Lagrangian constraint analysis of first-order classical field theories with an application to gravity

Verónica Errasti Díez[1,*] Markus Maier[1,2,†] Julio A. Méndez-Zavaleta[1,‡] and Mojtaba Taslimi Tehrani[1,§]

[1]*Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), Föhringer Ring 6, 80805 Munich, Germany*
[2]*Universitäts-Sternwarte, Ludwig-Maximilians-Universität München,*
Scheinerstrasse 1, 81679 Munich, Germany

We present a method that is optimized to explicitly obtain all the constraints and thereby count the propagating degrees of freedom in (almost all) manifestly first-order classical field theories. Our proposal uses as its only inputs a Lagrangian density and the identification of the *a priori* independent field variables it depends on. This coordinate-dependent, purely Lagrangian approach is complementary to and in perfect agreement with the related vast literature. Besides, generally overlooked technical challenges and problems derived from an incomplete analysis are addressed in detail. The theoretical framework is minutely illustrated in the Maxwell, Proca and Palatini theories for all finite $d \geq 2$ spacetime dimensions. Our novel analysis of Palatini gravity constitutes a noteworthy set of results on its own. In particular, its computational simplicity is visible, as compared to previous Hamiltonian studies. We argue for the potential value of both the method and the given examples in the context of generalized Proca and their coupling to gravity. The possibilities of the method are not exhausted by this concrete proposal.

I. INTRODUCTION

It is hard to overemphasize the importance of field theory in high-energy physics. Suffice it to recall that each and every of the fundamental interactions we are aware of as of yet—the gravitational, electromagnetic, strong and weak interactions—are described in terms of fields. Correspondingly, their dynamics are studied by means of field theory. Most often, this is done by writing a Lagrangian (or a Hamiltonian) density that is a real smooth function of the field components (and their conjugate momenta) and that is then subjected to the principle of stationary action. It is customary to encounter the situation where not all of the *a priori* independent quantities—field components and/or conjugate momenta—are conferred a dynamical evolution through the equations of motion. In such a case, the field theory is said to be singular or constrained. For instance, it is well known that all gauge theories are singular.

In this work, we focus on singular classical field theories that are manifestly first order and analyze them employing exclusively the Lagrangian formalism. Nonsingular theories are also in (trivial) reach. Throughout the paper, manifest first order shall stand for a Lagrangian that depends only on the field variables and their first derivatives. This implies the equations of motion are guaranteed to be second order at most. Within this framework, we present a systematic methodology that is optimized to determine the number of field components that do propagate, which we denominate physical and propagating modes and degrees of freedom. To do so, we explicitly obtain the constraints: specific functional relations among the field variables and their time derivatives that avoid the propagation of the remaining field components. Our approach is complementary to the similarly aimed procedures in [1–3] and is markedly distinct from, yet equivalent to, that in [4].

Apart from the intrinsic relevance of understanding and characterizing the constraint structure of those theories satisfying our postulates, an ulterior motivation for this investigation is to pave the way toward a consistent theory building principle. Indeed, theoretical physics is currently in need of new fundamental and effective field theories that are capable of accounting for experimental data—the strong *CP* problem, neutrino masses and the nature of the dark sector, to mention but a few of the most relevant examples. A recurrent and challenging obstacle in the development of well-defined field theories consists in

*veroerdi@mppmu.mpg.de
†maier@mppmu.mpg.de
‡julioamz@mpp.mpg.de
§motaslimi@gmail.com

guaranteeing the correct number of physical modes. In this context, most effort is devoted to avoiding the propagation of Ostrogradsky instabilities [5]—additional unphysical degrees of freedom, which we shall denote ghosts for short. The general problem is delineated in [6] and numerous realizations of this idea can be found, e.g., [7]. However, it is equally important to ensure the theory is not overconstrained, i.e., there are fewer than required propagating modes. Our subsequent prescription provides a solid footing to this (double) end and is presented in a comprehensible and ready-to-be-used manner, with the goal of being useful to communities such as, but not limited to, theoretical cosmology and black hole physics. We describe how to convert the analytical tool here exposed into a constructive one, but the concrete realization of this idea is postponed to future investigations.

A specific materialization of the preceding general discussion (and the one we later on employ to ground our conversion proposal) is as follows. We recall that an earlier version of the method here augmented and refined already allowed for the development of the most general nonlinear multivector field theory over four-dimensional flat spacetime: the Maxwell-Proca theory [8,9]. There, the inclusion of a dynamical gravitational field was beyond scope. The present work provides a sound footing for the study of singular field theories defined over curved backgrounds. Thus, it paves the way for the ghost-free coupling of Maxwell-Proca to gravity.

Bearing in mind the above future objective and in order to clarify the formal presentation of the method, we (re)analyze the simplest spin-one and -two theories by means of our proposed procedure: Maxwell, Proca and Einstein's gravity. While the former two are manifestly first order, the latter is not. Indeed, gravity, cast in the Einstein-Hilbert way, is a second-order Lagrangian for the metric, up to a noncovariant boundary term. As such, it exceeds the domain of applicability of our approach. Favorably, this property can be circumvented taking advantage of the deluge of reformulations available for the theory. Among them, we single out the Palatini formalism—see [10] for a historical overview—which considers the metric and the affine connection as *a priori* independent fields.

Our determination of the explicit constraints present in Palatini, while not yielding novel information about the theory, conforms a remarkable piece of work. Not only is it carried out minutely and can be readily seen to be computationally easier and shorter than the previously performed Hamiltonian studies, e.g., [11–14], it also provides the basis for a consistent inclusion of matter fields. As such, we regard this comprehensive analysis as an intrinsically valuable result.

A. Organization of the paper

In the following Sec. II, we introduce the Lagrangian methodology we shall use throughout the paper. Our

approach is complementary to the existing literature. In particular, it is equivalent to the recent proposal in [4], as argued and exemplified in Sec. V A.

We proceed to employ it to analyze various well-known theories: Maxwell electromagnetism, together with the (hard) Proca action in Sec. III and the Palatini formulation of gravity in Sec. IV. Their study is cornerstone to understand the Maxwell-Proca theory [8,9] and paves the way to its consistent coupling to gravity. This is discussed in Sec. V B.

We conclude in Sec. VI, restating the instances when our method is most convenient and emphasizing two crucial aspects that are sometimes overlooked.

B. Conventions

We work on a d-dimensional spacetime manifold $\mathcal{M}$ of the topology $\mathcal{M} \cong \mathbb{R} \times \Sigma$. Namely, we assume $\mathcal{M}$ admits a foliation along a timelike direction. This is true for all (pseudo-)Riemannian manifolds. For simplicity, we consider Σ has no boundary. The dimension d is taken to be arbitrary but finite, with the lower bound $d \geq 2$. Spacetime indices are denoted by the Greek letters $(\mu, \nu, \ldots)$ and raised or lowered with the metric $g_{\mu\nu}$ and its inverse $g^{\mu\nu}$. We employ the standard short-hand notation $\partial_\mu := \frac{\partial}{\partial x^\mu}$, where $x^\mu := (x^0, x^1, \ldots, x^{d-1}) \equiv (x^0, x^i)$, with $i = 1, 2, \ldots, d-1$, are spacetime local coordinates, naturally adapted to the foliation $\mathbb{R} \times \Sigma$. The dot stands for derivation with respect to time, so that for local functions $f \colon \mathcal{M} \to \mathbb{R}$, we write $\dot{f} := \partial_0 f$ and $\ddot{f} := \partial_0^2 f$. Brackets indicating symmetrization and antisymmetrization of indices are defined as $T_{(\mu\nu)} := (T_{\mu\nu} + T_{\nu\mu})/2$ and $T_{[\mu\nu]} := (T_{\mu\nu} - T_{\nu\mu})/2$, respectively. As is customary, summation over repeated indices should be understood at all times.

II. EXPOSITION OF THE METHOD

We begin by putting forward a coordinate-dependent, i.e., nongeometrical, Lagrangian approach to obtain all the constraints present in a manifestly first-order classical field theory. Needless to say, there exists a vast literature on the topic: some standard references are [15]; but for its elegance and concision, we particularly recommend [16]. This section serves us to fix the notation used throughout the paper and provide a self-contained derivation of all our results. We stress that, although the method is not new *per se*, we are not aware of any reference where this material is comprehensively presented in a ready-to-be-used manner and keeping the technicalities at a bare minimum, as we do here.

Our only assumptions shall be the principle of stationary action and finite reducibility. The first assumption is rather obviously a very mild one, but it is worth noting that this is not an essential requirement; for instance, see [17]. We will explain the second assumption shortly. For the time being, it suffices to note that, to our knowledge, the only known

example of a classical field theory (of the kind here considered) not satisfying it is bosonic string field theory, both in its open [18] and closed [19] variants.

Given a Lagrangian density $\mathcal{L}$ within the above postulates, our analysis yields the *constraint structure characterizing triplet*

$$t^{(N)} := (l, g, e). \qquad (2.1)$$

We stress that this is a purely Lagrangian statement, since it collects the outcome of our subsequently proposed purely Lagrangian method. Here, N is the number of *a priori independent field variables* in terms of which $\mathcal{L}$ is written. As such, N is equal to the dimension of the theory's configuration space, which we shall shortly introduce. The other numbers l, g and e are defined below.

On shell, we obtain l: the total number of *functionally independent Lagrangian constraints*. Our analysis elaborates on the iterative algorithm presented in [1] and employed in Appendix A of [8]. It is the suitable generalization to field theory of the coordinate-dependent method used in [2] for particle systems, which is in turn based on [20]. The nontrivial geometric extension to field theory of [2] was carried out in [3], where the discussion was extended to the treatment of off-shell constraints as well. Thus, our discussion is complementary to all these references [1–3].

Off shell, we shall obtain g and e: the number of *gauge identities* and *effective gauge parameters*, respectively. Gauge identities are to be understood in the usual sense, as (differential) relations between certain functional variations of the action that identically vanish. By effective gauge parameters we mean the number of independent gauge parameters plus their successive time derivatives that explicitly appear in the gauge transformations. We determine g and e for theories where the gauge transformations are known *a priori* and provide suitable references that deal with the treatment of theories where the gauge transformations are unknown beforehand. Notice that knowledge of the gauge transformations for the field theory is not a necessary assumption, unlike the principle of stationary action and finite reducibility. However, this information considerably shortens the analysis and, being a feature of all the theories we shall explicitly consider, we have opted for only developing in detail such case.

Given the triplet $t^{(N)}$, the *physical degrees of freedom* n_{dof} in the theory under study can be counted, employing the result derived in [2]:

$$n_{\text{dof}} = N - \frac{1}{2}(l + g + e). \qquad (2.2)$$

We will refer to (2.2) as the *master formula*, the way the authors of [2] themselves do. The remarkable feature about the previous counting is that it is purely Lagrangian, as opposed to the usually employed Hamiltonian formula

$$n_{\text{dof}} = N - N_1 - \frac{1}{2}N_2, \qquad (2.3)$$

attributed to Dirac. Here, (N_1, N_2) denote the number of first- and second-class constraints, respectively. As a reminder, first- (second-) class constraints are those which do (not) have a weakly vanishing Poisson bracket with all of the constraints present in a given theory.

Needless to say, the proven equivalence between the Lagrangian and Hamiltonian formulations of classical theories [1,21] is a most celebrated body of work. The two given prescriptions for the degree of freedom count in (2.2) and (2.3) are a particular materialization of this equivalence, which was further exploited in [2] to develop a one-to-one mapping between the Lagrangian parameters (l, g, e) and their Hamiltonian counterparts:

$$l = N_1 + N_2 - N_1^{(\text{P})}, \qquad g = N_1^{(\text{P})}, \qquad e = N_1, \qquad (2.4)$$

where $N_1^{(\text{P})}$ stands for the number of so-called primary first-class constraints, those first-class constraints that hold true off shell. Using this information, the triplet $t^{(N)}$ defined in (2.1) can be readily seen to admit the following equivalent Hamiltonian parametrization:

$$t^{(N)} = (N_1^{(\text{P})}, N_1, N_2). \qquad (2.5)$$

An important comment is in order here. Our subsequently proposed Lagrangian approach to determine $t^{(N)}$ does not guarantee $n_{\text{dof}} \in \mathbb{N} \cup \{0\}$. This means that, even though all l, g and e in (2.2) are integers by definition, their sum need not be an even number. The reason is simple: we put forward an analytical tool, not a mechanism to detect (or even correct) ill-posed theories. If, for some Lagrangian density $\mathcal{L}$, a half-integer number of physical degrees of freedom is found upon correctly employing our prescription for $t^{(N)}$ together with (2.2), then it must be concluded that the theory is unphysical. The (possibly nontrivial) modifications required on $\mathcal{L}$ for it to propagate an integer number of physical modes are a question beyond the scope of this manuscript.[1]

For the renowned examples in Secs. III and IV, we shall minutely determine the triplet $t^{(N)}$ defined in (2.1) and then use (2.2) to explicitly count physical modes. As such, we shall perform various countings solely in Lagrangian terms. Afterward, we shall (partially) verify our results by comparing them to a representative subset of the Hamiltonian-based literature via (2.3) and (2.4). Additionally, the examples of Sec. III shall be worked out in two different (but dynamically equivalent) Lagrangian formulations, based on distinct values N and $\bar{N} \neq N$ of

---

[1]This should not alarm the reader. The same is true on the standard Hamiltonian formalism. In (2.3), N_2 is not necessarily an even number, unless demands are made on the Hamiltonian.

the dimension of the configuration space. We will then see that, even though the constraint structure characterizing triplets do not coincide, the number of propagating modes $n_{\rm dof}$ does match for both descriptions:

$$t^{(N)} := (l, g, e) \neq t^{(N)} := (\boldsymbol{l}, \boldsymbol{g}, \boldsymbol{e}),$$

$$N - \frac{1}{2}(l + g + e) = n_{\rm dof} = \boldsymbol{N} - \frac{1}{2}(\boldsymbol{l} + \boldsymbol{g} + \boldsymbol{e}). \qquad (2.6)$$

This is because $n_{\rm dof}$ is a physical observable, while (N, l, g, e) are not. Obviously, the same situation arises in the Hamiltonian picture as well, which we briefly illustrate at the end of Sec. IV.

In the following, we explain how to obtain the constraint structure characterizing triplet $t^{(N)}$ in (2.1).

A. On-shell Lagrangian constraints

Let $\mathcal{C}$ be the configuration space of a classical field theory. As usual, we take $\mathcal{C}$ to be a differentiable Banach manifold whose points are labeled by N real field variables Q^A:

$$\mathcal{C} = {\rm span}\{Q^A\}, \qquad A = 1, 2, \dots, N. \qquad (2.7)$$

We stress that A comprises all possible discrete indices that the real field variables have. For instance, if one considers Yang-Mills theory, A consists of both spacetime indices and color indices. If one wishes to entertain complex Yang-Mills, then the real and imaginary parts of each and every Yang-Mills field component must be counted separately in A. So, for $SU(2)$ complex Yang-Mills theory in four spacetime dimensions, we would have that $N = 2(4 \cdot 3) = 24$. Notice that Q^A are real smooth functions of spacetime $Q^A = Q^A(x^\mu)$, but we will suppress this dependence all along, so as to alleviate notation. Thus, our notation matches that in [3] and leaves out the spacetime argument compared to the condensed notation introduced by DeWitt in [22] and extensively used in the literature, e.g., [23]. Then, $T\mathcal{C}$ is the tangent bundle of $\mathcal{C}$, which is spanned by $\{Q^A, \dot{Q}^A\}$. We refer to $(Q^A, \dot{Q}^A, \ddot{Q}^A)$ as the generalized coordinates, velocities and accelerations of the theory, respectively.

As already stated and common to most field theories, we assume that the dynamics are derivable from a principle of stationary action. In other words, the *Euler-Lagrange equations* $E_A \overset{!}{=} 0$ for the field theory follow from the requirement that the action functional

$$S = S[Q^A] = \int_{\mathcal{M}} d^d x \mathcal{L} = \int_{t_1}^{t_2} dx^0 \int_\Sigma d^{d-1} x \mathcal{L} \qquad (2.8)$$

remains stationary under arbitrary functional variations $\delta Q^A = \delta Q^A(x^0, x^i)$ that vanish at times t_1 and t_2 on the spatial slice Σ:

$$\delta S = \frac{\delta S}{\delta Q^A} \delta Q^A \equiv \int_{\mathcal{M}} d^d x E_A \delta Q^A \overset{!}{=} 0, \qquad (2.9)$$

with $\delta Q^A(t_1, x^i) = 0 = \delta Q^A(t_2, x^i)$. The above variational derivative is defined as

$$E_A := \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu Q^A)} \right) - \frac{\partial \mathcal{L}}{\partial Q^A} \overset{!}{=} 0, \qquad (2.10)$$

where the latter equality is the *on-shell* demand. This on-shell requirement commences the iterative algorithm we shall employ to determine the Lagrangian constraints present in the theory. Here, $\mathcal{L} = \mathcal{L}[Q^A]$ is the Lagrangian density. Observe that we have already restricted attention to *manifestly first-order* field theories; i.e., we consider $\mathcal{L}$ depends only on Q^A and its first derivatives $\partial_\mu Q^A$. The study of higher-order field theories[2]—where $\mathcal{L}$ explicitly depends on $\partial_\mu^n Q^A$, with $n \geq 2$—lies beyond the scope of our present investigations. We omit the possible dependence of $\mathcal{L}$ on nondynamical field variables, such as the spacetime metric in any special relativistic theory. The said dependence can be easily incorporated to our analysis, but it does not arise in the theories we discuss in this work.

An important remark on notation follows. As introduced in (2.7), Q^A is an ordered set of *a priori* independent field variables; it is neither a row nor a column vector. The same is true for E_A in (2.10): this is the ordered set of Euler-Lagrange equations for the Q^A field variables, not a vector. We have opted for a notation where the set indices are always assigned the same position when ascribed to a certain ordered set (for instance, upper position for the field variables Q^A and lower position for the Euler-Lagrange equations E_A). The assignation is such that the Einstein summation convention employed throughout the paper is apparent. The only quantities that will show up in this section which have a definite character within matrix calculus are the following. The various Hessians, their Moore-Penrose pseudoinverses and the Jacobians are all matrices. The null vectors of the Hessians are row vectors. Their transposed column vectors also show up. The row or column character of the ordered sets is then straightforwardly fixed according to dimensional analysis in all formulas.

As a practical starting point for our iterative method, it is convenient to recast the Euler-Lagrange equations (2.10) in the form

---

[2]One may be tempted to evade the higher-order character of a theory via the Ostrogradsky prescription, i.e., introducing additional generalized coordinates in a manner that results in a manifestly first-order Lagrangian density. Such alteration of $T\mathcal{C}$ must be compensated through the inclusion of Lagrange multipliers that preserve the equivalence to the original setup. To do so consistently, one needs to either verify the so-called Ostrogradsky nonsingularity condition or exploit alternative methods, as detailed in [24]. In view of these nontrivial subtleties, we restrict ourselves to the study of manifestly first-order theories.

$$E_B = \ddot{Q}^A W_{AB} + \alpha_B \overset{!}{=} 0, \qquad (2.11)$$

where we have defined the so-called *primary Hessian* $W_{AB} := \partial_{\dot{A}} \partial_{\dot{B}} \mathcal{L}$, as well as

$$\alpha_B := (\partial_{\dot{B}} \partial_A^i \mathcal{L} + \partial_B^i \partial_{\dot{A}} \mathcal{L}) \partial_i \dot{Q}^A + (\partial_{\dot{B}} \partial_A^j \mathcal{L}) \partial_i \partial_j Q^A$$
$$+ (\partial_B^i \partial_A \mathcal{L}) \partial_i Q^A + (\partial_{\dot{B}} \partial_A \mathcal{L}) \dot{Q}^A - \partial_B \mathcal{L}. \qquad (2.12)$$

To alleviate notation, we have introduced the following shorthand:

$$\partial_{\dot{A}} := \frac{\partial}{\partial \dot{Q}^A}, \qquad \partial_A^i := \frac{\partial}{\partial(\partial_i Q^A)}, \qquad \partial_A := \frac{\partial}{\partial Q^A}, \qquad (2.13)$$

which we shall extensively employ henceforth.

We focus on singular (or constrained) field theories next.[3] That is, we look at field theories described by a Lagrangian density whose primary Hessian has a vanishing determinant $\det(W_{AB}) = 0$. This means that the rank of W_{AB} (the number of linearly independent rows or columns) is not equal to its dimension N; instead, it is reduced.

By definition it follows that, for singular Lagrangians, the N number of Euler-Lagrange equations in (2.11) can be split into two types. First, *primary equations of motion*: these are the $\mathcal{R}_1 := \text{rank}(W_{AB})$ number of on-shell second-order differential equations that explicitly involve the generalized accelerations $\ddot{Q}^A$. Second, *primary Lagrangian constraints*: these are the $M_1 := \dim(W_{AB}) - \text{rank}(W_{AB}) = N - \mathcal{R}_1$ number of on-shell relations between the generalized coordinates Q^A and their generalized velocities $\dot{Q}^A$. We stress an explicit dependence on $\dot{Q}^A$ (Q^A) is not necessary for the primary Lagrangian constraints; they can be relations between the Q^A's ($\dot{Q}^A$'s) only. Consistency requires that these constraints are preserved under time evolution.

In the following, we obtain the said constraints and ensure the consistency of the field theory by means of an iterative algorithm. We refer to each iteration in the algorithm as a *stage*. In every stage, the above specified notions of equations of motion and Lagrangian constraints will arise. The algorithm closes when the preservation under time evolution of all Lagrangian constraints is guaranteed. Equivalently, when all nth stage Lagrangian constraints are stable, for some finite integer $n \geq 2$. An nth stage Lagrangian constraint is said to be *stable* if its time derivative does not lead to a new (i.e., functionally independent) Lagrangian constraint in the subsequent $(n + 1)$th stage. Below, we explain in detail the different

manners in which the necessary stability of the functionally independent Lagrangian constraints may manifest itself.

1. Primary stage

In order to determine the subset of M_1 number of primary Lagrangian constraints out of the set of all N number of Euler-Lagrange equations in (2.11), we first introduce a set of M_1 number of linearly independent null vectors γ_I associated to the primary Hessian W_{AB}:

$$(\gamma_I)^A W_{AB} = 0, \qquad I = 1, 2, \ldots, M_1. \qquad (2.14)$$

We require that these form an orthonormal basis of the kernel of W_{AB}, which amounts to imposing the normalization condition

$$(\gamma_I)^A (\gamma^J)_A = \delta_I{}^J, \quad \text{where } \gamma^I := (\gamma_I)^T, \qquad (2.15)$$

with T denoting the transpose operation. We stress that, even though in all the examples considered in Secs. III and IV we have chosen null vectors that are constant, this is not a required feature for our formalism. Rather, this is just a possible choice in all the given examples that has been opted for due to its computational convenience. Only the normalization (2.15) is an essential requirement for the null vectors. In full generality, the null vectors of all stages can have an explicit dependence on the field variables Q^A and their first derivatives $\partial_\mu Q^A$.

Then, the M_1 primary Lagrangian constraints are obtained by contracting the Euler-Lagrange equations E_A in (2.11) with the above null vectors.[4] Namely, by performing the contraction with γ_I:

$$\varphi_I \equiv (\gamma_I)^A E_A = (\gamma_I)^A \alpha_A \overset{!}{=} 0. \qquad (2.16)$$

Notice that the last equality is a direct consequence of the on-shell demand in (2.10) or equivalently in (2.11). Hence, the primary Lagrangian constraints are on-shell constraints by definition. One can also see this through equivalence to the more familiar Hamiltonian analysis. It is common knowledge, e.g., [25], that primary Lagrangian constraints relate to secondary constraints in the Hamiltonian framework, which are on-shell constraints by definition.

The primary Lagrangian constraints in (2.16) need not be functionally independent from each other.[5] When they are, the field theory is said to be *irreducible* at the primary stage. Otherwise, the theory is reducible at the primary

---

[3]We leave out nonsingular field theories because the subsequent analysis is redundant for them: in this case $\det(W_{AB}) \neq 0$, which implies $l = 0$ and one can directly move on to Sec. II B. Within our framework, scalar field theories in flat spacetime constitute a prominent example of nonsingularity.

[4]The complementary subset of $\mathcal{R}_1 = N - M_1$ primary equations of motion can be obtained by contracting E_A with the basis vectors of the image of W_{AB}. Here, we concentrate only on the Lagrangian constraints.

[5]This is in contrast to the primary equations of motion, which are guaranteed by construction to be functionally independent among themselves.

stage. Before we carry on, we must restrict attention to the functionally independent primary Lagrangian constraints $\varphi_{I'} \overset{!}{=} 0$, where $I' = 1, 2, \ldots M'_1 \leq M_1$. Their number is given by $M'_1 = \text{rank}(J_{I\Lambda})$, where the *Jacobian matrix* $J_{I\Lambda}$ is defined as

$$J_{I\Lambda} := \frac{\partial \varphi_I}{\partial X^\Lambda}, \qquad X^\Lambda = \{Q^A, \dot{Q}^A\}. \qquad (2.17)$$

This test can be easily related to the standard Hamiltonian framework: it is the pullback of the phase space regularity conditions in [26]. For the theories we are concerned with in this work, we verify $M'_1 = M_1$. Hence, all of the primary Lagrangian constraints in (2.16) must be considered in the following.[6]

The vanishing of all the functionally independent primary Lagrangian constraints defines the so-called *primary constraint surface* $T\mathcal{C}_1$, which is a subspace of the *moduli space* $T\mathcal{C}_0$ of the field theory:

$$T\mathcal{C}_1 := \{(Q^A, \dot{Q}^A) \in T\mathcal{C}_0 | \varphi_I = 0\} \subseteq T\mathcal{C}_0,$$
$$T\mathcal{C}_0 := \{(Q^A, \dot{Q}^A) \in T\mathcal{C} | E_A = 0\} \subset T\mathcal{C}. \qquad (2.18)$$

For brevity, we write

$$\varphi_I : \underset{1}{\overset{!}{\approx}} 0. \qquad (2.19)$$

Equalities that hold true in $T\mathcal{C}_1$ (and not in the entire of the moduli space) shall be denoted $\underset{1}{\approx}$ and referred to as *primary weak equalities*.

As previously noted, consistency requires us to not only enforce the primary Lagrangian constraints (2.19), but also to ensure that these are preserved under time evolution. Explicitly, $\tilde{E}_J := \dot{\varphi}_J \overset{!}{\approx} 0$. This requirement starts the second iteration in the algorithm.

2. Secondary stage

The freshly introduced demands $\tilde{E}_J \overset{!}{\approx} 10^7$ are known as the secondary Euler-Lagrange equations. In order to split

them into secondary equations of motion and secondary Lagrangian constraints, it is convenient to write them as

$$\tilde{E}_J = \ddot{Q}^A(\gamma^I)_A \tilde{W}_{IJ} + \tilde{\alpha}_J \underset{1}{\overset{!}{\approx}} 0, \qquad (2.20)$$

where we have defined

$$\tilde{W}_{IJ} := (\gamma_I)^A \partial_{\dot{A}} \varphi_J,$$
$$\tilde{\alpha}_J := (-\alpha_A M^{AB} \partial_{\dot{B}} + \dot{Q}^A \partial_A + (\partial_i \dot{Q}^A) \partial_{\dot{A}}^i) \varphi_J. \qquad (2.21)$$

We point out that, in obtaining this expressions, we have employed the on-shell statement (2.11), so as to eliminate from (2.20) as much dependence on the generalized accelerations $\ddot{Q}^A$ as possible.[8] Here, $\tilde{W}_{IJ}$ is the so-called secondary Hessian and the auxiliary matrix M^{AB} is the *Moore-Penrose pseudoinverse* (as detailed in [27]) of the primary Hessian. The latter is ensured to always exist and be unique. Its defining relations are[9]

$$M^{AB} W_{BC} - \delta^A{}_C + (\gamma^I)_C (\gamma_I)^A = 0, \quad M^{AB}(\gamma^I)_B = 0. \quad (2.22)$$

To gain some more intuition into M^{AB}, we note that it constitutes a generalization of the standard matrix inverse. It is introduced so that $W_{AB} M^{BC}$ and $M^{AB} W_{BC}$ are orthogonal projections onto the image of W_{AB} and M^{AB}, respectively. For regular square matrices, the Moore-Penrose pseudoinverse is equivalent to the standard matrix inverse: $M = W^{-1}$ iff $\det(W) \neq 0$.

If $\text{rank}(\tilde{W}_{IJ}) = \dim(\tilde{W}_{IJ}) = M_1$, no secondary Lagrangian constraints arise and thus the primary Lagrangian constraints are stable. In this case, we say that the consistency of the primary Lagrangian constraints (2.19) under time evolution is *dynamically ensured*, by a set of M_1 (necessarily functionally independent) secondary equations of motion $\tilde{E}_J = \tilde{E}_J(\ddot{Q}^A)$. As a result, the total number of functionally independent Lagrangian constraints present in such field theories is $l = M'_1$. However, this is not what happens in the theories of our interest.

Generically, the rank of the secondary Hessian is smaller than its dimension. Consequently, $M_2 := \dim(\tilde{W}_{IJ}) - \text{rank}(\tilde{W}_{IJ})$ of the equations in (2.20) are secondary Lagrangian constraints, whose consistency under time evolution must be ensured. This is done exactly as in

---

[6] If $M'_1 < M_1$ and the functionally independent constraints are not straightforwardly identifiable, more work is required. Indeed, there exists an iterative algorithm to extract the functionally independent subset of Lagrangian constraints from (2.16). This is explained in Sec. II D of [3] and subsequently exemplified. When the said algorithm requires a(n) finite (infinite) number of iterations, we face a(n) finitely (infinitely) reducible theory. As already pointed out, the procedure here described requires, at the very least, the closure of the reducibility algorithm to proceed. Thus, infinitely reducible theories cannot be studied with the present formalism. We restate bosonic string field theory [18,19] is the only physically relevant example of an infinitely reducible theory we are aware of.

[7] For clarity, we will use a notation where tilde quantities belong to the secondary stage and hat quantities pertain to the tertiary stage. This will be particularly helpful in Sec. IV A.

[8] In the equivalent and more familiar Hamiltonian approach, this corresponds to solving as many generalized velocities as possible in terms of generalized coordinates and conjugate momenta: $\dot{Q}^A = \dot{Q}^A(Q^A, \Pi_A)$.

[9] In [1], the first relation is referred to as the completeness relation. There, both equations in (2.22) are further used to obtain the explicit form of the functionally independent secondary equations of motion. Unlike at the primary stage, functional independence is not guaranteed by construction. As in the first iteration earlier on, our interest lies in the form of the secondary Lagrangian constraints exclusively.

the primary stage before. In other words, the analysis from Eq. (2.14) onward is to be repeated.

In details, the M_2 number of linearly independent null vectors $\tilde{\gamma}_R$ of the secondary Hessian must be obtained:

$$(\tilde{\gamma}_R)^I \tilde{W}_{IJ} = 0, \qquad R = 1, 2, \ldots, M_2, \qquad (2.23)$$

and chosen so that the normalization condition

$$(\tilde{\gamma}_R)^I (\tilde{\gamma}^S)_I = \delta_R{}^S, \quad \text{with} \quad \tilde{\gamma}^S := (\tilde{\gamma}_S)^T, \qquad (2.24)$$

is satisfied. Then, these must be contracted with the secondary Euler-Lagrange equations in (2.20) to yield the secondary Lagrangian constraints in the theory:

$$\tilde{\varphi}_R \equiv (\tilde{\gamma}_R)^I \tilde{E}_I = (\tilde{\gamma}_R)^I \tilde{\alpha}_I \overset{!}{\underset{1}{\approx}} 0. \qquad (2.25)$$

If the secondary Lagrangian constraints vanish when evaluated on the first constraint surface $\tilde{\varphi}_R \underset{1}{\approx} 0$, then the total number of functionally independent Lagrangian constraints is $l = M_1'$. Again, this is not what happens in (all of) the theories of our interest.

As a consequence, we must proceed with the algorithm. First, we need to obtain the (subset of) $\tilde{\varphi}_R$'s which are functionally independent among themselves when evaluated on the first constraint surface. Their number $M_2' \leq M_2$ is given by

$$M_2' = \text{rank}(\tilde{J}_{R\Lambda}), \quad \text{where} \quad \tilde{J}_{R\Lambda} := \frac{\partial}{\partial X^\Lambda}(\tilde{\varphi}_R|_{TC_1}) \qquad (2.26)$$

and X^Λ was introduced in (2.17). When $M_2' \neq 0$, we verify $M_2' = M_2$ for the theories we shall consider—so that they are irreducible theories at the secondary stage. Thus, all secondary Lagrangian constraints in (2.25) must be considered subsequently.[10]

The vanishing of the functionally independent secondary Lagrangian constraints defines the secondary constraint surface $TC_2 \subset TC_1$, which we write as $\tilde{\varphi}_R : \overset{!}{\underset{2}{\approx}} 0$. Equalities holding true in TC_2 shall be denoted $\underset{2}{\approx}$ and referred to as secondary weak equalities. It should be obvious that the secondary Lagrangian constraints are on-shell constraints by definition.

3. Tertiary stage

Let $\hat{W}_{RS} := (\tilde{\gamma}_R)^I (\gamma_I)^A \partial_{\dot{A}} \tilde{\varphi}_S$ be the tertiary Hessian. When the tertiary Hessian's rank does not match its dimension, the consistency under time evolution of $M_3 := \dim(\hat{W}_{RS}) - \text{rank}(\hat{W}_{RS})$ number of the functionally

---

[10]When $0 < M_2' < M_2$, the iterative algorithm referenced in footnote 6 must be employed to extract the functionally independent secondary Lagrangian constraints from (2.25).

independent secondary Lagrangian constraints is not (dynamically) guaranteed. Instead, it must be enforced through a third iteration of the just described procedure. We stress that it is essential to close the iterative algorithm in order to find the correct number l of functionally independent Lagrangian constraints.

For completeness, we provide the explicit expressions for all relevant quantities at some arbitrary stage of the algorithm in the Appendix. These have not appeared in the literature, as far as we know.

4. Closure of the algorithm

In full generality and as already anticipated, our algorithm stops when all functionally independent Lagrangian constraints have been stabilized. This can happen in either of the following different manners:

(i) *Dynamical closure.*—Firstly, it may happen when $M_n := \dim(W^{(n)}) - \text{rank}(W^{(n)}) = 0$ for some nth stage Hessian $W^{(n)}$, with $n \geq 2$. This implies that no Lagrangian constraints arise at the nth stage, since in this case $W^{(n)}$ has full rank and hence admits no null vector. Here, the consistency under time evolution of the previous stage's functionally independent Lagrangian constraints $\varphi^{(n-1)}$ is dynamically ensured, i.e., through the (necessarily functionally independent) nth stage equations of motion. In other words, the functionally independent $\varphi^{(n-1)}$'s are stable. This closure of the algorithm is exemplified in Sec. III B.

(ii) *Nondynamical closure.*—Secondly, it may happen when $M_n > 0$, but $M_n' = 0$, again with $n \geq 2$. This implies that the nth stage functionally independent Lagrangian constraints $\varphi^{(n)}$'s do not define a new constraint surface, so that $TC_n \equiv TC_{n-1}$. We differentiate two algebraically distinct scenarios:

(ii)(a) The $\varphi^{(n)}$'s vanish identically in the $(n-1)$th constraint surface: $\varphi^{(n)} \underset{n-1}{\approx} 0$. Such $\varphi^{(n)}$'s are known as *Lagrangian identities*. Clearly, Lagrangian identities are trivially stable. The example of Sec. III A illustrates this closure of the algorithm.

(ii)(b) The $\varphi^{(n)}$'s functionally depend on the $(n-1)$th stage functionally independent Lagrangian constraints. Schematically, $\varphi^{(n)} \underset{n-2}{\approx} (f_1 + f_2 \partial_i)\varphi^{(n-1)}$, where (f_1, f_2) are arbitrary real smooth functions of the generalized coordinates and velocities $(Q^A, \dot{Q}^A)$, such that (f_1, f_2) are naturally defined in TC_{n-2}. Then, it readily follows that $\varphi^{(n)} \underset{n-1}{\approx} 0$ and it is obvious that such Lagrangian constraints are stable. This closure happens in both of the examples in Sec. IV.

In all the detailed cases, the total number of functionally independent Lagrangian constraints is given by

$$l = \sum_{a=1}^{n-1} M'_a, \qquad (2.27)$$

where M'_a counts the number of functionally independent ath stage Lagrangian constraints and $n \geq 2$. We are not aware of any physically relevant example of a field theory where n is infinite.

5. Noteworthy considerations

We restate that it is of utmost importance to close the iterative procedure in order to determine l. If the algorithm is not closed (only some or none of the constraints are stabilized), one can only give a lower bound on l. While this may be enough to ensure the absence of Ostrogradsky instabilities [5] in the field theory, it is insufficient to guarantee the propagation of a definite number of degrees of freedom. In such case, one can only infer an upper bound on n_{dof}. This observation is further discussed and exemplified in Sec. VI.

We also point out that, in general, the different stabilizations of the functionally independent Lagrangian constraints that we listed are all present in a given field theory. Namely, some functionally independent Lagrangian constraints in the theory are stabilized dynamically, while others are stabilized nondynamically. This is indeed what happens in our examples of Secs. III C and IV.

Besides, we warn the readers against deceiving themselves regarding the ease of the exposed iterative algorithm. Even though our methodology is sound and rigorous and its logic is easy to follow, there can be no misapprehension as to the algebraic complexity of its implementation in concrete theories, most significantly those involving gravity. From this point of view, the examples in Sec. III are uninvolved, while that in Sec. IV A is quite challenging. The example in Sec. IV B constitutes an intermediate difficulty case. We comment further on this important (from a practical point of view) topic in Sec. VI.

At last, we remark that the algorithm just exposed does not break covariance. Namely, if a field theory within our postulates is covariant, its study under the outlined iterative methodology will preserve this feature. Nonetheless, a suitable space and time decomposition of the *a priori* independent field variables and an evaluation of the Lagrangian constraints in the various constraint surfaces will generically break manifest covariance. This should not be confused with the loss of covariance.

B. Off-shell gauge identities

We now obtain g and e, the two remaining numbers in the triplet $t^{(N)}$ defined in (2.1) of our interest. To begin with, we notice that in the principle of stationary action (2.9), we have so far only considered that $\delta S = 0$ follows from the

E_A piece. However, $\delta S = 0$ may also follow from the δQ^A piece. Subsequently, we briefly review the latter scenario: how the vanishing of δS may be a consequence of off-shell identities stemming from a strict symmetry of the action. This kind of symmetry—gauge invariance—is only manifest through specific field variations $\delta_\theta Q^A$, in contrast to our previous consideration in Sec. II A of arbitrary δQ^A's. Correspondingly, we will differentiate between $\delta_\theta S$ and δS as well.

There are different methods to obtain the said off-shell identities, but it is not our goal to provide an overview of them here. Our subsequent discussion summarizes and employs the approach put forward in [28] and later on adapted to exhibit manifest covariance in [14]. This adaptation makes it straightforward to apply [28] to any manifestly first-order classical field theory, which is our framework.

Consider the field transformations $Q^A \rightarrow Q^A + \delta_\theta Q^A$. Let the changes $\delta_\theta Q^A$ be of the form

$$\delta_\theta Q^A = \sum_{s=0}^{n} (-1)^s (\partial_{\mu_1} \partial_{\mu_2} \ldots \partial_{\mu_s} \theta^\beta)(\Omega_\beta{}^A)^{\mu_1 \mu_2 \ldots \mu_s}, \quad (2.28)$$

where $n \in \mathbb{N} \cup \{0\}$, β is an (possibly collective) index that is to be summed over and the θ^β's and $\Omega_\beta{}^A$'s are known as the *gauge parameters* and *gauge generators* of the transformation, respectively. The θ^β's are real smooth functions of the spacetime coordinates x^μ, while the $\Omega_\beta{}^A$'s are defined in $T\mathcal{C}$ and as such are real smooth functions of $(Q^A, \dot{Q}^A)$. The former are unspecified, while the latter are to be determined. Introducing the above in (2.9) and operating, one finds that

$$\delta_\theta S = \int_{\mathcal{M}} d^d x \, \theta^\beta \varrho_\beta,$$

$$\varrho_\beta := \sum_{s=0}^{n} \partial_{\mu_1} \ldots \partial_{\mu_s} [E_A (\Omega_\beta{}^A)^{\mu_1 \mu_2 \ldots \mu_s}]. \qquad (2.29)$$

If, under the field variations (2.28) for some $\Omega_\beta{}^A$'s, the action remains invariant $\delta_\theta S \equiv 0$, then we have that

$$\varrho_\beta \equiv 0 \qquad (2.30)$$

holds true off shell (i.e., without making use of $E_A \overset{!}{=} 0$). In such a case, (2.28) and (2.30) are known as the *gauge transformations* and *gauge identities* in the theory, respectively.

Given (2.28), g is equal to the number of different θ parameters there present. Equivalently, g is the number of linearly independent gauge identities (2.30). On the other hand, e is equal to the total number of distinct parameters plus their successive time derivatives $(\theta, \dot{\theta}, \ddot{\theta}, \ldots)$ that appear in (2.28). Obviously, $e \geq g$.

The recursive construction of the gauge generators $\Omega_\beta{}^A$ has been a subject of vivid interest for decades.
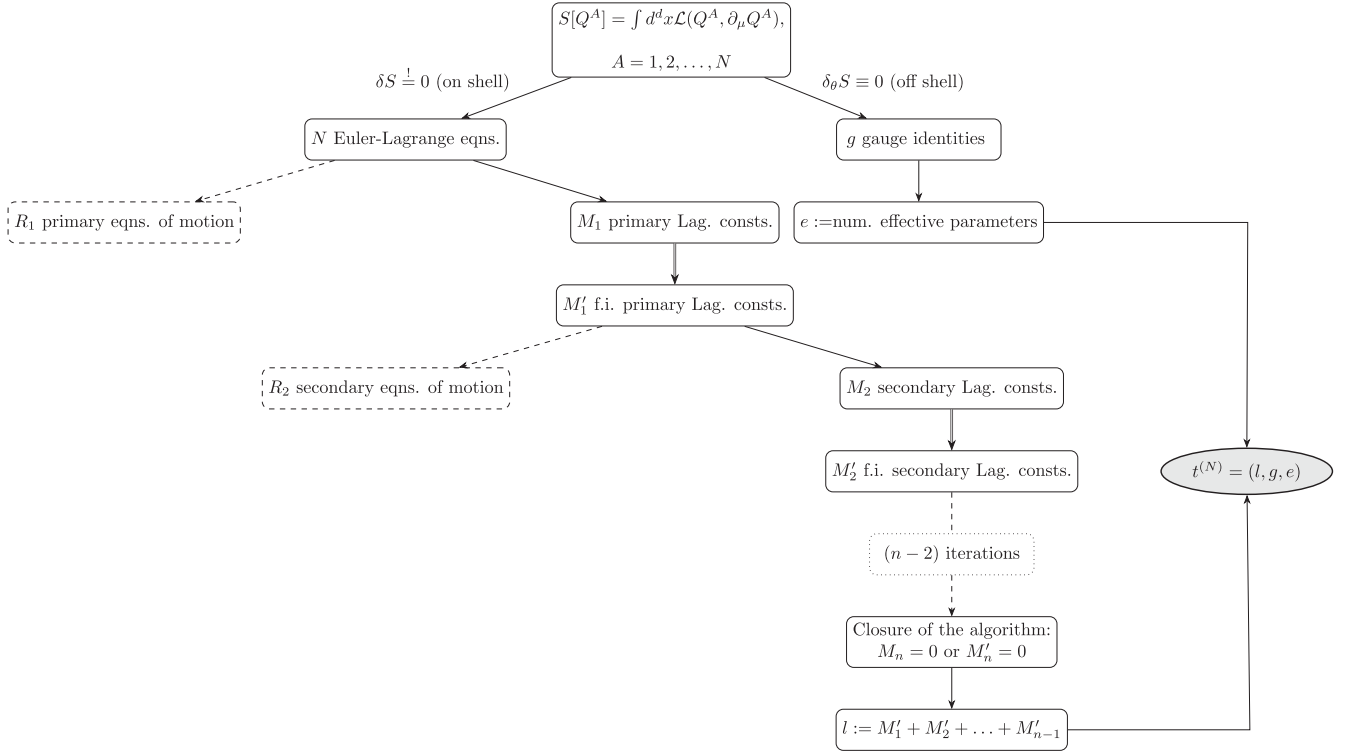
FIG. 1. Schematics of Sec. II. Here, eqns., Lag. consts., f.i., and num. stand for equations, Lagrangian constraints, functionally independent, and number, respectively. The computational challenge of the steps relating Lagrangian constraints to functionally independent Lagrangian constraints (represented with a double arrow), as well as the relevance of closing the iterative algorithm, are further discussed in Sec. VI.

The approach in [29] is perhaps the most befitting to our own exposition, requiring only a suitable adaptation from particle systems to manifestly first-order field theories that is devoid of conceptual subtleties. We shall not present the corresponding discussion here because, for the theories at hand, the explicit form of the gauge transformations is already known. This *a priori* knowledge allows us to effortlessly infer the generators $\Omega_\beta{}^A$ in all the subsequent examples.

We stress that the determination of g and e is possible and has been made systematic in theories for which the gauge transformations are unknown from the onset. The calculations in such theories are more involved, but there is no theoretical obstacle that has to be overcome. To illustrate this point, the reader can consult [28] for the explicit derivation of the gauge generators in Yang-Mills theory and both the metric and Palatini formulations of general relativity, by means of the formalism put forward in [29].

For the ease of the reader, we have schematically depicted the main line of reasoning behind this Sec. II in Fig. 1.

III. SIMPLE EXAMPLES: VECTOR FIELD THEORIES

This section is devoted to the study of some of the constraint structure characterizing triplets $t^{(N)}$ that are

possible for the theories (within the framework of Sec. II) describing the dynamics of a single vector field. Recall there are only two distinct types of vector fields that one can entertain classically: massless and massive. For simplicity, we will restrict to real Abelian vector fields and focus on their most elementary actions: Maxwell electromagnetism and the (hard) Proca theory, respectively. We shall consider two equivalent formulations of each of these theories, based on different numbers N and $\mathbf{N} \neq N$ of *a priori* independent field variables. Our forthcoming detailed analyses are based on the purely Lagrangian method described in the previous Sec. II and thus serve to illustrate it.

Besides and as we shall explain in Sec. V B, our forthcoming elementary calculations turn out to be enough to understand the complete set of manifestly first-order (self-) interactions among an arbitrary number of both Maxwell and (generalized) Proca [30] fields in four-dimensional flat spacetime [8,9]. This hints to the convenience of the proposed method, compared to other possible approaches, a point that shall be reinforced in the more elaborate examples of the next Sec. IV and discussed in the concluding Sec. VI.

In the remaining of this section, we shall work on d-dimensional Minkowski spacetime, still for finite $d \geq 2$. We will choose Cartesian coordinates with the mostly positive signature, so that $g_{\mu\nu} = \eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, \ldots, 1)$.

Subsequently, all spacetime indices shall be raised or lowered by $\eta_{\mu\nu}$ and its inverse $\eta^{\mu\nu}$.

A. Maxwell electromagnetism

This renowned manifestly first-order singular field theory describes an Abelian massless vector field and its linear interactions with sources in terms of $N = d$ number of *a priori* independent field variables. As already stated, we take the Maxwell vector field (which we denote A_μ) to be real and consider the particularly simple case when there are no sources.

1. Lagrangian constraints

The canonically normalized Lagrangian density of sourceless classical electromagnetism is

$$\mathcal{L}_{\mathrm{M}} = -\frac{1}{4} A_{\mu\nu} A^{\mu\nu}, \quad \text{with} \quad A_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu = -A_{\nu\mu}. \tag{3.1}$$

The components of the Maxwell field constitute the generalized coordinates for this theory: $Q^A = \{A_\mu\}$, so that $A = 1, 2, \ldots, d = \dim(\mathcal{C}) \equiv N$, as already announced. As is well known and can be easily calculated by means of (2.10), the Euler-Lagrange equations following from (3.1) are

$$E_A \equiv -\partial_\mu A^{\mu\nu} \overset{!}{=} 0. \tag{3.2}$$

If we decompose the Maxwell field into its space and time components $A_\mu := (A_0, A_i) \equiv (A, A_i)$ with $i = 1, 2, \ldots, d-1$, the Lagrangian (3.1) can be conveniently rewritten as

$$\mathcal{L}_{\mathrm{M}} = \frac{1}{2} [\dot{A}_i^2 + (\partial_i A)^2 - 2\dot{A}_i \partial_i A] - \frac{1}{4} A_{ij}^2, \tag{3.3}$$

where sum over repeated indices is to be understood and we have been careful to lower all indices with the flat metric $\eta_{\mu\nu}$. It is then easy to see that the primary Hessian following from (3.3) is $W_{AB} = \delta_{AB} - \delta_A{}^1 \delta_B{}^1$ and therefore manifestly possesses the symmetry dictated by its very definition: $W_{AB} = W_{BA}$. Further, its Moore-Penrose pseudoinverse is given by $M^{AB} = \delta^{AB} - \delta_1{}^A \delta_1{}^B$. Since the primary Hessian takes such an uncomplicated form, it readily follows that $\mathcal{R}_1 = 3$ and thus $M_1 = 1(= M'_1)$ in this case. A convenient choice for the null vector of W_{AB} amounts to $(\gamma_1)^A = \delta_1{}^A$. Then, the one and only primary Lagrangian constraint for the theory can be effortlessly calculated to take the explicit form

$$\varphi_1 = \partial_i A_{i0} \underset{1}{:} \approx 0. \tag{3.4}$$

This is the familiar Gauss law, telling us that, in the absence of sources, the electric field is divergenceless. Note that this is an on-shell statement by construction.

The Gauss law constraint straightforwardly yields a vanishing secondary Hessian $\tilde{W}_{11} \equiv 0$, so that $M_2 = M_1 = 1$ and we choose $(\tilde{\gamma}_1)^1 = 1$. With all this information, it is a matter of easy algebra to find the only secondary Lagrangian constraint:

$$\tilde{\varphi}_1 \underset{1}{\approx} 0. \tag{3.5}$$

Therefore, $M'_2 = 0$ and the end of the iterative algorithm is signalled according to the nondynamical prescription in case (iia). We have thus found that the total number of Lagrangian constraints for Maxwell electromagnetism is just $l = M'_1 + M'_2 = 1$.

2. Gauge identities

Maxwell's theory enjoys an apparent $U(1)$ gauge symmetry. Indeed, under the transformation $A_\mu \to A_\mu + \partial_\mu \theta$, the Lagrangian (3.1) remains invariant. Here, θ is the only gauge parameter, while $(\theta, \dot{\theta})$ are the sole two effective gauge parameters present in the fields' transformation. Consequently, we have that $g = 1$ and $e = 2$.

For completeness, we point out that the said transformation, when compared to (2.28) immediately allows us to read off the gauge generator of the symmetry. This is $(\Omega^A)^\nu = -\delta^{A\nu}$. When combined with the primary Euler-Lagrange equations (3.2) as indicated in (2.29), we can right away verify the off-shell gauge identity we counted: $\varrho = \partial_\mu \partial_\nu A^{\mu\nu} \equiv 0$.

3. Physical degrees of freedom

According to our prior analysis, which shows that the constraint structure of classical electromagnetism in its standard formulation with $N = d$ is

$$t_{\mathrm{M}}^{(N)} = (l = 1, g = 1, e = 2), \tag{3.6}$$

and making use of the master formula (2.2), we count $n_{\mathrm{dof}} = d - 2$ propagating modes. In $d = 4$, these correspond to the two polarizations of the photon. Exploiting the equalities in (2.4), we check that our counting corresponds to two first-class constraints, one primary and one secondary. Therefore, our purely Lagrangian investigation is in perfect agreement with the standard literature, e.g., [31]. It also matches the Hamiltonian definition of the Maxwell field given in [8]: "a real Abelian vector field [...] associated with two first-class constraints." This latter correspondence will play a role in Sec. V B.

B. The (hard) Proca theory

We turn our attention to the Proca theory next, in the modern formulation of the original proposal in [32]. Namely, we focus on the (manifestly first-order) field theory of a real Abelian vector field of mass m in the

absence of any source described by $N = d$ *a priori* independent field variables. The remark (hard) is to avoid ambiguity with respect to the generalized Proca theory, discussed in Sec. V B. We refer to the Proca field as B_μ.

1. Lagrangian constraints

The Lagrangian density of the said Proca theory is

$$\mathcal{L}_P = -\frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{2} m^2 B_\mu B^\mu, \quad \text{with}$$
$$B_{\mu\nu} := \partial_\mu B_\nu - \partial_\nu B_\mu = -B_{\nu\mu}. \quad (3.7)$$

As in the Maxwell case earlier on, the components of the Proca field are the generalized coordinates: $Q^A = \{B_\mu\}$. We thus see that $A = 1, 2, …, d = \dim(\mathcal{C}) \equiv N$ here as well. The Euler-Lagrange equations following from (3.7) can be easily obtained as indicated in (2.10). The result is

$$E_A \equiv -\partial_\mu B^{\mu\nu} + m^2 B^\nu \overset{!}{=} 0. \quad (3.8)$$

At this point, it is straightforward to see that the primary Hessian—and hence also its Moore-Penrose pseudoinverse—is the same as for the Maxwell theory earlier on. This implies $M_1 = 1 (= M_1')$ and the associated null vector can again be chosen as $(\gamma_1)^A = \delta_1{}^A$. The primary Lagrangian constraint differs, though:

$$\varphi_1 = \partial_i B_{i0} - m^2 B \underset{1}{\overset{!}{:\approx}} 0, \quad (3.9)$$

where we have introduced $B_\mu := (B_0, B_i) \equiv (B, B_i)$.

The above once more leads to a vanishing secondary Hessian, so that $M_2 = M_1 = 1$ and $(\tilde{\gamma}_1)^1 = 1$. The secondary Lagrangian constraint in this case takes the form

$$\tilde{\varphi}_1 = -m^2 \dot{B} \underset{2}{\overset{!}{:\approx}} 0. \quad (3.10)$$

Contrary to the Maxwell theory, (3.10) is obviously not a Lagrangian identity, so the algorithm is not closing here according to the prescription in case (iia). Notice as well that φ_1 and $\tilde{\varphi}_1$ are functionally independent from each other, so that we are not in case (iib) of the general method either. Instead, we have $M_2' = M_2 = 1$ and we must move on to the tertiary stage.

It is easy to check that the tertiary Hessian following from (3.10) is $\hat{W}_{11} = -m^2$. As such, its dimension and rank match ($M_3 = 0 = M_3'$) and the algorithm closes according to the dynamical prescription in case (i). Namely, the consistency of (3.10) under time evolution is ensured via a tertiary equation of motion and there are no tertiary constraints. As a result, we have obtained $l = M_1' + M_2' + M_3' = 2$ functionally independent Lagrangian constraints in the (hard) Proca theory.

2. Gauge identities

The mass term for the Proca field explicitly breaks the $U(1)$ gauge invariance of Maxwell electromagnetism. In our conventions, this means that there is no field transformation of the form (2.28) that leaves the action invariant. Therefore, there are no off-shell identities associated to (3.7) and we have $g = 0 = e$.

3. Physical degrees of freedom

Using the (hard) Proca constraint structure for $N = d$

$$t_P^{(N)} = (l = 2, g = 0, e = 0) \quad (3.11)$$

obtained before in the master formula (2.2), we count $n_{\text{dof}} = d - 1$ degrees of freedom in the theory. By means of (2.4), it is immediate to certify that this corresponds to two second-class constraints, as explicitly shown, for instance, in [33]. As with the Maxwell field before, we thus find agreement with the Proca field's definition given in [8]: "a real Abelian vector field […] associated with two second-class constraints." We will further comment on this connection in Sec. V B later on.

C. The Schwinger-Plebanski reformulation of Maxwell and Proca

In this section, we reanalyze the constraint structures of the above massless and massive vector field theories in a formulation with $N \neq N = d$ *a priori* degrees of freedom. Specifically, we entertain the reformulation of source-less classical electromagnetism originally proposed by Schwinger [34] and later on popularized by Plebanski [35] and employ it for the (hard) Proca theory simultaneously. In this setup, the real Abelian (covariant) vector field C_μ—be it massless or massive—and its antisymmetric (contravariant) field strength $F^{\mu\nu}$ are regarded as independent at the onset:

$$Q^A = \{C \equiv C_0, F^{ij} = -F^{ji}, F^i \equiv F^{0i} = -F^{i0}, C_i\},$$
$$A = 1, 2, …, N = d(d + 1)/2. \quad (3.12)$$

The aim of this Sec. III C is to determine the constraint structure characterizing triplets $t_M^{(N)}$ and $t_P^{(N)}$, so as to illustrate in a simple double example the general claim in (2.6). Namely, these triplets differ from the previously determined ones $t_M^{(N)}$ and $t_P^{(N)}$ but yield the same number of propagating degrees of freedom.

A clarifying remark follows. Classical electromagnetism as written in [34] is commonly called the manifestly first-order formulation of electrodynamics. This refers to the order of its primary Euler-Lagrange equations, contrarily to our convention here, where the order refers to the Lagrangian density. For us, all examples in Secs. III and IV are manifestly first order and as such can be

investigated by means of the methodology in Sec. II. In view of this dissonance, we can already anticipate that there will be no primary equations of motion in our subsequent examples. The primary Euler-Lagrange equations, being first order, will not involve the generalized accelerations $\ddot{Q}^A$ and so they will all be primary Lagrangian constraints. Further, this is possible iff the primary Hessian of the theories identically vanishes, as we shall see it does.

1. Lagrangian constraints

Inspired by [34], we take the Lagrangian density

$$
\mathcal{L}_V = \mathcal{L}_V[C_\mu, F^{\mu\nu}] = -\frac{1}{2}(\partial_\mu C_\nu - \partial_\nu C_\mu)F^{\mu\nu}
$$
$$
+\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}m^2 C_\mu C^\mu \quad \text{with} \quad m \geq 0 \quad (3.13)
$$

as our starting point. When $m = 0$, (3.13) describes classical electromagnetism. For $m \neq 0$, the (hard) Proca theory is portrayed. The Euler-Lagrange equations following from (3.13) are

$$
E_{(C_\nu)} := -\partial_\mu F^{\mu\nu} + m^2 C^\nu \overset{!}{=} 0,
$$
$$
E_{(F^{\mu\nu})} := F_{\mu\nu} - \partial_\mu C_\nu + \partial_\nu C_\mu \overset{!}{=} 0. \quad (3.14)
$$

Solving the latter for $F_{\mu\nu}$ and substituting the result into the former, we recover Maxwell's (3.2) or Proca's (3.8) equations of motion, depending on the value of m. Then, we say both formulations, in (3.13) and in (3.1) or (3.7) as pertinent, are dynamically equivalent, as foretold.

We proceed to explicitly confirm our predictions. The primary Hessian following from (3.13) vanishes identically $W_{AB} \equiv 0$, so $\mathcal{R}_1 = 0$ and $M_1 = N$. We can choose its appropriate null vectors as $(\gamma_I)^A = \delta_I{}^A$. As a result, the primary Lagrangian constraints coincide with the primary Euler-Lagrange equations. These can be readily seen to be functionally independent among themselves. Consequently, the first constraint surface $T\mathcal{C}_1$ coincides with the moduli space in this case. This set of circumstances can be summarized as

$$
0 \overset{!}{=} E_A = \alpha_A = (\gamma_I)^A \alpha_A = \varphi_I \underset{1}{:}{\approx} 0 \quad (3.15)
$$

or simply as $M_1' = M_1 = N$. Notice that $W_{AB} \equiv 0$ immediately makes its Moore-Penrose pseudoinverse vanish as well: $M^{AB} = 0$. We encounter this same situation of a zero primary Hessian in both of the theories analyzed in Sec. IV.

We briefly depart from the application of the iterative algorithm in order to introduce an extremely useful notation that will be recurrent from now on. We wish to be able to refer to each kind of field variables in (3.12) individually. To this aim, we shall henceforth understand that the index A therein decomposes into two distinct sets of indices

$A \equiv \mathcal{A}_1 \mathcal{A}_2$, the first referring to the type of field variable and the second to the spacetime structure of each type of field variable. In this way, $\mathcal{A}_1 = 1, 2, ..., 4$ and we have

$$
[Q^1] \equiv C, \quad [Q^2]^{ij} \equiv F^{ij}, \quad [Q^3]^i \equiv F^i, \quad [Q^4]_i \equiv C_i. \quad (3.16)
$$

Observe that we have employed the symbol $[\cdot]$ to visually split the $\mathcal{A}_1$ index from the $\mathcal{A}_2$ one.

Back to the algorithm and putting into practice the above notation, we write the primary Lagrangian constraints as

$$
[\varphi_1] = \partial_i F^i - m^2 C, \qquad [\varphi_2]_{ij} = F_{ij} - 2\partial_{[i} C_{j]},
$$
$$
[\varphi_3]_i = F_i + \dot{C}_i - \partial_i C, \qquad [\varphi_4]^i = -\dot{F}^i - \partial_j F^{ij} + m^2 C^i. \quad (3.17)
$$

Notice that $[\varphi_2]_{ij} = -[\varphi_2]_{ji}$, as required by definition.

We go on to the secondary stage next. The secondary Hessian $\tilde{W}_{IJ} = \partial_i \varphi_J$ can be portrayed in our recently introduced notation as follows:

$$
\tilde{W}_{IJ} = \begin{pmatrix}
[\tilde{W}_{11}] & [\tilde{W}_{12}]_{ij} & [\tilde{W}_{13}]_i & [\tilde{W}_{14}]^i \\
{}_{ij}[\tilde{W}_{21}] & {}_{ij}[\tilde{W}_{22}]_{kl} & {}_{ij}[\tilde{W}_{23}]_k & {}_{ij}[\tilde{W}_{24}]^k \\
{}_i[\tilde{W}_{31}] & {}_i[\tilde{W}_{32}]_{jk} & {}_i[\tilde{W}_{33}]_j & {}_i[\tilde{W}_{34}]^j \\
{}^i[\tilde{W}_{41}] & {}^i[\tilde{W}_{42}]_{jk} & {}^i[\tilde{W}_{43}]_j & {}^i[\tilde{W}_{44}]^j
\end{pmatrix}, \quad (3.18)
$$

where, for each entry of the secondary Hessian, we have placed the spacelike tensorial indices of the field variables (primary Lagrangian constraints) labeled by I (J) to the left (right). A few explicit examples that should clarify our notation are

$$
[\tilde{W}_{11}] := \frac{\partial [\varphi_1]}{\partial [Q^1]} \equiv \frac{\partial [\varphi_1]}{\partial C},
$$
$$
{}_{ij}[\tilde{W}_{21}] := \frac{\partial [\varphi_1]}{\partial [Q^2]^{ij}} \equiv \frac{\partial [\varphi_1]}{\partial F^{ij}},
$$
$$
{}^i[\tilde{W}_{42}]_{jk} := \frac{\partial [\varphi_2]_{jk}}{\partial [Q^4]_i} \equiv \frac{\partial [\varphi_2]_{jk}}{\partial C_i}. \quad (3.19)
$$

The only nonzero components in (3.18) are

$$
{}^j[\tilde{W}_{43}]_i = \delta_i^j = -{}_i[\tilde{W}_{34}]^j, \quad (3.20)
$$

which lead to a simple secondary Moore-Penrose pseudoinverse $\tilde{M}_{AB}$ with nonzero elements ${}_j[\tilde{M}_{43}]^i = -\delta_j^i = {}^i[\tilde{M}_{34}]_j$. This corresponds to the transpose of (3.18). It is easy to see that $\mathcal{R}_2 := \text{rank}(\tilde{W}_{IJ}) = 2(d-1)$, which in turn implies that $M_2 = (d^2 - 3d + 4)/2$. We choose the

suitably normalized linearly independent null vectors for (3.18) as $(\tilde{\gamma}_R)^I = \delta_R{}^I$.

The above results can be employed to determine the functionally independent secondary Lagrangian constraints $\tilde{\varphi}_R \underset{2}{:} \approx 0$. First, we calculate $\tilde{\varphi}_R = (\tilde{\gamma}_R)^I \dot{\varphi}_I$ and obtain

$$[\tilde{\varphi}_1] = \partial_i \dot{F}^i - m^2 \dot{C}, \qquad [\tilde{\varphi}_2]_{ij} = \dot{F}_{ij} - 2\partial_{[i}\dot{C}_{j]}, \quad (3.21)$$

where again the antisymmetry property $[\tilde{\varphi}_2]_{ij} = -[\tilde{\varphi}_2]_{ji}$ required by definition is apparent. Evaluation on the first constraint surface then gives

$$[\tilde{\varphi}_1] \underset{1}{\approx} m^2(\partial_i C^i - \dot{C}), \qquad [\tilde{\varphi}_2]_{ij} \underset{1}{\approx} \dot{F}_{ij} + 2\partial_{[i}F_{j]}, \quad (3.22)$$

which respects the noted symmetry, as it must. In more detail, the evaluation has been carried out as follows: by setting to zero all φ's in (3.17), solving for $(\dot{F}_i, \dot{C}_i)$ and plugging the resulting expressions into (3.21). Next, we need to select only the functionally independent secondary constraints. It is obvious that the mass m plays a crucial role here, as could easily be anticipated in view of our results in the previous Secs. III A and III B. Indeed, if $m = 0$, then one constraint identically vanishes in the first constraint surface $[\tilde{\varphi}_1] \underset{1}{\approx} 0$. It is thus a Lagrangian identity, meaning that $[\varphi_1]$ is nondynamically (trivially) stabilized at the secondary stage in this case. We therefore see that

$$M_2' = \begin{cases} M_2 - 1 & \text{if } m = 0, \\ M_2 & \text{if } m > 0. \end{cases} \quad (3.23)$$

We turn to the time evolution of the functionally independent secondary constraints; i.e., we commence the tertiary stage. The tertiary Hessian can be succinctly expressed as

$$\hat{W}_{RS} = \begin{pmatrix} [\hat{W}_{11}] & [\hat{W}_{12}]_{ij} \\ {}_{ij}[\hat{W}_{21}] & {}_{ij}[\hat{W}_{22}]_{kl} \end{pmatrix}, \quad (3.24)$$

where we have made use of the same notation as in (3.18) earlier on, so that

$$[\hat{W}_{11}] \equiv \frac{\partial[\tilde{\varphi}_1]}{\partial \dot{C}}, \qquad [\hat{W}_{12}]_{ij} \equiv \frac{\partial[\tilde{\varphi}_2]_{ij}}{\partial \dot{C}},$$
$${}_{ij}[\hat{W}_{21}] \equiv \frac{\partial[\tilde{\varphi}_1]}{\partial \dot{F}^{ij}}, \qquad {}_{ij}[\hat{W}_{22}]_{kl} \equiv \frac{\partial[\tilde{\varphi}_2]_{kl}}{\partial \dot{F}^{ij}}. \quad (3.25)$$

Notice that, for $m = 0$, the first row $[\hat{W}_{1R}]$ should not be present, as $[\tilde{\varphi}_1]$ weakly vanishes in this case. However, we keep it along here, so that both Maxwell and (hard) Proca theories can be reanalyzed simultaneously. The nonzero components are explicitly given by

$$[\hat{W}_{11}] = -m^2, \qquad {}_{ij}[\hat{W}_{22}]_{kl} = 2\delta_{i[k}\delta_{l]j}. \quad (3.26)$$

Hence, the tertiary Hessian has full rank $\mathcal{R}_3 = M_2'$ and consequently $M_3 = 0$. Observe that this is true for both the $m = 0$ and the $m > 0$ cases. The functionally independent secondary constraints' consistency under time evolution is at this point dynamically ensured and the algorithm closes according to the prescription in case (i). The total number of functionally independent Lagrangian constraints is

$$l = M_1' + M_2' = \begin{cases} d(d-1) + 1 & \text{if } m = 0, \\ d(d-1) + 2 & \text{if } m > 0. \end{cases} \quad (3.27)$$

We see that the mass m gives rise to one more functionally independent Lagrangian constraint, exactly as in the previous sections, where we found that $l = 1$ for electromagnetism, while $l = 2$ for the (hard) Proca theory.

Here, the remaining constraints that l counts are associated to the field strength $F^{\mu\nu}$, as a result of having promoted it to a set of *a priori* independent field variables. Notice that there are $d(d-1)$ number of such supplementary constraints, 2 times the number of independent components in F^{ij}. This duplicity makes it manifest that these fields are superfluous when describing the dynamics of the theory. In other words, no initial data are needed for them: F^{ij} and $\dot{F}^{ij}$ need not be specified at some initial time t_1 when solving their associated equations of motion. Yet another way to understand this is to map them to the Hamiltonian picture, where they correspond to second-class constraints, as we shall shortly see.

2. Gauge identities

Consider the following transformations of the field variables: $C_\mu \rightarrow C_\mu + \delta_\theta C_\mu$ and $F^{\mu\nu} \rightarrow F^{\mu\nu} + \delta_\theta F^{\mu\nu}$, with

$$\delta_\theta C_\mu = \partial_\mu \theta, \qquad \delta_\theta F^{\mu\nu} = 0. \quad (3.28)$$

Here, θ is an arbitrary parameter. It can be easily checked that, under the said transformations, the Lagrangian (3.13) remains invariant iff $m = 0$. Therefore, these are the very same gauge transformations of the massless theory that we noted in Sec. III A, while the massive theory does not exhibit any kind of gauge symmetry. Straightforwardly, we count

$$g = \begin{cases} 1 & \text{if } m = 0, \\ 0 & \text{if } m > 0, \end{cases} \qquad e = \begin{cases} 2 & \text{if } m = 0, \\ 0 & \text{if } m > 0. \end{cases} \quad (3.29)$$

For completeness, we provide the gauge identity and generators for $m = 0$ next. Comparing (2.28) and (3.28), we can immediately read off the nonzero generators:

$$\delta_\theta C_\mu = -(\partial_{\mu_1}\theta)(\Omega_\mu)^{\mu_1}, \qquad (\Omega_\mu)^{\mu_1} = -\delta_\mu^{\mu_1}. \quad (3.30)$$

Notice that here we have dropped the, in this case, single-valued β index from (2.28). Putting together (3.14) and (3.30) as indicated in (2.29), we readily confirm the gauge identity:

$$\varrho = \partial_{\mu_1}[E_{(C_\mu)}(\Omega_\mu)^{\mu_1}] \equiv 0. \qquad (3.31)$$

3. Physical degrees of freedom

We have now achieved our goal. Namely, we have shown that the constraint structure characterizing triplet for (3.13) is

$$t_V^{(N)} = \begin{cases} t_M^{(N)} = (\boldsymbol{l} = d(d-1)+1, \boldsymbol{g} = 1, \boldsymbol{e} = 2) & \text{if } m = 0, \\ t_P^{(N)} = (\boldsymbol{l} = d(d-1)+2, \boldsymbol{g} = 0, \boldsymbol{e} = 0) & \text{if } m > 0. \end{cases}$$
$$(3.32)$$

Substituting the quantities (3.32) into the master formula (2.2), we count

$$n_{\text{dof}} = \begin{cases} d-2 & \text{if } m = 0, \\ d-1 & \text{if } m > 0, \end{cases} \qquad (3.33)$$

propagating degrees of freedom. This counting coincides with the ones performed in Secs. III A and III B, where appropriate. We have thus verified (2.6) in two simple examples. Exploiting the equalities in (2.4), we see the following relation to the Hamiltonian side. The massless theory exhibits two first-class constraints, one of which is a primary first-class constraint, and $d(d-1)$ second-class constraints. On the other hand, the massive theory has only second-class constraints, $d(d-1)+2$ of them. Our purely Lagrangian investigation is thus in perfect agreement with the standard Hamiltonian literature, e.g., [36].

IV. A COMPREHENSIVE CONSTRAINT ANALYSIS OF PALATINI THEORIES

In the following, we apply the general framework presented in Sec. II to the Palatini action. We split our calculations into the $d > 2$ and the $d = 2$ cases, as these are physically distinct theories. As we shall see, the former case is much more algebraically involved than the latter. However, compared to their equivalent Hamiltonian investigations, our Lagrangian approach shall prove much simpler in both instances.

For concreteness, we specify our framework to be that of the metric-affine Palatini formulation of general relativity, ordinarily ascribed to Palatini but firstly suggested by Einstein himself [10,37]. As such, we shall study a manifestly first-order formulation of gravity based on $N = d(d+1)^2/2$ number of *a priori* independent degrees of freedom. Even though alternative manifestly first-order formulations do exist, such as the tetradic-Palatini action (for example, see [38] and its recent canonical study [39]),

inconvenient subtleties to our aims arise in those frameworks due to their geometric construction. For instance, unlike the metric, vielbeine are not required to be invertible. In such a scenario, the strict equivalence between Palatini and Einstein's gravity is lost due to a singular vielbein and, in general, ends up in a dynamical manifestation of torsion [40]. Similar situations might arise in other manifestly first-order formulations, like the Barbero-Holst action [41] or background field (BF)-like models [42], which happen to enclose the most celebrated Plebanski action. For a complete review on these topics, we refer the interested reader to [43].

A. Palatini in $d > 2$

The Palatini action in $d > 2$ is a well-known (re)formulation of the Einstein-Hilbert action, which is dynamically equivalent to it. This is explained shortly. Most significantly for us, Palatini is a manifestly first-order formulation of general relativity, which treats the spacetime metric $g_{\mu\nu} = g_{\nu\mu}$ and the affine connection $\Gamma^\rho_{\mu\nu} = \Gamma^\rho_{\nu\mu}$ as *a priori* independent variables. As such, and unlike Einstein-Hilbert, it readily allows for the application of the methodology introduced in Sec. II.

1. Lagrangian constraints

The Palatini action is of the general form given in (2.8) and its Lagrangian density can be written as [44]

$$\mathcal{L}_{\text{Pa}} = -(\partial_\rho h^{\mu\nu})G^\rho_{\mu\nu} + h^{\mu\nu}(cG^\rho_{\rho\mu}G^\sigma_{\sigma\nu} - G^\rho_{\sigma\mu}G^\sigma_{\rho\nu}),$$
$$\text{with} \quad c := \frac{1}{d-1}. \qquad (4.1)$$

Here, the independent variables $h^{\mu\nu}$ and $G^\rho_{\mu\nu}$ are defined exclusively in terms of the spacetime metric and affine connection, respectively:

$$h^{\mu\nu} := \sqrt{-\det(g_{\mu\nu})}g^{\mu\nu}, \qquad G^\rho_{\mu\nu} := \Gamma^\rho_{\mu\nu} - \delta^\rho_{(\mu}\Gamma^\sigma_{\nu)\sigma}, \qquad (4.2)$$

and thus inherit their symmetry properties.

The primary Euler-Lagrange equations for $h^{\mu\nu}$ and $G^\lambda_{\mu\nu}$ following from (4.1) are

$$E_{(h^{\mu\nu})} := \partial_\rho G^\rho_{\mu\nu} + cG^\rho_{\rho\mu}G^\sigma_{\sigma\nu} - G^\rho_{\sigma\mu}G^\sigma_{\rho\nu} \overset{!}{=} 0,$$
$$E_{(G^\rho_{\mu\nu})} := -\partial_\rho h^{\mu\nu} + 2cG^\lambda_{\lambda\sigma}h^{\sigma(\mu}\delta^{\nu)}_\rho - 2G^{(\mu}_{\rho\sigma}h^{\nu)\sigma} \overset{!}{=} 0. \qquad (4.3)$$

Notice that these vanishings are on-shell statements. Multiplying the second set of field equations by $h_{\mu\nu}$ and employing the identity $h_{\mu\nu}h^{\nu\rho} = \delta_\mu{}^\rho$, one finds that

$$2(c-1)G^\mu_{\mu\rho} - h_{\mu\nu}\partial_\rho h^{\mu\nu} \overset{!}{=} 0. \qquad (4.4)$$

Solving (4.4) implies that $G^\rho_{\mu\nu}$ is fixed (on shell) to be a function of $h^{\mu\nu}$ and its first derivatives. The substitution of

the resulting expression into (4.1) yields the second-order formulation of general relativity and we say $d > 2$ Palatini is dynamically equivalent to it.

It is natural and convenient to decompose the variables in (4.2) as follows:

$$h \equiv h^{00}, \qquad h^i \equiv h^{0i}, \qquad G \equiv G_{00}^0,$$

$$G_i \equiv G_{0i}^0, \qquad G_{ij} \equiv G_{ij}^0, \qquad \mathcal{G}^i \equiv G_{00}^i,$$

$$\mathcal{G}_j^i \equiv G_{0j}^i, \qquad \mathcal{G}_{jk}^i \equiv G_{jk}^i. \qquad (4.5)$$

The explicit form of the Lagrangian (4.1) in terms of the above variables is

$$\mathcal{L}_{\mathrm{Pa}} = -\dot{h}G - 2\dot{h}^i G_i - \dot{h}^{ij} G_{ij} - (\partial_i h)\mathcal{G}^i - 2(\partial_i h^j)\mathcal{G}_j^i$$
$$- (\partial_i h^{jk})\mathcal{G}_{jk}^i + h[(c-1)G^2 + 2(cG\mathcal{G}_i^i - G_i\mathcal{G}^i)$$
$$+ c\mathcal{G}_i^i\mathcal{G}_j^j - \mathcal{G}_j^i\mathcal{G}_i^j] + 2h^i[(c-1)GG_i + cG\mathcal{G}_{ij}^j$$
$$+ cG_i\mathcal{G}_j^j - G_j\mathcal{G}_i^j - G_{ij}\mathcal{G}^j + c\mathcal{G}_j^j\mathcal{G}_{ik}^k - \mathcal{G}_j^k\mathcal{G}_{ik}^j]$$
$$+ h^{ij}[(c-1)G_iG_j + 2(cG_{(i}\mathcal{G}_{j)k}^k - G_{k(i}\mathcal{G}_{j)}^k)$$
$$+ c\mathcal{G}_{ik}^k\mathcal{G}_{jl}^l - \mathcal{G}_{il}^k\mathcal{G}_{jk}^l]. \qquad (4.6)$$

We express the generalized coordinates of the Palatini Lagrangian in (4.6) as

$$Q^A = \{h, h^i, h^{ij}, G, G_i, G_{ij}, \mathcal{G}^i, \mathcal{G}_j^i, \mathcal{G}_{jk}^i\}. \qquad (4.7)$$

Notice that A comprises all possible indices of our chosen field variables, so that

$$A = 1, 2, \ldots, d(d+1)^2/2 = \dim(\mathcal{C}) \equiv N. \qquad (4.8)$$

Henceforth, we shall employ the notation $[\cdot]$ introduced in Sec. III C for the collective index A above. In particular, see

(3.16) and explanations around. This notation shall prove of utmost convenience. For instance, in this way, it is obvious that

$$G \equiv [Q^4] \neq Q^4 = \begin{cases} h^{11} & \text{if } d = 3, \\ h^3 & \text{otherwise.} \end{cases} \qquad (4.9)$$

The primary Hessian following from (4.6) vanishes identically: $W_{AB} \equiv 0$, as a result of having promoted the affine connection to a set of *a priori* independent field variables. This parallels the reformulations of classical electromagnetism and the (hard) Proca theory in Sec. III C. In passing, we note that the primary Hessian is symmetric $W_{AB} = W_{BA}$, as it should by definition. Obviously, $\mathrm{rank}(W_{AB}) = 0$ and we have $M_1 = N = d(d+1)^2/2$. This trivialization of the primary Hessian has a number of direct implications. First, it allows us to straightforwardly pick its suitably normalized null vectors to be $(\gamma_I)^A = \delta_I^A$. Second, it immediately makes its Moore-Penrose pseudoinverse vanish as well: $M^{AB} = 0$. Third, it becomes apparent that the primary Euler-Lagrange equations coincide with the primary Lagrangian constraints. All of these constraints turn out to manifestly be functionally independent from each other in this specific theory. In other words, the moduli space is the primary constraint surface in this case and we have $M_1' = M_1$. Thus,

$$0 \stackrel{!}{=} E_A = \alpha_A = (\gamma_I)^A \alpha_A = \varphi_I : \stackrel{!}{\underset{1}{\approx}} 0, \qquad (4.10)$$

exactly as in our examples of Sec. III C before; see (3.15). By means of the notation employed in (3.17), the explicit form of the φ_I's is

$$[\varphi_1] = -[\dot{G} + \partial_i\mathcal{G}^i + \boldsymbol{G}G + 2(cG\mathcal{G}_i^i - G_i\mathcal{G}^i) + \mathcal{G}_j^i\mathcal{G}_i^j],$$

$$\frac{[\varphi_2]_i}{2} = -[\dot{G}_i + \partial_j\mathcal{G}_i^j + \boldsymbol{G}G_i + cG\mathcal{G}_{ij}^j + G_j\mathcal{G}_i^j - G_{ij}\mathcal{G}^j + \mathcal{G}_k^j\mathcal{G}_{ij}^k],$$

$$[\varphi_3]_{ij} = -[\dot{G}_{ij} + \partial_k\mathcal{G}_{ij}^k + \boldsymbol{G}_iG_j + 2(cG_{(i}\mathcal{G}_{j)k}^k - G_{k(i}\mathcal{G}_{j)}^k) + \mathcal{G}_{ik}^l\mathcal{G}_{lj}^k],$$

$$[\varphi_4] = \dot{h} - 2[h\boldsymbol{G} + ch\mathcal{G}_i^i + h^i\boldsymbol{G}_i + ch^i\mathcal{G}_{ij}^j],$$

$$\frac{[\varphi_5]^i}{2} = \dot{h}^i + h\mathcal{G}^i - [h^i\boldsymbol{G} + h^j\boldsymbol{\mathcal{G}}_j^i + h^{ij}\boldsymbol{G}_j + ch^{ij}\mathcal{G}_{jk}^k],$$

$$[\varphi_6]^{ij} = \dot{h}^{ij} + 2(h^{(i}\mathcal{G}^{j)} + h^{k(i}\mathcal{G}_k^{j)}),$$

$$[\varphi_7]_i = \partial_i h + 2(hG_i + h^jG_{ij}),$$

$$\frac{[\varphi_8]_i^j}{2} = \partial_i h^j - [c(hG + h^kG_k)\delta_i^j + h\mathcal{G}_i^j + h^k\mathcal{G}_{ik}^j] + h^jG_i + h^{jk}G_{ik},$$

$$[\varphi_9]_i^{jk} = \partial_i h^{jk} - 2[h^{(j}(cG\delta_i^{k)} + \mathcal{G}_i^{k)}) + h^{l(j}(cG_l\delta_i^{k)} + \mathcal{G}_{il}^{k)})], \qquad (4.11)$$

where we have defined

$$\boldsymbol{G} := (c-1)G, \qquad \boldsymbol{G}_i := (c-1)G_i,$$
$$\boldsymbol{\mathcal{G}}_j^i := c\mathcal{G}_k^k\delta_j^i - \mathcal{G}_j^i, \qquad \boldsymbol{\mathcal{G}}_{jk}^i := c\mathcal{G}_{kl}^l\delta_j^i - \mathcal{G}_{jk}^i. \qquad (4.12)$$

Note that the φ_I's in (4.11) are manifestly symmetric where appropriate, i.e., $[\varphi_3]_{ij} = [\varphi_3]_{ji}$, $[\varphi_6]^{ij} = [\varphi_6]^{ji}$ and $[\varphi_9]_i^{jk} = [\varphi_9]_i^{kj}$.

We now turn to the secondary stage, where we inspect the consistency under time evolution of the functionally independent primary Lagrangian constraints. The secondary Hessian is given by $\tilde{W}_{IJ} = \partial_I\varphi_J$. With the conventions introduced below (3.18), we may succinctly write it as

$$\tilde{W}_{IJ} = \begin{pmatrix} [\tilde{W}_{11}] & \cdots & [\tilde{W}_{19}]_i^{jk} \\ \vdots & \ddots & \vdots \\ {}_i^{jk}[\tilde{W}_{91}] & \cdots & {}_i^{jk}[\tilde{W}_{99}]_l^{mn} \end{pmatrix}. \qquad (4.13)$$

For clarity, we provide a few examples of what is meant by our notation:

$$[\tilde{W}_{11}] \equiv \frac{\partial[\varphi_1]}{\partial\dot{h}}, \qquad {}_i[\tilde{W}_{23}]_{jk} \equiv \frac{\partial[\varphi_3]_{jk}}{\partial\dot{h}^i},$$
$$[\tilde{W}_{45}]^i \equiv \frac{\partial[\varphi_5]^i}{\partial\dot{G}}, \qquad {}_i^{jk}[\tilde{W}_{99}]_l^{mn} \equiv \frac{\partial[\varphi_9]_l^{mn}}{\partial\dot{\mathcal{G}}_{jk}^i}. \qquad (4.14)$$

In (4.13), the only nonzero components are

$$[\tilde{W}_{14}] = -1 = -[\tilde{W}_{41}], \quad {}_i[\tilde{W}_{25}]^j = -2\delta_i^j = -{}^j[\tilde{W}_{52}]_i,$$
$${}_{ij}[\tilde{W}_{36}]^{kl} = -\delta_{(i}^k\delta_{j)}^l = -{}^{kl}[\tilde{W}_{63}]_{ij}. \qquad (4.15)$$

Notice that the secondary Hessian is antisymmetric $\tilde{W}_{IJ} = -\tilde{W}_{JI}$, as it should by definition. It is easy to see that $\mathcal{R}_2 := \mathrm{rank}(\tilde{W}_{IJ}) = d(d+1)$, thus yielding $M_2 = d(d^2-1)/2$. This means that $\mathcal{R}_2$ number of the functionally independent primary Lagrangian constraints are being dynamically stabilized at the secondary stage, while the remaining M_2 primary Lagrangian constraints are not stable: they lead to secondary Lagrangian constraints, which we proceed to determine.

To this aim, we first choose the suitably normalized linearly independent null vectors associated to the secondary Hessian as

$$(\tilde{\gamma}_R)^I = (0, 0, ..., 0, 1, 0, 0, ..., 0), \qquad (4.16)$$

where the nonvanishing vector component is at $I = \mathcal{R}_2 + R$. Notice that all M_2 null vectors have length M_1 and their first $\mathcal{R}_2$ components are zero.

All our results so far can be used to obtain the secondary Lagrangian constraints $\tilde{\varphi}_R \underset{1}{\approx} 0$. These can readily be seen to be functionally independent from each other, as well as with respect to the primary constraints, so that $M_2' = M_2$. This means their vanishing defines the secondary constraint surface: $\tilde{\varphi}_R \overset{!}{\underset{2}{\approx}} 0$. In our $[\cdot]$ notation, we have

$$\frac{[\tilde{\varphi}_1]_i}{2} \underset{1}{\approx} h(\tilde{\tau}_1)_i + h^j(\tilde{\tau}_2)_{ij},$$
$$\frac{[\tilde{\varphi}_2]_i^j}{2} \underset{1}{\approx} h^j(\tilde{\tau}_1)_i + h^{jk}(\tilde{\tau}_2)_{ik} + h(\tilde{\tau}_3)_i^j + h^k(\tilde{\tau}_4)_{ik}^j,$$
$$\frac{[\tilde{\varphi}_3]_i^{jk}}{2} \underset{1}{\approx} h^{(j}(\tilde{\tau}_3)_i^{k)} + h^{l(j}(\tilde{\tau}_4)_{il}^{k)}, \qquad (4.17)$$

where we have defined

$$(\tilde{\tau}_1)_i := \boldsymbol{\partial}_i G + \partial_{k\cdot i}^l \mathcal{G}_l^k + \boldsymbol{G}\mathbb{G}_i - \boldsymbol{G}G_i + (c-2)G_k\boldsymbol{\mathcal{G}}_i^k + \mathcal{G}_l^k\tilde{\boldsymbol{\mathcal{G}}}_{ik}^l,$$
$$(\tilde{\tau}_2)_{ij} := \boldsymbol{\partial}_i G_j + \partial_{k\cdot i}^l \mathcal{G}_{lj}^k + \boldsymbol{G}_j\mathbb{G}_i - \boldsymbol{G}G_{ij} + \boldsymbol{G}_k\boldsymbol{\mathcal{G}}_{ij}^k - G_{jk}\boldsymbol{\mathcal{G}}_i^k + \mathcal{G}_{jl}^k\tilde{\boldsymbol{\mathcal{G}}}_{ik}^l,$$
$$(\tilde{\tau}_3)_i^j := -\delta_{i\cdot m}^{j\cdot l}\dot{\mathcal{G}}_l^m + \partial_{i\cdot l}^j \mathcal{G}^l + \boldsymbol{G}\boldsymbol{\mathcal{G}}_i^j - G_k\boldsymbol{\mathcal{G}}_i^{k\cdot j} + \mathcal{G}^l\boldsymbol{\mathcal{G}}_{il}^j - \mathcal{G}_l^m\boldsymbol{\mathcal{G}}_{m\cdot i}^{l\cdot j},$$
$$(\tilde{\tau}_4)_{ik}^j := -\delta_{i\cdot m}^{j\cdot l}\dot{\mathcal{G}}_{lk}^m + \partial_{i\cdot l}^j \mathcal{G}_k^l - G_{kl}\boldsymbol{\mathcal{G}}_i^{l\cdot j} + \mathcal{G}_k^l\boldsymbol{\mathcal{G}}_{il}^j - \mathcal{G}_{kl}^m\boldsymbol{\mathcal{G}}_{m\cdot i}^{l\cdot j} \qquad (4.18)$$

in terms of (4.12) as well as the following quantities:

$$\boldsymbol{\partial}_i := (c-1)\partial_i, \qquad \partial_{i\cdot j}^k := c\delta_i^k\partial_j - \delta_j^k\partial_i, \qquad \delta_{i\cdot j}^{k\cdot l} := c\delta_i^k\delta_j^l - \delta_j^k\delta_i^l, \qquad \mathbb{G}_i := c(G_i + \mathcal{G}_{ik}^k),$$
$$\boldsymbol{\mathcal{G}}_{k\cdot l}^{i\cdot j} := c\mathcal{G}_k^i\delta_l^j - \mathcal{G}_k^j\delta_l^i, \qquad \boldsymbol{\mathcal{G}}_k^{i\cdot j} := c\mathcal{G}^i\delta_k^j - \mathcal{G}^j\delta_k^i, \qquad \tilde{\boldsymbol{\mathcal{G}}}_{ij}^k := c(c-1)\mathcal{G}_{il}^l\delta_j^k - \mathcal{G}_{ij}^k. \qquad (4.19)$$

Observe that the appropriate symmetry $[\tilde{\varphi}_3]_i^{jk} = [\tilde{\varphi}_3]_i^{kj}$ is manifest. To obtain the above, we have first computed $\tilde{\varphi}_R = (\tilde{\gamma}_R)^I\dot{\varphi}_I$. Then, we have evaluated the result in the first constraint surface. In practice, this means that we have substituted $(\dot{G}, \dot{G}_i, ..., \partial_i h^{jk})$ for their suitable weak

expressions in terms of the generalized coordinates Q^A, which follow from setting to zero (4.11).

To conclude the secondary stage, we calculate the Moore-Penrose pseudoinverse of $\tilde{W}_{IJ}$. It can be easily checked that this is $\tilde{M}^{IJ} = -(\tilde{W}_{IJ})^T$.

Next, the consistency under time evolution of the above functionally independent secondary Lagrangian constraints is to be inspected at the tertiary stage. The first step is to calculate the tertiary Hessian $\hat{W}_{RS} = (\tilde{\gamma}_R)^I \partial_i \tilde{\varphi}_S$. Employing the same conventions as in (4.13) before, we write

$$\hat{W}_{RS} = \begin{pmatrix} {}_i[\hat{W}_{11}]_j & {}_i[\hat{W}_{12}]_j^k & {}_i[\hat{W}_{13}]_j^{kl} \\ {}_i^j[\hat{W}_{21}]_l & {}_i^j[\hat{W}_{22}]_l^m & {}_i^j[\hat{W}_{23}]_l^{mn} \\ {}_i^{jk}[\hat{W}_{31}]_l & {}_i^{jk}[\hat{W}_{32}]_l^m & {}_i^{jk}[\hat{W}_{33}]_l^{mn} \end{pmatrix}, \qquad (4.20)$$

where the nonzero components are

$${}_i^j[\hat{W}_{22}]_k^l \equiv \frac{\partial[\tilde{\varphi}_2]_k^l}{\partial \dot{\mathcal{G}}_j^i} = -2h(c\delta_i^j \delta_k^l - \delta_i^l \delta_k^j),$$

$${}_i^{jk}[\hat{W}_{33}]_l^{mn} \equiv \frac{\partial[\tilde{\varphi}_3]_l^{mn}}{\partial \dot{\mathcal{G}}_{jk}^i} = 2(\delta_i^{(m} h^{n)(j} \delta_l^{k)} - c\delta_l^{(m} h^{n)(j} \delta_i^{k)}),$$

$${}_i^j[\hat{W}_{23}]_k^{lm} \equiv \frac{\partial[\tilde{\varphi}_3]_k^{lm}}{\partial \dot{\mathcal{G}}_j^i} = -2h^{(l}(c\delta_i^j \delta_k^{m)} - \delta_i^{m)} \delta_k^j)$$

$$= \frac{\partial[\tilde{\varphi}_2]_i^j}{\partial \dot{\mathcal{G}}_{lm}^k} \equiv {}_k^{lm}[\hat{W}_{32}]_i^j. \qquad (4.21)$$

Therefore, the tertiary Hessian obviously satisfies componentwise the symmetry properties that ensure $\hat{W}_{RS} = \hat{W}_{SR}$, as required by definition. It is not hard to see that $\mathcal{R}_3 := \text{rank}(\hat{W}_{RS}) = d(d^2 - 3)/2$ and hence $M_3 = d$. In more detail, the rank is equal to all nonzero rows (equivalently, columns) of the Hessian, minus one. There is the obvious set of $d - 1$ number of zero rows given by ${}_i[\hat{W}_{1R} = 0$. But the rank of the Hessian is further reduced by one because the linear combination of rows ${}_i^i[\hat{W}_{2R}$ is zero. This vanishing is a direct consequence of the velocity independence of $[\tilde{\varphi}_2]_i^i$ in TC_1, as can be readily verified from (4.17) and (4.18). It follows that $\mathcal{R}_3$ number of the functionally independent secondary Lagrangian constraints are being dynamically stabilized at the tertiary stage. The remaining M_3 secondary Lagrangian constraints are not stable: they lead to the tertiary Lagrangian constraints that we shall find next.

The suitably normalized linearly independent null vectors of the tertiary Hessian can be chosen as follows. Associated to ${}_i[\hat{W}_{1R} = 0$, we pick

$$(\hat{\gamma}_U)^R = (0, 0, …, 0, 1, 0, 0, …, 0),$$
$$\text{with} \quad U = 1, 2, …, M_3 - 1, \qquad (4.22)$$

where the nonvanishing vector component is at $R = U$. These null vectors have length M_2 and their last $\mathcal{R}_3$ components are all zero. Corresponding to ${}_i^i[\hat{W}_{2R} = 0$, we select the null vector

$$(\hat{\gamma}_{U=M_3})^R = \frac{1}{\sqrt{d-1}}(0, …, 0, 1, 0, …, 0, 1, 0, …, 0), \qquad (4.23)$$

where the $(d - 1)$ number of nonvanishing vector components are at $R = M_3, 2M_3, …, (d - 1)M_3$.[11]

The tertiary Lagrangian constraints are given by the requirement $\hat{\varphi}_U = (\hat{\gamma}_U)^R \dot{\tilde{\varphi}}_R \overset{!}{\underset{2}{\approx}} 0$. In this case, the derivation with respect to time is particularly simple and coincides with the naively expected one, so that

$$\hat{\varphi}_U = (\hat{\gamma}_U)^R [\partial_i \dot{Q}^A \partial_A^i + \dot{Q}^A \partial_A] \tilde{\varphi}_R \overset{!}{\underset{2}{\approx}} 0. \qquad (4.24)$$

In our shorthand notation, we find it convenient to express these constraints as follows:

$$\frac{[\hat{\varphi}_1]_i}{2} = \dot{h}(\tilde{\tau}_1)_i + \dot{h}^j(\tilde{\tau}_2)_{ij} + \mathcal{O}[h(\tilde{\tau}_1)_i + h^j(\tilde{\tau}_2)_{ij}],$$

$$\frac{[\hat{\varphi}_2]}{2} = \dot{h}^i(\tilde{\tau}_1)_i + \dot{h}^{ij}(\tilde{\tau}_2)_{ij} + \mathcal{O}[h^i(\tilde{\tau}_1)_i + h^{ij}(\tilde{\tau}_2)_{ij}], \qquad (4.25)$$

where the operator $\mathcal{O}$ is defined as

$$\mathcal{O} := (\partial_k \dot{G}) \frac{\partial}{\partial(\partial_k G)} + (\partial_k \dot{G}_l) \frac{\partial}{\partial(\partial_k G_l)} + (\partial_k \dot{\mathcal{G}}_m^l) \frac{\partial}{\partial(\partial_k \mathcal{G}_m^l)}$$
$$+ (\partial_k \dot{\mathcal{G}}_{mn}^l) \frac{\partial}{\partial(\partial_k \mathcal{G}_{mn}^l)} + \dot{G} \frac{\partial}{\partial G} + \dot{G}_k \frac{\partial}{\partial G_k}$$
$$+ \dot{G}_{kl} \frac{\partial}{\partial G_{kl}} + \dot{\mathcal{G}}_l^k \frac{\partial}{\partial \mathcal{G}_l^k} + \dot{\mathcal{G}}_{lm}^k \frac{\partial}{\partial \mathcal{G}_{lm}^k}. \qquad (4.26)$$

Recall that $(\tilde{\tau}_1)_i$ and $(\tilde{\tau}_2)_{ij}$ are as introduced in (4.18).

Following the procedure described under (4.19), the tertiary Lagrangian constraints in (4.25) can be evaluated on the first constraint surface TC_1. After tedious algebraic manipulations, the weak tertiary Lagrangian constraints can be written exclusively in terms of the functionally independent secondary constraints as

$$[\hat{\varphi}_1]_i \underset{1}{\approx} - (\delta_i^j G - 2\mathcal{G}_i^j)[\tilde{\varphi}_1]_j - [\delta_i^j(G_k + \partial_k)$$
$$- \frac{1}{2}\delta_k^j\partial_i - \mathcal{G}_{ik}^j][\tilde{\varphi}_2]_j^k - G_{jk}[\tilde{\varphi}_3]_i^{jk},$$

$$\frac{[\hat{\varphi}_2]}{2} \underset{1}{\approx} 2\mathcal{G}^i[\tilde{\varphi}_1]_i + \mathcal{G}_i^j[\tilde{\varphi}_2]_j^i + \partial_j[\tilde{\varphi}_3]_i^{ij}. \qquad (4.27)$$

The above is a nontrivial result. Indeed, it becomes increasingly computationally challenging to evaluate Lagrangian constraints on constraint surfaces as one goes to higher stages. We elaborate on this topic and advice on how to handle the evaluations in Sec. VI.

---

[11]To determine the null vectors of $\hat{W}_{RS}$ in (4.20), we considered the ansatz $(\hat{\gamma}_U)^R = (a^i, a_j^i, a_{jk}^i)$, with $a_{jk}^i = a_{kj}^i$. Then, the equation $(\hat{\gamma}_U)^R \hat{W}_{RS} = 0$ results in $a_i^j = ca_l^j\delta_i^l$ and $a_{il}^j = ca_{kl}^k\delta_i^j$ but does not impose any condition on a^i. The first equation implies $a_j^i = 0$ for $i \neq j$. Setting $j = l$ in the second equation yields $a_{ij}^i = 0$, which in turn implies $a_{jk}^i = 0$ for all i, j, k.

Our results in (4.27) must be further evaluated on the second constraint surface TC_2, namely, in the subspace of TC_1 defined by the vanishing of (4.17). We thus see that

$$[\hat{\varphi}_1]_i \underset{2}{\approx} 0, \qquad [\hat{\varphi}_2] \underset{2}{\approx} 0, \qquad (4.28)$$

which implies $TC_3 \equiv TC_2$ and there are no functionally independent tertiary constraints $M_3' = 0$. Consequently, the algorithm closes nondynamically, according to case (iib).

We are finally able to obtain the result of interest from the analysis here presented. The number of functionally independent Lagrangian constraints for the Palatini theory in $d > 2$, when described in terms of $N = d(d + 1)^2/2$ number of *a priori* independent field variables, is equal to

$$l = M_1' + M_2' + M_3' = M_1 + M_2 + 0$$
$$= \frac{d}{2}(d + 1)^2 + \frac{d}{2}(d^2 - 1) + 0 = d^2(d + 1). \quad (4.29)$$

2. Gauge identities

It is well known (for instance, see [14]) that the Palatini action corresponding to the Lagrangian density (4.1) remains invariant under the following transformations of its independent variables: $h^{\mu\nu} \to h^{\mu\nu} + \delta_\theta h^{\mu\nu}$ and $G^\rho_{\mu\nu} \to G^\rho_{\mu\nu} + \delta_\theta G^\rho_{\mu\nu}$, with

$$\delta_\theta h^{\mu\nu} = 2h^{\rho(\mu}\partial_\rho\theta^{\nu)} - \partial_\rho(h^{\mu\nu}\theta^\rho),$$
$$\delta_\theta G^\rho_{\mu\nu} = -\partial_\mu\partial_\nu\theta^\rho + \delta^\rho_{(\mu}\partial_{\nu)}\partial_\sigma\theta^\sigma - \theta^\sigma\partial_\sigma G^\rho_{\mu\nu}$$
$$+ G^\sigma_{\mu\nu}\partial_\sigma\theta^\rho - 2G^\rho_{\sigma(\mu}\partial_{\nu)}\theta^\sigma, \quad (4.30)$$

where θ^μ are the (unspecified) gauge parameters. Notice that the pertinent symmetries $\delta_\theta h^{\mu\nu} = \delta_\theta h^{\nu\mu}$ and $\delta_\theta G^\rho_{\mu\nu} = \delta_\theta G^\rho_{\nu\mu}$ are apparent in the precedent expressions. Of course, (4.30) is just the Palatini (re)formulation of the renowned diffeomorphism invariance of the Einstein-Hilbert action. This holds true off shell.

It is easy to see in (4.30) that the gauge parameters θ^μ appear explicitly in all the gauge transformations $\forall \mu$. Similarly, we note that the effective gauge parameters $(\theta^\mu, \dot{\theta}^\mu, \ddot{\theta}^\mu)$ are manifestly present in the gauge transformations $\forall \mu$ as well. By definition, it follows that

$$g = d, \qquad e = 3d, \qquad (4.31)$$

which are the off-shell parameters we aimed to obtain in this short analysis.

For completeness, we provide the gauge generators and confirm the gauge identities of $d > 2$ Palatini next. A direct comparison between (2.28) and (4.30) allows us to rewrite the latter as

$$\delta_\theta h^{\mu\nu} = \theta^\beta[(\Omega_\beta)^{\mu\nu}] - (\partial_{\mu_1}\theta^\beta)[(\Omega_\beta)^{\mu\nu}]^{\mu_1},$$
$$\delta_\theta G^\rho_{\mu\nu} = \theta^\beta[(\Omega_\beta)^\rho_{\mu\nu}] - (\partial_{\mu_1}\theta^\beta)[(\Omega_\beta)^\rho_{\mu\nu}]^{\mu_1}$$
$$+ (\partial_{\mu_1}\partial_{\mu_2}\theta^\beta)[(\Omega_\beta)^\rho_{\mu\nu}]^{\mu_1\mu_2}, \quad (4.32)$$

where we have introduced a bracket $(\cdot)$ to visually split the (in general collective) indices β and A for latter convenience. In view of these transformations, the gauge generators can easily be identified to be

$$[(\Omega_\beta)^{\mu\nu}] = -\partial_\beta h^{\mu\nu}, \qquad [(\Omega_\beta)^{\mu\nu}]^{\mu_1} = -2h^{\mu_1(\mu}\delta^{\nu)}_\beta + h^{\mu\nu}\delta^{\mu_1}_\beta,$$
$$[(\Omega_\beta)^\rho_{\mu\nu}] = -\partial_\beta G^\rho_{\mu\nu}, \qquad [(\Omega_\beta)^\rho_{\mu\nu}]^{\mu_1} = 2G^\rho_{\beta(\mu}\delta^{\mu_1}_{\nu)} - G^{\mu_1}_{\mu\nu}\delta^\rho_\beta,$$
$$[(\Omega_\beta)^\rho_{\mu\nu}]^{\mu_1\mu_2} = \delta^\rho_{(\mu}\delta^{(\mu_1}_{\nu)}\delta^{\mu_2)}_\beta - \delta^{\mu_1}_{(\mu}\delta^{\mu_2}_{\nu)}\delta^\rho_\beta. \quad (4.33)$$

Combining (4.3) with the above as prescribed in (2.29) and working through, the gauge identities are obtained:

$$\rho_\beta = E_{(h^{\mu\nu})}[(\Omega_\beta)^{\mu\nu}] + \partial_{\mu_1}(E_{(h^{\mu\nu})}[(\Omega_\beta)^{\mu\nu}]^{\mu_1})$$
$$+ E_{(G^\rho_{\mu\nu})}[(\Omega_\beta)^\rho_{\mu\nu}] + \partial_{\mu_1}(E_{(G^\rho_{\mu\nu})}[(\Omega_\beta)^\rho_{\mu\nu}]^{\mu_1})$$
$$+ \partial_{\mu_1}\partial_{\mu_2}(E_{(G^\rho_{\mu\nu})}[(\Omega_\beta)^\rho_{\mu\nu}]^{\mu_1\mu_2}) \equiv 0. \quad (4.34)$$

3. Physical degrees of freedom

Putting everything together, we can finally count the number of propagating modes present in the theory. Namely, employing (4.8), (4.29) and (4.31) in the master formula (2.2), we get

$$n_{\text{dof}} = \frac{d}{2}(d - 3). \qquad (4.35)$$

When $d = 4$, we have that $n_{\text{dof}} = 2$, corresponding to the two massless tensor's polarizations of the graviton. For $d = 3$, the widely known triviality is recovered, with no physical degrees of freedom being propagated.

Our result is in perfect agreement with the counting performed in [13,14], where a purely Hamiltonian analysis was done. We have thus carried out another (nontrivial) explicit verification of the already noted equivalence between (2.2) and (2.3). This equivalence can be further verified as follows. It is explicitly shown in [13,14] that $N_1 = 3d$, $N_2 = d(d - 1)(d + 2)$ and $N_1^{(P)} = d$ for the $d > 2$ Palatini theory when (4.8) holds true. Substitution of these results in (2.4) readily confirms our own counting in (4.29) and (4.31). Besides, a direct comparison between the calculations in [13,14] and those presented in this Sec. IV A unequivocally shows that our purely Lagrangian computation is an algebraically much simpler way to derive (4.36), from which the number of physical modes follows readily.

To sum up, we have derived the constraint structure characterizing triplet $t_{\text{Pa}}^{(N)}$, with $N = d(d + 1)^2/2$, of the

Palatini theory in $d > 2$ dimensions in a purely Lagrangian approach and ratified its equivalence with a representative Hamiltonian analysis performed in the past. Mathematically,

$$t_{\text{Pa}}^{(N)} = (l = d^2(d+1), g = d, e = 3d) \quad (4.36)$$

in the Lagrangian picture, while

$$t_{\text{Pa}}^{(N)} = (N_1^{(\text{P})} = d, N_1 = 3d, N_2 = d(d-1)(d+2)) \quad (4.37)$$

in the Hamiltonian side—recall (2.5)—both of which imply (4.35).

B. A special case: Palatini in $d = 2$

General relativity, in its standard second-order formulation, behaves drastically different in two dimensions. Specifically, it can be shown that

$$S_{\text{EH}} = \int_{\mathcal{M}_2} d^2x \sqrt{-g} R \propto \chi(\mathcal{M}_2). \quad (4.38)$$

Namely, the Einstein-Hilbert action is proportional to the Euler characteristic χ of the spacetime manifold $\mathcal{M}_2$; see e.g., [45]. The above implies that general relativity is a topological theory in $d = 2$ and, accordingly, propagates no degrees of freedom, a fact that we shall explicitly verify in the following.

Turning to the Palatini Lagrangian in (4.1) for $d = 2$, we restate that this is not dynamically equivalent to two-dimensional Einstein's gravity (4.38). To see this, consider its corresponding Euler-Lagrange equations in (4.3). These are valid for $d \geq 2$. However, recall that $c := (d-1)^{-1}$, so that $c = 1$ in two dimensions. In this particular case, it is obvious that (4.4) cannot be solved as we said, i.e.,

$$G_{\mu\nu}^{\rho} \overset{!}{=} G_{\mu\nu}^{\rho}(h^{\mu\nu}, \partial_\rho h^{\mu\nu}). \quad \text{As a result, the dynamical}$$

equivalence to Einstein's gravity is lost. A more general yet detailed argumentation can be found in [46].

Correspondingly, the dynamics of the two-dimensional Palatini action does not constitute a smooth limit of its higher-dimensional counterpart. Namely, the Lagrangian (4.1) in $d = 2$ does not describe the evolution of the same family of fields as that very same Lagrangian in $d > 2$: these are two physically different theories. The easiest way to ratify this second inequivalence is to note that the counting of degrees of freedom in (4.35), when we set $d = 2$, yields a negative number of propagating modes, which is an unphysical result. Thus, a different constraint structure characterizing triplet

$$t_{2\text{Pa}}^{(N=9)} \neq \lim_{d \to 2} t_{\text{Pa}}^{(N)}, \quad \text{where } N = d(d+1)^2/2, \quad (4.39)$$

is then to be expected. We proceed to determine this $t_{2\text{Pa}}^{(N=9)}$ next.

1. Lagrangian constraints

As a starting point, we express the generalized coordinates of the Palatini theory in $d = 2$ as

$$Q^A = \{h, h^1, h^{11}, G, G_1, G_{11}, \mathcal{G}^1, \mathcal{G}_1^1, \mathcal{G}_{11}^1\},$$
$$A = 1, 2, \ldots, 9 = \dim(\mathcal{C}) \equiv N, \quad (4.40)$$

in direct analogy to (4.7) earlier on. Next, we compute the first stage quantities associated to the $d = 2$ version of the Palatini Lagrangian (4.6). One can verify that the set of primary Lagrangian constraints thus obtained matches the consistent $d = 2$ evaluation of those for a generic dimension in (4.10) and (4.11). Comparatively, these two-dimensional constraints have a much simpler form, given by the vanishing of

$$[\varphi_1] = -[\dot{G} + \partial_1 \mathcal{G}^1 + 2(G\mathcal{G}_1^1 - G_1 \mathcal{G}^1)], \qquad \frac{[\varphi_2]_1}{2} = -[\dot{G}_1 + \partial_1 \mathcal{G}_1^1 + G\mathcal{G}_{11}^1 - G_{11}\mathcal{G}^1],$$

$$[\varphi_3]_{11} = -[\dot{G}_{11} + \partial_1 \mathcal{G}_{11}^1 + 2(G_1 \mathcal{G}_{11}^1 - G_{11}\mathcal{G}_1^1)], \qquad [\varphi_4] = \dot{h} - 2(h\mathcal{G}_1^1 + h^1 \mathcal{G}_{11}^1),$$

$$\frac{[\varphi_5]^1}{2} = \dot{h}^1 + h\mathcal{G}^1 - h^{11}\mathcal{G}_{11}^1, \qquad [\varphi_6]^{11} = \dot{h}^{11} + 2(h^1 \mathcal{G}^1 + h^{11}\mathcal{G}_1^1),$$

$$[\varphi_7]_1 = \partial_1 h + 2(hG_1 + h^1 G_{11}), \qquad \frac{[\varphi_8]_1^1}{2} = \partial_1 h^1 - hG + h^{11}G_{11}, \quad [\varphi_9]_1^{11} = \partial_1 h^{11} - 2(h^1 G + h^{11}G_1). \quad (4.41)$$

The demand that the above be zero constitutes a set of nine scalar primary Lagrangian constraints ($M_1 = 9$), whose functional independence is rather obvious ($M_1' = 9$)—and can be ratified through the Jacobian test in (2.17). Therefore, such vanishing defines the primary constraint surface $T\mathcal{C}_1$ of the theory, which coincides

with the moduli space due to the primary Hessian being zero, as in the $d > 2$ case before. In other words, (4.10) holds true here as well.

The progress to the subsequent stage parallels that of the $d > 2$ case. The secondary Hessian is given by the skew-symmetric constant matrix

$$\tilde{W}_{IJ} = \begin{pmatrix} 0 & -\omega & 0 \\ \omega & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \text{with} \quad \omega := \text{diag}(1,2,1).$$

(4.42)

The Hessian (4.42) has rank $\mathcal{R}_2 = 6$ and so $M_2 = 3$. This means that six of the primary Lagrangian constraints are dynamically stabilized by the (functionally independent) secondary equations of motion. For the remaining three primary Lagrangian constraints, the algorithm must be pursued.

We choose the suitably normalized linearly independent null vectors of (4.42) as

$$(\tilde{\gamma}_R)^I = \delta_{R+6}{}^I, \quad \text{with} \quad R = 1,2,3. \quad (4.43)$$

Using (4.41) and (4.43), we obtain the three secondary Lagrangian constraints as the vanishing of

$$[\tilde{\varphi}_1]_1 = \partial_1 \dot{h} + 2(\dot{h}G_1 + \dot{h}^1 G_{11} + h\dot{G}_1 + h^1\dot{G}_{11}),$$
$$[\tilde{\varphi}_2]_1^1 = 2(\partial_1 \dot{h}^1 - \dot{h}G + \dot{h}^{11}G_{11} - h\dot{G} + h^{11}\dot{G}_{11}),$$
$$[\tilde{\varphi}_3]_1^{11} = \partial_1 \dot{h}^{11} - 2(\dot{h}^1 G + \dot{h}^{11}G_1 + h^1\dot{G} + h^{11}\dot{G}_1). \quad (4.44)$$

Notice that the above are the total time derivatives of $[\varphi_7]_1$, $[\varphi_8]_1^1$ and $[\varphi_9]_1^{11}$ in (4.41), respectively. It is easy to check that the secondary Lagrangian constraints are functionally dependent on the primary Lagrangian constraints. Specifically,

$$[\tilde{\varphi}_1]_1 = \partial_1[\varphi_4] - h[\varphi_2]_1 - 2h^1[\varphi_3]_{11} + 2G_1[\varphi_4]$$
$$\quad + G_{11}[\varphi_5]^1 + 2\mathcal{G}_1^1[\varphi_7]_1 + \mathcal{G}_{11}^1[\varphi_8]_1^1,$$
$$[\tilde{\varphi}_2]_1^1 = \partial_1[\varphi_5]^1 + 2(h[\varphi_1] - h^{11}[\varphi_3]_{11} - G[\varphi_4] + G_{11}[\varphi_6]^{11}$$
$$\quad - \mathcal{G}^1[\varphi_7]_1 + \mathcal{G}_{11}^1[\varphi_9]_1^{11}),$$
$$[\tilde{\varphi}_3]_1^{11} = \partial_1[\varphi_6]^{11} + 2h^1[\varphi_1] + h^{11}[\varphi_2]_1$$
$$\quad - (G[\varphi_5]^1 + 2G_1[\varphi_6]^{11} + \mathcal{G}^1[\varphi_8]_1^1 + 2\mathcal{G}_1^1[\varphi_9]_1^{11}).$$

(4.45)

Therefore, the secondary constraints vanish on $T\mathcal{C}_1$:

$$[\tilde{\varphi}_1]_1 \underset{1}{\approx} 0, \qquad [\tilde{\varphi}_2]_1^1 \underset{1}{\approx} 0, \qquad [\tilde{\varphi}_3]_1^{11} \underset{1}{\approx} 0, \quad (4.46)$$

and so $T\mathcal{C}_2 \equiv T\mathcal{C}_1$. This in turn implies that there are no functionally independent secondary constraints: $M_2' = 0$. Here, the algorithm closes nondynamically, as described in case (iib). It follows that the total number of functionally independent Lagrangian constraints is $l = M_1' + M_2' = 9$. We note that this result does not correspond to setting $d = 2$ in (4.29).

2. Gauge identities

Given the already pointed out inequivalence between the $d > 2$ and $d = 2$ Palatini theories, it is not too surprising that the gauge transformations (4.30) preserving the action (4.6) meet a nonsmooth limit for $d = 2$. The argument is more subtle than that of the purely on-shell inequivalence; for example, see [11]. We shall touch upon it shortly.

It has been proven, e.g., [11,47], that the two-dimensional Palatini action is invariant under the field transformations $h^{\mu\nu} \to h^{\mu\nu} + \delta_\theta h^{\mu\nu}$ and $G^\rho_{\mu\nu} \to G^\rho_{\mu\nu} + \delta_\theta G^\rho_{\mu\nu}$, with

$$\delta_\theta h^{\mu\nu} = 2\epsilon^{\rho(\mu} h^{\nu)\sigma} \theta_{\rho\sigma},$$
$$\delta_\theta G^\rho_{\mu\nu} = \epsilon^{\rho\sigma} \partial_\sigma \theta_{\mu\nu} + 2\epsilon^{\sigma\lambda} G^\rho_{\sigma(\mu} \theta_{\nu)\lambda}. \quad (4.47)$$

Here, $\epsilon^{\mu\nu}$ stands for the two-dimensional Levi-Civita symbol (we work with the convention $\epsilon^{01} = 1$) and $\theta_{\mu\nu} = \theta_{\nu\mu}$, so there are three arbitrary gauge parameters that characterize the transformation. It is obvious that all the gauge parameters explicitly appear in the gauge transformations. Hence, $g = 3$. Their first time derivatives also show up, adding to a total number of effective gauge parameters $e = 6$. We point out that g does not match the value predicted in (4.31) for $d = 2$. There is a match for e, but this is purely coincidental.

These numbers ($g = 3$, $e = 6$), in contrast to the naively expected ones ($g = 2$, $e = 6$) from the diffeomorphism transformations (4.31) in $d > 2$ Palatini, reflect the fact that $d = 2$ Palatini is associated to a comparatively larger symmetry group. Its connection to the $d > 2$ gauge group is not obvious but finds its origin in the underlying two-dimensional geometry. Briefly recall the conformal flatness of two-dimensional spacetimes, i.e.,

$$g_{\mu\nu} = \Omega^2 \eta_{\mu\nu}, \quad g^{\mu\nu} = \Omega^{-2}\eta^{\mu\nu}, \quad \text{with} \quad \mu,\nu = 0,1, \quad (4.48)$$

for some conformal factor $\Omega = \Omega(x^\mu)$. Given this property, the variable $h^{\mu\nu}$ introduced in (4.2) simplifies to

$$h^{\mu\nu} := \sqrt{\det(-g_{\mu\nu})}g^{\mu\nu} = \sqrt{\det(-\Omega^2\eta_{\mu\nu})}\Omega^{-2}\eta^{\mu\nu} = \eta^{\mu\nu}.$$

(4.49)

Consequently, in the conformal frame, $h^{\mu\nu}$ is flat and $\det(h^{\mu\nu}) = -1$, independent of $\det(g^{\mu\nu})$. This latter equality can be expressed as an algebraic constraint:

$$hh^{11} - (h^1)^2 + 1 = 0, \quad (4.50)$$

referred to as the metricity condition. We will soon get back to such condition. For a richer discussion on this topic, though, we refer the reader to [12].

In analogy to the higher dimensional case before, we provide the gauge generators and identities of $d = 2$

Palatini next. Comparing (2.28) to (4.47), we can conveniently rewrite the latter as

$$\delta_\theta h^{\mu\nu} = \theta_{\alpha\beta}[(\Omega^{\alpha\beta})^{\mu\nu}],$$

$$\delta_\theta G^\rho_{\mu\nu} = \theta_{\alpha\beta}[(\Omega^{\alpha\beta})^\rho_{\mu\nu}] - (\partial_{\mu_1}\theta_{\alpha\beta})[(\Omega^{\alpha\beta})^\rho_{\mu\nu}]^{\mu_1}, \quad (4.51)$$

with the gauge generators readily recognized as

$$[(\Omega^{\alpha\beta})^{\mu\nu}] = 2(\epsilon^{\alpha(\mu}h^{\nu)\beta} + \epsilon^{\beta(\mu}h^{\nu)\alpha}),$$

$$[(\Omega^{\alpha\beta})^\rho_{\mu\nu}] = 4\epsilon^{\sigma(\alpha|}G^\rho_{\sigma(\mu}\delta^{|\beta)}_{\nu)}, \quad [(\Omega^{\alpha\beta})^\rho_{\mu\nu}]^{\mu_1} = 2\epsilon^{\rho\mu_1}\delta^\alpha_{(\mu}\delta^\beta_{\nu)},$$

$$(4.52)$$

where the bar | notation delimits the symmetrized indices. Besides the generators, the other element needed to determine the gauge identities are the primary Euler-Lagrange equations. These are given by the straightforward evaluation of (4.3) for $d = 2$. Let us refer to them as $E^{(2)}_{(h^{\mu\nu})}$ and $E^{(2)}_{(G^\rho_{\mu\nu})}$, respectively. Then, their merging together with (4.52) as indicated in (2.29) yields the gauge identities for $d = 2$ Palatini we were seeking, after some tedious yet elementary algebra:

$$\varrho^{\alpha\beta} = E^{(2)}_{(h^{\mu\nu})}[(\Omega^{\alpha\beta})^{\mu\nu}] + E^{(2)}_{(G^\rho_{\mu\nu})}[(\Omega^{\alpha\beta})^\rho_{\mu\nu}]$$

$$+ \partial_{\mu_1}(E^{(2)}_{(G^\rho_{\mu\nu})}[(\Omega^{\alpha\beta})^\rho_{\mu\nu}]^{\mu_1}) \equiv 0. \quad (4.53)$$

Observe that the manifest symmetry under the exchange $\alpha \leftrightarrow \beta$ makes the number of independent gauge identities coincide with our earlier counting: $g = 3$.

3. Physical degrees of freedom

Altogether, we have now obtained the constraint structure characterizing triplet of Palatini in $d = 2$ in terms of Lagrangian quantities:

$$t^{(N)}_{2\text{Pa}} = (l = 9, g = 3, e = 6), \quad (4.54)$$

where $N = 9$. Employing (2.4) and (2.5), it is immediate to rewrite this triplet in Hamiltonian terms:

$$t^{(N)}_{2\text{Pa}} = (N^{(\text{P})}_1 = 3, N_1 = 6, N_2 = 6). \quad (4.55)$$

Plugging (4.54) in the master formula (2.2), we confirm the well-known fact that there are no physical degrees of freedom propagated by the theory: $n_{\text{dof}} = 0$. We restate that the above cannot be obtained by simply setting $d = 2$ in (4.35).

To wrap up this section, we check our results are in good agreement with some of the previously carried out Hamiltonian calculations. We begin our comparisons by looking into the approach closest to our own, the one in [11]. There, the quantities $(h^{\mu\nu}, G^\rho_{\mu\nu})$ were regarded as the

$N = 9$ *a priori* independent field variables for $d = 2$ Palatini, exactly as we did here. Following the Dirac-Bergmann procedure, it was shown that $N^{(\text{P})}_1 = 3$, $N_1 = 6$ and $N_2 = 6$, which readily confirms our own independent findings. In [12], the metricity condition (4.50) was taken into account from the onset. As a result of incorporating this information in the form of additional terms preceded by two Lagrange multipliers in the Hamiltonian, their setup had $N = 11 \neq N = 9$ *a priori* independent field variables. It was there shown that, in such formulation, $N^{(\text{P})}_1 = 5$, $N_1 = 7$ and $N_2 = 8$, yielding no propagating degrees of freedom and thus corroborating again our own final result. This latter comparison provides another example for our point around (2.6), this time in Hamiltonian terms. Namely,

$$t^{(N)}_{2\text{Pa}} = (N^{(\text{P})}_1 = 3, N_1 = 6, N_2 = 6) \neq t^{(N)}_{2\text{Pa}}$$

$$= (N^{(\text{P})}_1 = 5, N_1 = 7, N_2 = 8), \quad (4.56)$$

but we find that $n_{\text{dof}} = 0$ for both sets of numbers upon employing (2.3). As a last remark, we notice that our calculations in this Sec. IV B are comparatively simpler than those in [11,12]. Namely, our approach is certainly to be preferred if the goal is to determine the constraint structure of the theory and thereby manifestly count its propagating modes.

V. CONTEXTUALIZATION AND POTENTIALITY OF OUR RESULTS

The study of constrained systems was initiated in the 1930s by Rosenfeld, in a sometimes overlooked work [48], nowadays acknowledged and revisited [49]. It was later greatly developed during the 1950s [50] and has since been a very active field of theoretical research. As such, one may have the impression that the investigation of manifestly first-order singular classical field theories must be an already closed subject. This is not true. There are ongoing advances in this fundamental topic, particularly within the Lagrangian picture. Besides the references already provided in Sec. II, the recent work [4] stands as a neat example. The methodology there put forward is equivalent to our own proposal, as we shall show in the next Sec. V A.

To further reassure the reader of the topicality of our formalism, in Sec. V B we explain how our method lends itself to a conversion from an analytic machinery to a constructive one. Indeed, the Lagrangian building principle originally put forward in [8,9] finds in the contents of Sec. II a solid footing for attempting the construction of novel theories. This argumentation is carried out in terms of a concrete application for clarity, but the general proposal is much broader. In particular, we explain that the less elaborated upon procedure in [8,9] was cornerstone for the development of the so-called Maxwell-Proca theory. This discussion justifies an interest in the calculations of

Sec. III well beyond a simple exemplification of the explicit usage of the proposed method. When gravity is to be involved, the examples in Sec. IV provide a useful possible basis.

A. On a recent equivalent Lagrangian approach

During the preparation of this manuscript, a novel Lagrangian approach to obtain the functionally independent Lagrangian constraints and count propagating modes in constrained systems (of the kind here considered) appeared [4]. The method therein is physically equivalent to that put forward in [2,3], which—as already mentioned—are complementary references to our own discussion in Sec. II. This equivalence can be easily verified, as both [4] and [2,3] provide a mapping between their proposed Lagrangian parameters and the usual numbers of different kinds of Hamiltonian constraints. We have checked this leads to a consistent mapping between their different Lagrangian parameters.

In our understanding, the method in [4] distinguishes itself because it introduces the notion of first- and second-class (functionally independent) Lagrangian constraints. In our language, these are easy to identify. They are the sum of the various functionally independent Lagrangian constraints arising at all prior stages whose algorithm finalizes nondynamically [as in cases (iia) and (iib)] and dynamically [as in case (i)], respectively. This abstract definition is clarified in the following, by classifying the functionally independent Lagrangian constraints we found in all the given examples into first- and second-class Lagrangian constraints.

In the case of Maxwell electromagnetism, the primary Lagrangian constraint (3.4) we found is a first-class Lagrangian constraint. This is because it leads to a secondary constraint (3.5) that is identically satisfied and so nondynamically stabilized by means of the closure (iia). In fact, this same example is worked out in [4] as well.

Next, consider the (hard) Proca theory. There, both the primary (3.9) and secondary (3.10) Lagrangian constraints we determined are second-class Lagrangian constraints, since the algorithm closes dynamically at the next stage by means of case (i). Such closure implies that the consistency under time evolution of the secondary constraint is determined through a tertiary equation of motion.

We move to Schwinger-Plebanski formulation of both electromagnetism and the (hard) Proca theory. In both cases, the velocity-independent primary constraints $[\varphi_3]_i : \underset{1}{\approx} 0$ and $[\varphi_4]^i : \underset{1}{\approx} 0$ in (3.17) are second-class Lagrangian constraints. This is because their stability is dynamically ensured, via the secondary equations of motion. Thus, the algorithm closes as in case (i) for them. Similarly, the secondary constraints $[\tilde{\varphi}_2]_{ij} : \underset{2}{\approx} 0$ in (3.22) are stabilized dynamically at the subsequent stage. Therefore, these and also their ascendant primary constraints $[\varphi_2]_{ij} : \underset{1}{\approx} 0$ in (3.17) are second-class Lagrangian constraints $\forall\, m \geq 0$. The same is true for $[\varphi_1] : \underset{1}{\approx} 0$ and $[\tilde{\varphi}_1] : \underset{2}{\approx} 0$ in the massive case. When $m = 0$, we see that $[\tilde{\varphi}_1] \underset{1}{\approx} 0$. This is a Lagrangian identity that closes the algorithm nondynamically, following case (iia). As a result, both the said Lagrangian identity and its ascendant $[\varphi_1] : \underset{1}{\approx} 0$ are first-class Lagrangian constraints.

Turning to $d = 2$ Palatini, we see it is rather simple to reclassify the nine functionally independent Lagrangian constraints we obtained into first and second class. At the primary level, we notice that there are six velocity-dependent Lagrangian constraints among the relations that follow from requiring the vanishing of (4.41). These are $[\varphi_a] \underset{1}{\overset{!}{\approx}} 0$, where $a = 1, 2, \ldots, 6$ and the tensorial indices outside the square brackets $[\cdot]$ have been omitted. Their stability is dynamically ensured (via the secondary equations of motion) and so these are second-class Lagrangian constraints. The three remaining primary Lagrangian constraints, $[\varphi_a] \underset{1}{\overset{!}{\approx}} 0$ with $a = 7, 8, 9$, are manifestly velocity independent. They show a trivial stability at the secondary stage; see (4.46). Accordingly, we identify these as three first-class Lagrangian constraints.

At last, we reclassify the functionally independent Lagrangian constraints we found for $d > 2$ Palatini into first- and second-class Lagrangian constraints. Recall that we obtained $M_1 = d(d+1)^2/2$ functionally independent primary Lagrangian constraints, given by the vanishing of (4.11). Notice now that we can straightforwardly split these primary constraints into

$$M_{1(\mathrm{v})} = d(d+1)\ \text{number of velocity-dependent constraints}\colon\ [\varphi_1] : \underset{1}{\overset{!}{\approx}} 0, [\varphi]_i : \underset{1}{\overset{!}{\approx}} 0, \ldots, [\varphi_6]^{ij} : \underset{1}{\overset{!}{\approx}} 0, \qquad (5.1)$$

$$M_{1(\mathrm{nv})} = \frac{d}{2}(d^2 - 1)\ \text{number of velocity-independent constraints}\colon\ [\varphi_7]_i : \underset{1}{\overset{!}{\approx}} 0, [\varphi_8]^j_i : \underset{1}{\overset{!}{\approx}} 0, [\varphi_9]^{jk}_i : \underset{1}{\overset{!}{\approx}} 0, \qquad (5.2)$$

where the subscripts (n)v stand for (non)velocity-dependent constraints. The consistency under time evolution of the velocity-dependent constraints is dynamically fixed at the secondary stage and so these are second-class Lagrangian constraints. The remaining velocity-independent constraints give rise to the $M_2 = M_{1(\mathrm{nv})}$ number of functionally

independent secondary Lagrangian constraints, which are equal to the vanishing of (4.17). Once more, it is trivial to differentiate between

$$M_{2(\text{v})} = \frac{d}{2}(d^2 - 3) \text{ number of velocity-dependent constraints:} \; [\tilde{\varphi}_2]_i^j \overset{!}{\underset{1}{:\approx}} 0 \quad \text{with} \quad i \neq j, [\tilde{\varphi}_3]_i^{jk} \overset{!}{\underset{1}{:\approx}} 0, \tag{5.3}$$

$$M_{2(\text{nv})} = d \text{ number of velocity-independent constraints:} \; [\tilde{\varphi}_1]_i \overset{!}{\underset{1}{:\approx}} 0, [\tilde{\varphi}_2]_i^i \overset{!}{\underset{1}{:\approx}} 0. \tag{5.4}$$

The consistency under time evolution of the former is ensured by the tertiary equations of motion. Equivalently, the algorithm closes according to the dynamical case (i) for them. As a result, they are second-class Lagrangian constraints. Further, the subset of $M_{2(\text{v})}$ number of velocity-independent primary constraints they follow from are second-class Lagrangian constraints as well. Specifically, $[\varphi_8]_i^j \overset{!}{\underset{1}{:\approx}} 0$ with $i \neq j$ and $[\varphi_9]_i^{jk} \overset{!}{\underset{1}{:\approx}} 0$ are second class. On the other hand, the above velocity-independent constraints are trivially stable, as their time evolution yields tertiary constraints that identically vanish in TC_2: recall (4.28). The algorithm closes nondynamically as in case (iib) for them. Consequently, they are first-class Lagrangian constraints and their ascendant primary Lagrangian constraints $[\varphi_7]_i \overset{!}{\underset{1}{:\approx}}$ and $[\varphi_8]_i^i \overset{!}{\underset{1}{:\approx}} 0$ are first class too.

B. Relation to the Maxwell-Proca theory and beyond

As we explicitly showed in Sec. III A, in a purely Lagrangian formulation with as many *a priori* independent field variables as the dimension of the underlying flat spacetime, the constraint structure of the simplest theory for a single Maxwell field can be characterized by the triplet $t_{\text{M}}^{(N)}$ in (3.6). The analogous investigation in Sec. III B of the most elementary theory of one (hard) Proca field yielded the constraint structure characterizing triplet $t_{\text{P}}^{(N)}$ in (3.11). Employing the results of [2], we also verified the corresponding Hamiltonian characterization of these two triplets. We thus checked that the Maxwell and (hard) Proca fields are associated with two first- and second-class constraints, respectively. Although usually Maxwell and Proca fields are defined in the latter Hamiltonian manner, in the following we take the former Lagrangian triplets as the vector fields' defining features. We stress both definitions are equivalent.

The manifestly first-order completions of the Maxwell and (hard) Proca theories analyzed in Secs. III A and III B are nonlinear electrodynamics (NLE) and the so-called generalized Proca (GP) or vector-Galileon theory,[12]

---

[12]We are aware of the recent proposal in [51]. However, the Lagrangian there put forward is not in a manifestly first-order form. The authors leave for further studies this result. In the lack of it, their theory lies beyond our framework and we cannot address it.

respectively. NLE encompasses a large class of theories. The celebrated Born-Infeld theory [52] is part of it, but also the more recently proposed exponential [53] and logarithmic [54] electrodynamics, among others. Schematically, the Lagrangian density for NLE can be written as

$$\mathcal{L}_{\text{NLE}} = \mathcal{L}_{\text{M}} + f(A_{\mu\nu}), \tag{5.5}$$

where $\mathcal{L}_{\text{M}}$ is the Maxwell Lagrangian as introduced in (3.3) and f is a smooth real function. Notice that the above depends on the Maxwell field A_μ exclusively through its field strength $A_{\mu\nu}$—up to boundary terms. Indeed, it is well known [35] that a more involved dependence is not possible, if the $U(1)$ gauge symmetry is to be respected. This feature remains true even when coupling the Maxwell field to general relativity [55]. Only a few fine-tuned terms that contract $A_{\mu\nu}$ with the Riemann tensor are possible in such a case. It is not hard to convince oneself that the constraint structure of NLE is characterized by the triplet $t_{\text{M}}^{(N)}$ in (3.6). In other words, it has the same constraint structure as classical electromagnetism, in its standard formulation of Sec. III A.

The GP theory was put forward in [30] and its complete Lagrangian was established in [56]. Again schematically, we may express it as

$$\mathcal{L}_{\text{GP}} = \mathcal{L}_{\text{P}} + g(B_\mu) + \sum_{n=1}^{d} \mathcal{T}^{\nu_1 \ldots \nu_n \rho_1 \ldots \rho_n} \partial_{\nu_1} B_{\rho_1} \ldots \partial_{\rho_n} B_{\mu_n} \tag{5.6}$$

in d dimensions, where $\mathcal{L}_{\text{P}}$ is the (hard) Proca Lagrangian in (3.7), g is a real smooth function and each $\mathcal{T}^{\nu_1 \ldots \nu_n \rho_1 \ldots \rho_n}$ is a certain smooth real object constructed out of the spacetime metric $\eta_{\mu\nu}$, the d-dimensional Levi-Civita tensor $\epsilon_{\mu_1 \ldots \mu_d}$ and the Proca field B_μ. Although GP has only been formulated for $d = 4$, its systematic construction allows for a straightforward inferring of (5.6). Here, the underlying key idea consists in supplementing the (hard) Proca Lagrangian with derivative self-interaction terms of the Proca field B_μ. This implies a nonlocal extension of the notion of mass for the vector field. As such, we regard GP

as an effective classical field theory.[13] It can be readily inferred from the calculations in [8] that the constraint structure of GP is characterized by the triplet $t_P^{(N)}$ in (3.11). Namely, GP has the same constraint structure as the (hard) Proca theory, when the latter is formulated as in Sec. III B.

Next, we consider a multifield scenario, including n_M number of Maxwell fields, as well as n_P number of (generalized) Proca fields. In four-dimensional Minkowski spacetime, the Maxwell-Proca (MP) theory [8,9] is the complete set of manifestly first-order (self-)interactions among an arbitrary number of real Abelian vector fields that propagates the correct number of degrees of freedom. These consistent interactions were derived by demanding that the constraint structure of each Maxwell and Proca field is characterized by the triplets $t_M^{(N)}$ and $t_P^{(N)}$, respectively. Let $\mathcal{N}_1 = (n_M + n_P)d$. We denote the constraint structure characterizing the triplet of MP as $t_{MP}^{(\mathcal{N}_1)}$. Then, we say that the building principle of the theory is based on the requirement

$$t_{MP}^{(\mathcal{N}_1)} \overset{!}{=} n_M \cdot t_M^{(N=d)} \oplus n_P \cdot t_P^{(N=d)}$$
$$= (l = n_M + 2n_P, g = n_M, e = 2n_M), \quad (5.7)$$

where in the last equality we have made use of (3.6) and (3.11).

The conversion of any of the above necessary conditions into a Lagrangian density building principle is an algebraically involved exercise beyond the scope of our present investigations. We thus leave it for future works.

A last remark is due. As we observed at the very end of Sec. II A and should be apparent from our calculations in Sec. IV A, it is in general a conceptually clear but algebraically nontrivial exercise to obtain the triplet $t^{(N)}$ of a given Lagrangian density $\mathcal{L}$ within our framework. It is even more challenging to determine the (exhaustive) form of $\mathcal{L}$ from the necessary condition that it should be

At this point, it should be clear that our calculations of $t_M^{(N)}$ and $t_P^{(N)}$ in Sec. III, elementary as they are, can be used as a basis for the construction of nontrivial theories. Having a ready-to-be-used method optimized to obtain such triplets (i.e., the method explained in Sec. II and graphically summarized in Fig. 1) is thus a powerful tool for the development of manifestly first-order classical field theories where multiple fields of different spins (self-)interact.

For instance, an interesting open question is that of the consistent coupling of the MP theory to gravity. It is in principle possible to combine our calculations in all the previous sections to attempt this ambitious goal as follows. Let $\mathcal{N}_2 = (n_M + n_P)d + d(d+1)^2/2$, with $d \geq 2$ the dimension of the spacetime. A manifestly first-order Lagrangian density $\mathcal{L}_{MP(2)Pa}$ that describes the dynamics of n_M number of Maxwell fields and n_P number of (generalized) Proca fields in the presence of Einstein's gravity in terms of $\mathcal{N}_2$ *a priori* independent field variables must be associated with a constraint structure characterizing triplet $t_{MP(2)Pa}^{(\mathcal{N}_2)}$ satisfying

$$t_{MP(2)Pa}^{(\mathcal{N}_2)} \overset{!}{=} n_M \cdot t_M^{(N)} \oplus n_P \cdot t_P^{(N)} \oplus t_{(2)Pa}^{(N)}, \quad (5.8)$$

where all the triplets on the right-hand side have already been calculated in this work; see (3.6), (3.11), (4.36) and (4.54). Substituting these results, we have that

$$t_{MP(2)Pa}^{(\mathcal{N}_2)} \overset{!}{=} \begin{cases} (l = n_M + 2n_P + d^2(d+1), g = n_M + d, e = 2n_M + 3d) & \text{if } d > 2, \\ (l = n_M + 2n_P + 9, g = n_M + 3, e = 2n_M + 6) & \text{if } d = 2. \end{cases} \quad (5.9)$$

associated to a certain triplet $t^{(N)}$. The reason is that such inversion in the logic requires solving sets of coupled nonlinear partial differential equations in most cases. Therefore, it is overwhelmingly convenient to use all freedom of choice available in order to simplify this task to the utmost. For instance, one is advised to choose constant null vectors for the Hessians at all stages, if possible. For the concrete research project here proposed, it may be the case that (5.8) is not the optimal starting point. It could happen that the equivalent demand

$$t_{MP(2)Pa}^{(\mathcal{N}_3)} \overset{!}{=} n_M \cdot t_M^{(N)} \oplus n_P \cdot t_P^{(N)} \oplus t_{(2)Pa}^{(N)} \quad \text{with}$$
$$\mathcal{N}_3 = (n_M + n_P + d + 1)d(d+1)/2, \quad (5.10)$$

with the right-hand side triplets as given in (3.32), (4.36) and (4.54), is a more befitting way to try to derive the set of consistent (self-)interactions of vector fields in a curved background. For the reasons given at the beginning of Sec. IV, we believe that $t_{(2)Pa}^{(N)}$ is indeed a beneficial basis for the gravity piece above.

---

[13]The quantization of nonlocal field theories generically leads to acausality, most often for phenomena beyond tree level. It is sometimes possible to circumvent this problem, especially in the presence of supersymmetry—for instance, see [57]. The consistent quantization of GP has been investigated in [58]. Although the results obtained so far justify an optimistic attitude, no complete and rigorous quantization scheme seems to have been proposed so far, along the lines of the BRST and path integral quantization of the (hard) Proca theory in [59,60], respectively. Therefore, we here adopt the conservative point of view that regards GP as an effective classical field theory.

VI. CONCLUSIONS

In the following, we summarize the results we have put forward in this manuscript. Then, we proceed to discuss their relevance and pertinence. At last, we comment on the increasing (in n) computational difficulty of evaluating Lagrangian constraints on constraint surfaces TC_n and concretize the pathologies a theory may suffer from when the algorithm of Sec. II A is not verified to close.

A. Summary of results

In Sec. II, we have collected and complemented results from the extensive literature on constrained systems and presented a self-contained and ready to be used method to determine all the constraints in a theory. By postulation, the theory is required to be described by a manifestly first-order Lagrangian. We make the mild assumptions of the principle of stationary action and finite reducibility. When the theory is covariant, the iterative algorithm presented for the determination of the functionally independent Lagrangian constraints does not contravene this feature. Nonetheless, manifest covariance is generically lost in our approach. In Secs. III and IV, we have minutely exemplified the usage of our said procedure. In Sec. V, we have argued for the pertinence and contemporaneity of both the general formalism and the given examples. Indeed, an equivalent but different methodology has been put forward lately [4]. The examples of Sec. III constitute the foundation of the also recent Maxwell-Proca theory [8,9] and those of Sec. IV can potentially form the basis for the consistent coupling of Maxwell-Proca to gravity.

B. Critical discussion of results

The procedure explained in Sec. II presents two main appealing features. First, it is a coordinate-dependent approach, as opposed to a geometrical one. It thus readily allows for its application, given a Lagrangian density satisfying the initial postulates, without having to work out any symplectic two-form. With pragmatism in mind, Sec. II has been written in a way that is (hopefully) accessible to a broad audience. Even though the method stands on a rigorous footing, the discussion has been made largely devoid of mathematical technicalities.

Second, it is an intrinsically Lagrangian procedure, as opposed to a Hamiltonian or a hybrid one. The appeal of this characteristic resides in the fact that, in many areas of high-energy theoretical physics, manifestly first-order classical field theories are predominantly posed and studied in their Lagrangian formulation. This is the case for instance in cosmology, astrophysics, black hole physics and holographic condensed matter. In all these disciplines, GP, MP and allied theories, especially in the presence of gravity, have been convincingly argued to be of significant interest, e.g., [9,30,61]. As such, our proposed procedure avoids non-negligible obstacles that typically arise in the

transformation from the Lagrangian to the Hamiltonian picture. Besides, as already noted in the end of Secs. IV A and IV B, our Lagrangian approach is a computationally faster and simpler way to obtain the constraint structures of these theories, compared to representative Hamiltonian analyses. (The examples in Sec. III are so effortless comparatively that they do not substantiate an analogous argumentation.)

In more detail, implementing our algorithm in Sec. II A is considerably easier than carrying out a Hamiltonian counterpart algorithm based on the Dirac-Bergman [50] procedure. As the attentive reader will have already noticed in our explicit examples of Sec. IV and we shall address shortly, the most demanding step in our approach consists in evaluating the nth stage Lagrangian constraints in the $(n-1)$th constraint surface, with $n \geq 1$. An analogous evaluation is necessary within the Hamiltonian picture as well, where two additional hurdles arise. On the one hand, one must classify the Dirac constraints into first and second class. This entails calculating the Poisson brackets of all Dirac constraints, a generically challenging task in field theory because nonlocal algebras usually arise,[14] e.g., [13,14]. On the other hand, in the standard Hamiltonian transition from one stage to the next, novel constraints emerge and must be consistently included via Lagrange multipliers. Closure of the algorithm requires the determination of as many Lagrange multipliers as possible, which in turn implies the resolution of algebraic or even differential equations. Even in the comparatively benign algebraic scenario, finding a solution is an increasingly (in stage) laborious and nontrivial task that involves inverting field-dependent matrices with complicated spatial index structures.

For a suggestive utility of the examples in Secs. III and IV, the reader is referred to Sec. V B. Recall that the proposal therein is illustrative of the general theory-construction idea outlined in the introduction Sec. I and at the beginning of Sec. V.

C. Two final observations

In the first of our observations, we bring to light a series of considerations that must be taken into account when applying our method. In particular, we wish to discuss the practical complications that field theories of the kind here considered commonly exhibit when their Lagrangian constraints are to be evaluated on the suitable constraint surface.

First, we debunk what naively may look like an ambiguity. Recall that any constraint surface TC_n for some finite $n \geq 1$ is defined by the weak vanishing of the

---

[14]This nonlocality is as a consequence of the distributional nature of fields. Dirac constraints must be smeared with suitable test functions and integrated over for a correct evaluation of the Poisson brackets.

functionally independent Lagrangian constraints at all prior stages:

$$\varphi_I(Q^A, \dot{Q}^A, \partial_i Q^A) \overset{!}{\underset{1}{\approx}} 0,$$

$$\tilde{\varphi}_R(Q^A, \dot{Q}^A, \partial_i Q^A) \overset{!}{\underset{2}{\approx}} 0,$$

$$\hat{\varphi}_U(Q^A, \dot{Q}^A, \partial_i Q^A) \overset{!}{\underset{3}{\approx}} 0, \qquad \text{etc.,} \qquad (6.1)$$

As a direct consequence of the above, one can determine a maximal set of functionally independent relations of the form

$$\dot{Q}^B \overset{!}{\underset{n}{\approx}} \dot{Q}^B(Q^A, \dot{Q}^A, \partial_i Q^A),$$

$$\partial_i \dot{Q}^B \overset{!}{\underset{n}{\approx}} \partial_i \dot{Q}^B(Q^A, \dot{Q}^A, \partial_i Q^A),$$

$$Q^B \overset{!}{\underset{n}{\approx}} Q^B(Q^A). \qquad (6.2)$$

Though it should be clear by now, we confirm the different role played by the generalized velocities $\dot{Q}^A$ and the spacelike derivatives of the generalized coordinates $\partial_i Q^A$. The former are independent coordinates on $T\mathcal{C}$, while the latter are functionally related to the generalized coordinates Q^A. This clarification becomes pertinent when evaluating the secondary Lagrangian constraints in $T\mathcal{C}_1$ already. At this point (and in subsequent stages), derivatives of the form $\partial_i \dot{Q}^A$ generically show up. In such expressions, one must first replace the primary weak expression for the generalized velocity $\dot{Q}^A$—if pertinent—and then apply the spatial derivative on it.

Having clarified this point, we notice that its consistent implementation leads to the following nested situation. Substitution of $\dot{Q}^A$ according to (6.2) in $\partial_i \dot{Q}^A$ normally leads to the presence of terms of the form $\partial_i Q^B$ in (6.2). These are again prone to be evaluated in $T\mathcal{C}_n$ and can in turn contribute terms depending on Q^B's in (6.2), etc. We emphasize that one must reach an expression where this nesting ceases to occur, before proceeding with the algorithm. Not doing so would imply a wrong evaluation of the Lagrangian constraints in $T\mathcal{C}_n$, may lead to a misidentification of the functionally independent Lagrangian constraints and will almost invariably yield wrong results at the following $(n+1)$th stage. In fact, a wrong evaluation will typically land the researcher in a physically inequivalent theory from the one he or she started with.

Additionally and normally, when evaluating some Lagrangian constraints in $T\mathcal{C}_n$, potentially contrived functions of the previous stages' functionally independent Lagrangian constraints also show up. To understand the difficulty their appearance implies, consider the tertiary Lagrangian constraints (4.25) we found for $d > 2$ Palatini.

Their raw expressions, prior to any evaluation in a constraint surface, contain quantities $f = f(Q^A, \dot{Q}^A, \partial_i Q^A)$ that vanish in $T\mathcal{C}_1$. However, recognizing such f's as primary weak zeros is a challenging task. Specifically,

$$[\hat{\varphi}_1]_i \supset \mathcal{G}^j f_{ij}, \quad f_{ij} = \partial_{[i}[\varphi_7]_{j]} - 2G_{(i}[\varphi_7]_{j)} - G_{k(i}[\varphi_8]_{j)}^k \underset{1}{\approx} 0,$$
$$(6.3)$$

where in the $\supset$ relation we have omitted numerical factors and the φ's are as given in (4.11). In the expression for f_{ij}, the first equality is nontrivial, while the subsequent primary weak equality is obvious. An analogous situation arises with other f's that are based on both functionally independent primary (4.11) and secondary (4.17) Lagrangian constraints. A brute force resolution to identify all such f's consists in putting forward the most general ansatz compatible with the tensorial character of each of the Lagrangian constraints one is trying to evaluate and comparing it to their explicit expressions. This is indeed how we laboriously arrived at (4.27).

For the second and last observation, the reader should heed (2.2) and (2.27). We already stressed the importance of closing the iterative algorithm for obtaining the functionally independent Lagrangian constraints toward the end of Sec. II A. Now, we are equipped to better grasp the implications of not doing so, mentioned in the introductory Sec. I. Most often, failure to close the algorithm will give rise to the propagation of unphysical modes. These are Ostrogradsky instabilities [5], but we shall loosely refer to them as ghosts. Even after ensuring ghost freedom, not closing the algorithm can lead to trouble: it may overconstrain the theory, so that fewer than the desired number of degrees of freedom are propagated.

Let us consider the MP theory [8,9] discussed in Sec. V B as a concrete framework to clarify the above two unwanted scenarios. For our present purposes, it will suffice to consider the case when there are no Maxwell fields $n_M = 0$ and there are an arbitrary but finite number of Proca fields n_P. Recall that, in the standard formulation, we already saw in Sec. III B that a Proca field is associated to $l = M_1' + M_2' = 1 + 1 = 2$ number of functionally independent Lagrangian constraints. Bear in mind that this is also true for a generalized Proca field.

We denote the natural generalization of the GP theory in (5.6) to a multifield setup as $\mathcal{L}_{PP}$. $\mathcal{L}_{PP}$ automatically leads to $M_1' = n_P$ number of functionally independent primary constraints. The consistency under time evolution of these constraints does not generically yield the $M_2' = n_P$ number of functionally independent secondary constraints one would naively expect. Only a fine-tuned subset of terms in $\mathcal{L}_{PP}$ does, precisely the terms that are part of the MP theory. For all those terms, it was shown that no tertiary constraints arise $(M_3 = 0)$ and the algorithm closes dynamically giving rise to $l = 2n_P$. Therefore, the correct

number of physical modes $n_{\mathrm{dof}} = d - n_{\mathrm{P}}$ are present in the theory. (To obtain this result, notice that, since there are no gauge identities, $g = 0 = e$.)

Notice that, if one studies only the primary stage for $\mathcal{L}_{\mathrm{PP}}$, one will be deceived into thinking that the theory is valid, as it suitably extends the primary stage of GP. However, $\mathcal{L}_{\mathrm{PP}}$ has $M_2' < n_{\mathrm{P}}$ in general and therefore $l < 2n_{\mathrm{P}}$ and $n_{\mathrm{dof}} > d - n_{\mathrm{P}}$. The additional propagating modes are precisely the ghosts of the first scenario we warn against.

If one studies both the primary and secondary stages for $\mathcal{L}_{\mathrm{PP}}$, then one can fine-tune the Lagrangian density so that $M_2' = n_{\mathrm{P}}$ as desired. But these functionally independent secondary constraints in the fine-tuned theory are at this point not necessarily stable. Their consistency under time evolution could in principle lead to further functionally independent tertiary constraints, so that $l > 2n_{\mathrm{P}}$ and $n_{\mathrm{dof}} < d - n_{\mathrm{P}}$. This would place us in the second unwelcome scenario. For the given example, it so happens that the fine-tuned $\mathcal{L}_{\mathrm{PP}}$ is associated to a full-rank tertiary Hessian. Consequently, the functionally independent secondary constraints are dynamically stabilized without further fine-tunings of the theory. However, this cannot be assumed; it has to be checked, so as to ensure the theory is not overconstrained.

It is interesting to point out that in [62] our very same admonition against the overconstrained scenario is made, albeit in a different context. The authors look into second-order field theories with no gauge symmetry and derive the necessary conditions for such Lagrangians to not propagate ghosts. They show that, in the presence of Lorentz symmetry, the existence of any number $M_1' > 0$ of functionally independent Lagrangian constraints automatically leads to the same number $M_2' = M_1'$ of functionally independent Lagrangian constraints. They unequivocally recognize our second scenario: those M_2' are not necessarily stable, so one could be facing an overconstrained theory.

ACKNOWLEDGMENTS

APPENDIX: FORMULAS AT AN ARBITRARY STAGE OF THE ALGORITHM

In this Appendix, we show the explicit expressions of all quantities involved in an arbitrary ath stage of the iterative algorithm for irreducible theories presented in Sec. II A. Needless to say, in the appropriate limit, the general expressions here given yield the primary and secondary stages' formulas there shown.

Let $\varphi_{A_a} : \overset{!}{\underset{a}{\approx}} 0$ be a set of M_a number of functionally independent Lagrangian constraints in the ath stage, with $A_a = 1, 2, \ldots, M_a$. These constraints are relations between the generalized coordinates Q^A and velocities $\dot{Q}^A$ of the field theory under consideration. They define the so-called ath constraint surface

$$
\begin{aligned}
T\mathcal{C}_a &:= \{ (Q^A, \dot{Q}^A) | \varphi_{A_a} \underset{a-1}{\approx} 0 \} \\
&\subseteq T\mathcal{C}_{a-1} \subset T\mathcal{C}_{a-2} \subset \cdots T\mathcal{C}_1 \subseteq T\mathcal{C}_0, \quad \text{(A1)}
\end{aligned}
$$

where $T\mathcal{C}_0$ is the moduli space of the theory, defined in (2.18). In order to ensure the preservation of the said constraints under time evolution, we demand

$$
E_{A_a} := \dot{\varphi}_{A_a} \overset{!}{\underset{a}{\approx}} 0. \quad \text{(A2)}
$$

We refer to E_{A_a} as the $(a+1)$th stage Euler-Lagrange equations. Next, we will explicitly write E_{A_a}. But to do so, we must first define the following objects.

Let $W_{A_a A_b}$ denote the $(a+1)$th stage Hessian. This is a square matrix of dimension M_a that allows us to define $M_{a+1} := \dim(W_{A_a A_b}) - \mathrm{rank}(W_{A_a A_b})$. We refer to the M_{a+1} number of linearly independent null vectors associated to $W_{A_a A_b}$ as $\gamma_{A_{a+1}}$. Explicitly, $(\gamma_{A_{a+1}})^{A_a} W_{A_a A_b} = 0$, with $A_{a+1} = 1, 2, \ldots, M_{a+1}$. We require them to fulfill the normalization condition

$$
(\gamma_{A_{a+1}})^{A_a} (\gamma^{A_{b+1}})_{A_a} = \delta_{A_{a+1}}{}^{A_{b+1}}, \quad \text{with} \quad \gamma^{A_{a+1}} := (\gamma_{A_{a+1}})^T,
$$

$$
\text{(A3)}
$$

so that they form a basis in the kernel of $W_{A_a A_b}$. Here, T stands for the transpose operation. With the help of the above null vectors, the $(a+1)$th stage Hessian can be expressed in terms of the functionally independent ath stage Lagrangian constraints as follows:

$$
W_{A_a A_b} = (\gamma_{A_a})^{A_{a-1}} (\gamma_{A_{a-1}})^{A_{a-2}} \ldots (\gamma_{A_1})^A \partial_{\dot{A}} \varphi_{A_b},
$$

$$
\text{with} \quad A = 1, 2, \ldots, N = \dim(\mathcal{C}). \quad \text{(A4)}
$$

Finally, we introduce the auxiliary matrix $M^{A_a A_b}$ (the Moore-Penrose pseudoinverse of $W_{A_a A_b}$), which always exists and is uniquely determined from the relations

$$M^{A_a A_b} W_{A_b A_c} - \delta^{A_a}{}_{A_c} + (\gamma^{A_{a+1}})_{A_c} (\gamma_{A_{a+1}})^{A_a} = 0,$$
$$M^{A_a A_b} (\gamma^{A_{a+1}})_{A_b} = 0. \quad (A5)$$

Using the above, the $(a+1)$th stage Euler-Lagrange equations in (A2) can be written as

$$E_{A_b} = \ddot{Q}^A (\gamma^{A_1})_A (\gamma^{A_2})_{A_1} \ldots (\gamma^{A_a})_{A_{a-1}} W_{A_a A_b} + \alpha_{A_b} \overset{!}{\underset{a}{\approx}} 0, \quad (A6)$$

where the expression (A4) is to be employed for $W_{A_a A_b}$ and where we have (recursively) defined

$$\alpha_{A_b} := [-\alpha_{A_{a-1}} M^{A_{a-1} A_{b-1}} (\gamma_{A_{b-1}})^{A_{a-2}} (\gamma_{A_{a-2}})^{A_{a-3}} \ldots (\gamma_{A_1})^A \partial_{\dot{A}} - \alpha_{A_{a-2}} M^{A_{a-2} A_{b-2}} (\gamma_{A_{b-2}})^{A_{a-3}} (\gamma_{A_{a-3}})^{A_{a-4}} \ldots (\gamma_{A_1})^A \partial_{\dot{A}}$$
$$- \cdots - \alpha_A M^{AB} \partial_{\dot{B}} + \dot{Q}^A \partial_A + (\partial_i \dot{Q}^A) \partial^i_{\dot{A}}] \varphi_{A_b}, \quad (A7)$$

with α_A as given in (2.12). To obtain the presented α_{A_b}, the previous ath stage Euler-Lagrange equations are employed. These in turn depend on the $(a-1)$th stage Euler-Lagrange equations and so on. This is the origin of the noted recursion. Notice we therefore explicitly employ the primary Euler-Lagrange equations and so the expression (A6) is an on-shell statement.

In order to reproduce the results in Sec. II A from the above discussion, the reader only needs to do the index replacements $(A_0 \equiv A, B, \ldots)$, $(A_1 \to I, J, \ldots)$, $(A_2 \to R, S, \ldots)$, etc., as well as take footnote 7 into account.

[1] K. Kamimura, Nuovo Cimento B **68**, 33 (1982).
[2] B. Díaz, D. Higuita, and M. Montesinos, J. Math. Phys. (N.Y.) **55**, 122901 (2014).
[3] B. Díaz and M. Montesinos, J. Math. Phys. (N.Y.) **59**, 052901 (2018).
[4] M. J. Heidari and A. Shirzad, arXiv:2003.13269.
[5] M. Ostrogradsky, Mem. Ac. St. Petersbourg **VI**, 385 (1850).
[6] F. Sbisa, Eur. J. Phys. **36**, 015009 (2015).
[7] C. Deffayet, G. Esposito-Farese, and A. Vikman, Phys. Rev. D **79**, 084003 (2009); C. Deffayet, S. Deser, and G. Esposito-Farese, Phys. Rev. D **82**, 061501 (2010); S. F. Hassan, R. A. Rosen, and A. Schmidt-May, J. High Energy Phys. 02 (2012) 026; L. Buoninfante and W. Li, Phys. Lett. B **779**, 485 (2018); L. Heisenberg, J. Cosmol. Astropart. Phys. 10 (2018) 054.
[8] V. E. Díez, B. Gording, J. A. Méndez-Zavaleta, and A. Schmidt-May, Phys. Rev. D **101**, 045009 (2020).
[9] V. E. Díez, B. Gording, J. A. Méndez-Zavaleta, and A. Schmidt-May, Phys. Rev. D **101**, 045008 (2020).
[10] M. Ferraris, M. Francaviglia, and C. Reina, Gen. Relativ. Gravit. **14**, 243 (1982).
[11] N. Kiriushcheva, S. V. Kuzmin, and D. G. C. McKeon, Mod. Phys. Lett. A **20**, 1895 (2005).
[12] N. Kiriushcheva, S. V. Kuzmin, and D. G. C. McKeon, Int. J. Mod. Phys. A **21**, 3401 (2006).
[13] R. N. Ghalati and D. G. C. McKeon, arXiv:0712.2861.
[14] D. G. C. McKeon, Int. J. Mod. Phys. A **25**, 3453 (2010).
[15] E. C. G. Sudarshan and N. Mukunda, *Classical Dynamics: A Modern Perspective* (John Wiley, New York, 1974); T. Regge and C. Teitelboim, *Constrained Hamiltonian Systems* (Academia Nazionale dei Lincei, Rome 1976);

N. Mukunda, Ann. Phys. (N.Y.) **99**, 408 (1976); Phys. Ser. **21**, 783 (1980).
[16] J. Lee and R. M. Wald, J. Math. Phys. (N.Y.) **31**, 725 (1990).
[17] J. D. Bekenstein and B. R. Majhi, Nucl. Phys. **B892**, 337 (2015).
[18] I. Bengtsson, Phys. Lett. B **172**, 342 (1986).
[19] C. Batlle and J. Gomis, Phys. Lett. B **187**, 61 (1987).
[20] A. Shirzad, J. Phys. A **31**, 2747 (1998); H. J. Rothe and K. D. Rothe, *Classical and Quantum Dynamics of Constrained Hamiltonian Systems* (World Scientific, Singapore, 2010).
[21] R. Sugano, Prog. Theor. Phys. **68**, 1377 (1982); J. M. Pons, J. Phys. A **21**, 2705 (1988); X. Gràcia and J. M. Pons, Ann. Phys. (N.Y.) **187**, 355 (1988).
[22] B. S. DeWitt, Am. J. Phys. **34**, 1209 (1966).
[23] L. Parker and D. Toms, *Cambridge Monographs on Mathematical Physics* (Cambridge University Press, Cambridge, England, 2009).
[24] K. Andrzejewski, J. Gonera, P. Machalski, and P. Maslanka, Phys. Rev. D **82**, 045008 (2010).
[25] J. F. Cariñena, C. López, and N. Román-Roy, J. Math. Phys. (N.Y.) **29**, 1143 (1988).
[26] O. Miskovic and J. Zanelli, arXiv:hep-th/0301256.
[27] T. L. Boullion and P. L. Odell, *Generalized Inverse Matrices* (John Wiley & Sons, New York, 1971); G. H. Golub and C. F. van Loan, *Matrix Computations* (John Hopkins University Press, Baltimore, 2013).
[28] S. Samanta, Int. J. Theor. Phys. **48**, 1436 (2009).
[29] R. Banerjee, H. J. Rothe, and K. D. Rothe, J. Phys. A **33**, 2059 (2000).
[30] G. Tasinato, J. High Energy Phys. 04 (2014) 067.

[31] K. Sundermeyer, *Constrained Dynamics*, Lecture Notes in Physics Vol. 169 (Springer-Verlag, New York, 1982); N. Kiriushcheva, S. V. Kuzmin, and D. G. C. McKeon, Can. J. Phys. **90,** 165 (2012).

[32] A. Proca, J. Phys. Radium **7,** 347 (1936); **9,** 61 (1939).

[33] F. Darabi and F. Naderi, Int. J. Theor. Phys. **50,** 3432 (2011).

[34] J. Schwinger, Philos. Mag. **44,** 1171 (1953).

[35] J. F. Plebanski, Lecture Notes, Nordita, Copenhagen, 1970.

[36] K. Sundermeyer, Lect. Notes Phys. **169,** 1 (1982).

[37] A. Palatini, Rendiconti del Circolo Matematico di Palermo (1884-1940), **43,** 203 (1919).

[38] L. Castellani, P. van Nieuwenhuizen, and M. Pilati, Phys. Rev. D **26,** 352 (1982).

[39] M. Montesinos, R. Escobedo, J. Romero, and M. Celada, Phys. Rev. D **101,** 024042 (2020).

[40] R. K. Kaul and S. Sengupta, Phys. Rev. D **94,** 104047 (2016).

[41] S. Holst, Phys. Rev. D **53,** 5966 (1996).

[42] J. Lewandowski and A. Okolow, Classical Quantum Gravity **17,** L47 (2000); R. Capovilla, M. Montesinos, V. A. Prieto, and E. Rojas, Classical Quantum Gravity **18,** L49 (2001).

[43] M. Celada, D. González, and M. Montesinos, Classical Quantum Gravity **33,** 213001 (2016).

[44] P. Horava, Classical Quantum Gravity **8,** 2069 (1991).

[45] D. Grumiller, W. Kummer, and D. V. Vassilevich, Phys. Rep. **369,** 327 (2002).

[46] S. Deser, Found. Phys. **26,** 617 (1996).

[47] D. G. C. McKeon, Can. J. Phys. **95,** 548 (2017).

[48] L. Rosenfeld, Ann. Phys. (Berlin) **397,** 113 (1930).

[49] D. Salisbury and K. Sundermeyer, Eur. Phys. J. H **42,** 23 (2017).

[50] P. A. M. Dirac, Can. J. Math. **2,** 129 (1950); J. L. Anderson and P. G. Bergmann, Phys. Rev. **83,** 1018 (1951); P. G. Bergmann and J. Goldberg, Phys. Rev. **98,** 531 (1955); P. G. Bergmann,

[51] C. de Rham and V. Pozsgay, arXiv:2003.13773.

[52] M. Born, Proc. R. Soc. A **143,** 410 (1934).

[53] S. H. Hendi, Ann. Phys. (Amsterdam) **333,** 282 (2013).

[54] P. Gaete and J. Helayël-Neto, Eur. Phys. J. C **74,** 2816 (2014).

[55] G. W. Horndeski, J. Math. Phys. (N.Y.) **17,** 1980 (1976).

[56] J. B. Jimenez and L. Heisenberg, Phys. Lett. B **757,** 405 (2016).

[57] A. Addazi and G. Esposito, Int. J. Mod. Phys. A **30,** 1550103 (2015).

[58] A. Amado, Z. Haghani, A. Mohammadi, and S. Shahidi, Phys. Lett. B **772,** 141 (2017); C. de Rham, S. Melville, A. J. Tolley, and S. Y. Zhou, J. High Energy Phys. 03 (2019) 182; M. S. Ruf and C. F. Steinwachs, Phys. Rev. D **98,** 025009 (2018); L. Heisenberg and C. F. Steinwachs, J. Cosmol. Astropart. Phys. 01 (2020) 014; 02 (2020) 031; L. Heisenberg and J. Zosso, arXiv:2005.01639.

[59] Y. W. Kim, M. I. Park, Y. J. Park, and S. J. Yoon, Int. J. Mod. Phys. A **12,** 4217 (1997).

[60] J. C. Su, arXiv:hep-th/9805196.

[61] E. Allys, P. Peter, and Y. Rodríguez, J. Cosmol. Astropart. Phys. 02 (2016) 004; E. Allys, J. P. B. Almeida, P. Peter, and Y. Rodríguez, J. Cosmol. Astropart. Phys. 09 (2016) 026; E. Allys, P. Peter, and Y. Rodríguez, Phys. Rev. D **94,** 084041 (2016); J. B. Jiménez and L. Heisenberg, Phys. Lett. B **770,** 16 (2017); Y. Rodríguez and A. A. Navarro, J. Phys. Conf. Ser. **831,** 012004 (2017); L. Heisenberg, J. B. Jiménez, C. de Rham, and L. Heisenberg, Phys. Lett. B **802,** 135244 (2020); J. Sanongkhun and P. Vanichchapongjaroen, Gen. Relativ. Gravit. **52,** 26 (2020).

[62] M. Crisostomi, R. Klein, and D. Roest, J. High Energy Phys. 06 (2017) 124.

Helv. Phys. Acta Suppl. **4,** 79 (1956); P. A. M. Dirac, Proc. R. Soc. A **246,** 326 (1958); Phys. Rev. **114,** 924 (1959).