are indeed fundamental. This approach answers the questions we raised earlier quite nicely:

- There *are* microscopic degrees of freedom ("atoms of spacetime") which we know nothing about. But just as thermodynamics worked even before we understood atomic structure, we can understand long wavelength gravity in a corpuscular spacetime by a thermodynamic approach.

- Einstein's equations *are* essentially thermodynamic identities valid for each and every local Rindler observer. In spacetimes with horizons and *high level of symmetry*, one can also consider virtual displacements of these horizons (like $r_H \to r_H + \epsilon$) and obviously we will again get $TdS = dE + PdV$.

- If the the flow of matter across a horizon costs entropy, the resulting gravitational entropy has to be related to the microscopic degree of freedom associated with the horizon surface. It follows that any dynamical description will require a holographic action with both surface and bulk encoding the same information. For the same reason, the surface term in the action will give the gravitational entropy.

Most importantly, we are not just reformulating Einstein's theory; Shifting the emphasis from Einstein's field equations to a broader picture of spacetime thermodynamics of horizons leads to a general class of field equations in Eq.(5) which includes Lanczos-Lovelock gravity. It is now no surprise that Lanczos-Lovelock action is also holographic, is related to entropy and has a thermodynamic interpretation.

References

[1] T Padmanabhan,*Class. Quan. Grav.*, **19**, 5387, (2002) [gr-qc/0204019].

[2] R-G Cai, L-M Cao, [gr-qc/0611071]; M Akbar, R-G Cai, [hep-th/0609128]; D. Kothawala, S. Sarkar, T. Padmanabhan, Phys. Letts, **B 652**, 338-342 (2007) [gr-qc/0701002].

[3] T. Padmanabhan, AIP Conference Proceedings **861**, 858-866, [astro-ph/0603114].

[4] A Paranjape, S Sarkar and T Padmanabhan, *Phys. Rev.* **D74**, 104015, (2006) [hep-th/0607240].

[5] A Mukhopadhyay, T Padmanabhan, *Phys. Rev.* **D74**,124023, (2006).

[6] T. Padmanabhan, *Gen.Rel.Grav.*, **40**, 529-564 (2008) [arXiv:0705.2533]; T. Padmanabhan, A. Paranjape, *Phys.Rev.* **D75** 064004, (2007).

[7] C Lanczos, *Z. Phys.* **73**, 147, (1932); *Annals Math.* **39**, 842, (1938); D Lovelock, *J. Math. Phys.*, **12**, 498 (1971).