

Gauge invariant derivation of zeroth and first laws of black hole thermodynamics

Kamal Hajian,^{1,2,*} M. M. Sheikh-Jabbari^{3,†} and Bayram Tekin^{1,‡}

¹*Department of Physics, Middle East Technical University, 06800 Ankara, Turkey*

²*Institute of Physics, University of Oldenburg, Postfach 2503, D-26111 Oldenburg, Germany*

³*School of Physics, Institute for Research in Fundamental Sciences (IPM),
P.O. Box 19395-5531, Tehran, Iran*



(Received 14 September 2022; accepted 21 October 2022; published 15 November 2022)

In gauge invariant theories, like Einstein-Maxwell theory, physical observables should be gauge invariant. In particular, mass, entropy, angular momentum, electric charge, and their respective chemical potentials, temperature, horizon angular velocity, and electric potential, which appear in the laws of black hole thermodynamics, should be gauge invariant. In the usual construction of the laws of black hole thermodynamics, gauge invariance of the intensive quantities is subtle; here, we remedy this and provide a gauge invariant derivation and the proof of the zeroth and first laws of black hole thermodynamics.

DOI: [10.1103/PhysRevD.106.104030](https://doi.org/10.1103/PhysRevD.106.104030)

I. INTRODUCTION

Symmetries have been the main guiding principle in formulating laws of physics. They help us identify conserved quantities, through the celebrated Noether's theorem, as well as formulate interactions, through the notion of local gauge symmetries. Maxwell's theory was the first such example, which was realized to be a gauge theory with the $U(1)$ gauge symmetry. This symmetry appeared to be the key to formulating quantum electrodynamics (QED). An extension of the same notion to other gauge groups laid the foundation of what then led to the standard model of particle physics. In a parallel line of development, Einstein's equivalence principle was formulated through diffeomorphism invariance of general relativity (GR), which is a local gauge symmetry under general coordinate transformations. Gauge symmetries introduce nonpropagating and generically nonphysical degrees of freedom into the description; nonetheless, all physical observables ought to be gauge invariant.

Defining (conserved) charges has been one of the outstanding questions in the context of GR, as the theory is not only diffeomorphism invariant, but also background independent. There have been several proposals to do so starting

with Komar [1] in 1958, followed by Arnowitt-Deser-Misner [2–4], Regge-Teitelboim Hamiltonian method [5], Brown-York [6], the ADT formulation [7–9], and more recently [10–12], to just name a few. Each of these formulations have their own advantages and disadvantages; see chapter 5 of [13] for a recent review. In particular, the covariant phase space formalism (CPSF), which was developed by many in particular by Wald and collaborators [14–19], has provided a suitable framework to compute diffeomorphism invariant conserved charges associated with spacetime symmetries.

Black holes are ubiquitous solutions to theories of gravity and have raised many questions at conceptual, theoretical, and observational level since the conception of Einstein's GR in 1915. After seminal works of Bekenstein, Hawking, and many others since early 1970s, it is now well accepted that black holes satisfy laws of thermodynamics. The idealized stationary black holes are classically in thermal equilibrium at a given temperature with an entropy and other charges and associated chemical potentials. There have been many works on proving laws of black hole thermodynamics from the first principles, starting from diffeomorphism invariant theories of gravity. The first steps were taken by Hawking in proving the constancy of surface gravity on the horizon of a stationary black hole as the zeroth law and the area theorem as a statement of the second law; see Ref. [13] for a detailed discussion and review. The next breakthroughs came by the seminal works of Wald, where it was shown that the entropy is the conserved charge à la Noether, associated with a Killing vector generating the horizon [20], and then a derivation of the first law was given as a result of diffeomorphism invariance of the theory [15].

*khajian@metu.edu.tr, kamalhajian@uni-oldenburg.de

†jabbari@theory.ipm.ac.ir

‡btekin@metu.edu.tr

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

Wald's derivations were based on CPSF, and despite the elegance and depth, the original analysis had some loose ends; e.g., it relied on notions of mass and angular momentum at asymptotically flat spacetimes, and the electric charge was not included. Some of these points were improved and addressed in later works. In particular, it was noted in [21] that for stationary black holes with a Killing horizon, one can relax the dependence on the asymptotic behavior of the spacetime geometry and define the first law at the horizon or any other codimension 2 slice surrounding the horizon.

In this work, we address another loose end of Wald's derivation: It is not manifestly gauge invariant. The early steps in this direction were taken in [21–26]; see also Ref. [13]. In this work, we provide a full proof of gauge invariance of each and every thermodynamical quantity appearing in the first and zeroth laws of black hole thermodynamics.

II. THE SETUP AND CONVENTIONS

Consider a generic D dimensional gravitational scalar-Maxwell theory described by the action

$$I = \int d^D x \sqrt{-g} \left(\mathcal{L}_0(g_{\mu\nu}, \phi_a, \dots) + \frac{1}{4} \sigma(\phi_a) F_{\mu\nu} F^{\mu\nu} \right) \\ = \int \left(\mathbf{L}_0 + \frac{1}{2} (\sigma \star F) \wedge F \right), \quad (1)$$

in which the dynamical fields are the metric $g_{\mu\nu}$, the gauge field A , scalar fields ϕ_a , and ellipsis refer to other possible dynamical fields which do not directly couple to A . $\mathbf{L}_0 = \star \mathcal{L}_0$ is a generic covariant gravitational Lagrangian, and $\sigma(\phi_a)$ is an arbitrary differentiable function of some scalar fields ϕ_a , which couples them to the electromagnetic field. In the second equality, we have adopted form notation of [15], where Lagrangian \mathbf{L}_0 is a D form, $F = dA$ is the $U(1)$ gauge field strength, and $F_{\mu\nu}$ denotes its components. $\star F$ is its Hodge dual, which is a $D - 2$ form. We are using the conventions in which the speed of light, Planck's constant, and Newton's constant are all set to one. These may be recovered upon dimensional analysis, if needed. For our analysis below, we need the explicit form of the gauge field equations of motion:

$$d(\sigma \star F) = 0. \quad (2)$$

The above action is invariant under diffeomorphisms $x^\mu \rightarrow x^\mu - \xi^\mu(x)$ as well as the gauge transformation $A \rightarrow A + d\lambda$. Under a diffeomorphism, a generic field X transforms as $X \rightarrow X + \mathcal{L}_\xi X$, and under a gauge transformation, only the gauge field transforms. We collectively denote the diffeomorphism and gauge generators by $\eta = \{\xi, \lambda\}$.

A. Charge computation

Covariant phase space formulation of charges [14–19] is a method that facilitates our discussion on the gauge invariance of charges (see reviews in [13,27,28]). Here, we recapitulate the basics of this method. The formulation begins with a Lagrangian density \mathcal{L} in D dimensions and the action $I = \int \mathbf{L} \equiv \int \star \mathcal{L}$ with some dynamical fields $\Phi(x^\mu)$. Variation of the Lagrangian reads

$$\delta \mathbf{L} = \mathbf{E}_\Phi \delta \Phi + d\Theta, \quad (3)$$

where $\mathbf{E}_\Phi = 0$ are the equations of motion. We will use \approx to denote on-shell equality, e.g., $\delta \mathbf{L} \approx d\Theta$. By another variation of the symplectic potential $\Theta(\delta \Phi, \Phi)$, we obtain the symplectic current ω

$$\omega(\delta_1 \Phi, \delta_2 \Phi, \Phi) = \delta_1 \Theta(\delta_2 \Phi, \Phi) - \delta_2 \Theta(\delta_1 \Phi, \Phi). \quad (4)$$

Since $\delta^2 \mathbf{L} = 0$ identically, recalling the identity $\delta d = d\delta$ and assuming the linearized field equations $\delta \mathbf{E}_\Phi \approx 0$, the symplectic current is locally conserved on-shell $d\omega \approx 0$ and by virtue of Poincaré Lemma, $\omega \approx dk(\delta_1 \Phi, \delta_2 \Phi, \Phi)$.

For a generic diffeomorphism and gauge transformation $\eta = \{\xi, \lambda\}$, a charge variation can be defined by

$$\delta Q_\eta = \int_\Sigma \omega(\delta \Phi, \delta_\eta \Phi, \Phi) \\ = \int_\Sigma (\delta \Theta(\delta_\eta \Phi, \Phi) - \delta_\eta \Theta(\delta \Phi, \Phi)) \quad (5)$$

$$= \oint_{\partial \Sigma} k_\eta(\delta \Phi, \Phi), \quad (6)$$

where Σ is a $D - 1$ dimensional spacelike hypersurface (a Cauchy surface), $\partial \Sigma$ is the $D - 2$ dimensional boundary of Σ , $\delta_\eta \Phi = \{\mathcal{L}_\xi \Phi, \delta_\lambda \Phi\}$, and $dk_\eta(\delta \Phi, \Phi) \approx \omega(\delta \Phi, \delta_\eta \Phi, \Phi)$. Note that each term in a Lagrangian $\mathbf{L} = \mathbf{L}_1 + \mathbf{L}_2 + \mathbf{L}_3 + \dots$ contributes to Θ , ω , and k_η , in an additive way, i.e., $\Theta = \Theta_1 + \Theta_2 + \Theta_3 + \dots$, $\omega = \omega_1 + \omega_2 + \omega_3 + \dots$, and $k_\eta = k_{1\eta} + k_{2\eta} + k_{3\eta} + \dots$.

B. Integrability of the charge

The charge variation δQ_η is not necessarily a variation of a charge Q_η on the phase space. The charge variation is integrable if $(\delta_1 \delta_2 - \delta_2 \delta_1) Q_\eta = 0$. In general, if the symmetry generators $\eta = \{\xi, \lambda\}$ are also field dependent, as is in the cases of our interest, this condition takes the form [13,21,29]

$$\oint_{\partial \Sigma} \xi \cdot \omega(\delta_1 \Phi, \delta_2 \Phi, \Phi) + k_{\delta_1 \eta}(\delta_2 \Phi, \Phi) - k_{\delta_2 \eta}(\delta_1 \Phi, \Phi) \approx 0. \quad (7)$$

III. GAUGE NONINVARIANCE OF CHARGE

While δQ_η in (6) is by construction diffeomorphism invariant $\mathcal{L}_{\tilde{\xi}}(\delta Q_\eta) = 0$ for a generic $\tilde{\xi}$, it is not invariant under a generic gauge transformation $A \rightarrow A + d\tilde{\lambda}$, $\delta_{\tilde{\lambda}}(\delta Q_\eta) \neq 0$. To see this, consider the Lagrangian in (1); i.e., $\mathbf{L} = \mathbf{L}_0 + \mathbf{L}_A$ where $\mathbf{L}_0 = \mathbf{L}_0(g_{\mu\nu}, \phi_a, \dots)$ and $\mathbf{L}_A = \frac{1}{2}(\sigma(\phi_a) \star F) \wedge F$. For this Lagrangian, $\mathbf{k}_\eta = \mathbf{k}_{0\eta} + \mathbf{k}_{A\eta}$. It is readily seen that $\mathbf{k}_{0\eta}$ is invariant under the $\tilde{\lambda}$ gauge transformation, and we hence focus on $\mathbf{k}_{A\eta}$. Starting from \mathbf{L}_A , its symplectic potential is found to be $\Theta_A(\delta A) = (\sigma \star F) \wedge \delta A$. Then

$$\begin{aligned} \omega_A(\delta, \delta_\eta) &= \delta((\sigma \star F) \wedge \delta_\eta A) - \delta_\eta((\sigma \star F) \wedge \delta A) \\ &= \delta((\sigma \star F) \wedge (\mathcal{L}_\xi A + d\lambda)) \\ &\quad - \mathcal{L}_\xi((\sigma \star F) \wedge \delta A). \end{aligned} \quad (8)$$

In the last equation, we replaced $\delta_\eta A = \mathcal{L}_\xi A + d\lambda$ and used $\delta_\lambda \delta A = 0$ as $\delta_\lambda(A_2 - A_1) = (A_2 + d\lambda) - (A_1 + d\lambda) - (A_2 - A_1) = 0$. Using the Cartan identity $\mathcal{L}_\xi X = \xi \cdot dX + d(\xi \cdot X)$ for an arbitrary form X , we obtain

$$\begin{aligned} \omega_A(\delta, \delta_\eta) &= \delta((\sigma \star F) \wedge (\xi \cdot dA + d(\xi \cdot A + \lambda))) \\ &\quad - \xi \cdot d((\sigma \star F) \wedge \delta A) - d(\xi \cdot (\sigma \star F) \wedge \delta A) \\ &= d[\delta(\sigma \star F \wedge (\xi \cdot A + \lambda)) - \xi \cdot (\sigma \star F \wedge \delta A)], \end{aligned} \quad (9)$$

where we used (2), $\delta d = d\delta$ and $\delta(\sigma \star F \wedge (\xi \cdot dA)) \approx \xi \cdot d(\sigma \star F \wedge \delta A)$ to obtain (9). So, we find $\omega_A(\delta, \delta_\eta) = d\mathbf{k}_{A\eta}$, where

$$\mathbf{k}_{A\eta} = \delta(\sigma \star F \wedge (\xi \cdot A + \lambda)) - \xi \cdot (\sigma \star F \wedge \delta A). \quad (10)$$

While the second term in (10) is gauge invariant, the first term is not: $\delta_{\tilde{\lambda}}(\mathbf{k}_{A\eta}) = \delta(\sigma(\xi \cdot d\tilde{\lambda}) \star F) \neq 0$. Note that the above analysis on gauge noninvariance of the charge variation also works for field-dependent gauge transformations when $\delta\lambda \neq 0$.

Although (10) is not gauge invariant for an arbitrary η , but if we augment a diffeomorphism ξ by the specific gauge transformation,

$$\lambda = -\xi \cdot A, \quad (11)$$

then the first term in (10) vanishes, and the result is manifestly gauge invariant. Explicitly, this happens because with this choice for λ ,

$$\delta_\eta A = \mathcal{L}_\xi A - d(\xi \cdot A) = \xi \cdot F, \quad F = dA, \quad (12)$$

is manifestly gauge invariant. So, δQ_η for the “augmented diffeomorphism” $\eta = \{\xi, -\xi \cdot A\}$ is both gauge and diffeomorphism invariant. This augmented diffeomorphism was first introduced by E. Bessel-Hagen in his work [30] and is

repeatedly used in the literature (e.g., [31]). It was dubbed “gauge covariant Lie derivative” (see, e.g., [32,33]) or “improved diffeomorphisms” [34,35]. See also discussions and references in [25]. We use “augmented diffeomorphism” in order to prevent confusion with a different procedure that is called “improved energy-momentum tensor” (see Refs. [35–37] to find more on their difference).

In what follows, we use η as the augmented diffeomorphism. We note that the CPSF and the charge variation work for a generic field-dependent diffeomorphisms or gauge transformations when $\delta\eta \neq 0$ [13]. The above readily generalizes to the cases with several $U(1)$ gauge fields $A^{(a)}$, as $\eta = (\xi, -\xi \cdot A^{(a)})$. We note that the generalized symmetry generators $\eta = \{\xi, -\xi \cdot A\}$ do not form an algebra, as $[\eta_1, \eta_2] = \{[\xi_1, \xi_2]_{\text{L.B.}}, -[\xi_1, \xi_2]_{\text{L.B.}} \cdot A + 2\xi_1 \cdot (\xi_2 \cdot F)\}$. Nonetheless, if ξ_1, ξ_2 are Killing vectors, $\xi_1 \cdot (\xi_2 \cdot F) = 0$, and we retain an algebra structure.

IV. GAUGE INVARIANT BLACK HOLE CHARGES

As the first step toward gauge invariant first law, we define gauge invariant black hole charges. Consider a generic stationary axisymmetric black hole solution to (1), with stationarity Killing vector (KV) ξ_M and axisymmetry KVs ξ_{J_i} . Assume that this black hole has a nondegenerate Killing horizon H generated by KV ξ_H and let κ and Ω_H^i , respectively, denote surface gravity and horizon angular velocities. For a stationary black hole, $\xi_H = \xi_M - \Omega_H^i \xi_{J_i}$. We can then define the augmented isometries,

$$\begin{aligned} \eta_S &= \frac{2\pi}{\kappa} \{\xi_H, -\xi_H \cdot A\}, & \eta_M &= \{\xi_M, -\xi_M \cdot A\}, \\ \eta_{J_i} &= \{\xi_{J_i}, -\xi_{J_i} \cdot A\}, \end{aligned} \quad (13)$$

where A is the gauge field configuration corresponding to the black hole solution. The above symmetry generators have hence both field dependence and parametric dependence (through κ, Ω_i) [13,21]. In particular, we note that

$$\eta_S = \frac{2\pi}{\kappa} \left(\eta_M - \sum_i \Omega_H^i \eta_{J_i} \right), \quad (14)$$

which was already mentioned and used in Sec. 5.4.3 of [13].

As usual, we can define the charge variations as

$$\begin{aligned} \delta M &:= \oint_{S_\infty} \mathbf{k}_{\eta_M}(\delta\Phi, \Phi), & \delta J_i &:= \oint_{S_\infty} \mathbf{k}_{\eta_{J_i}}(\delta\Phi, \Phi), \\ \delta S &:= \oint_B \mathbf{k}_{\eta_S}(\delta\Phi, \Phi), \end{aligned} \quad (15)$$

where B is the bifurcation horizon and S_∞ denotes a constant time slice at the spatial infinity i^0 . We note that M, J_i are the usual mass and angular momenta, and S is the

Wald entropy. These charges as shown in [15,20] are integrable. The virtue of our new definition, using η instead of ξ , is that these charges are now gauge invariant by construction.

V. GAUGE INVARIANT FIRST LAW

Consider the symplectic current $\omega(\delta\Phi, \delta_{\hat{\eta}}\Phi; \Phi)$ where $\hat{\eta} = \{\xi_H, 0\}$. Integrate the charge expression over a (partial) Cauchy surface Σ with two end points (codimension 2 surfaces) B and S_∞ , and recall that $d\omega \approx 0$, to find

$$\int_{\Sigma} \omega(\delta\Phi, \delta_{\hat{\eta}}\Phi; \Phi) = \oint_{S_\infty} k_{\hat{\eta}}(\delta\Phi; \Phi) - \oint_B k_{\hat{\eta}}(\delta\Phi; \Phi). \quad (16)$$

Next, recall that

$$\hat{\eta} = \{\xi_M, -\xi_M \cdot A\} - \Omega_H^i \{\xi_{J_i}, -\xi_{J_i} \cdot A\} - \{0, -\xi_H \cdot A\} \quad (17a)$$

$$= \eta_M - \Omega_H^i \eta_{J_i} - \{0, -\xi_H \cdot A\} \quad (17b)$$

$$= \frac{\kappa}{2\pi} \eta_S - \{0, -\xi_H \cdot A\}, \quad (17c)$$

where in the last line, we used (14).

We now calculate the left- and right-hand sides of (16). For the lhs, we note that ξ_H is a KV; hence, δ_{Ham} acting on $g_{\mu\nu}$ and the scalar fields ϕ_a vanish. Therefore, only the Maxwell part of the action contributes, cf. (8),

$$\omega(\delta\Phi, \delta_{\hat{\eta}}\Phi; \Phi) \approx \delta((\sigma \star F) \wedge (\mathcal{L}_{\xi_H} A)) - \mathcal{L}_{\xi_H}((\sigma \star F) \wedge \delta A). \quad (18)$$

Using $\mathcal{L}_{\xi_H}(\sigma \star F) = 0$ and $\mathcal{L}_{\xi_H} \delta A = \delta \mathcal{L}_{\xi_H} A$, we find

$$\int_{\Sigma} \omega(\delta\Phi, \delta_{\hat{\eta}}\Phi; \Phi) \approx \int_{\Sigma} \delta(\sigma \star F) \wedge (\mathcal{L}_{\xi_H} A). \quad (19)$$

The rhs of (16) can be computed using (17b) and (17c) in the first and second terms, respectively,

$$\begin{aligned} & \oint_{S_\infty} k_{\hat{\eta}}(\delta\Phi; \Phi) - \oint_B k_{\hat{\eta}}(\delta\Phi; \Phi) \\ &= \delta M - \Omega_H^i \delta J_i - \frac{\kappa}{2\pi} \delta S - \oint_{S_\infty} \delta((\sigma \star F) \wedge (\xi_H \cdot A)) \\ &+ \oint_B \delta((\sigma \star F) \wedge (\xi_H \cdot A)), \end{aligned} \quad (20)$$

where we used the fact that κ, Ω_H^i are constants and pulled them out of the integrals and

$$k_{\{0, -\xi_H \cdot A\}}(\delta\Phi; \Phi) = \delta((\sigma \star F) \wedge (-\xi_H \cdot A)). \quad (21)$$

Equating (19) and (20), we arrive at the first law

$$\delta M = T_H \delta S + \Omega_H^i \delta J_i + \delta \Psi, \quad (22)$$

where $T_H = \frac{\kappa}{2\pi}$ is the Hawking temperature, and $\delta \Psi$ is explicitly given as

$$\begin{aligned} \delta \Psi &= \oint_{S_\infty} \delta(\sigma \star F)(\xi_H \cdot A) - \oint_B \delta(\sigma \star F)(\xi_H \cdot A) \\ &+ \int_{\Sigma} \delta(\sigma \star F) \wedge (\mathcal{L}_{\xi_H} A) \\ &\approx \int_{\Sigma} \delta(\sigma \star F) \wedge (\xi_H \cdot F), \end{aligned} \quad (23)$$

where in the last equation, we used (2) and the identity $\mathcal{L}_{\xi_H} A = \xi_H \cdot F + d(\xi_H \cdot A)$. (As a comment, recall that $\xi_H \cdot F = \delta_{\eta_H} A$.) While $\delta \Psi$ is manifestly gauge invariant, it is not necessarily integrable. Note also that, as we discussed $\delta M, \delta J_i, \delta S$ by construction and T_H, Ω_H^i , which are determined only from the metric $g_{\mu\nu}$, are all gauge invariant. So, (22) is the promised gauge invariant first law.

Let us discuss the $\delta \Psi$ term a bit more. Recall that the electric charge is given by the Gauss law $\delta Q = \oint_{S_\infty} \delta(\sigma \star F)$, which is obviously gauge invariant. We may then define the gauge invariant electric chemical potential μ_Q by

$$\delta \Psi := \mu_Q \delta Q. \quad (24)$$

So, the first law can be written as

$$\delta M = T_H \delta S + \Omega_H^i \delta J_i + \mu_Q \delta Q. \quad (25)$$

While μ_Q is gauge invariant, one can present its explicit expression in specific gauges. We will discuss this further in the next section.

We close this part pointing out that our analysis and discussions, and hence, the result, extend to the cases with any number of $U(1)$ gauge fields [see comments below (11)] as well as generic p -form gauge fields in any dimension D [24] and, in particular, to the case of a cosmological constant if it is viewed as the charge associated with a $D-1$ -form field [38,39].

VI. ZEROth LAW

To complete the gauge invariant description of black hole thermodynamics, we also need to discuss the zeroth law. The zeroth law is a statement about a system to be in thermal equilibrium when the flow of charges is zero. Since the flow of charges is proportional to the gradient of the corresponding chemical potential, the zeroth law implies

the constancy of the chemical potentials in the thermal system. With the same token, the zeroth law of black hole thermodynamics requires constancy of the temperature and other chemical potentials on the horizon. In our case, that is constancy of T_H, Ω_H^i, μ_Q . There are established textbook arguments proving constancy of these for stationary black holes with a Killing horizon; see chapter 5 of [13,40] for analysis in beyond Einstein gravity theories. Here, we only focus on the interplay of gauge invariance and the constancy of chemical potentials. Since T_H, Ω_H^i and the usual proofs for their constancy are manifestly gauge invariant, we only discuss the case of μ_Q .

Since we have already established the gauge invariance of μ_Q , to argue for its constancy over the horizon, without loss of generality, we can work in a specific gauge. To this end, note that (23) is not localized on the horizon; in particular, it has a term which is an integral over Σ . Note also that since ξ_H is a KV and a symmetry of the black hole solution, $\mathcal{L}_{\xi_H} F = 0$. We now show that there exists a gauge in which $\mathcal{L}_{\xi_H} A = 0$. First, we note that $\mathcal{L}_{\xi_H} F = 0$ implies $\mathcal{L}_{\xi_H} dA = d\mathcal{L}_{\xi_H} A = 0$. Then, by the Poincaré Lemma $\mathcal{L}_{\xi_H} A = d\lambda$ for a scalar λ and if we take $\lambda = \mathcal{L}_{\xi_H} \tilde{\lambda}$, which we can always do so, $\mathcal{L}_{\xi_H} \tilde{A} = 0$ in the $\tilde{A} = A - d\tilde{\lambda}$ gauge. In this gauge, (23) reduces to integrals over the boundaries,

$$\begin{aligned} \delta\Psi &= \oint_{S_\infty} \Phi_{S_\infty} \delta(\sigma \star F) + \oint_B \Phi_H \delta(\sigma \star F), \\ \Phi_H &= \xi_H \cdot A|_B, \quad \Phi_{S_\infty} = \xi_H \cdot A|_{S_\infty}. \end{aligned} \quad (26)$$

We now show the constancy of Φ_H over the horizon and Φ_{S_∞} over S_∞ . In the $\mathcal{L}_{\xi_H} A = 0$ gauge, constancy of Φ_H over the horizon amounts to showing $\zeta^\alpha \partial_\alpha \Phi_H = 0$ for any horizon tangent vector ζ . To see the latter, recall that horizon is a null hypersurface generated by the orbits of ξ_H , and for $\zeta = \xi_H$, we have $\xi_H^\alpha \partial_\alpha \Phi_H = \mathcal{L}_{\xi_H} \Phi_H = 0$. It hence remains to show $\zeta^\alpha \partial_\alpha \Phi_H = 0$ for any vector ζ tangent to the bifurcation surface B , $(\zeta \cdot \xi_H)_B = 0$:

$$\begin{aligned} \zeta^\alpha \partial_\alpha \Phi_H &= \zeta \cdot d(\xi_H \cdot A)|_B = \zeta \cdot (\mathcal{L}_{\xi_H} A - \xi_H \cdot dA)|_B \\ &= -\zeta \cdot (\xi_H \cdot F)|_B = 0, \end{aligned} \quad (27)$$

where in the second equality, we used the Cartan identity; the third equality follows from the gauge choice $\mathcal{L}_{\xi_H} A = 0$, and the last equality is a result of the vanishing of ξ_H on the bifurcation horizon B where the field strength F has finite components. In fact, using gauge field equation (2), one can show that $\xi_H \cdot F|_B \approx dX$ for a gauge invariant scalar function X defined at B . There can't be any such function X , and hence, $\xi_H \cdot F|_B \approx 0$; see Ref. [25] for more formal proof. (As a side comment, recalling (12), that is $\delta_{\eta_H} A|_B \approx 0$.) Therefore, Φ_H is constant over the horizon H . A similar analysis also shows that Φ_{S_∞} is a constant over S_∞ . Therefore, we may rewrite (26) as

$$\delta\Psi = (\Phi_H - \Phi_{S_\infty})\delta Q, \quad (28)$$

where we used $Q = \oint_{S_\infty} (\sigma \star F) = -\oint_B (\sigma \star F)$. With the above, we recover the standard term in the first law.

VII. CONCLUDING REMARKS

We have given an explicitly gauge invariant derivation of the zeroth and first laws of black hole thermodynamics. We achieved this by extending the notion of Killing vectors to symmetries that include certain field-dependent gauge transformations. At the heart of our proposal is the gauge invariant transformation for the gauge field (12); our analysis here completes the discussions in [21,25]. Despite the fact that we limited our presentation to a specific class of theories, as our arguments are within the covariant phase space formalism (CPSF), our analyses and results hold for any diffeomorphism and gauge invariant theory. In particular, our analysis extends to cases with higher p forms and/or asymptotic (A)dS black holes.

It is well known that in the context of CPSF, there are certain “ambiguities” in the definition of charge variations [15]. Also, there are ambiguities in Wald’s definition of the entropy [16,20]. These ambiguities are ultimately defining the boundary theory, its symplectic form and its Lagrangian [41]; here, the boundary theory resides at H or i^0 . Therefore, the gauge invariance of the full theory, which is our starting point, implies gauge invariance of these ambiguities.

CPSF leads to charge variations. While the first law deals with charge variations, we need to make sure the charges are well defined; i.e., they are integrable over the solution space. It has been argued that the charge integrability in general depends on the slicing of the solution phase space and can be achieved with appropriate symmetry generators [42,43]. For the case of our interest here, stationary black holes and their thermodynamical laws, we need to make sure of the integrability of charges under parametric variations, variations in parameter space of black hole solutions [13,21]. One can show that our augmented symmetry generators η ’s yield integrable charges, while the ξ symmetry generators do not; see the Appendix for an explicit example.

Besides the zeroth and first laws, one may naturally ask about gauge (non)invariance of the usual derivations of second and third laws of black hole thermodynamics. Hawking’s area theorem [44] (and its extensions to generic modified gravity theories) provides the basis for generalized second law arguments; see, e.g., [45,46]. If the notion of the entropy (of black hole and the surrounding systems) are gauge invariant, they lead to the gauge invariant notion of the generalized second law. Here, we have provided such a gauge invariant notion for the entropy. The most rigorous derivation of the third law, to our knowledge, is due to Israel [47]. If the thermodynamical quantities entering the

analysis are gauge invariant, which here we established they are, the arguments in [47] seem to be gauge invariant. Establishing these more explicitly and in some clarifying examples is a well worth exercise to take.

Here, we limited our analysis to stationary black holes with a bifurcate Killing horizon, while the laws of black hole thermodynamics can hold in general gravitational systems where some of these conditions are relaxed or replaced by less restrictive assumptions. For instance, it is known that the near-horizon-extremal geometries that are not black holes in the usual sense also obey laws of thermodynamics, albeit at zero temperature [48]. One may check that our analysis here readily extends to such cases as well as to extremal black holes, which do not have a bifurcate Killing horizon. It is also interesting to study gauge invariance by relaxing stationarity or explore in more detail black holes in asymptotically (A)dS cases or in more general theories like those discussed in [49].

ACKNOWLEDGMENTS

We would like to thank Kartik Prabhu for detailed discussions on his related work [25] and Ibrahim

Shehzad for comments. This work has been supported by TÜBİTAK international researchers program No. 2221. M. M. Sh. J. acknowledges partial support by SarAmadan Grant No. ISEF/M/401332.

APPENDIX: AN EXAMPLE: KERR-NEWMAN BLACK HOLE

To clarify more the abstract analysis in the main text, we present an example for which we also discuss the integrability of charges, which we skipped in the general discussions. As our example, we choose the Kerr-Newman black hole solution to the four-dimensional Einstein-Maxwell theory

$$\mathcal{L} = \frac{1}{16\pi} (R - F^2), \quad (\text{A1})$$

and calculate δQ_η in (6). For an arbitrary $\eta = \{\xi, \lambda\}$, \mathbf{k}_η is found (see Ref. [50] for the details)

$$k_\eta^{\mu\nu}(\delta g, g) = \frac{1}{16\pi} \left(\left[\xi^\nu \nabla^\mu h - \xi^\nu \nabla_\alpha h^{\mu\alpha} + \xi_\alpha \nabla^\nu h^{\mu\alpha} + \frac{1}{2} h \nabla^\nu \xi^\mu - h^{\alpha\nu} \nabla_\alpha \xi^\mu \right] - [\mu \leftrightarrow \nu] \right) + \frac{1}{4\pi} \left(\left(-\frac{h}{2} F^{\mu\nu} - F^{\rho[\mu} h_{\rho}^{\nu]} - \delta F^{\mu\nu} \right) (\xi^\alpha A_\alpha + \lambda) - F^{\mu\nu} \xi^\alpha \delta A_\alpha - F^{\alpha[\mu} \xi^{\nu]} \delta A_\alpha \right), \quad (\text{A2})$$

in which $h^{\mu\nu} := g^{\mu\beta} g^{\nu\rho} \delta g_{\beta\rho}$ and $h := h_\mu^\mu$. Note that $k_\eta^{\alpha\beta} := \epsilon^{\mu\nu\alpha\beta} (\mathbf{k}_\eta)_{\mu\nu}$, and the charge variation is diffeomorphism invariant only if it acts also on the η (see more details in, e.g., [51,52]).

The Kerr-Newman (KN) metric and the gauge field in coordinates $x^\mu = (t, r, \theta, \varphi)$ and in an arbitrary gauge is [53]

$$ds^2 = -(1-f)dt^2 + \frac{\rho^2}{\Delta_r} dr^2 + \rho^2 d\theta^2 - 2fa \sin^2 \theta dt d\varphi + (r^2 + a^2 + fa^2 \sin^2 \theta) \sin^2 \theta d\varphi^2, \\ A = \frac{qr}{\rho^2} (dt - a \sin^2 \theta d\varphi) + d\tilde{\lambda}, \quad (\text{A3})$$

where m , a , and q are free parameters,

$$\rho^2 := r^2 + a^2 \cos^2 \theta, \quad \Delta_r := r^2 + a^2 - 2mr + q^2, \\ f := \frac{2mr - q^2}{\rho^2}, \quad (\text{A4})$$

and $\tilde{\lambda} = \tilde{\lambda}(t, r, \theta, \varphi)$ is an arbitrary scalar function. Moreover,

$$r_H = m + \sqrt{m^2 - a^2 - q^2}, \quad \kappa = \frac{r_H^2 - a^2}{2r_H(r_H^2 + a^2)}, \\ \Omega_H = \frac{a}{r_H^2 + a^2}, \quad \Phi_H = \frac{qr_H}{r_H^2 + a^2} \\ \xi_M = \partial_t, \quad \xi_J = -\partial_\varphi, \quad \xi_S = \frac{2\pi}{\kappa} (\xi_M - \Omega_H \xi_J). \quad (\text{A5})$$

The first issue we want to check by this example is whether η_M , η_J , and η_S in (13) generate the mass, angular momentum, and entropy, respectively. To this end, as discussed in [21], we need to consider the parametric variations. That is, consider set of KN solutions (A3) and vary the solution parameters m , a , q :

$$\hat{\delta} g_{\mu\nu} = \frac{\partial g_{\mu\nu}}{\partial m} \delta m + \frac{\partial g_{\mu\nu}}{\partial a} \delta a + \frac{\partial g_{\mu\nu}}{\partial q} \delta q, \\ \hat{\delta} A_\mu = \frac{\partial A_\mu}{\partial m} \delta m + \frac{\partial A_\mu}{\partial a} \delta a + \frac{\partial A_\mu}{\partial q} \delta q. \quad (\text{A6})$$

After performing the calculation using these variations in

$$\begin{aligned}
\hat{\mathcal{Q}}_{\eta_M} &= \oint_{S_\infty} d\theta d\varphi \sqrt{-g} k_{\eta_M}^{tr}, \\
\hat{\mathcal{Q}}_{\eta_J} &= \oint_{S_\infty} d\theta d\varphi \sqrt{-g} k_{\eta_J}^{tr}, \\
\hat{\mathcal{Q}}_{\eta_S} &= \oint_B d\theta d\varphi \sqrt{-g} k_{\eta_S}^{tr},
\end{aligned} \tag{A7}$$

we find

$$\begin{aligned}
\hat{\delta}\mathcal{Q}_{\eta_M} &= \delta m, \quad \hat{\delta}\mathcal{Q}_{\eta_J} = \delta m a + a \delta m, \\
\hat{\delta}\mathcal{Q}_{\eta_S} &= \frac{2\pi}{\kappa} \left(\delta m - \Omega_H (m \delta a + a \delta m) - \left(\frac{q r_H}{r_H^2 + a^2} \right) \delta q \right),
\end{aligned} \tag{A8}$$

which are manifestly integrable if we take m , a , q to be independent parameters and produce the standard charges

$$M := \mathcal{Q}_{\eta_M} = m, \quad J := \mathcal{Q}_{\eta_J} = m a, \quad S = \pi(r_H^2 + a^2), \tag{A9}$$

respectively. The result is independent of $\tilde{\lambda}$, confirming that the augmented generators in (13) produce gauge invariant charges.

The second issue to investigate is the gauge noninvariance and nonintegrability of the charge variations associated to ξ_M , ξ_J , and ξ_S and show that the charge variations depend on $\tilde{\lambda}$ and are not integrable. To this end, it suffices to focus on a subset of $\tilde{\lambda}$, which produces finite and coordinate-independent charge variations.

Let us begin with ξ_M and consider $\tilde{\lambda} = \tilde{\lambda}(t, r, \theta)$ such that at S_∞ , it tends to a linear function in time $\tilde{\lambda}_\infty(m, a, q)t$ [23]. If we insert $\{\partial_t, 0\}$ and the parametric variations (A6) into (A2), we find

$$\delta\mathcal{Q}_{\xi_M} = \delta m + \tilde{\lambda}_\infty(m, a, q) \delta q. \tag{A10}$$

Presence of $\tilde{\lambda}$ in the result clearly shows its gauge dependence. Moreover, it is not integrable for arbitrary $\tilde{\lambda}_\infty$; i.e., it is not $\delta M(m, a, q)$ for a function $M(m, a, q)$.

To check the gauge noninvariance of the charge variation of ξ_J , we can choose $\tilde{\lambda} = \tilde{\lambda}(r, \theta, \varphi)$ such that at S_∞ , it tends to a constant $\tilde{\lambda}_\infty(m, a, q)$ times φ . A similar calculation shows

$$\delta\mathcal{Q}_{\xi_J} = \delta(ma - q\tilde{\lambda}_\infty), \tag{A11}$$

while integrable, it is gauge dependent for arbitrary $\tilde{\lambda}_\infty$.

To show the nonintegrability of $\xi_S = \frac{2\pi}{\kappa} \xi_H$, one can follow similar steps while integrating over B instead of S_∞ . However, there is a shortcut, which we follow here. The point is that even for the case of $\tilde{\lambda} = 0$, ξ_S has nonintegrable charge variation. This important result was first reported in [21], and we summarize it here. Let us set $\tilde{\lambda} = 0$ in (A3). In this gauge, the KVs ξ_M and ξ_J are exact symmetries, and ω in (16) vanishes for these. So by Stokes' theorem, their charge variations can be calculated not only at S_∞ but also on any other smooth closed codimension-2 surface that encompasses B (including the B itself) producing the same result. Therefore, by $\xi_S = \frac{2\pi}{\kappa} (\xi_M - \Omega_H \xi_J)$ and by setting $\tilde{\lambda} = 0$ in (A10) and (A11), we find

$$\delta\mathcal{Q}_{\xi_S} = \frac{2\pi}{\kappa} (\delta m - \Omega_H (m \delta a + a \delta m)), \tag{A12}$$

which is not integrable over the whole set of parameters m , a , q .

-
- [1] A. Komar, Covariant conservation laws in general relativity, *Phys. Rev.* **113**, 934 (1959).
 - [2] R. L. Arnowitt, S. Deser, and C. W. Misner, Dynamical structure and definition of energy in general relativity, *Phys. Rev.* **116**, 1322 (1959).
 - [3] R. L. Arnowitt, S. Deser, and C. W. Misner, Canonical variables for general relativity, *Phys. Rev.* **117**, 1595 (1960).
 - [4] R. L. Arnowitt, S. Deser, and C. W. Misner, The dynamics of general relativity, *Gen. Relativ. Gravit.* **40**, 1997 (2008).
 - [5] T. Regge and C. Teitelboim, Role of surface integrals in the Hamiltonian formulation of general relativity, *Ann. Phys. (N.Y.)* **88**, 286 (1974).
 - [6] J. D. Brown and J. W. York, Jr., Quasilocal energy and conserved charges derived from the gravitational action, *Phys. Rev. D* **47**, 1407 (1993).
 - [7] L. F. Abbott and S. Deser, Stability of gravity with a cosmological constant, *Nucl. Phys.* **B195**, 76 (1982).
 - [8] S. Deser and B. Tekin, Gravitational Energy in Quadratic Curvature Gravities, *Phys. Rev. Lett.* **89**, 101101 (2002).
 - [9] S. Deser and B. Tekin, Energy in generic higher curvature gravity theories, *Phys. Rev. D* **67**, 084009 (2003).
 - [10] G. Barnich and F. Brandt, Covariant theory of asymptotic symmetries, conservation laws and central charges, *Nucl. Phys.* **B633**, 3 (2002).
 - [11] G. Barnich and G. Compere, Surface charge algebra in gauge theories and thermodynamic integrability, *J. Math. Phys. (N.Y.)* **49**, 042901 (2008).
 - [12] P. N. Chen, M. T. Wang, Y. K. Wang, and S. T. Yau, Supertranslation invariance of angular momentum, *Adv. Theor. Math. Phys.* **25**, 777 (2021). Conserved quantities in general relativity—the view from null infinity, [arXiv: 2204.04010](https://arxiv.org/abs/2204.04010).
 - [13] D. Grumiller and M. M. Sheikh-Jabbari, *Black Hole Physics: From Collapse to Evaporation*, Springer's Graduate

- Texts in Physics (Springer Nature Switzerland, 2022), ISBN: 978-3-031-10342-1.
- [14] J. Lee and R. M. Wald, Local symmetries and constraints *J. Math. Phys. (N.Y.)* **31**, 725 (1990).
 - [15] 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).
 - [16] R. M. Wald and A. Zoupas, A general definition of ‘conserved quantities’ in general relativity and other theories of gravity, *Phys. Rev. D* **61**, 084027 (2000).
 - [17] A. Ashtekar, L. Bombelli, and R. Koul, Phase space formulation of general relativity without a $3+1$ splitting, *Lect. Notes Phys.* **278**, 356 (1987).
 - [18] A. Ashtekar, L. Bombelli, and O. Reula, The covariant phase space of asymptotically flat gravitational fields, in *Mechanics, Analysis and Geometry: 200 Years after Lagrange*, edited by M. Francaviglia (Elsevier Science Publishers B.V., Amsterdam, 1990), pp. 417–450.
 - [19] C. Crnkovic and E. Witten, Covariant description of canonical formalism in geometrical theories, in *Three Hundred Years of Gravitation*, edited by S. W. Hawking and W. Israel (Cambridge University, Melbourne, Australia, 1987), pp. 676–684.
 - [20] R. M. Wald, Black hole entropy is the Noether charge, *Phys. Rev. D* **48**, R3427(R) (1993).
 - [21] K. Hajian and M. M. Sheikh-Jabbari, Solution phase space and conserved charges: A general formulation for charges associated with exact symmetries, *Phys. Rev. D* **93**, 044074 (2016).
 - [22] S. Gao and R. M. Wald, The ‘Physical process’ version of the first law and the generalized second law for charged and rotating black holes, *Phys. Rev. D* **64**, 084020 (2001).
 - [23] S. Gao, The first law of black hole mechanics in Einstein-Maxwell and Einstein-Yang-Mills theories, *Phys. Rev. D* **68**, 044016 (2003).
 - [24] G. Compere, Note on the first law with p-form potentials, *Phys. Rev. D* **75**, 124020 (2007).
 - [25] K. Prabhu, The first law of black hole mechanics for fields with internal gauge freedom, *Classical Quantum Gravity* **34**, 035011 (2017).
 - [26] Z. Elgood, P. Meessen, and T. Ortín, The first law of black hole mechanics in the Einstein-Maxwell theory revisited, *J. High Energy Phys.* **09** (2020) 026.
 - [27] K. Hajian, On thermodynamics and phase space of near horizon extremal geometries, Ph.D thesis, 2015 [arXiv:1508.03494].
 - [28] A. Seraj, Conserved charges, surface degrees of freedom, and black hole entropy, Ph.D thesis, 2016 [arXiv:1603.02442].
 - [29] G. Compère, P. Mao, A. Seraj, and M. M. Sheikh-Jabbari, Symplectic and killing symmetries of AdS_3 gravity: Holographic vs boundary gravitons, *J. High Energy Phys.* **01** (2016) 080.
 - [30] E. Bessel-Hagen, Über die Erhaltungssätze der Elektrodynamik, *Math. Ann.* **84**, 258 (1921).
 - [31] R. Jackiw, Gauge Covariant Conformal Transformations, *Phys. Rev. Lett.* **41**, 1635 (1978).
 - [32] J. P. Bourguignon and P. Gauduchon, Spinors, Dirac operators and metric variations (in French), *Commun. Math. Phys.* **144**, 581 (1992).
 - [33] T. Jacobson and A. Mohd, Black hole entropy and Lorentz-diffeomorphism Noether charge, *Phys. Rev. D* **92**, 124010 (2015).
 - [34] M. Banados, L. J. Garay, and M. Henneaux, The dynamical structure of higher dimensional Chern-Simons theory, *Nucl. Phys.* **476**, 611 (1996).
 - [35] M. Bañados and I. A. Reyes, A short review on Noether’s theorems, gauge symmetries and boundary terms, *Int. J. Mod. Phys. D* **25**, 1630021 (2016).
 - [36] D. N. Blaschke, F. Gieres, M. Reboud, and M. Schweda, The energy–momentum tensor(s) in classical gauge theories, *Nucl. Phys.* **B912**, 192 (2016).
 - [37] F. Gieres, Improvement of a conserved current density versus adding a total derivative to a Lagrangian density, *Fortschr. Phys.* **70**, 2200078 (2022).
 - [38] D. Chernyavsky and K. Hajian, Cosmological constant is a conserved charge, *Classical Quantum Gravity* **35**, 125012 (2018).
 - [39] K. Hajian, H. Özşahin, and B. Tekin, First law of black hole thermodynamics and Smarr formula with a cosmological constant, *Phys. Rev. D* **104**, 044024 (2021).
 - [40] R. Ghosh and S. Sarkar, Black hole zeroth law in higher curvature gravity, *Phys. Rev. D* **102**, 101503 (2020).
 - [41] H. Adami, P. Mao, M. M. Sheikh-Jabbari, V. Taghiloo, and H. Yavartanoo, Symmetries at causal boundaries in 2D and 3D gravity, *J. High Energy Phys.* **05** (2022) 189.
 - [42] H. Adami, M. M. Sheikh-Jabbari, V. Taghiloo, H. Yavartanoo, and C. Zwickel, Symmetries at null boundaries: Two and three dimensional gravity cases, *J. High Energy Phys.* **10** (2020) 107.
 - [43] H. Adami, D. Grumiller, M. M. Sheikh-Jabbari, V. Taghiloo, H. Yavartanoo, and C. Zwickel, Null boundary phase space: Slicings, news & memory, *J. High Energy Phys.* **11** (2021) 155.
 - [44] J. M. Bardeen, B. Carter, and S. W. Hawking, The four laws of black hole mechanics, *Commun. Math. Phys.* **31**, 161 (1973).
 - [45] J. D. Bekenstein, Generalized second law of thermodynamics in black hole physics, *Phys. Rev. D* **9**, 3292 (1974).
 - [46] A. C. Wall, Ten proofs of the generalized second law, *J. High Energy Phys.* **06** (2009) 021.
 - [47] W. Israel, Third Law of Black-Hole Dynamics: A Formulation and Proof, *Phys. Rev. Lett.* **57**, 397 (1986).
 - [48] K. Hajian, A. Seraj, and M. M. Sheikh-Jabbari, NHEG mechanics: Laws of near horizon extremal geometry (thermo)dynamics, *J. High Energy Phys.* **03** (2014) 014.
 - [49] K. Hajian, S. Liberati, M. M. Sheikh-Jabbari, and M. H. Vahidinia, On black hole temperature in Horndeski gravity, *Phys. Lett. B* **812**, 136002 (2021).
 - [50] M. Ghodrati, K. Hajian, and M. R. Setare, Revisiting conserved charges in higher curvature gravitational theories, *Eur. Phys. J. C* **76**, 701 (2016).
 - [51] E. Altas and B. Tekin, New approach to conserved charges of generic gravity in AdS spacetimes, *Phys. Rev. D* **99**, 044016 (2019).
 - [52] E. Altas and B. Tekin, Conserved charges in AdS: A new formula, *Phys. Rev. D* **99**, 044026 (2019).
 - [53] E. T. Newman, R. Couch, K. Chinnapared, A. Exton, A. Prakash, and R. Torrence, Metric of a rotating, charged mass, *J. Math. Phys. (N.Y.)* **6**, 918 (1965).