Horizon entropy and higher curvature equations of state

Raf Guedens, Ted Jacobson, Sudipta Sarkar

Abstract. The Clausius relation between entropy change and heat flux has previously been used to derive Einstein’s field equations as an equation of state. We examine here whether a similar derivation can be given for extensions beyond Einstein gravity to include higher derivative and higher curvature terms.

The origin of the thermodynamic behavior of spacetime horizons can be traced to local physics, and in a sense arises from the nature of the vacuum. This led to the observation that the Einstein equation can be derived as an equation of state for local causal horizons in the neighborhood of a point ‘p’ in spacetime, by imposing the Clausius relation between their entropy change and the energy flux across them [1]. In this paper we present an approach [2], which generalizes this equation of state derivation to allow for higher derivative contributions to the entropy and field equations. The Einstein-Hilbert Lagrangian is only the lowest order term (other than the cosmological constant) in a derivative expansion of generally covariant actions for a metric theory, and the presence of higher derivative terms is presumably inevitable. Hence, it is interesting to examine whether the thermodynamic derivation of the field equations transcends beyond general relativity (GR).

Our approach differs in several ways from the original derivation for GR, among which are: (i) the entropy is compared on two horizon slices that share a common boundary, (ii) the bifurcation surface lies to the past of the terminal point p and, (iii) the entropy depends on the approximate Killing vector. In particular, it has the same dependence on the approximate Killing vector as have the Noether charges associated with a Lagrangian, in analogy with the Wald entropy [3, 4] for stationary black holes. It makes some sense that the entropy depends on the approximate Killing vector, because the latter determines the notion of stationarity and enters the definition of the heat flux. However, at present we can offer no statistical interpretation for this form of the entropy.

The equation of state we can derive is consistent with local energy-momentum conservation only if the leading order term in the entropy satisfies an integrability condition. This condition is satisfied if the entropy arises from variation of a generally covariant function with respect to curvature. In other words, the entropy coincides with a Noether charge associated with a (particular type of) gravitational Lagrangian. The need for such an integrability condition was anticipated in Ref. [1], since it was known that the entropy of stationary black hole horizons has this form[4, 5, 6]. We have not been able to ascertain whether this is the only way to satisfy
the integrability condition, and in particular whether field equations for Lagrangians involving derivatives of curvature can be obtained.

To present the derivation, let us first define the notion of a local causal horizon \( H \) as follows. Consider any spacetime point \( p \) in a \( D \)-dimensional spacetime, and let \( \Sigma_p \) be any small patch of spacelike \((D - 2)\)-surface through \( p \). The boundary of the past of \( \Sigma_p \) in the neighborhood of \( p \) has two components, each of which is a null surface generated by a congruence of null geodesics orthogonal to \( \Sigma_p \). The local causal horizon (LCH), denoted by \( H \) is defined as one of these components.

Next we define the approximate Killing vector \( \xi^a \) that plays a central role. We will refer to this vector field as the “local Killing vector”. Of course a general curved spacetime has no Killing vectors. Nevertheless, in a small enough neighborhood of any point, any spacetime is approximately flat. Then, there exits a notion of an approximate Killing vector \( \xi^a \), such that it obeys properties of a true Killing vector in an approximate way. A true Killing vector \( \xi^a \) obeys the Killing equation, \( \nabla_a \xi^b = 0 \) and also the Killing identity,

\[
\nabla_a \nabla_c \xi_b = R^d_{\ abc} \xi_d, \tag{1}
\]

which follows from the Killing equation. Our computations will rely on this identity being satisfied at a certain order, but for the approximate Killing field, the Killing identity will generally not hold at that order. Then, narrowing our sights, \( \xi^b \) can be chosen so that the Killing identity holds to the required order on the single horizon generator \( \Gamma \) that ends on the terminal point \( p \), and this turns out to be good enough. In effect, the local Killing symmetry can be extended away from \( p \) to a better approximation along a single null generator \( \Gamma \) than across the whole LCH.

Given that, we confine the local causal horizon to a narrow region surrounding the central horizon generator \( \Gamma \), the integrals appearing in the Clausius relation will be dominated by their integrands evaluated on \( \Gamma \). As such, any condition that \( \xi^a \) may need to satisfy in order for the Clausius relation to lead to a consistent equation of state will be conditions imposed on \( \Gamma \). The local Killing vector \( \xi^a \) is taken to have a bifurcation surface \( \Sigma_0 \)—or at least a bifurcation point—at \( p_0 \) to the past of \( p \), where it vanishes and where its covariant derivative generates boosts in the plane orthogonal to the bifurcation surface. Then, we use the null normal coordinate (NNC) system introduced in [7] to express the full set of requirements that must be imposed on \( \xi^a \) and we can show by explicit construction in NNC that both Killing equation and Killing identity are satisfied by the approximate Killing vector \( \xi^a \) at the required order as long as we concentrate on a narrow patch near the central null generator \( \Gamma \) of LCH. For details, see Ref. [2].

We also need to define the heat flux through the LCH as,

\[
\delta Q = \int_{H} (-T_{ab}) \xi^b dH_a \tag{2}
\]

The integration measure is \( dH_b = -k_b dV dA \), where \( V \) is the affine parameter along the generators of LCH, \( k^b = (\partial V)^b \) is tangent to the generators, and \( dA \) is the area element of a constant \( V \) slice. The heat flux integrand is proportional to the local Killing vector which vanishes at \( p_0 \) and is thus of \( O(x) \) in NNCs in the neighborhood of \( p_0 \). Also, the relevant temperature is the Unruh boost temperature, \( T = \hbar/2\pi \).

The Clausius relation \( T \delta S = \delta Q \), applied to a horizon refers to the entropy change \( \delta S \) between two times. Those times are spacelike hypersurfaces, one to the future of the other, which intersect the horizon in two slices. For a black hole with a compact horizon, the two slices may bound a cylindrical region of the horizon. For a local causal horizon, the considered process must be local, since the LCH is not even well defined except in a small neighborhood of the terminal point \( p \). To localize the process we can restrict attention to cases where, rather
than pushing the spacelike hypersurface forward in time everywhere, it is deformed to the future only in a small neighborhood of the a bifurcation point \( p_0 \). Then the two corresponding horizon slices also coincide everywhere except in a small region. \( \Sigma_0 \) corresponds to the \( V = V_0 \) surface in NNC of [7] and \( \Sigma \) lies to the future. If we truncate the horizon slices outside the region where they differ, their union \( \Sigma_0 \cup \Sigma \) forms the closed boundary of a patch of the horizon. This will allow the difference of the entropies on the two slices to be computed using Stokes’ theorem.

In the context of horizon thermodynamics it is natural to expect that the entropy is an extensive quantity. Thus we assume that it can be expressed as an integral of a \((D-2)\)-form over a spacelike slice \( \Sigma \) of the horizon. We adopt the dual description, and express the entropy as the integral

\[
S = \int_{\Sigma} s^{ab} N_{ab} dA,
\]

where the entropy density \( s^{ab} \) is an antisymmetric tensor, \( N^{ab} \) is the binormal to the slice and \( dA \) is the area element on the slice. The change of this entropy between the two horizon slices is

\[
\delta S = S - S_0 = \int_{\Sigma \cup \Sigma_0} s^{ab} N_{ab} dA = -2 \int_{H} \nabla_b s^{ab} k_a dV dA,
\]

where \( \Sigma_0 \) is taken with the opposite orientation to \( \Sigma \), and in the last step, we have used Stokes’ theorem.

The Clausius relation is imposed in the limit \( p_0 \to p \), which means that the entropy change integrand of (4), multiplied by \( T \), must be equal to the \( O(x) \) heat flux integrand, up to \( O(x^2) \) terms. That is,

\[-\left(\frac{\hbar}{\pi}\right) \nabla_b s^{ab} k_a = T^{ab} \xi^b k_a + O(x^2).
\]

This relation is imposed at all points \( p \) and for all null directions at \( p \).

Next, we propose a local entropy density of the same form as the Noether potential [3, 4] and assume the entropy density takes the form

\[
s^{ab} = \frac{2\pi}{\hbar} \left( W^{abc} \xi_c + X^{abcd} \nabla_c \xi_d \right),
\]

where \( W \) and \( X \) are, so far, unspecified tensors, constructed locally from the dynamical fields, antisymmetric in the first two indices, and in addition \( X \) is antisymmetric in the last two indices. Then, the validity of Clausius relation \( T \delta S = \delta Q \) requires that \( W \) is a combination of divergences of \( X \) as,

\[
W^{arb} = \nabla_s \left( X^{sarb} + X^{sbra} + X^{srba} \right).
\]

For details, see Ref. [2]. Using this result and imposing the validity of the Clausius relation (5) for all \( k^a \) ultimately imply,

\[
R^{(a}_{rst}X^{b)rst} + 2\nabla_r \nabla_s X^{(a|s|b)r} + \Phi g^{ab} = -\frac{1}{2} T^{ab},
\]

where \( \Phi \) is some scalar function that may depend on the metric and curvature. Note that (8) follows from the Clausius relation irrespective of the values of expansion and shear anywhere on the horizon patch.

We now impose local conservation of energy-momentum tensor, \( T_{ab} \) to determine the function \( \Phi \). Then Eqn. (8) leads to

\[
\nabla_a \Phi = -\nabla_b \left( R^{(a}_{rst}X^{b)rst} + 2\nabla_r \nabla_s X^{(a|s|b)r} \right).
\]
In order for such a $\Phi$ to exist, the right hand side must be the gradient of a scalar. This integrability condition further constrains the nature of $X$, which so far is only required to be anti-symmetric in both the first and second pair of indices.

If the integrability condition is satisfied, then the left hand side of (8) is a divergence free tensor constructed from the metric and its derivatives. One way to obtain such a tensor is from the variational derivative of a scalar action functional with respect to the metric, $\delta I_g / \delta g_{ab}$, which is automatically divergence free. In fact, it was argued in Ref. [8] that all such tensors arise in this way. If so, then the ‘Clausius equation’ (8) is precisely the equation of motion that derives from the action $I_g + I_{\text{matter}}$, with an undetermined cosmological constant. More specifically, if a gravitational Lagrangian $L[g_{ab}, R^{a}{}_{bcd}]$ is a scalar formed algebraically from the metric and Riemann tensor, then the corresponding equation of motion is precisely (8), with

$$X^{abcd} = - \frac{\partial L}{\partial R_{abcd}},$$  \hspace{1cm} (10)

and $\Phi = L/2$. This completes the derivation of the field equations from the Clausius relationship applied to LCH.

To summarize, we assume that the entropy of LCH depends on an approximate local Killing vector in a way that mimics the diffeomorphism Noether charge that yields the entropy of a stationary black hole. We showed that by a careful choice of the nature of the horizon patch to which the Clausius relation is applied, it is possible to derive the field equation of a gravity theory whose lagrangian is a algebraic function of metric and curvature. In particular, the Clausius relation must refer to the change of entropy between two slices of the horizon that together form the complete boundary of a patch, and this patch must be narrow enough to neglect violations of the Killing identity.

Together with matter energy conservation, the Clausius relation applied to all such local horizon patches leads to an integrability condition on the assumed horizon entropy density. This condition can be satisfied if the latter is in fact a Noether potential associated with a Lagrangian constructed algebraically from the metric and Riemann tensor. In that case the Clausius relation implies that the field equation for that Lagrangian holds. We have not proved that this is the only way to satisfy the integrability condition, but that may be the case. In particular, the field equation for a theory with derivatives of curvature in the Lagrangian is unlikely to be obtained in this way using for the entropy density a Noether potential derived from the Lagrangian, although it might conceivably arise from a different entropy density.

It is therefore not clear to us whether the derivation has any thermodynamic significance. We regard it as a positive answer to a technical question, but its physical interpretation remains obscure. Then again, it may just be that the contribution of higher curvature terms to a gravitational field equation cannot be sensibly captured at a local thermodynamic level.

References
[1] T. Jacobson, “Thermodynamics of space-time: The Einstein equation of state,” Phys. Rev. Lett. 75, 1260 (1995) [arXiv:gr-qc/9504004].
[2] R. Guedens, T. Jacobson and S. Sarkar, Phys. Rev. D 85, 064017 (2012) [arXiv:1112.6215 [gr-qc]].
[3] R. M. Wald, “Black hole entropy is the Noether charge,” Phys. Rev. D48, 3427-3431 (1993). [gr-qc/9307038].
[4] V. Iyer and R. M. Wald, “Some properties of Noether charge and a proposal for dynamical black hole entropy,” Phys. Rev. D 50, 846 (1994) [arXiv:gr-qc/9403028].
[5] M. Visser, “Dirty black holes: Entropy as a surface term,” Phys. Rev. D48, 5697-5705 (1993). [hep-th/9307194].
[6] T. Jacobson, G. Kang, R. C. Myers, “On black hole entropy,” Phys. Rev. D49, 6587-6598 (1994). [gr-qc/9312023].
[7] R. Guedens, arXiv:1201.0542 [gr-qc].
[8] E. Curiel, “On Tensorial Concomitants and the Non-Existence of a Gravitational Stress-Energy Tensor,” arXiv:0908.3322 [gr-qc].