Gauge-invariant quadratic approximation of quasi-local mass and its relation with Hamiltonian for gravitational field

Jacek Jezierski\textsuperscript{1,}\textsuperscript{*}, Jerzy Kijowski\textsuperscript{2} and Piotr Waluk\textsuperscript{2}

\textsuperscript{1} Faculty of Physics, University of Warsaw, ul. Pasteura 5, 02-093 Warsaw, Poland
\textsuperscript{2} Center for Theoretical Physics, Polish Academy of Sciences, Aleja Lotników 32/46, 02-668 Warsaw, Poland

E-mail: Jacek.Jezierski@fuw.edu.pl, kijowski@cft.edu.pl and waluk@cft.edu.pl

Received 25 August 2020, revised 7 January 2021
Accepted for publication 13 January 2021
Published 8 April 2021

Abstract

Gauge invariant, Hamiltonian formulation of field dynamics within a compact region $\Sigma$ with boundary $\partial\Sigma$ is given for the gravitational field linearized over a Kottler metric. The boundary conditions which make the system autonomous are discussed. The corresponding Hamiltonian functional $\mathcal{H}_{\text{Inv}}$ uniquely describes the energy carried by the (linearized) gravitational field. It is shown that, under specific boundary conditions, the quasi-local Hawking mass $\mathcal{H}_{\text{Haw}}$ reduces to $\mathcal{H}_{\text{Inv}}$ in the weak field approximation. This observation is a quasi-local version of the classical Brill–Deser result (Brill and Deser 1968 \textit{Ann.Phys.(N.Y.)} 50 3) and enables us to declare Hawking mass as the correct expression (at least up to quadratic terms in the Taylor expansion) for the quasi-local mass, which correctly describes energy carried by the gravitational field.

Keywords: quasi-local energy, Hawking mass, Hamiltonian, linearized gravity

(Some figures may appear in colour only in the online journal)
Dynamics of the linear theory of gravity can be formulated in terms of two gauge-invariant, non-constrained degrees of freedom. For this purpose one can use e.g. selected components of the (linearized) Weyl tensor [2–5]. The two degrees of freedom contain the entire information about dynamics: knowing them, the complete field configuration can be uniquely (up to gauge transformations) reconstructed in a quasi-local way. The phase space of (complete) initial data on a given Cauchy surface \( C \) carries a canonical ADM-symplectic structure. The value of its Hamiltonian functional is uniquely determined by this structure and the field dynamics: it is a positive, quadratic form of gauge invariants [3, 5]. In case of a flat background, this quantity satisfies an important consistency test: it is equal to the second variation (the leading term in the Taylor expansion) of the total ADM energy, commonly accepted as the global Hamiltonian function (i.e. the total field energy) in the complete, nonlinear theory [1].

Any reasonable definition of a quasi-local mass should also satisfy a quasi-local version of the above test in the following sense. Given a compact region \( \Sigma \subset C \) with boundary \( \partial \Sigma \) on a Cauchy surface, the field dynamics within \( \Sigma \) of the linearized gravitational field is not an autonomous system because of the influence of the exterior \( \partial \Sigma \) on the interior of \( \Sigma \). But, there is a unique way to impose boundary conditions on \( \partial \Sigma \) which act like an adiabatic insulation, i.e. the dynamics becomes an autonomous Hamiltonian system and the corresponding Hamiltonian functional \( \mathcal{H}_{\text{inv}} \) is positive (cf e.g. [6, 7]). This way, the field energy acquires a quasi-local (and not only global) meaning.

In the present paper we check the consistency of the (quasi)-local Hawking mass with the local energy content of the linear theory for a general spherically-symmetric background: an arbitrary Kottler metric. Our main result is that the second variation of the Hawking mass assigned to topological two-dimensional (2D) spheres constituting a boundary \( \partial \Sigma \) of a compact region \( \Sigma \) agrees with the amount of (gauge-independent) field energy of the linearized gravity contained within, modulo a certain gauge-dependent boundary term. The complete agreement can be obtained if we impose an appropriate gauge condition at the boundary \( \partial \Sigma \) which annihilates the undesirable term in the Taylor expansion.

The gauge-dependence of the field energy is not a paradox. We stress that the imposed gauge condition plays a role which is much more fundamental than merely a ‘convenient gauge’ used to annihilate unwanted terms in the expansion. Indeed, the 2D surface \( \partial \Sigma \) plays a double role in the definition of quasi-local mass \( E_{\partial \Sigma} \). The first, obvious one, is to demarcate the region \( \Sigma \) whose energy content we want to measure\(^4\). But the second role, that of a ‘reference frame’, is related to the very notion of field energy, which is not a scalar quantity: it is always measured with respect to a reference frame. In special relativity theory, reference frame can be identified with a vector field, say \( T \), which must be a symmetry field of the spacetime geometry. Field evolution consisting in shifting the field configuration along this field becomes an autonomous Hamiltonian system, with the Hamiltonian function provided by the Noether theorem. The same procedure works not only for the total energy, but also for local energy contained in a bounded 3D region \( \Sigma \), provided appropriate boundary conditions at the boundary of the

---

\(^3\) The term ‘quasi-local’ refers to the fact that the obtained notion of gravitational field energy, although no longer dependent on the whole global field configuration, is not a local object. It can be defined for a finite, but extended region of spacetime, and to compute it for a given bounded region \( \Sigma \) one requires not only the knowledge of the whole field configuration on \( \Sigma \), but also of the boundary conditions on \( \partial \Sigma \).

\(^4\) In General relativity theory, two 3D regions \( \Sigma_1 \) and \( \Sigma_2 \) having the same boundary \( \partial \Sigma_1 = \partial \Sigma_2 \) contain the same amount of energy due to the diffeomorphism invariance of the dynamics.
world tube $\Sigma \times \mathbb{R}$ are satisfied, which assure the adiabatic insulation of its interior from the exterior.\(^5\)

In principle, nothing prevents us from using the same construction for an arbitrary vector field $T$. But the resulting Hamiltonian system is no longer autonomous if $T$ is not a symmetry field. The value of the corresponding Hamiltonian is no longer conserved and cannot be interpreted as the field energy. If $T$ is a combination of time translation, space translation, rotation, boost etc, the resulting Hamiltonian function is a strange combination of energy, momentum, angular momentum, static moment etc, but for a generic $T$ it is difficult to find any reasonable interpretation of such a quantity (see [8]).

In general relativity, Noether theorem does not provide any valuable ‘energy density’ and Penrose introduced the notion of a ‘quasi-local mass’, cf [9], to stress the fact that energy contained in $\Sigma$ must depend only upon geometric properties of its boundary $\partial \Sigma$.\(^6\) Nevertheless, the field evolution can still be interpreted as a Hamiltonian system, provided the interior of the tube $\Sigma \times \mathbb{R}$ is adiabatically insulated by appropriate boundary conditions [6, 8, 12]. Again, it is hard to call ‘energy’ a Hamiltonian functional obtained via such a procedure, unless we choose the field $T$ in a way that eliminates all those unwanted ingredients (like rotation, boost, space translation etc). To do that, the only ‘reference frame’ being at our disposal when defining the quasi-local energy $E_{\partial \Sigma}$ is the surface itself: whenever the extrinsic curvature vector $H^\mu$ of the boundary is spacelike there is a geometrically preferred timelike vector field $T^\mu$ which is orthogonal to $\partial \Sigma$ and to $H^\mu$, i.e. satisfying $g_{\mu \nu} T^\mu H^\nu = 0$.

In the flat Minkowski spacetime this procedure produces a useful, self-parallel vector field $T$ only when the surface $\partial \Sigma$ is sufficiently ‘rigid’, i.e. contained in a flat 3D hyperplane. Otherwise, the value of the Hamiltonian function obtained this way cannot be interpreted as a field energy. In particular, one cannot expect any reasonable properties (like positivity, convexity etc) that usually characterize the energy functional and enable us to prove important properties of the field evolution (e.g. stability).

It turns out that the above rigidity condition can be generalized to a generic, curved spacetime (see [13, 14], where the existence of the eight-parameter family of rigid spheres has been proved). A natural hypothesis arises that the quasi-local Hamiltonian function, defined by imposing appropriate boundary conditions, can be interpreted as the field energy only for such ‘rigid spheres’. The main result of our paper supports this hypothesis. Indeed, the curious gauge condition, which is necessary to obtain equality between the quadratic term in the expansion of the Hawking mass and the field energy of the linear gravity, is just the linearized version of the rigidity condition of the surface $\partial \Sigma$.

In our opinion, this result shows that the Hawking mass is a viable candidate for the correct expression for the quasi-local gravitational field energy. It also becomes a sort of ‘reference gauge’—any other proposed expression should not differ from the Hawking mass up to second order approximation to be a sensible option.

\(^5\)Here, again, two 3D regions $\Sigma_1$ and $\Sigma_2$ having the same boundary contain the same amount of energy due to: (1) ‘conservation laws’ satisfied by the Noether energy–momentum tensor and (2) boundary conditions. The latter assure the uniqueness of the evolution.

\(^6\)Although we pursue the way of defining quasi-local mass as a boundary object, one should note that competitive approaches exist, in which the quasi-local mass is being assigned to a three-dimensional region or even to whole four-dimensional domains of dependence [10]. Various approaches naturally coincide in certain special cases, eg. the Tolman mass, originally defined as an integral of the stress-energy tensor over a three-dimensional volume can be rewritten as a boundary integral of generalized surface gravity in stationary spacetimes [11].
1. Technical setup

We will work within the framework of the Cauchy problem for the Einstein equation in vacuum, in the ADM formulation thereof. The background for linearization of the theory will be a Kottler metric:

$$\eta_{\mu\nu} dx^\mu dx^\nu = -f dt^2 + \frac{1}{f} dr^2 + r^2 [d\theta^2 + \sin^2 \theta d\varphi^2].$$ (1)

This is a general spherically symmetric solution, which encompasses Minkowski, (anti)de Sitter and Schwarzschild solutions as special cases. The function $f$ has the form: $f(r) := 1 - \frac{2m}{r} - \frac{r^2}{2} \Lambda$. Our choice of coordinates: $(x^0, x^1, x^2, x^3) = (t, \theta, \varphi, r)$ is fixed by this form of the background metric.

We consider a compact region on a Cauchy surface, foliated by a family of 2D spheres:

$$\Sigma = \{ x^0 = t_0, r_0 \leq x^3 \leq r_1 \} = \bigcup_{r \in [r_0, r_1]} S(r), \quad S(r) = \{ x \in \Sigma : x^3 = r \}. \quad (2)$$

We assume that $\Sigma$ lies within the domain of positive $f$—outside of the black hole horizon and within the cosmological horizon (if it exists). One should note that apart from its impact on the location and existence of horizons, the value of the cosmological constant $\Lambda$ does not influence our derivations. Our formalism and reasoning remain valid whether $\Lambda$ is positive, negative or zero, and since we are dealing with a compact region—the changes in asymptotic behaviour of the metric due to the presence and sign of the cosmological constant $\Lambda$ do not come into play. We are therefore able to consider all three cases simultaneously.

A following indexing convention will be used to denote dimensionality of geometric objects: lowercase Greek, lowercase Latin and uppercase Latin indices will denote full space-time (\( \mu, \nu = 0, 1, 2, 3 \)), Cauchy surface (\( k, l = 1, 2, 3 \)) and 2D sphere (\( A, B = 1, 2 \)) coordinates respectively. The coordinate derivative and the 2D covariant derivative on $S(r)$ will be denoted by a comma and the symbol ‘\( \parallel \)’. The volume form on the spheres defined by the Kottler metric is $\Pi := r^2 \sin \theta$ and $\lambda$ denotes the counterpart volume form for $g_{AB}$. The symbol $\hat{\Delta}$ is the Laplace–Beltrami operator on a unit sphere (which is a negative-definite operator). We use a geometric set of units, in which both the speed of light and the gravitational constant are equal to one, $c = 1 = G$.

1.1. Reduced variables. Hamiltonian formulation of linearized theory

The complete description of the canonical structure of linear gravity on the Kottler background has been given in paper [5] (a generalization of earlier results [3] to the case with cosmological...
constant). We summarize crucial results below. Initial (Cauchy) data for Einstein equation can be represented in the form of two symmetric tensors—the induced metric and the ADM momentum (trace-corrected extrinsic curvature of $\Sigma$):

$$(g_{kl}, p^{kl}), \quad g_{kl} = g_{\mu\nu}|_{\Sigma}, \quad p^{kl} = \sqrt{g}(g^{kl} K - K^{kl}), \quad g := \det g_{kl}. \quad (3)$$

In linearized theory, we consider vacuum solutions of the form $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ and obtain corresponding Cauchy data as perturbations of the point $(\eta_{kl}, p_{kl}(\eta_{\mu\nu}) = 0)$.

The linearized data set $(h_{kl}, p^{kl})$ is constrained by four Gauss–Codazzi constraint equations and partially redundant due to a four-parameter family of gauge transformations. This redundancy can be removed by condensing the data into a set of two pairs of mutually conjugate (like positions and momenta) observables. As a preliminary step, we perform a $(2 + 1)$ decomposition of the ADM data with respect to the geometry of the spherical foliation. The tensors are therefore split into components that behave like scalar, vector-like and tensor objects with respect to the 2D sphere geometry (e.g. $h_{33}, h_{3A}$ and $h_{AB}$). We then extract traces from the 2D tensors:

$$H := \eta^{AB} h_{AB}, \quad \chi_{AB} := h_{AB} - \frac{1}{2} \eta_{AB} H, \quad (4)$$

$$S := \eta^{AB} p_{AB}, \quad S_{AB} := p_{AB} - \frac{1}{2} \eta_{AB} S.$$

Finally, by acting upon tensor and vector-like components with 2D divergence and rotation operators we obtain a set of 12 scalar (and pseudo-scalar) functions, from which we compose our reduced variables:

$$y := 2\Pi^{-1} r^2 p^{3A|B} \varepsilon_{AB}, \quad Y := \Pi(\Delta + 2) h_{3A|B} e^{AB} - \Pi(r^2 \chi_{A|C} e^{AB} 3)|_3, \quad (5)$$

$$x := r^2 \chi_{AB} |_{AB} - \frac{1}{2} (\Delta + 2) H + f B Q,$$

$$X := 2 r^2 S_{AB} |_{AB} + B \Xi,$$

where $\varepsilon_{AB} := \Pi e_{AB}$ is the Levi–Civita tensor on the 2D leaves ($e_{AB}$ is the Levi–Civita symbol, $\epsilon_{12} = 1$) and $\Xi$ and $Q$ denote the following expressions:

$$\Xi := 2 r p^{3A|A} + \overset{\circ}{\Delta} p^{A}, \quad (6)$$

$$Q := 2 h_{3A} + 2 r h_{3A}|_{A} - r H_{3}, \quad (7)$$

and we have introduced a following quasi-local (non-local on individual spheres $S(r)$) operator:

$$B := (\Delta + 2) \left( \Delta + 2 - \frac{6m}{r} \right)^{-1}. \quad (8)$$

The set of four scalar functions $(x, X, y, Y)$:

(a) is gauge invariant,

(b) is no longer restricted by any constraint,

(c) diagonalizes the ADM symplectic structure of the phase space of Cauchy data (i.e. is a set of canonical variables).
(d) carries the entire physical information about the gravitational field, i.e. \((h_{ij}, P^{ij})\) can be uniquely reconstructed up to gauge transformations from this set, (e) reduces field dynamics to a set of four equations of motion:

\[
\begin{align*}
\frac{\partial}{\partial t} \mathbf{x} &= \frac{f}{\Pi} \mathbf{X}, \\
\frac{\partial}{\partial t} \mathbf{X} &= \frac{\Pi}{r^2} \left\{ \left( f r^2 \mathbf{x}_3 \right)_3 + \left[ \Delta + f (1 - 2B) + 1 - r^2 \Delta \right] \mathbf{E} \right\}, \\
\frac{\partial}{\partial t} \mathbf{y} &= \frac{f}{\Pi} \mathbf{Y}, \\
\frac{\partial}{\partial t} \mathbf{Y} &= \frac{\Pi}{r^2} \left\{ \left[ f r^2 (r^2 \mathbf{y})_3 ]_3 + \frac{1}{r^2} \left( \Delta + 2 \right) \mathbf{y} \right\} .
\end{align*}
\]

One subtle point needs to be discussed here: when decomposed into spherical harmonics on \(S(r)\), the monopole and dipole parts of \((\mathbf{x}, \mathbf{X}, \mathbf{y}, \mathbf{Y})\) are not dynamical—they encode conserved charges. There are up to ten charges in the perturbation, depending on the symmetry of the background metric. The monopole part of \(\mathbf{x}\) is the mass (energy) of the perturbation. For \(m = 0\) (pure de Sitter or Minkowski background) the dipole parts of \(\mathbf{X}\) and \(\mathbf{x}\) describe the linear momentum and the static moment (i.e. information about center of mass) respectively, producing six charges altogether. When \(m \neq 0\) we lose the translational symmetry of the background and these charges vanish. Finally, the dipole part of \(\mathbf{y}\) (denoted dip \(\mathbf{y}\)) describes the angular momentum, providing the last three charges. The remaining mono-dipole parts of the canonical variables vanish identically.

In the regime of weak fields, the first seven charges can be easily eliminated, as the splitting of \(g_{\mu\nu}\) into background and perturbation is not unique. By changing the background to a Kottler metric with a different mass parameter and/or acting upon it with a small boost or translation we can modify the value of the charges and shift them from the dynamical field \(h\) to the background \(\eta\). To eliminate the angular momentum, however, one would need to use a background with a non-vanishing angular momentum (Kerr–de Sitter). For the sake of simplicity we keep the spherical symmetry of the background and describe angular momentum on the level of perturbation.

The space of Cauchy data of the complete (non-linear) theory is endowed with the canonical symplectic form \(\Omega\), known as the ADM structure. In case of a bounded region \(\Sigma\) with boundary \(\partial \Sigma\), the symplectic form \(\Omega\) contains not only the ADM bulk term \(\int_{\Sigma} \delta P^{ij} \wedge \delta h_{ij}\), but also an extra boundary term, which makes it gauge invariant (see [6, 8]). It turns out that for linearized theory only the dynamical, ‘mono-dipole-free’ part, denoted by underscored symbols \((\underline{\mathbf{x}}, \underline{\mathbf{X}}, \underline{\mathbf{y}}, \underline{\mathbf{Y}})\)\(^8\) of the initial data remains in the bulk integral. The ADM symplectic structure assumes a canonical form:

\[
\Omega = -\frac{1}{16\pi} \int_{\Sigma} \delta P^{ij} \wedge \delta h_{ij}
= \frac{1}{16\pi} \int_{\Sigma} \delta \mathbf{X} \wedge A \delta \mathbf{x} + \delta \mathbf{Y} \wedge A \delta \mathbf{y} + \text{boundary terms},
\]

\(^8\)By the ‘mono-dipole-free’ part of a function \(q\) we mean the function with its projections onto monopole and dipole harmonics removed, \(\underline{q} := q - \text{mon}(q) - \text{dip}(q)\).
where \( A := \frac{1}{\Delta} \Delta^{-1} (\Delta + 2)^{-1} \) and the boundary terms can be killed by appropriate gauge conditions.

Field dynamics (9), symbolically represented as \( \frac{\partial}{\partial t} \), uniquely defines the gauge-invariant Hamiltonian functional through formula:

\[
\Omega \left( \frac{\partial}{\partial t}, \right) \sim -16\pi \delta H_{\text{inv}} + \int_{\partial \Sigma} \frac{\Pi f}{r} \left[ (r\gamma)_3 A \delta x + (r\gamma)_3 A \delta \tilde{X} \right],
\]

where the symbol \( \sim \) indicates the fact that we omit the gauge-dependent boundary terms. If the values of \( x \) and \( y \) are controlled on the boundary \( (\delta x)|_{\partial \Sigma} = 0 = (\delta y)|_{\partial \Sigma} \) then the system becomes autonomous and is governed by the Hamiltonian \( H_{\text{inv}} \), whose explicit form is:

\[
16\pi H_{\text{inv}} = \frac{1}{2} \int_{\Sigma} \frac{\Pi f}{r^2} \left[ x AX + x AY \right] + \frac{1}{2} \int_{\Sigma} \frac{\Pi f}{r^2} \left[ f(r\gamma)_3 A (r\gamma)_3 + x r^2 V^{(+) A} \right] + \left[ \frac{1}{2} \int_{\Sigma} \frac{\Pi f}{r^2} \left[ f(r\gamma)_3 A (r\gamma)_3 + x r^2 V^{(-) A} \right] \right].
\]

\( V^{(+)} \) and \( V^{(-)} \) are quasi-local, positive-definite potential operators (cf [5]):

\[
V^{(-)} := - \frac{f}{r^2} \left( \Delta + \frac{6m}{r} \right),
\]

\[
V^{(+)} := - \frac{f}{r^2} \left[ (\Delta + 2)^2 \left( \Delta + \frac{6m}{r} \right) + \frac{36m^2}{r^2} \left( \Delta + 2 - \frac{2m}{r} + \frac{2}{3} \right) \right] \left( \Delta + 2 - \frac{6m}{r} \right)^2.
\]

### 1.2. Extrinsic curvature and torsion of a 2D manifold. Rigid sphere condition

For a metric manifold \( M \) with a submanifold \( S \), the extrinsic curvature \( H^\mu_{\ AB} \) of \( S \) in \( M \) is defined by:

\[
H^\mu_{\ AB} X^A Y^B := \mathcal{P}^\mu (\nabla_X Y),
\]

where \( X, Y \in T S, \nabla \) is the metric connection on \( M \) and \( \mathcal{P}^\mu \) denotes the orthogonal projection onto the space of vectors orthogonal to \( S \) (the normal bundle \( NS \)). The trace of \( H^\mu_{\ AB} \) with respect to the intrinsic metric of \( S \) is called the mean extrinsic curvature \( H^\mu \) and is a vector in \( NS \).

If \( \dim NS = 1 \), we identify elements of \( NS \) with scalar functions on \( S \) by choosing a unit normal vector field as a basis in \( NS \).

For \( S \) of codimension two and a spacelike \( H^\mu \) we can find \( T^\mu \in NS \) complementing \( H^\mu / ||H|| \) to an orthonormal basis and define the extrinsic torsion \( t_A \in T^* S \):

\[
t_A X^A := T^\mu \nabla_X \left( \frac{H^\mu}{||H||} \right).
\]

The symbol \( ||H|| \) denotes the length of the vector \( H^\mu \).

If \( S \) is a 2D topological sphere, then it is conformally equivalent to a standard unit sphere \( S^2 \). This conformal equivalence can be used to define a notion of spherical harmonics on \( S \). Consider a sufficiently smooth isomorphism \( \phi : S^2 \rightarrow S \), such that \( \phi^* g_{AB} = p \cdot \eta_{AB} \), where \( \eta_{AB} \) is the standard metric on a unit sphere and \( p \) is a strictly positive conformal factor. The decomposition of functions on \( S \) into a series of spherical multipoles (i.e. eigenfunctions of the operator
Δ) is now defined by their pullback with respect to φ. Imposing a condition dip(ρ) = 0 on φ makes this decomposition unique (whereas the representation of these multipole components in terms of spherical harmonic functions is not unique: it is unique up to rotations).

Using this apparatus, we may now formulate the rigid sphere conditions [13, 14]. A topological sphere S is called a rigid sphere if:

(a) Its mean curvature $H^\nu$ is everywhere spacelike,

(b) The mono-dipole-free parts of the length of $H^\nu$ and of the divergence of extrinsic torsion vanish:

$$\|H\| = 0, \quad \|A^A\| = 0.$$  \hspace{1cm} (16)

This conditions were created to distinguish, in a generic spacetime, a set of spheres which in some way resemble the standard round spheres in Minkowski spacetime and could therefore provide some kind of analogue to the group of Poincaré symmetries. In the Minkowski metric, the solution to the above equations is exactly the eight-parameter family of round spheres. Likewise, an eight-parameter family of solutions exists in a generic spacetime, if it is sufficiently close to the Minkowski metric [14].

2. Relation of the Hawking mass to a volume integral

The Hawking mass [10, 15] of a region surrounded by a topological 2D sphere S is defined by:

$$\mathcal{H}_{\text{Haw}} := \sqrt{\frac{\text{Area } S}{16\pi}} \left( 1 - \frac{1}{16\pi} \int_S \left( H_\nu H^\nu + \frac{4}{3} \Lambda \right) \, da \right),$$ \hspace{1cm} (17)

where $H^\nu$ is the mean extrinsic curvature of the boundary S in the enveloping four-dimensional spacetime. Formula (17) includes a correction term $\frac{4}{3} \Lambda$, which accounts for presence of the cosmological constant.

We wish to compare the Hawking mass, defined in terms of a boundary integral, to the Hamiltonian function of the linearized theory, which is a volume integral of linearized ADM data over the whole region $\Sigma$. We will therefore begin by finding a way to express the Hawking boundary integral in terms of a volume integral of ADM data in the full theory. Such relation can be obtained from the scalar Gauss–Codazzi constraint:

$$g^R - 2\Lambda g = P^i|jP_{ij} - \frac{1}{2}p^2,$$ \hspace{1cm} (18)

where $g := \det g_{ij}$. We also remind the reader that we are working in vacuum, hence no matter density term. The Gauss–Codazzi geometric identities allow one to express the three-dimensional Ricci scalar by objects related to the 2D spherical foliation:

$$\sqrt{g^R} = \sqrt{g(R + k^2 - k_{AB}k^{AB}) + 2\left[ \sqrt{g}(M'k + M'|M'|) \right]},$$ \hspace{1cm} (19)

Here $M'$ is the unit normal vector to the leaves of the 2D spherical foliation, $k_{AB}$ is the extrinsic curvature of these spheres in $\Sigma$, $k := k_{AB}g^{AB}$ is the mean curvature, and $g^{AB}$ is the inverse of $g_{AB}$. $\sqrt{g}, R$ and $\sqrt{g^R}$ are the Ricci scalars of the Cauchy surface and of the leaves of the 2D foliation within. The symbol $|$ denotes the covariant derivative with respect to the three-dimensional geometry of $\Sigma$. 

8
Substituting (19) into (18) and taking into account that $\sqrt{g} = \lambda / \sqrt{g^{33}}$, we can obtain the following equation:

$$
\lambda \tilde{R} - \frac{1}{2} \lambda k^2 - 2 \lambda \Lambda + \frac{\sqrt{g^{33}}}{k} (w'^2)_{\alpha} + \left[ \lambda \tilde{g}^{AB}(\log g^{33})_{,B} \right]_{,A} = \frac{g^{33}}{\lambda} \left( P_{,\alpha} P^{\alpha} - \frac{1}{2} P^2 \right) + \lambda \left( k_{AB} k^{AB} - \frac{1}{2} k^2 + \frac{1}{2} \tilde{g}^{AB}(\log g^{33})_{,A}(\log g^{33})_{,B} \right),
$$

(20)

where $w'^2 := \lambda \tilde{w}'^2$. We will now integrate both sides of this equation over the region $\Sigma$. The last term in the first line will then vanish, since it is a total divergence on each of the 2D spheres of the foliation of $\Sigma$. A similar argument proves that $\int_{\Sigma} \lambda \lambda_3 = \int_{\Sigma} (w'')_3 = \int_{\Sigma} (w')_3$. Using this identity, the Gauss–Bonnet theorem, and the relation $w''_{,\alpha} = -\lambda k / \sqrt{g^{33}}$, one can show that:

$$
\int_{\partial \Sigma} r \left( \lambda \tilde{R} - \frac{1}{2} \lambda k^2 - \frac{2}{3} \lambda \Lambda \right) = \int_{\Sigma} \left[ r \left( \lambda \tilde{R} - \frac{1}{2} \lambda k^2 - \frac{2}{3} \lambda \Lambda \right) \right]_{,3} = \int_{\Sigma} r \lambda \left( \frac{\sqrt{g^{33}}}{k} + \frac{r}{3} \lambda w'' \right)_{,\alpha} \left( \frac{r}{2} + \frac{\sqrt{g^{33}}}{k} \right).
$$

(21)

With this identity, we can replace the volume integral of the first three terms in (20) with a boundary integral. After transferring the remaining volume integrals to the right-hand side, we arrive at the desired relation between a boundary data integral and a volume integral of ADM data:

$$
\int_{\partial \Sigma} r \lambda \left( \lambda \tilde{R} - \frac{1}{2} \lambda k^2 - \frac{2}{3} \lambda \Lambda \right) = - \int_{\Sigma} \left( \frac{\sqrt{g^{33}}}{k} + \frac{r}{3} \lambda w'' \right) \left( k^2 + \frac{4}{3} \lambda \right)_{,\alpha} + \int_{\Sigma} \frac{g^{33}}{\lambda} \left( P_{,\alpha} P^{\alpha} - \frac{1}{2} P^2 \right)
$$

$$
+ \int_{\Sigma} \lambda \left( k_{AB} k^{AB} - \frac{1}{2} k^2 + \frac{1}{2} \tilde{g}^{AB}(\log g^{33})_{,A}(\log g^{33})_{,B} \right).
$$

(22)

The boundary integral on the left-hand side of (22) resembles the definition of the Hawking mass (17), since due to Gauss–Bonnet theorem the expression $\lambda \tilde{R}$ integrates to a constant, $8\pi$, on both spheres constituting the boundary, and the coefficient $r$ can be thought to roughly correspond to the square root of area of the sphere. However, we must make note of one serious discrepancy—instead of $H^2 H_\alpha$, the square of mean curvature of the 2D spheres with respect to the four-dimensional spacetime, in formula (22) we have $k^2$, a square of mean curvature with respect to the 3D Cauchy hypersurface. In fact, the left-hand side of (22) is an expression for the Geroch energy [10], a modified version of the Hawking mass. We could, in principle, further manipulate the above expression to explicitly obtain the Hawking mass integral on the left-hand side. It turns out, however, that it is much simpler to introduce the adjustment to the left-hand side later, after performing the approximation of the right-hand side, since a suitable correction term will then appear by itself.

### 3. The second order approximation

We now perform a second order approximation of the right-hand side of equation (22). After grouping appropriate terms together, the result turns out to be the volume integral for the Hamiltonian function of the linearized theory (12) with a correction from the angular momentum
charge and several boundary terms:
\[
\int_{\partial\Omega} r\lambda \left( \frac{2}{5}k^2 - \frac{2}{3}\Lambda \right) \approx 16\pi H_{\text{int}} + \int_{\Sigma} \frac{\Pi}{\sqrt{-g}} \left[ \text{dip}(y)(-\Delta)^{-1}\text{dip}(y) \right] \quad \text{(square of angular momentum)}
\]
\[
- \frac{1}{2} \int_{\partial\Omega} \frac{rf}{\Pi} (P^3)^2 + \int_{\partial\Omega} \frac{2m}{r^2} (\Pi - \lambda) + \frac{1}{2} \int_{\partial\Omega} \frac{f\Pi}{r} (\lambda A_y + \lambda(\mathbf{B} - 1)A_\mathbf{x})
\]
\[
+ \frac{1}{2} \int_{\partial\Omega} r\lambda A_B \mathbf{x} - \frac{1}{2} \int_{\partial\Omega} \frac{f\Pi}{r} B_{Q}\Delta A \left[ \lambda + \frac{1}{4} \Delta Q - \frac{1}{2} f(\mathbf{B} - 1)Q \right].
\]
\[(23)\]

The calculation is rather straightforward, but tedious in its length. The discussion of techniques used in the derivation of (23) was therefore moved to appendix A. We will instead focus our attention on moulding the above result into the final form. The desired Hamiltonian of linear gravity has already appeared on one side. What remains is to correct the left-hand side to vanish into the Hawking mass integral. Simultaneously, we will examine each of the boundary corrections on the right-hand side to understand their role in the relation and determine whether they can be removed.

3.1. Correcting the mean curvature

We will begin by looking at the first term in the second line of (23): \(-\frac{1}{2} \int_{\partial\Omega} \frac{rf}{\Pi} (P^3)^2\). Note that due to vanishing of the ADM momentum of the background metric, the term \((P^3)^2\) is by itself of second order in the perturbation. Using this and the ADM momentum definition (3), we see that up to second order corrections:

\[
\frac{rf}{\Pi} (P^3)^2 \approx r\Pi(K_{AB}i)^2 \approx r\lambda(K_{AB}i)^2.
\]

We can now transfer this term to the left-hand side of (23), which results in the following expression:

\[
\int_{\partial\Omega} r\lambda \left( \frac{2}{5}k^2 - (K_{AB}i)^2 - \frac{2}{3}\Lambda \right).
\]

\[(25)\]

Note that we have not yet performed an approximation of the left-hand side of (23). All objects in the expression above, except for \((K_{AB}i)^2\), are fully non-linear. We should remember, however, that we are ultimately interested only in the value of this integral up to second order corrections in \(h_{\mu\nu}\). Within this level of precision, a following identity holds:

\[
k^2 - (K_{AB}i)^2 \approx H^\mu H_\mu,
\]

which allows us to rewrite (25) as:

\[
\int_{\partial\Omega} r\lambda \left( \frac{2}{5}k^2 - (K_{AB}i)^2 - \frac{2}{3}\Lambda \right) \approx \int_{\partial\Omega} \frac{r}{2} \lambda \left( \frac{2}{5}k^2 - (K_{AB}i)^2 - \frac{2}{3}\Lambda \right).
\]

\[(27)\]

3.2. The radial coordinate

Next, we would like to address the second term in the second line of (23): \(\int_{\partial\Omega} \frac{2m}{r^2} (\Pi - \lambda)\). It is a difference between boundary integrals of volume densities defined by the full and the background metric. Since the volume density \(\lambda\) is approximated by:
\[ \lambda = \sqrt{\det g_{AB}} \approx \Pi \left( 1 + \frac{1}{2} H - \frac{1}{4} \chi_{AB} \chi^{AB} \right), \]  

(28)

this term is actually a combination of first and second order corrections.

To obtain an equality between the Hawking mass and \( H_{\text{inv}} \) we need for this term to vanish. This could be achieved simply by treating it as a boundary condition on the perturbation. However, one may also look at it from a different perspective. Let us return for a moment to the basic conceptual construction of the linearized theory. We may either start from the background metric and introduce some perturbation field on it, or go the other way around—begin with a ‘perturbed’ metric and try to compare it with a ‘background’ spacetime by means of an appropriate diffeomorphism. The second approach, although a bit more convoluted, strongly highlights the amount of freedom we possess when aligning the two metrics. Being given the perturbed Cauchy surface \( \Sigma \) with its intrinsic topologically spherical foliation, we are not only able to choose the mass and cosmological constant parameters of the background, but also the precise way in which we identify the points of \( \Sigma \) with the points in the ‘reference’ spacetime. Although some of this freedom is already fixed, since we wish to identify the leafs of the spherical foliation of \( \Sigma \) with the natural spherical foliation in the Kottler metric, we still have some leeway left in the way we pair the leafs of the foliations. This pairing is naturally represented by the labelling of the 2D spheres in \( \Sigma \) by the radial coordinate. Different choices of pairings will in turn result in monotonous rescaling of this coordinate. We may use this freedom to demand that on the two boundary spheres \( r \) is equal to the areal radius. Since the radial coordinate we use in Kottler is the areal radius, by choosing the identification of boundary spheres in this way we obtain the desired equality:

\[ r_{0,1} = \sqrt{\int_{S(0,1)} \frac{\Pi}{4\pi}} = \sqrt{\int_{S(0,1)} \frac{\lambda}{4\pi}} \implies \int_{\partial \Sigma} (\lambda - \Pi) = 0. \]  

(29)

We note that both approaches to the construction of the linearization lead to the same final result. They just differ in the way in which we interpret equation (29)—we may either view it as a constraint on the perturbation, or as a recipe for the identification of the region boundaries between the two spacetimes, as shown in the graphic.
Finally, we need to observe that fulfilling (29) has another important consequence. If we multiply this equation by a factor of $1/2$, we obtain:

$$\frac{1}{2}r_{0,1} = \frac{1}{2} \sqrt{\int_{S(r_{0,1})} \frac{\lambda}{4\pi}} = \sqrt{\frac{\text{Area}(r_{0,1})}{16\pi}}, \quad (30)$$

which is exactly the relation required to turn the left-hand side of (23), in its corrected form (27) into the Hawking mass expression (17). In fact, we could also start from the demand that expression (27) become the Hawking mass integral, deduce condition (29) from this and obtain the vanishing of the right-hand side correction term as a consequence.

3.3. Final result

With these observations and assumption (29), we may rewrite (23) in a cleaned up, final form:

$$16\pi H_{\text{Haw}} \approx 16\pi H_{\text{Inv}} + \int_{\Sigma} \frac{\Pi}{2r} \text{dip}(y) (-\Delta)^{-1} \text{dip}(y)$$

$$+ \frac{1}{2} \int_{\partial \Sigma} f \frac{\Pi}{r} (\lambda_{\mathbf{x}} A_{\mathbf{x}} + \lambda_{\mathbf{y}} (B - 1) A_{\mathbf{y}})$$

$$+ \frac{1}{2} \int_{\partial \Sigma} r f \frac{\Pi}{r} A_{\mathbf{y}} B_{\mathbf{x}} = -\frac{f \Pi}{r} R Q \Delta A \left[ x + \frac{1}{4} \Delta Q - \frac{1}{2} f (B - 1) Q \right].$$

(31)

We would now like to interpret the remaining boundary integrals on the right-hand side. The quadratic expressions in $\mathbf{x}$ and $\mathbf{y}$ are constant if we control the value of these ‘true degrees of freedom’ at the boundary and can then be neglected: once we calculate the variation of (31), these expressions will turn into terms of the form $\mathbf{y} \delta \mathbf{x}$ and $\mathbf{y} \delta \mathbf{y}$, which vanish if $\delta \mathbf{x}|_{\partial \Sigma} = 0 = \delta \mathbf{y}|_{\partial \Sigma}$. This is, in fact, the same boundary condition that appeared during the derivation of $H_{\text{Inv}}$, cf (11). Mathematically, controlling the Dirichlet data on $\partial \Sigma$ for the ‘true degrees of freedom’ is necessary whenever we want to describe field evolution within the domain $\Sigma$ as a Hamiltonian system (see [6, 8, 12, 16]). Physically, such control defines an adiabatic insulation of the physical system we want to describe (i.e. the field contained in the interior of $\Sigma$) from the ‘rest of the World’. This belongs to the standard repertoire of the Hamiltonian field theory.

The meaning of the last line in (31) is a bit more puzzling. The quantities $\xi$ and $Q$ are gauge dependent and the whole integral could be just brushed off as some artefact of the gauge freedom of linearized theory. However, they might actually posses a deeper meaning. If we look at the first order approximations of the mean curvature vector length and the divergence of extrinsic torsion:

$$\|H\| \approx \frac{2\sqrt{T}}{r} - \frac{\sqrt{T}}{2r} Q \approx \frac{2\sqrt{T}}{r} + \delta(\|H\|),$$

$$t^{\|A\|} \approx \frac{1}{2\Pi r} \xi \approx \delta(t^{\|A\|}),$$

(32)

(33)

we observe that the last line of (31) could actually be rewritten in terms of perturbations of these functions:
\[
+2 \int_{\partial \Sigma} r^3 f \Pi (t_{A} \not \parallel A)_{AB} \delta (t_{A} \not \parallel A) \\
-2 \int_{\partial \Sigma} r \Pi (\not \parallel H) \Delta_{AB} \left[ \left( \frac{1}{4} \Delta - \frac{1}{2} f (B - 1) \right) \delta (\not \parallel H) - \sqrt{f \not \parallel r} \right].
\]

(34)

We remind the reader that these are the same functions that are used to define the rigid sphere conditions (16). The spheres \(S(r)\) in the Kottler metric satisfy these conditions. Therefore, the gauge choice forced upon us by the last line in (31) can be interpreted as a linearized rigid sphere condition. We stress that it arises spontaneously in our calculation, as a result of demanding that the approximation of the Hawking mass and the linear gravity Hamiltonian coincide.

The gauge choice can also be interpreted as a condition imposed on the vector field \(T = \frac{\partial}{\partial t}\), i.e. on the reference frame. As noted in the introduction, a Hamiltonian functional generating evolution with respect to a generic reference frame does not have the properties which we expect from a true energy, namely positivity, convexity etc. But the quadratic expression \(H_{\text{Haw}}\) does fulfill these properties! Thus, the Hawking mass measured on a rigid sphere is positive and convex, at least in the regime of weak fields, when its quadratic approximation prevails. It represents the field energy measured with respect to a reference frame \(T\) satisfying the condition: \(g_{\mu \nu} T^\mu T^\nu = 0\). If the boundary of \(\Sigma\) does not satisfy the rigidity condition \(\Xi = 0 = \Omega\), the corresponding quantity \(H_{\text{Haw}}\) is not necessarily positive even in the weak field regime and cannot be identified with the field energy contained in \(\Sigma\) because the corresponding field \(T\) cannot be treated as a ‘time translation’ in any reasonable sense.

4. Conclusions

We used our formalism of reduced variables to investigate the behaviour of Hawking mass for weak gravitational perturbations of the Kottler metric. It turns out that quadratic approximation of the quasi-local mass is related to the gauge-invariant Hamiltonian (the generator of dynamics of linear theory) provided that the boundary of the considered region is composed of spheres satisfying a certain gauge condition. This condition is a linearization of the ‘rigid sphere condition’ from the complete, non-linear theory. The rigid sphere condition, in turn, characterizes spheres which provide a physically reasonable ‘reference frame’ for defining time translations of a region in spacetime—which is necessary to talk about energy. This observation supports our hypothesis that the quasi-local energy can be reasonably defined only for 2D surfaces \(\partial \Sigma\) which satisfy extra conditions—those for which such time translations are well defined.

The relation (31) contains an extra bulk term: the ‘square of angular momentum’. But, due to the vector constraint, this term can be rewritten as a boundary term, see [17].

Acknowledgments

This work was supported in part by Narodowe Centrum Nauki (Poland) under Grant No. 2016/21/B/ST1/00940. One of the authors (PW) was also supported by a special internal Grant for young researchers, provided by Center for Theoretical Physics, PAS, Warsaw, Poland.

Data availability statement

No new data were created or analysed in this study.
Appendix A. Performing the second order approximation

Deriving the approximation of the right-hand side of equation (22) is a lengthy calculation, consisting mostly of simple algebraic manipulations. Therefore, we do not present it in detail. Instead, we will just outline the crucial points of the procedure.

To briefly summarize the calculation: we begin by replacing the objects under the integral on the right-hand side of (22) with their expansions in terms of the perturbation, then trim the results so only corrections up to second order remain, and finally express as much of the outcome as possible in terms of our reduced variables \((x, X, y, Y)\) and auxiliary objects \(\Xi\) and \(Q\) (as can be seen in the second row of (23), some leftover terms will remain). We also use the emergent radial derivatives to change volume integrals into boundary ones.

The first two parts of the task are greatly simplified by the way in which the terms of equation (22) are already arranged. Observe that the second and third row is comprised of squares of objects which vanish for the unperturbed background metric: \(P_{\alpha\beta}, P, \partial_A(\log g^{\alpha\beta})\), and finally \(k_{AB} - \frac{1}{2}k\eta_{AB}\). This means that the first terms in the perturbative expansion of these rows will already be of second order and all we need to do is just replace the above expressions with their first order approximations. Unfortunately, the product of brackets in the first row of (22) does not possess this useful property and to calculate its approximation we need to compute second order corrections of its factors. Still, even in this case some effort can be saved by performing the calculations in an appropriate order and noticing that the left bracket in the product is again an object which vanishes for the unperturbed background.

The next step is to rewrite the result in terms of the gauge-invariant variables. For this, we must perform a \((2 + 1)\) decomposition in a manner similar to the original construction of the reduced variables in section 1.1. The less obvious part is perhaps generating the 2D rotations and divergences of the tensors. This can be done by employing the following integral identities:

For two vector fields \(v^A\) and \(w^A\), tangent to a 2D sphere \(S(r)\), the following holds:

\[
\int_{S(r)} v^A w^A = \int_{S(r)} \left( (r v^A)^{\parallel A} \Delta^{-1} (r w^A) - (r v^A)^{AB} \epsilon_{AB}^{\parallel A} \Delta^{-1} (r w^A)^{\parallel AB} \right). \tag{35}
\]

Similarly, for a pair of traceless, symmetric tensor fields \(p_{AB}, t_{AB}\) on a 2D sphere:

\[
\int_{S(r)} t_{AB} p^{AB} = \int_{S(r)} 2(r^2 t_{AB})^{\parallel AB} A (r^2 p_{AB}^{\parallel AB}) + 2(r^2 \epsilon^{AC} t_A^{\parallel BC}) A (r^2 \epsilon^{AC} p_A^{\parallel BC}). \tag{36}
\]

Once the decomposition has been performed, we may use the system of linear equations consisting of definitions of gauge-invariant variables, definitions of auxiliary variables \(Q\) and \(\Xi\), and the set of linearized Gauss–Codazzi constraints (cf [5]):

\[
\begin{align*}
\frac{1}{\sqrt{f}} (\sqrt{f} P^3)_{,3} + P^A_{,A} - r^{-1} S &= 0, \\
(r^2 P^{3A})_{,3} + r^2 S^{AB}_{,AB} + \frac{1}{2} \Delta S &= 0, \tag{37} \\
(r^2 P^{3B})_{,3} + r^2 S^{AB}_{,AB} &= 0, \\
f^B Q - N (r^2 N)_{,3} + (\Delta + 2f - r f') h^3_{,3} - x &= 0,
\end{align*}
\]

to replace the ADM data components by the gauge-invariant variables through algebraic substitution. Applying the linearized Gauss–Codazzi constraints will produce some radial derivatives.
in the approximated formula. As mentioned earlier, these should be combined into total diver-
gences in the radial direction and turned into boundary integrals, until the remaining terms in
the volume integral contain either no radial derivatives or squares of them (which become part
of the Hamiltonian of the linearized theory (12)). This concludes the derivation.

ORCID iDs

Jacek Jezierski  https://orcid.org/0000-0003-3306-9453
Jerzy Kijowski  https://orcid.org/0000-0001-7652-6744
Piotr Waluk  https://orcid.org/0000-0003-0245-6987

References

[1] Brill D R and Deser S 1968 Variational methods and positive energy in general relativity Ann. Phys.,
NY 50 548–70
[2] Jezierski J and Kijowski J 1990 The localization of energy in gauge field theories and in linear
gravitation Gen. Relativ. Gravit. 22 1283–307
[3] Jezierski J 1999 Energy and angular momentum of the weak gravitational waves on the
Schwarzschild background–quasilocal gauge-invariant formulation Gen. Relativ. Gravit. 31
1855–90
[4] Waluk P and Jezierski J 2017 Degrees of freedom of weak gravitational field on a spherically
symmetric background Acta Phys. Pol. B Proc. Suppl. 10 391–5
[5] Waluk P and Jezierski J 2019 Gauge-invariant description of weak gravitational field on a spherically
symmetric background with cosmological constant Class. Quantum Grav. 36 215006
[6] Kijowski J 1997 A simple derivation of canonical structure and quasilocal Hamiltonians in general
relativity Gen. Relativ. Gravit. 29 307–43
[7] Jezierski J, Kijowski J and Wiatr M 2020 Localizing energy in Fierz–Lanczos theory Phys. Rev. D
102 024015
[8] Chruściel P T, Jezierski J and Kijowski J 2002 Hamiltonian Field Theory in the Radiating Regime
(Lecture Notes in Physics vol 70) (Berlin: Springer)
[9] Penrose R 1982 Quasilocal mass and angular momentum in general relativity Proc. R. Soc. A 381
53–63
[10] Szabados L 2009 Quasilocal energy–momentum and angular momentum in GR Living Rev. Relat.
12 4
[11] Abreu G and Visser M 2010 Tolman mass, generalized surface gravity, and entropy bounds Phys. Rev. Lett. 105 041302
[12] Deser S 2019 Energy in gravitation and Noether’s theorems J. Phys. A: Math. Theor. 52 381001
[13] Gittel H-P, Jezierski J, Kijowski J and Łe¸ski S 2013 Rigid spheres in Riemannian spaces Class. Quantum Grav. 30 175019
[14] Gittel H-P, Jezierski J and Kijowski J 2016 On the existence of rigid spheres in four-dimensional
spacetime manifolds Vietnam J. Math. 44 231–49
[15] Hawking S W 1968 Gravitational radiation in an expanding universe J. Math. Phys. 9 598–604
[16] Kijowski J and Tulczyjew W M 1979 A Symplectic Framework for Field Theories (Lecture Notes in Physics vol 107) (Berlin: Springer)
[17] Jezierski J and Smolka T 2017 Examination of quasi-local mass for asymptotically Kerr spacetimes
Acta Phys. Pol. B Proc. Suppl. 10 397