How the non-metricity of the connection arises naturally in the classical theory of gravity

Bartłomiej Bąk*
Faculty of Physics University of Warsaw,
Department of Mathematical Methods in Physics, Poland
and
Jerzy Kijowski†
Center for Theoretical Physics
Polish Academy of Sciences, Warsaw, Poland

September 14, 2022

Abstract

Three equivalent variational formulations of General Relativity Theory: 1) metric, 2) metric-affine (Palatini) and 3) purely affine, are discussed. The classical "Palatini method of variation" is thoroughly analysed. It implies non-metricity of the spacetime connection, as soon as the matter Lagrangian density is connection-sensitive. This conclusion is not new, but was never seriously discussed. We show, that accepting this non-metricity, i.e. treating both geometric structures of spacetime: the metric tensor and the connection, as being *a priori* independent, but interacting with each other *via* field equations, considerably simplifies the conceptual structure of the present Gravity Theory and, maybe, provides a better starting point for its generalisations from our Solar System scale to the Universe scale (dark matter).

*Email: Bartlomiej.Bak@fuw.edu.pl
†Email: kijowski@cft.edu.pl

1 Introduction

Gravity is believed to describe the universal interaction between all possible kinds of matter. Mathematically, the above conjecture can be formulated as the following rule (called often the *minimal coupling* rule): for every matter field ϕ, whose local dynamics in the flat Minkowski space, equipped with the flat Lorentzian metric $\eta_{\mu\nu}$, is described by the *matter* Lagrangian density:

$$\mathcal{L}_{matt} = \mathcal{L}_{matt}(\phi, \nabla\phi, \eta) \, , \tag{1}$$

its *global* dynamics in an arbitrary spacetime (equipped with a – possibly non-flat – metric $g_{\mu\nu}$), together with interaction between matter and geometry, is described by the following *total* (i.e. "matter + gravity"), metric Lagrangian density:

$$\mathcal{L}_g := \mathcal{L}_H + \mathcal{L}_{matt} \, . \tag{2}$$

Here, by $\mathcal{L}_H$ we denote the Hilbert Lagrangian density[1] (see [1]):

$$\mathcal{L}_H = \mathcal{L}_H(g, \partial g, \partial^2 g) := \frac{\sqrt{|\det g|}}{16\pi} \, \overset{\circ}{R} \, , \tag{3}$$

and $\overset{\circ}{R}$ denotes the scalar curvature of the metric g or, more precisely, of the metric Levi-Civita connection

$$\overset{\circ}{\Gamma}{}^{\kappa}{}_{\lambda\mu} := \frac{1}{2} \, g^{\kappa\sigma} \left(g_{\sigma\lambda,\mu} + g_{\sigma\mu,\lambda} - g_{\lambda\mu,\sigma} \right) \, . \tag{4}$$

Geometric structure of the Minkowski spacetime in formula (1) (both the metric η and its flat connection) has to be replaced by the actual, possibly non-flat, structure $(g, \overset{\circ}{\Gamma})$. This means that $\mathcal{L}_H$ depends upon the metric components $g_{\mu\nu}$, together with their first and second derivatives, according to (3), whereas matter Lagrangian depends upon metric and its *first derivatives* only, contained in the Levi-Civita connection $\overset{\circ}{\Gamma}$:

$$\mathcal{L}_{matt} = \mathcal{L}_{matt}(\phi, \overset{\circ}{\nabla} \phi, g) = \mathcal{L}_{matt}(\phi, \partial\phi, g, \partial g) \, . \tag{5}$$

Physical intuitions concerning the fundamental conceptual structure of General Relativity Theory were built on specific examples: electrodynamics, scalar field

[1] In this paper we use geometric system of physical units, where $c = 1$ and $G = 1$ are dimensionless numbers. To rewrite formula (3) in an arbitrary system of units one has to multiply the dimensionless constant "$\pi \approx 3,14\ldots$" appearing here, by the fundamental physical constant "$\frac{G}{c^4}$".

and, especially, on mechanics of continuous media (see, e.g., [2], [3]). In all these cases the matter Lagrangian (5) **does not** depend upon connection (and, therefore, upon derivatives of the metric tensor), i.e. instead of (5), we have:

$$\mathcal{L}_{matt} = \mathcal{L}_{matt}(\phi, \partial\phi, g)\,. \tag{6}$$

This, very special, property of the matter Lagrangian was *explicitly* assumed in most papers discussing the fundamental structures of the theory, like, e.g., the famous Palatini article [4]. If the matter field belongs to this exceptional category, both matter and gravitational field equations (i.e. Euler-Lagrange equations derived from the total Lagrangian density (2)) can be rewritten as follows:

$$\partial_\lambda p^\lambda = \frac{\partial \mathcal{L}_{matt}}{\partial \phi}\,, \tag{7}$$

$$\text{where} \quad p^\lambda := \frac{\partial \mathcal{L}_{matt}}{\partial \phi_{,\lambda}}\,, \qquad \phi_{,\lambda} := \partial_\lambda \phi\,, \tag{8}$$

$$\mathcal{G}^{\mu\nu} = 8\pi\, \mathcal{T}^{\mu\nu}\,, \tag{9}$$

$$\text{where} \quad \mathcal{T}^{\mu\nu} := 2\,\frac{\partial \mathcal{L}_{matt}}{\partial g_{\mu\nu}}\,, \tag{10}$$

with

$$\mathcal{G}_{\mu\nu} = \sqrt{|\det g|}\, G_{\mu\nu} \tag{11}$$

denoting density of the Einstein tensor:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}\, g_{\mu\nu}\, R\,.$$

Here, $R_{\mu\nu}$ denotes the Ricci tensor, whereas its trace $R := g^{\mu\nu} R_{\mu\nu}$ denotes the scalar curvature.

But what is most important in these exceptional examples, is the fact that the quantity (10), defined in this way, can really be interpreted as the energy-momentum tensor of the matter field, because its "time-time-component" correctly describes the special-relativistic energy density of the matter field evolving over a non-dynamical (i.e. fixed *a priori*) geometric spacetime structure $(g, \overset{\circ}{\Gamma})$. This observation appears in the classical literature under the name of the so called Belinfante-Rosenfeld theorem (cf. [5], [6] and [7]).

Unfortunately, this simple scheme fails for a generic matter field (e.g. vectorial, spinorial or tensorial), when *covariant* (and not just *partial*) derivatives are necessary to define the *covariant* matter Lagrangian, i.e. when, instead of (6), we have:

$$\mathcal{L}_{matt} = \mathcal{L}_{matt}(\phi, \overset{\circ}{\nabla} \phi, g) = \mathcal{L}_{matt}(\phi, \partial\phi, g, \overset{\circ}{\Gamma}) = \mathcal{L}_{matt}(\phi, \partial\phi, g, \partial g)\,. \qquad (12)$$

In this case variation of (2) with respect to the metric $g_{\mu\nu}$ produces an extra term on the right-hand-side of Einstein equation (9). Consequently, remaining field equations (7) – (8) of the theory do not change, but the partial derivative (10) of the matter Lagrangian density, which provides the right-hand-side of Einstein equation (9), must be replaced now by the so called “variational derivative”:

$$\mathcal{T}^{\mu\nu} = 2\,\frac{\delta\mathcal{L}_{matt}}{\delta g_{\mu\nu}} = 2\left\{\frac{\partial\mathcal{L}_{matt}}{\partial g_{\mu\nu}} - \partial_\kappa \frac{\partial\mathcal{L}_{matt}}{\partial g_{\mu\nu,\kappa}}\right\}, \qquad (13)$$

where we use the following notation: $g_{\mu\nu,\kappa} := \partial_\kappa g_{\mu\nu}$. In a generic case, the second (extra) term contains second order derivatives of the metric. This changes considerably the structure of Einstein equations, makes it extremely complicated and, above all, ruins the standard physical interpretation of the theory: “mass generates curvature”, which was proposed already by Riemann and Clifford (cf. [8]). This is due to the fact that (13) does not represent neither energy (mass) nor momentum of the matter field, because *Belinfante-Rosenfeld-theorem* does not apply in this case and, consequently, (13) differs from the energy-momentum tensor, defined properly as the Hamiltonian density generating evolution of the matter field in the Hamiltonian picture. Moreover, small perturbations of the metric (i.e. *gravitational waves*) propagate differently than the electromagnetic waves (i.e. they do not follow the light-cones of the metric). The simplest example of such a matter field, namely a vector field ϕ^μ, is discussed in Section 3. Hence, the traditional, heuristic interpretation of Einstein equation (mass causes the spacetime curvature) fails in this case. To avoid this discrepancy, one should probably treat matter and gravity on a more equal footing, as already suggested strongly by Albert Einstein, who stressed many times in his papers (see, e.g. [9]), that dividing physical reality into “matter” and “geometry” is, from the fundamental point of view, probably not justified.

A considerable simplification of the conceptual framework of the General Relativity Theory is due to the discovery of its purely affine variational formulation (see [10]), where variation of the affine Lagrangian density

$$\mathcal{L}_A = \mathcal{L}_A(R_{\mu\nu}, \phi, \nabla\phi) \qquad (14)$$

is performed with respect to the connection variable only, whereas the metric tensor arises as the “momentum canonically conjugate” to the connection. Such a purely affine Lagrangian density depends[2] upon connection coefficients $\Gamma^{\lambda}{}_{\mu\nu}$ (contained in the curvature tensor, but also in the matter covariant derivatives $\nabla\phi$), together with their first (and not the *second* !) derivatives contained in the curvature tensor, (cf. [10], [11], [12] and a recent review article [13]). A specific sensitivity of the covariant derivatives towards the connection:

$$\nabla\phi = \partial\phi + \text{``}\ \phi\ \cdot\ \Gamma'' \ , \tag{15}$$

is implied by geometric properties of the matter field ϕ (e.g. vector, tensor or other) and will be decisive for its gravitational properties.

One of the unexpected consequences of the affine formulation of the General Relativity Theory is that, in a generic case of a Lagrangian density (12), the resulting connection Γ is (possibly) non-metric, i.e. differs (possibly) from the Levi-Civita connection (4) in a way depending upon above sensitivity of the matter field. In fact, the metricity condition for Γ is obtained in the affine picture only for specific matter Lagrangians, namely those fulfilling (6). One can say shortly that, whereas matter causes the spacetime curvature, its specific sensitivity (15) towards the connection causes the non-metricity of the connection.

This **is not** a new phenomenon: it arises already when applying the classical, well established "Palatini method of variation" (cf. [4]), to a generic matter Lagrangian (5), but is usually left unsaid by most authors.

In the present paper we show in a simple and natural way, how and why the non-metricity arises in a generic case of a matter Lagrangian (5). For this purpose, we analyse the validity of the Palatini principle and prove that it leads, in a generic case, to the non-metric connection: the only exception are theories given by specific matter Lagrangian densities, namely those fulfilling (6), where the Palatini equation implies, indeed, the conventional metricity condition for the connection.

Naively, one could think that the theory with a non-metric connection will not be equivalent with the conventional Einstein theory, where the connection's metricity is assumed *a priori*. We show in this paper that such a naive conclusion is false. Indeed, when rewritten in terms of the metric connection, the non-metric theories described here assume their conventional, Einsteinian form, the only difference being the geometric interpretation of various mathematical (computational) terms arising in field equations. We stress, therefore, that theories discussed here, do not belong

[2]A non-symmetric connection **is not** an irreducible object. It splits into two irreducible parts: 1) a symmetric connection and 2) the torsion tensor. Being a tensor field, the latter can be interpreted as a part of the matter fields. Thus, without losing generality, we limit ourselves to the case of a symmetric connection: $\Gamma^{\lambda}{}_{\mu\nu} = \Gamma^{\lambda}{}_{\nu\mu}$.

to any of the "generalisations of General Relativity Theory", but to the standard, Einsteinian gravity theory!

The goal of this paper is to convince the reader that such a reformulation of non-metric connection, back to the metric connection, is neither necessary nor useful. We show that formulation of the conventional gravity theory in terms of the (possibly) non-metric connection, arising in a natural way from the consequent use of the Palatini "method of variation", simplifies considerably the conceptual framework of the theory and its entire mathematical structure. Being geometrically natural and algebraically useful, the above non-metric connection deserves, maybe, more attention on the side of the physical interpretation of the theory as the "true geometric structure" of spacetime . . .

The paper is organised as follows. In Section 2 we present a simple technique which enables one to treat a variational principle as a "symplectic relation". We show that the usual treatment, based on "imposing spacetime-boundary-conditions" contradicts the hyperbolic structure of field equation. The techniques presented here enable us to easily manipulate various variational principles and greatly simplify the proofs of their equivalence. In Section 3 we first analyse thoroughly the content of the simplified version of the "Palatini method of variations", limiting ourselves to the case of "connection-non-sensitive" matter fields. This was the case considered by Palatini himself in his outstanding paper [4]. Next, we show that – in case of a generic matter field – it leads to a non-metric spacetime connection. Consequently, we provide its complete formulation as a correctly defined symplectic relation between the two geometric spacetime structures: the metric one (defined by the metric tensor) and the affine one (defined by connection). In Section 4 we discuss the transition from the metric formulation to the affine formulation in terms of the simple Legendre transformation (i.e. change of the control mode of the theory). In Section 5 we construct the universal Palatini formulation of the present General Relativity theory and in Section 6 we briefly discuss the inverse Legendre transformation: from affine to the metric picture. Finally we discuss possible physical consequences (Section 7) of the mathematical results obtained here. The most complex calculations have been shifted to Appendixes.

2 The role of variational principles in hyperbolic field theories. Affine variational principle for gravitational field

Consider a Lagrangian density $\mathcal{L} = \mathcal{L}(\varphi^K, \varphi^K_\lambda)$ depending upon a one-parameter-family of fields

$$\varphi^K = \varphi^K(x^\mu, \epsilon)\,,$$

where (x^μ) are spacetime coordinates. We denote by

$$\varphi^K_\lambda := \partial_\lambda \varphi^K\,,$$

the field's partial, spacetime derivatives. Traditionally, derivative with respect to the parameter ϵ is denoted by δ:

$$\delta := \frac{\mathrm{d}}{\mathrm{d}\epsilon}\,. \tag{16}$$

Operator δ commutes, obviously, with spacetime derivatives:

$$\delta\varphi^K_\mu = \partial_\mu \delta\varphi^K\,,$$

(this trivial observation is, in many textbooks, upgraded to the level of "the fundamental Lemma of the calculus of variations"). The following, obvious identity:

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\varphi^K}\,\delta\varphi^K + \frac{\partial\mathcal{L}}{\partial\varphi^K_{,\lambda}}\,\delta\varphi^K_{,\lambda} = \left(\frac{\partial\mathcal{L}}{\partial\varphi^K} - \partial_\lambda \frac{\partial\mathcal{L}}{\partial\varphi^K_{,\lambda}}\right)\delta\varphi^K + \partial_\lambda\left(\frac{\partial\mathcal{L}}{\partial\varphi^K_{,\lambda}}\,\delta\varphi^K\right)\,, \tag{17}$$

is the starting point of the calculus of variations. The first term of (17) is called the *volume part* (the *bulk term*) and the second one the *boundary part*. Traditionally, one neglects the boundary part because, when integrated over a spacetime volume $\mathcal{O}$, imposing boundary conditions on its boundary $\partial\mathcal{O}$, implies $\delta\varphi^K|_{\partial\mathcal{O}} = 0$. This way one derives the second-order partial differential equations for unknown functions φ^K, called *Euler-Lagrange equations*:

$$\frac{\partial\mathcal{L}}{\partial\varphi^K} - \partial_\lambda \frac{\partial\mathcal{L}}{\partial\varphi^K_{,\lambda}} = 0\,, \tag{18}$$

as the necessary condition for the extremum of the functional:

$$\mathcal{F} := \int_{\mathcal{O}} \mathcal{L}\,, \tag{19}$$

within the functional-analytic space of fields fulfilling the fixed boundary conditions $\varphi^K|_{\partial\mathcal{O}} = f^K$.

For example, C.Misner, K.Thorne and J.A.Wheeler in their monograph [14], otherwise excellent as an introduction to General Relativity Theory, calculate only the volume part of the variation of the Hilbert Lagrangian density and neglect the boundary part. As a justification of such a careless procedure (see [14], page 520, just above their formula 21.86) they write:

"Variation of the geometry interior to the boundary make no difference in the value of the surface term. Therefore, it has no influence on the equations of motion to drop the term (21.85)".

The term, which is dropped there, is precisely the surface term. It was never calculated in this monograph (cf. also [15] for more details).

Such a procedure, whose origin goes back to Johann Bernoulli and his classical *brachistochrone problem* (1696), works perfectly for "optimisation problems", where Euler-Lagrange equations are elliptic. Unfortunately, it is entirely false in case of dynamical theories, governed by hyperbolic – not elliptic – field equations. Most theoretical physicists have already learned long time ago that there is no extremum, but only a "saddle point", in the hyperbolic case, and believe that the above observation represents the pinnacle of human understanding of the principles of variation.

For example, the "Feynman integral" quantisation method is based precisely on the observation that the classical trajectory – i.e. the action's stationary point – gives the main contribution to the integral over classical trajectories. But the "extremum *versus* saddle point" dichotomy is not enough to describe the entire complexity of the variational problems, because the real difficulty lies elsewhere! Namely: **no matter whether we expect extremum or a saddle point**, imposing the spacetime-boundary conditions is *strictly forbidden* in hyperbolic theories. This means, that there is no solution of (18) for a generic choice of boundary data! To convince oneself that this is the case, it is enough to consider the "mother of all hyperbolic theories", that is, the wave equation

$$\left(\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial t^2} \right) \varphi = 0 \, , \tag{20}$$

in two dimensional spacetime $\mathbb{R}^2 = \{(t, x)\}$. Implying advanced and retarded coordinates $(u, v) = (t - x, t + x)$ and twice integrating it over the rectangle

$$\mathcal{R} = \{(u, v) \in \mathbb{R}^2 : \, u_0 \leq u \leq u_0 + 2\delta, \, v_0 \leq v \leq v_0 + 2\epsilon\} \, ,$$

is easy to prove that field equation (20) for the function $\varphi(u,v)$ is equivalent to the following identity

$$\varphi(u_0+2\delta,v_0+2\epsilon)-\varphi(u_0+2\delta,v_0)-\varphi(u_0,v_0+2\epsilon)+\varphi(u_0,v_0)=0\,,$$

for any choice of four numbers: $\{u_0,v_0,\epsilon,\delta\}$. Above equation could be simply re-transformed to $(t,x)=(\frac{v+u}{2},\frac{v-u}{2})$ coordinates. Then:

$$\varphi(t_0+\epsilon+\delta,x_0+\epsilon-\delta)-\varphi(t_0+\delta,x_0-\delta)-\varphi(t_0+\epsilon,x_0+\epsilon)+\varphi(t_0,x_0)=0\,. \qquad (21)$$

Putting $t_0=0$, $x_0=x$, $\delta=x$ and $\epsilon=1-x$ we obtain an identity which must be fulfilled for any $0\leq x\leq 1$:

$$\varphi(1,1-x)-\varphi(x,0)-\varphi(1-x,1)+\varphi(0,x)=0\,, \qquad (22)$$

Consider now the spacetime volume $\mathcal{O}=[0,a]\times[0,a]$, i.e:

$$\mathcal{O}=\{(t,x)|0\leq t\leq a\ ;\ 0\leq x\leq a\}\ . \qquad (23)$$

We see (unfortunately, few physicists are aware of that!) that field equation implies here a constraint in space of boundary data: the value of the field on the upper wall (i.e.: $\varphi(1,\cdot)$) is uniquely given by its value on the remaining three walls (i.e.: $\varphi(\cdot,0)$, $\varphi(\cdot,1)$ and $\varphi(0,\cdot)$). There is no solution of the wave equation if the boundary data do not belong to the subspace $\mathcal{C}$ defined by equation (22)! Moreover: field equation (20) is *equivalent* to this constraint!

Hence, the "brachistochrone" philosophy enables us to derive field equation as soon as we already know field equation (choosing boundary data "at random", there is zero probability, that there is any solution satisfying this choice).

We stress, that the above constraint (22) is still "relatively manegeable" for the simple spacetime rectangle (23), whereas for a generic spacetime volume $\mathcal{O}$ (e.g., a time slice $\{a\leq t\leq b; x\in\mathbb{R}\}$) it is a much worse, very singular, non-close subspace in any reasonable topology of boundary data.

We conclude that, as a method of *deriving* field equation, the above "brachistochrone philosophy" breaks down completely in case of hyperbolic field equations. We stress, however, that this method works perfectly for purposes of optimisation problems governed by elliptic field equations, because boundary conditions can be imposed without any restriction in those cases.

Our conclusion does not mean that the formulae used in the Lagrangian field theory are false and useless! Below we will give them a coherent mathematical meaning, which replaces the nonsensical heuristics based on the "spacetime boundary conditions".

For this purpose we propose to work "on shell", i.e. to restrict oneself only to those field configurations (and their jets $(\varphi^K, \varphi^K_{,\lambda})$) which fulfil field equations (18). This means, that – instead of neglecting the boundary part of (17) – we neglect its volume part. This way, formula (17) is no longer an identity, but becomes an equation imposed on the first jet of the field configuration:

$$\delta\mathcal{L}(\varphi^K, \varphi^K_{,\lambda}) = \partial_\lambda \left(\frac{\partial \mathcal{L}}{\partial \varphi^K_{,\lambda}} \, \delta\varphi^K \right) = \left(\partial_\lambda p_K^{\ \lambda}\right) \, \delta\varphi^K + p_K^{\ \lambda} \, \delta\varphi^K_{,\lambda} \, , \tag{24}$$

where the canonical momentum $p_K^{\ \lambda}$ has been introduced as a shortcut notation for the following expression:

$$p_K^{\ \lambda} := \frac{\partial \mathcal{L}}{\partial \varphi^K_{,\lambda}} \, . \tag{25}$$

We see that the system of first-order partial differential equations (24) for the variables $\left(\varphi^K, p_K^{\ \lambda}\right)$ is equivalent to the second-order Euler-Lagrange equation (18), written in the form:

$$\partial_\lambda p_K^{\ \lambda} = \frac{\partial \mathcal{L}}{\partial \varphi^K} \, , \tag{26}$$

together with definition (25) of momenta. This way, at each spacetime point $\mathbf{m} = (x^\mu)$, field equations (24) can be considered as a symplectic relation (i.e. a Lagrangian submanifold) in a symplectic space $\mathcal{P}_\mathbf{m}$ parameterized by the following "generalized jets" of fields: $(\varphi^K$, $\varphi^K_{,\lambda}$, $p_K^{\ \lambda}$, $j_K := \partial_\lambda p_K^{\ \lambda})$. Mathematically, this approach was rigorously defined in [11], [16] and [17], but its strength consists in the fact that it is very well adapted for practical calculations in both the Lagrangian and Hamiltonian formalisms (especially when constraints are present) and avoids the nonsensical procedure of "imposing the spacetime-boundary conditions". Practically, this approach is based on splitting canonical field variables into two groups: the "control parameters" (those, who appear under the sign "δ" – in case of (24) these are configuration variables φ^K and their "velocities" $\varphi^K_{,\lambda}$) and the "response parameters" (in case of (24) these are momenta $p_K^{\ \lambda}$ and their "currents" $j^K = \partial_\lambda p_K^{\ \lambda}$). Field equations are then treated as the "control – response relation". We shall use this formalism in the sequel.

This techniques was informally present already in classical texts, written by Lagrange, Caratheodory and other pioneers of the calculus of variations, and also in classical thermodynamics. As an example, consider the classical, thermodynamical formula:

$$\delta U(V, S) = -p \, \delta V + T \, \delta S \, ,$$

equivalent to:

$$p = -\frac{\partial U}{\partial V} \, , \qquad T = \frac{\partial U}{\partial S} \, ,$$

which selects the two-dimensional subspace of all the physically admissible states of a simple thermodynamical body, as a Lagrangian submanifold within a four-dimensional symplectic manifold parameterized by (V, S, p, T) (volume, entropy, pressure, temperature) and equipped with the canonical symplectic form:

$$\omega = -\delta p \wedge \delta V + \delta T \wedge \delta S \,. \tag{27}$$

Similarly, classical mechanics can be formulated as a symplectic relation:

$$\delta L(q, \dot{q}) = \frac{\mathrm{d}}{\mathrm{d}t}\,(p\,\delta q) = \dot{p}\,\delta q + p\,\delta\dot{q}\,,$$

(equivalent to: $\dot{p} = \frac{\partial L}{\partial q}$, $p = \frac{\partial L}{\partial \dot{q}}$) with respect to the canonical symplectic form:

$$\omega = \frac{\mathrm{d}}{\mathrm{d}t}\,(\delta p \wedge \delta q) = \delta\dot{p} \wedge \delta q + \delta p \wedge \delta\dot{q}\,. \tag{28}$$

Legendre transformations, like transition from adiabatic to the thermostatic insulation in thermodynamics, or from the Lagrangian to the Hamiltonian picture in mechanics, are simply described in this formalism as an exchange between control and response parameters: T *versus* S in (27) and p *versus* $\dot{q}$ in (28).

In case of the gravitational field, the missing boundary term in the Wheeler-Misner-Thorn formula (21.86) was calculated in paper [15] and will be presented below. Following V.A. Fock (see [18]), who realized that the substantial simplification of the canonical structure of gravity theory is obtained when, instead of the covariant tensor $g_{\mu\nu}$, one represents the metric structure of spacetime by its contravariant density, we introduce the following notation[3]:

$$\pi^{\mu\nu} := \frac{1}{16\pi}\sqrt{|\det g|}\; g^{\mu\nu} = \frac{\partial \mathcal{L}_H}{\partial R_{\mu\nu}}\,. \tag{29}$$

(With respect to the Fock's book [18], our modest contribution here consists in incorporating the gravitational constant – i.e. $\frac{1}{16\pi}$ in geometric units – into the "momentum" variable π.) Moreover, the following object arises automatically in the formula for the variation:

$$\pi_\lambda{}^{\mu\nu\kappa} = \frac{\partial \mathcal{L}_H}{\partial \Gamma^\lambda{}_{\mu\nu,\kappa}} = \pi^{\mu\nu}\,\delta^\kappa_\lambda - \delta^{(\mu}_\lambda\,\pi^{\nu)\kappa}\,. \tag{30}$$

The Hilbert Lagrangian density assumes now the following form:

$$\mathcal{L}_H = \mathcal{L}_H(g, \Gamma, \partial\Gamma) := \frac{\sqrt{|\det g|}}{16\pi} R = \pi^{\mu\nu} R^\lambda{}_{\mu\lambda\nu} = \pi^{\mu\nu} R_{\mu\nu}\,, \tag{31}$$

[3] See footnote on page (1) for physical units used here.

where

$$R^{\lambda}{}_{\mu\nu\kappa} \;:=\; \Gamma^{\lambda}{}_{\mu\kappa,\nu} - \Gamma^{\lambda}{}_{\mu\nu,\kappa} + \Gamma^{\lambda}{}_{\alpha\nu}\,\Gamma^{\alpha}{}_{\mu\kappa} - \Gamma^{\lambda}{}_{\alpha\kappa}\,\Gamma^{\alpha}{}_{\mu\nu}\,, \tag{32}$$

$$R_{\mu\nu} \;:=\; R^{\lambda}{}_{\mu\lambda\nu}\,. \tag{33}$$

The missing boundary term in the variation of $\mathcal{L}_H$ follows from an identity, which is universally valid (see [15] for the proof) for an arbitrary metric tensor g and an arbitrary symmetric connection Γ (not necessarily metric):

$$\delta\mathcal{L}_H = -\,\frac{1}{16\pi}\,\mathcal{G}^{\mu\nu}\,\delta g_{\mu\nu} - \left(\nabla_\kappa \pi_\lambda{}^{\mu\nu\kappa}\right)\,\delta\Gamma^{\lambda}{}_{\mu\nu} + \partial_\kappa\left(\pi_\lambda{}^{\mu\nu\kappa}\;\delta\Gamma^{\lambda}_{\mu\nu}\right)\;. \tag{34}$$

We see, that variation of the Hilbert Lagrangian (3) with respect to the connection can be easily calculated:

$$\frac{\delta\mathcal{L}_H}{\delta\Gamma^{\lambda}{}_{\mu\nu}} = -\nabla_\kappa \pi_\lambda{}^{\mu\nu\kappa}\,, \tag{35}$$

where ∇ denotes the covariant derivative with respect to Γ (see again [15]). An obvious algebraic identity:

$$\delta\left(\sqrt{|\det g|}\right) = \frac{1}{2}\sqrt{|\det g|}\;g^{\alpha\beta}\;\delta g_{\alpha\beta}\,, \tag{36}$$

implies yet another representation of variation of the Hilbert Lagrangian density:

$$\begin{aligned} R_{\mu\nu}\,\delta\pi^{\mu\nu} &= \frac{1}{16\pi}\,R_{\mu\nu}\,\delta\left(\sqrt{|\det g|}\,g^{\mu\nu}\right) = \frac{\sqrt{|\det g|}}{16\pi}\left(R_{\mu\nu}\,\delta g^{\mu\nu} + \frac{1}{2}\,R\,g^{\alpha\beta}\,\delta g_{\alpha\beta}\right) = \\ &= -\frac{\sqrt{|\det g|}}{16\pi}\left(R^{\mu\nu} - \frac{1}{2}\,R\,g^{\mu\nu}\right)\,\delta g_{\mu\nu} = -\frac{1}{16\pi}\,\mathcal{G}^{\mu\nu}\,\delta g_{\mu\nu}\,, \end{aligned} \tag{37}$$

which enables us to rewrite (34) in an equivalent form:

$$\delta\mathcal{L}_H = R_{\mu\nu}\,\delta\pi^{\mu\nu} - \left(\nabla_\kappa \pi_\lambda{}^{\mu\nu\kappa}\right)\,\delta\Gamma^{\lambda}{}_{\mu\nu} + \partial_\kappa\left(\pi_\lambda{}^{\mu\nu\kappa}\,\delta\Gamma^{\lambda}{}_{\mu\nu}\right)\;. \tag{38}$$

In case of the purely metric Hilbert Lagrangian (3) without any matter, the second term in both (34) and (38) vanishes automatically – see definitions (29), (30) – which implies the following field equation:

$$\nabla_\kappa \pi_\lambda{}^{\mu\nu\kappa} = 0 \iff \nabla_\kappa g_{\mu\nu} = 0 \iff \Gamma^{\lambda}{}_{\mu\nu} = \mathring{\Gamma}^{\lambda}{}_{\mu\nu}\,. \tag{39}$$

Hence, we end up with:

$$\delta \mathcal{L}_H(g,\,\partial g,\,\partial^2 g) = -\frac{1}{16\pi}\,\overset{\circ}{\mathcal{G}}{}^{\mu\nu}\,\delta g_{\mu\nu} + \partial_\kappa\left(\pi_\lambda{}^{\mu\nu\kappa}\,\delta\,\overset{\circ}{\Gamma}{}^\kappa{}_{\lambda\mu}\right) = \tag{40}$$

$$= \overset{\circ}{R}_{\mu\nu}\,\delta\pi^{\mu\nu} + \partial_\kappa\left(\pi_\lambda{}^{\mu\nu\kappa}\,\delta\,\overset{\circ}{\Gamma}{}^\kappa{}_{\lambda\mu}\right)\,, \tag{41}$$

where the last, boundary term in both (equivalent) formulae: (40) and (41), represents the missing term in the Wheeler-Misner-Thorn formula (21.86), page 520 (a circle above geometric objects denotes their metricity, but the formula is valid for an arbitrary symmetric connection $\Gamma^\kappa{}_{\lambda\mu}$, too).

The boundary term $\partial_\kappa\left(\pi_\lambda{}^{\mu\nu\kappa}\delta\Gamma^\lambda{}_{\mu\nu}\right)$ in the variational formula provides a strong argument for the affine approach, where connection $\Gamma^\kappa{}_{\lambda\mu}$ plays role of the gravitational field configuration, whereas the metric tensor, encoded by the tensor-density π, plays role of its canonically conjugate momentum, according to formulae (30) and (29).

There is also a strong *physical* argument, based on the Newton's First Law, for choosing the connection Γ (instead of the metric tensor) as the fundamental configuration variable of the gravitational field (see [19]).

3 Mathematical structure of the Palatini variational principle. Emergence of the non-metricity

Below, we are going to use above formalism of "symplectic relations" (in contrast to the "least action principle", based on the "spacetime boundary value problem", which is strictly forbidden by the mathematical structure of the theory).

The modern Palatini approach is based on the following observation: derivatives of the metric enter linearly into the Levi-Civita connection (4) and, whence, to calculate variation of (2) it is useful to "change variables" in space of second jets of the metric: from $(g,\,\partial g,\,\partial^2 g)$ to $(g,\,\Gamma,\,\partial\Gamma)$. Next, one can treat both the metric tensor and the connection coefficients as independent quantities: we do not assume *a priori* the metricity condition (39) of the connection, but derive it as one of the Euler-Lagrange equations. This simple trick is often called "the Palatini method of variation", although it has been earlier used by Hilbert, Weyl and Einstein himself. In fact, the originality of the Palatini paper [4] with respect to these authors consists in the fact, that when calculating variation $\delta\mathcal{L}_H$, he was able to select properly the contribution due to the variation $\delta\Gamma$ of connection. However, in those days, the connection was not regarded as an independent geometric object: only "Christoffel symbols" were known, equivalent to our "metric connection" $\overset{\circ}{\Gamma}$. This fact obscures considerably the understanding of the Palatini's contribution.

We see that the contravariant density of metric $\pi^{\mu\nu}$ (see formula (29)) plays role of the "momentum canonically conjugate to the connection" (derivative of the Lagrangian density with respect to the derivatives of the connection). As already mentioned in the previous Section (see formula (35)), variation of the Hilbert Lagrangian with respect to the connection is following:

$$\frac{\delta \mathcal{L}_H}{\delta \Gamma^\lambda{}_{\mu\nu}} = -\nabla_\kappa \pi_\lambda{}^{\mu\nu\kappa} \,, \tag{42}$$

where ∇ denotes the covariant derivative with respect to the symmetric connection Γ. Consequently, variation of the total metric Lagrangian density $\mathcal{L}_g = \mathcal{L}_{matt} + \mathcal{L}_H$ equals:

$$0 = \frac{\delta \mathcal{L}_g}{\delta \Gamma^\lambda{}_{\mu\nu}} \;:=\; \frac{\partial \mathcal{L}_g}{\partial \Gamma^\lambda{}_{\mu\nu}} - \partial_\kappa \frac{\partial \mathcal{L}_g}{\partial \Gamma^\lambda{}_{\mu\nu\kappa}} = \frac{\partial \mathcal{L}_{matt}}{\partial \Gamma^\lambda{}_{\mu\nu}} - \nabla_\kappa \pi_\lambda{}^{\mu\nu\kappa} \,, \tag{43}$$

$$\begin{aligned} 0 = \frac{\delta \mathcal{L}_g}{\delta g_{\mu\nu}} \;&:=\; \frac{\partial \mathcal{L}_g}{\partial g_{\mu\nu}} - \partial_\kappa \frac{\partial \mathcal{L}_g}{\partial g_{\mu\nu\kappa}} = \frac{\partial \mathcal{L}_g}{\partial g_{\mu\nu}} = \\ &= \frac{1}{16\pi} \left(8\pi \mathcal{T}^{\mu\nu} - \mathcal{G}^{\mu\nu} \right) \,, \end{aligned} \tag{44}$$

equivalently:

$$\frac{\partial \mathcal{L}_{matt}}{\partial \Gamma^\lambda{}_{\mu\nu}} \;=\; \nabla_\kappa \pi_\lambda{}^{\mu\nu\kappa} \,, \tag{45}$$

$$\mathcal{T}^{\mu\nu} =: 2\,\frac{\partial \mathcal{L}_{matt}}{\partial g_{\mu\nu}} \;=\; \frac{1}{8\pi}\,\mathcal{G}^{\mu\nu} \,. \tag{46}$$

We see that, in particular case of a connection - independent matter Lagrangian (6), we have

$$\frac{\partial \mathcal{L}_{matt}}{\partial \Gamma^\lambda{}_{\mu\nu}} = 0 \,, \tag{47}$$

and, whence, Euler-Lagrange equation (43) (variation with respect to the connection) reduces (see eqn. (39)) to metricity condition $\Gamma = \overset{\circ}{\Gamma}$. Hence, what was assumed *a priori* in the purely metric approach, here, in the Palatini approach (i.e. in the mixed – metric-affine approach), is obtained as one of field equations, i.e. as a result of the variational principle.

Example: In electrodynamics, which was historically the first example, carefully analysed by Hilbert (see [1]), we have:

$$\mathcal{L}_{matt} = -\frac{1}{4}\sqrt{|\det g|}\, f_{\mu\nu}\, f_{\alpha\beta}\, g^{\mu\alpha}\, g^{\nu\beta} \,,$$

where

$$f_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = 2A_{[\nu,\mu]}$$

and the electromagnetic four-potential A_μ plays role of the matter field ϕ. We see that this matter Lagrangian fulfils (47) and, whence, metricity condition can be either assumed *a priori* (metric picture) or obtained as one of the field equations (Palatini picture).

According to equation (8), the role of the momentum "$p^{\mu\lambda}$", canonically conjugate to the "matter variable" A_μ, is assumed by the Faraday contravariant tensor-density:

$$\mathcal{F}^{\mu\lambda} := \frac{\partial \mathcal{L}_{matt}}{\partial A_{\mu,\lambda}} = \sqrt{|\det g|}\, f_{\alpha\beta}\, g^{\alpha\mu}\, g^{\beta\lambda} = \sqrt{|\det g|}\, f^{\mu\lambda} \,.$$

Hence, Euler-Lagrange equation (7) encodes the Maxwell equations $\partial_\nu \mathcal{F}^{\mu\nu} = 0$, whereas (10) becomes the Maxwell (symmetric!) energy-momentum tensor density

$$\mathcal{T}^{\mu\nu} = 2\,\frac{\partial \mathcal{L}_{matt}}{\partial g_{\mu\nu}} = \sqrt{|\det g|}\, \left[f^{\mu}{}_{\beta}\, f^{\nu\beta} - \frac{1}{4}\, g^{\mu\nu}\, f_{\alpha\beta}\, f^{\alpha\beta} \right] . \qquad (48)$$

This means that

$$\mathcal{T}^{\mu\nu} = \sqrt{|\det g|}\; T^{\mu\nu} \,,$$

with:

$$T^{\mu\nu} = \frac{2}{\sqrt{|\det g|}}\, \frac{\partial \mathcal{L}_{matt}}{\partial g_{\mu\nu}} = f^{\mu}{}_{\beta}\, f^{\nu\beta} - \frac{1}{4}\, g^{\mu\nu}\, f_{\alpha\beta}\, f^{\alpha\beta} \,.$$

There is no doubt that, indeed, this quantity describes the energy-momentum density carried by the Maxwell field (cf. [14], [15]).

Unfortunately, the naive implementation of the above "Palatini method" fails when the matter Lagrangian depends upon connection coefficients Γ, which are contained in covariant derivatives of the matter fields. This happens for generic matter fields, like vector, spinor or tensor fields, where covariant derivatives are necessary as the "building blocks" of the coordinate-invariant matter Lagrangian density. Hence, in a generic case we have:

$$\mathcal{L}_{matt} = \mathcal{L}_{matt}(\phi, \partial\phi, g, \Gamma) \,, \qquad (49)$$

and, consequently:

$$\mathcal{P}^{\mu\nu}{}_{\lambda} := \frac{\partial \mathcal{L}_{matt}}{\partial \Gamma^{\lambda}{}_{\mu\nu}} \neq 0\,. \tag{50}$$

We see, that the metricity of the connection (equation (39)) **is not** recovered! Instead, we would have obtained the following value of the non-metricity of the connection:

$$\nabla_{\kappa}\pi_{\lambda}{}^{\mu\nu\kappa} = \mathcal{P}^{\mu\nu}{}_{\lambda} := \frac{\partial \mathcal{L}_{matt}}{\partial \Gamma^{\lambda}{}_{\mu\nu}} \neq 0\,. \tag{51}$$

Given the value of the (newly defined above) field $\mathcal{P}^{\mu\nu}{}_{\lambda}$, this equation can be easily solved with respect to the connection Γ, i.e. Γ can be uniquely reconstructed in the form:

$$\Gamma^{\lambda}{}_{\mu\nu} = \overset{\circ}{\Gamma}{}^{\lambda}{}_{\mu\nu} + N^{\lambda}{}_{\mu\nu}\,. \tag{52}$$

This way, the covariant derivative of the metric: $\nabla\pi$, calculated with respect to the connection Γ, splits into the sum: the covariant derivative $\overset{\circ}{\nabla}\pi$, which vanishes identically, plus a combination of N's multiplied by π. Finally, we obtain the following, linear equation for $N^{\lambda}{}_{\mu\nu}$:

$$\mathcal{P}^{\mu\nu}{}_{\lambda} = \pi^{\mu\alpha}N^{\nu}{}_{\lambda\alpha} + \pi^{\nu\alpha}N^{\mu}{}_{\lambda\alpha} - \pi^{\mu\nu}N^{\alpha}{}_{\lambda\alpha} - \frac{1}{2}\left(\delta^{\mu}_{\lambda}N^{\nu}{}_{\alpha\beta} + \delta^{\nu}_{\lambda}N^{\mu}{}_{\alpha\beta}\right)\pi^{\alpha\beta}\,, \tag{53}$$

which can easily be solved:

$$\begin{aligned} N^{\kappa}{}_{\lambda\mu} \;=\;& \frac{8\pi}{\sqrt{|\det g|}}\, g^{\kappa\sigma} \times \Big[\mathcal{P}_{\sigma\lambda\mu} + \mathcal{P}_{\sigma\mu\lambda} - \mathcal{P}_{\lambda\mu\sigma} + g_{\lambda\mu}\left(\mathcal{P}_{\sigma\nu}{}^{\nu} - \frac{1}{2}\mathcal{P}^{\nu}{}_{\nu\sigma}\right) + \\ & + \mathcal{P}^{\nu}{}_{\nu(\lambda}\, g_{\mu)\sigma} - \frac{2}{3}\, g_{\sigma(\lambda}\, \mathcal{P}_{\mu)\nu}{}^{\nu}\Big]\,. \end{aligned} \tag{54}$$

(proof in the Appendix F, see also Appendix B).

In the present paper we show that it is worthwhile to use the above non-metric connection (52), arising from the naive implementation (51) of the "Palatini method of variation", because it simplifies considerably the standard description of the canonical structure of General Relativity Theory. In particular, it is precisely the one, which describes properly the field energy (cf. [5], [6], [7]) and also arises in its "purely affine" formulation (cf. [10], [15], [12]).

3.1 Simplified version of the Palatini principle

To analyse better the mathematical structure of the Palatini principle, let us begin with the special case (6), where the matter Lagrangian density $\mathcal{L}_{matt}$ does not depend upon derivatives of the metric tensor (i.e. upon connection $\overset{\circ}{\Gamma}$). We see that the "on shell" variation of the matter Lagrangian density can be written as:

$$\delta\mathcal{L}_{matt}(\phi, \partial\phi, g) = \frac{\partial\mathcal{L}_{matt}}{\partial g_{\mu\nu}}\, \delta g_{\mu\nu} + \partial_\lambda \left(p^\lambda\, \delta\phi\right) . \tag{55}$$

Because $\mathcal{L}_{matt}$ does not depend upon derivatives of the metric tensor (i.e. upon connection), we can add an extra, trivial term to formula (55):

$$\delta\mathcal{L}_{matt}(\phi, \partial\phi, g) = \frac{\partial\mathcal{L}_{matt}}{\partial g_{\mu\nu}}\, \delta g_{\mu\nu} + \mathcal{P}^{\mu\nu}{}_{\lambda}\, \delta\Gamma^{\lambda}{}_{\mu\nu} + \partial_\lambda \left(p^\lambda\, \delta\phi\right) , \tag{56}$$

because it is equivalent to an extra, trivial field equation:

$$\mathcal{P}^{\mu\nu}{}_{\lambda} := \frac{\partial\mathcal{L}_{matt}}{\partial\Gamma^{\lambda}{}_{\mu\nu}} = 0 . \tag{57}$$

Formulae (56) and (57) provide the simplest proof of the validity of the "Palatini method of variation" for a Lagrangian density of the type (6): instead of the second order variation of (2) with respect to the metric g, we may equivalently perform the first order variation with respect to the two (*a priori* independent) geometric fields: g and Γ. Indeed, denoting

$$\mathcal{L}_P(\phi, \partial\phi, g, \Gamma, \partial\Gamma) := \mathcal{L}_{matt}(\phi, \partial\phi, g) + \frac{\sqrt{\det g}}{16\pi}\, R(g, \Gamma, \partial\Gamma) , \tag{58}$$

and using (34) together with (56), we obtain:

$$\begin{aligned} \delta\mathcal{L}_P \;=\; & \frac{1}{16\pi}\left(8\pi\mathcal{T}^{\mu\nu} - \mathcal{G}^{\mu\nu}\right)\delta g_{\mu\nu} + \left(\mathcal{P}^{\mu\nu}{}_{\lambda} - \nabla_\kappa \pi_{\lambda}{}^{\mu\nu\kappa}\right)\delta\Gamma^{\lambda}{}_{\mu\nu} + \\ & + \partial_\kappa \left(p^\kappa\, \delta\phi + \pi_{\lambda}{}^{\mu\nu\kappa}\, \delta\Gamma^{\lambda}{}_{\mu\nu}\right) , \end{aligned} \tag{59}$$

which simply means that its "variational derivatives" are equal to (43) – (44). Moreover, because quantity $\mathcal{P}$ vanishes (see (57)), equation (43) reduces to (39). This means that the metricity condition of the connection – which was not assumed *a priori* – is obtained as one of the Euler-Lagrange equations of theory, namely equation (39).

We conclude, that in case of a special matter Lagrangian $\mathcal{L}_{matt}$, fulfilling (6) (like electrodynamics), the original, purely metric, second order variational principle (2), where variation is performed with respect to the metric tensor exclusively, is equivalent to the first order "Palatini variational principle", where variation is performed with respect to both geometric quantities: connection and metric, independently.

3.2 Generic case

The above equivalence of the two "methods of variation" **does not hold** in a generic case of a Lagrangian density $\mathcal{L}_{matt}(\phi, \partial\phi, g, \overset{\circ}{\Gamma})$, i.e. where (57) is no longer true, because the corresponding Euler-Lagrange equation, resulting from variation with respect to Γ:

$$\nabla_\kappa \pi_\lambda{}^{\mu\nu\kappa} = \mathcal{P}^{\mu\nu}{}_\lambda := \frac{\partial \mathcal{L}_{matt}}{\partial \overset{\circ}{\Gamma}{}^\lambda{}_{\mu\nu}} \neq 0 \,, \tag{60}$$

would imply the non-metricity of the connection i.e. the theory which is – naively – nonequivalent to the original Einstein theory.

Example: If $\phi = (\phi^\alpha)$ is a vector field, then unique way to construct an invariant scalar out of derivatives of ϕ is to use covariant derivatives:

$$\overset{\circ}{\nabla}_\beta \phi^\alpha = \partial_\beta \phi^\alpha + \overset{\circ}{\Gamma}{}^\alpha{}_{\beta\sigma}\, \phi^\sigma \,,$$

and, whence, according to (60) we have the non-metricity tensor, which does not vanish identically:

$$\mathcal{P}^{\lambda\mu}{}_\kappa = \frac{\partial \mathcal{L}_{matt}}{\partial \overset{\circ}{\Gamma}{}^\kappa{}_{\lambda\mu}} = \frac{\partial \mathcal{L}_{matt}}{\partial \left(\overset{\circ}{\nabla}_\beta \phi^\alpha\right)} \frac{\partial \left(\overset{\circ}{\nabla}_\beta \phi^\alpha\right)}{\partial \overset{\circ}{\Gamma}{}^\kappa{}_{\lambda\mu}} = \frac{\partial \mathcal{L}_{matt}}{\partial \phi^\alpha{}_{,\beta}}\, \delta^\alpha_\kappa\, \delta^{(\lambda}_\beta\, \delta^{\mu)}_\sigma\, \phi^\sigma = p_\kappa{}^{(\lambda}\, \phi^{\mu)} \,. \tag{61}$$

We see, that the naive (i.e. straightforward) implementation of the "Palatini method of variation" implies the non-metricity of the connection and, consequently, field equations which are *a priori* nonequivalent with the original metric theory.

However, we are going to prove in this paper, that the universal "Palatini variational principle", consisting in varying with respect to independent geometric fields Γ and g *does exist*, if we only accept the above non-metricity, defined by equation (60). The resulting field theory **is not** a new physical theory, but merely a mathematically equivalent reformulation of the standard *purely metric* theory, based on Lagrangian density (2). Even if formulated in terms of the non-metric connection, when recalculated back in terms of the metric g and their derivatives, it is perfectly equivalent to the original, metric theory, as will be proved in the sequel. The essence of our construction can be formulated simply as follows: it is worth combining the information about the matter field contained in the "non-metricity-tensor N" (given by formula (54)) with the metric connection $\overset{\circ}{\Gamma}$, because the use of the resulting non-metric connection $\Gamma = \overset{\circ}{\Gamma} + N$ significantly simplifies the structure of the theory.

Technically, the simplest way to find such a universal Palatini formulation goes through the purely affine theory. As will be proved in the next Section, affine picture arises naturally and provides the easiest method to simplify the canonical structure of the theory.

4 The universal affine formulation of General Relativity Theory

Let us, therefore, begin with the purely metric formulation of General Relativity Theory in case of a generic matter Lagrangian density (5). Its variation has the following form:

$$\delta\mathcal{L}_{matt}(\phi,\, \partial\phi,\, g,\, \overset{\circ}{\Gamma}) = \frac{\partial \mathcal{L}_{matt}}{\partial g_{\mu\nu}}\, \delta g_{\mu\nu} + \mathcal{P}^{\mu\nu}{}_{\lambda}\, \delta\, \overset{\circ}{\Gamma}{}^{\lambda}{}_{\mu\nu} + \partial_\lambda \left(p^\lambda\, \delta\phi\right) , \tag{62}$$

where, in general, the field $\mathcal{P}$ does not vanish identically:

$$\mathcal{P}^{\mu\nu}{}_{\lambda} := \frac{\partial \mathcal{L}_{matt}}{\partial\, \overset{\circ}{\Gamma}{}^{\lambda}{}_{\mu\nu}} \neq 0\,. \tag{63}$$

We have the following

Lemma 4.1. *The term* $\mathcal{P}^{\mu\nu}{}_{\lambda}\, \delta\, \overset{\circ}{\Gamma}{}^{\lambda}{}_{\mu\nu}$ *in formula* (62) *can be rewritten as follows:*

$$\mathcal{P}^{\mu\nu}{}_{\lambda}\, \delta\, \overset{\circ}{\Gamma}{}^{\lambda}{}_{\mu\nu} = \partial_\kappa \left(\mathcal{R}^{\mu\nu\kappa}\, \delta g_{\mu\nu}\right) - \left(\overset{\circ}{\nabla}{}_{\kappa} \mathcal{R}^{\mu\nu\kappa}\right)\, \delta g_{\mu\nu}\,,$$

where:

$$\mathcal{R}^{\mu\nu\kappa} := \frac{1}{2} \left(\mathcal{P}^{\kappa\mu\nu} + \mathcal{P}^{\kappa\nu\mu} - \mathcal{P}^{\mu\nu\kappa}\right) \tag{64}$$

and $\overset{\circ}{\nabla}$ *is the covariant derivative with respect to the Levi-Civita connection* $\overset{\circ}{\Gamma}$.

The proof of this simple Lemma is given in Appendix A. Hence, formula (62), generating field dynamics, can be rewritten as:

$$\delta\mathcal{L}_{matt} = \left(\frac{\partial \mathcal{L}_{matt}}{\partial g_{\mu\nu}} - \overset{\circ}{\nabla}_{\kappa} \mathcal{R}^{\mu\nu\kappa} \right) \delta g_{\mu\nu} + \partial_\kappa \left(\mathcal{R}^{\mu\nu\kappa}\, \delta g_{\mu\nu} + p^\kappa \delta\phi \right) . \tag{65}$$

We see that (as mentioned already in the Introduction, cf. formula (13)) derivative of the matter Lagrangian with respect to the metric in (62), has been replaced now by its variational derivative:

$$\frac{\delta \mathcal{L}_{matt}(\phi, \partial\phi, g, \partial g)}{\delta g_{\mu\nu}} = \frac{\partial \mathcal{L}_{matt}}{\partial g_{\mu\nu}} - \partial_\kappa \frac{\partial \mathcal{L}_{matt}}{\partial g_{\mu\nu,\kappa}} = \frac{\partial \mathcal{L}_{matt}}{\partial g_{\mu\nu}} - \overset{\circ}{\nabla}_{\kappa} \mathcal{R}^{\mu\nu\kappa} . \tag{66}$$

Together with variation of the Hilbert Lagrangian given by (40), we obtain this way the following, universal formula for the variation of the metric Lagrangian density $\mathcal{L}_g = \mathcal{L}_H + \mathcal{L}_{matt}$:

$$\begin{aligned} \delta\mathcal{L}_g(\phi,\, \partial\phi,\, g,\, \partial g,\, \partial^2 g) \;&=\; \left[\frac{\partial \mathcal{L}_{matt}}{\partial g_{\mu\nu}} - \frac{1}{16\pi} \overset{\circ}{\mathcal{G}}{}^{\mu\nu} - \overset{\circ}{\nabla}_{\kappa} \mathcal{R}^{\mu\nu\kappa} \right] \delta g_{\mu\nu} + \\ &\quad + \partial_\kappa \left(\mathcal{R}^{\mu\nu\kappa}\, \delta g_{\mu\nu} + p^\kappa\, \delta\phi + \pi_\lambda{}^{\mu\nu\kappa} \delta\, \overset{\circ}{\Gamma}{}^\lambda{}_{\mu\nu} \right) , \end{aligned} \tag{67}$$

(we have put the circle above the Einstein tensor density $\overset{\circ}{\mathcal{G}}$, in order to stress that it is calculated for the Levi-Civita metric connection $\overset{\circ}{\Gamma}$.) Above formula splits into the "volume part" (describing Euler-Lagrange field equations, vanishing "on shell") and the "boundary part", describing the *Lagrangian control-mode* of the theory.

As we show in Appendix E, the troublesome term $\overset{\circ}{\nabla}_{\kappa} \mathcal{R}^{\mu\nu\kappa}$ combines, together with the metric Einstein-tensor-density $\overset{\circ}{\mathcal{G}}{}^{\mu\nu}$ to its non-metric analogue $\mathcal{G}^{\mu\nu}$ (plus an extra term, which also finds a nice geometric interpretation and is discussed in the sequel).

Implementing our "*on shell*" philosophy – see Section 2 – we obtain the volume term producing Einstein equations:

$$\frac{\partial \mathcal{L}_{matt}}{\partial g_{\mu\nu}} - \frac{1}{16\pi} \overset{\circ}{\mathcal{G}}{}^{\mu\nu} - \overset{\circ}{\nabla}_{\kappa} \mathcal{R}^{\mu\nu\kappa} = 0 \tag{68}$$

and the "*on shell*" – variation of $\mathcal{L}_g$:

$$\delta\mathcal{L}_g = \partial_\kappa \left(\mathcal{R}^{\mu\nu\kappa}\, \delta g_{\mu\nu} + p^\kappa\, \delta\phi + \pi_\lambda{}^{\mu\nu\kappa}\, \delta\, \overset{\circ}{\Gamma}{}^\lambda{}_{\mu\nu} \right) . \tag{69}$$

Unfortunately, above formula is not yet fully satisfactory because the metric g appears here in a double role: as a control parameter ($\delta g_{\mu\nu}$) and as the response parameter ($\pi_\lambda{}^{\mu\nu\kappa}$). Mathematically, this means that the underlying symplectic structure is degenerate: some combinations of the "momenta" $\pi^{\mu\nu}$ (response) are equal to some combinations of the "configuration" $g_{\mu\nu}$ (control). Now, we are going to reduce this degeneracy, rewriting the formula in terms of independent parameters. First, we rewrite the term $\partial_\kappa \left(\mathcal{R}^{\mu\nu\kappa}\, \delta g_{\mu\nu}\right)$ according to following lemma:

Lemma 4.2. *The following identity holds:*

$$\partial_\kappa \left(\mathcal{R}^{\mu\nu\kappa}\, \delta g_{\mu\nu}\right) = \partial_\nu \left[\pi_\kappa{}^{\lambda\mu\nu}\, \delta N^\kappa{}_{\lambda\mu}\right] + \delta \left[\overset{\circ}{\nabla}_\kappa \mathcal{R}_\sigma{}^{\sigma\kappa}\right] , \tag{70}$$

where

$$N^\kappa{}_{\lambda\mu} \; := \; \frac{16\pi}{\sqrt{|\det g|}} \left[\mathcal{R}_{\lambda\mu}{}^{\kappa} - \frac{1}{2}\mathcal{R}_\sigma{}^{\sigma\kappa}\, g_{\lambda\mu} - \frac{2}{3}\left(\delta^\kappa_{(\lambda}\, \mathcal{R}_{\mu)\sigma}{}^{\sigma} - \frac{1}{2}\mathcal{R}^\sigma{}_{\sigma(\lambda}\, \delta^\kappa_{\mu)}\right)\right] , \tag{71}$$

is simply equation (54) *written in terms of the auxiliary quantity* $\mathcal{R}$ *defined by* (64)*, instead of the original object* $\mathcal{P}$*.*

Proof of this Lemma is presented in Appendix B. Using this identity we rewrite the metric variation (69) as follows:

$$\delta\mathcal{L}_g = \partial_\kappa \left[p^\kappa\, \delta\phi + \pi_\lambda{}^{\mu\nu\kappa} \delta\left(\overset{\circ}{\Gamma}{}^\lambda{}_{\mu\nu} + N^\lambda{}_{\mu\nu}\right)\right] + \delta \left[\overset{\circ}{\nabla}_\kappa \mathcal{R}_\sigma{}^{\sigma\kappa}\right] . \tag{72}$$

We see, that our non-metric connection $\Gamma = \overset{\circ}{\Gamma} + N$, already defined in (52), arises here in a natural way:

$$\delta\mathcal{L}_g = \partial_\kappa \left[p^\kappa\, \delta\phi + \pi_\lambda{}^{\mu\nu\kappa}\, \delta\Gamma^\lambda{}_{\mu\nu}\right] + \delta \left[\overset{\circ}{\nabla}_\kappa \mathcal{R}_\sigma{}^{\sigma\kappa}\right] . \tag{73}$$

Now, we perform the Legendre transformation between metric and the connection. As the first step, we put the last term (the complete variation) on the left-hand side and obtain finally the universal affine Lagrangian:

$$\mathcal{L}_A := \mathcal{L}_g - \overset{\circ}{\nabla}_\kappa \mathcal{R}_\sigma{}^{\sigma\kappa} , \tag{74}$$

where now only the symmetric (but not necessarily metric!) connection Γ and the matter fields ϕ play role of independent configuration variables:

$$\begin{aligned} \delta\mathcal{L}_A \;&=\; \partial_\kappa \left[p^\kappa\, \delta\phi + \pi_\lambda{}^{\mu\nu\kappa}\, \delta\Gamma^\lambda{}_{\mu\nu}\right] = \left(\partial_\kappa p^\kappa\right) \delta\phi + p^\kappa\, \delta\phi_\kappa + \\ &\quad + \left(\partial_\kappa \pi_\lambda{}^{\mu\nu\kappa}\right) \delta\Gamma^\lambda{}_{\mu\nu} + \pi_\lambda{}^{\mu\nu\kappa}\, \delta\Gamma^\lambda{}_{\mu\nu,\kappa} . \end{aligned} \tag{75}$$

The metric $g_{\mu\nu}$ (represented here by the momentum $\pi^{\mu\nu}$) and its derivatives (represented by $\partial_\kappa \pi_\lambda{}^{\mu\nu\kappa}$) are now shifted to the level of "response parameters": they arise as derivatives of the new affine Lagrangian with respect to $\Gamma^\lambda{}_{\mu\nu}$ and its derivatives $\Gamma^\lambda{}_{\mu\nu,\kappa}$. But, to be consistent, the second step of the Legendre transformation must follow the first one: the function $\mathcal{L}_A$ must be expressed in terms of the new control parameters, whereas the old control parameters must be eliminated with the help of the appropriate field equations. We have an analogous situation in thermodynamics where, to perform transition from the adiabatic to the thermostatic mode, it is not sufficient to replace formula $\delta U(V,S) = -p\delta V + T\delta S$ by $\delta\,(U - TS) = -p\delta V - S\delta T$, but the new generating function $H := U - TS$ (the Helmholtz "free energy") must be expressed in terms of the new control parameters: $H = H(V,T)$. A similar phenomenon occurs during any Legendre transformation which we commonly use in theoretical physics. For example, formula

$$H(p,q,\dot{q}) = p\,\dot{q} - L(q,\dot{q})\,, \qquad p := \frac{\partial L}{\partial \dot{q}}\,, \tag{76}$$

describing transformation from the Lagrangian control mode to the Hamiltonian one in classical mechanics means, that we have to eliminate the Lagrangian variable $\dot{q}$ and express it in terms of the "true Hamiltonian control parameters" (p,q). Mathematically rigorous definition of these geometric structures can be found in [11], [17]. We stress, however, that no assumption about the non-degeneracy of the original Lagrangian density is necessary here: theories with constraints, where only a subspace of control parameters is physically accesible and, consequently, the "control – response" relation is no longer equivocal, (i.e.: non-physical "gauge parameters" arise), fit perfectly into this framework.

To perform the second step of the Legendre transformation between the metric and the affine picture in the gravity theory, we observe that the numerical value of the affine Lagrangian $\mathcal{L}_A$ equals:

$$\mathcal{L}_A := \mathcal{L}_g - \overset{\circ}{\nabla}_\kappa \mathcal{R}_\sigma{}^{\sigma\kappa} = \mathcal{L}_{matt}(\phi, \partial\phi, g, \overset{\circ}{\Gamma}) + \mathcal{L}_H(g_{\mu\nu}, \overset{\circ}{R}_{\mu\nu}) - \overset{\circ}{\nabla}_\kappa \mathcal{R}_\sigma{}^{\sigma\kappa}\,. \tag{77}$$

Hence, we have *a priori*:

$$\mathcal{L}_A = \mathcal{L}_A(\phi,\, \partial\phi,\, g_{\mu\nu},\, \partial_\kappa g_{\mu\nu},\, \Gamma^\lambda{}_{\mu\nu},\, R_{\mu\nu})\,. \tag{78}$$

But now, the variables $(g_{\mu\nu},\, \partial_\kappa g_{\mu\nu})$ are no longer "control parameters" – see variational formula (75) – and have to be eliminated by virtue of dynamical equations.

For this purpose we can use two equations which enable us to reduce the number of control parameters: the non-metricity equation (60) and the Einstein equation (68):

$$\nabla_\kappa \pi_\lambda{}^{\mu\nu\kappa} = \mathcal{P}^{\mu\nu}{}_\lambda := \frac{\partial \mathcal{L}_{matt}}{\partial \overset{\circ}{\Gamma}{}^\lambda{}_{\mu\nu}} , \tag{79}$$

$$\frac{1}{16\pi} \overset{\circ}{\mathcal{G}}{}^{\mu\nu} = \frac{\partial \mathcal{L}_{matt}}{\partial g_{\mu\nu}} - \overset{\circ}{\nabla}_\kappa \mathcal{R}^{\mu\nu\kappa} . \tag{80}$$

The complete Legendre transformation from the metric picture ("control mode") to the affine picture is, therefore, given by definition (77) of the affine Lagrangian density, together with formulas (79) and (80), the latter being necessary to express "response parameters" $(g, \overset{\circ}{\Gamma})$ in terms of the new control parameters: $(\Gamma, \partial\Gamma)$.

As will be seen in the Appendix E, derivatives of the connection enter here *via* the non-metric Ricci tensor $R_{\mu\nu}$ only:

$$R_{\mu\nu} := -\Gamma^\kappa{}_{\kappa\mu,\nu} + \Gamma^\kappa{}_{\mu\nu,\kappa} - \Gamma^\sigma{}_{\kappa\mu} \Gamma^\kappa{}_{\nu\sigma} + \Gamma^\sigma{}_{\mu\nu} \Gamma^\kappa{}_{\kappa\sigma} . \tag{81}$$

and even less, namely *via* its symmetric part (the non-metric Ricci may contain a skew-symmetric part!):

$$K_{\mu\nu} := R_{(\mu\nu)} = -\Gamma^\kappa{}_{\kappa(\mu,\nu)} + \Gamma^\kappa{}_{\mu\nu,\kappa} - \Gamma^\sigma{}_{\kappa\mu} \Gamma^\kappa{}_{\nu\sigma} + \Gamma^\sigma{}_{\mu\nu} \Gamma^\kappa{}_{\kappa\sigma} , \tag{82}$$

similarly as it was the case in the metric picture. Formula expressing $K_{\mu\nu}$ in terms of its metric analogue $\overset{\circ}{R}_{\mu\nu}$ is given in Appendix E.

Practically, to solve equations (79), (80), and to express explicitly metric g and the metric connection $\overset{\circ}{\Gamma}$ in terms of K and Γ (together with the matter field ϕ) can be computationally difficult (see example in Appendix G). But, theoretically, the implicit relation between the metric and the affine control parameters is sufficient. It might well be that "what is simple in the metric picture, is complicated in the affine one and *vice versa*". The affine picture being conceptually much simpler that the metric picture, we are deeply convinced that when looking for fundamental laws of nature (e.g. description of the dark matter) one should begin with geometric structures which are simple rather in affine picture, even if their metric counterparts would be computationally complicated. From this point of view, attempts to describe dark matter by modifying slightly the metric Hilbert Lagrangian density do not look very convincing. However, there are important examples of the matter fields, where the above Legendre transformation can be performed explicitly.

Examples:

- $$\mathcal{L}_A(K_{\mu\nu}) = C \cdot \sqrt{|\det K_{\mu\nu}|}\,. \tag{83}$$

  The resulting field equations are the Einstein equations for empty space with cosmological constant $\Lambda := \frac{1}{8\pi C}$.

- $$\mathcal{L}_A(K_{\mu\nu}, \phi, \partial_\alpha \phi) = \\ = \frac{2}{m^2\phi^2}\sqrt{\left|\det\left(\frac{1}{8\pi}K_{\mu\nu} - \phi_{,\mu}\phi_{,\nu}\right)\right|}\,. \tag{84}$$

  It describes the Klein-Gordon-Einstein theory of the scalar field ϕ, having mass m and interacting with gravity.

- $$\mathcal{L}_A(K_{\mu\nu}, \partial_\alpha A_\beta) = \\ = -\frac{1}{4}\sqrt{|\det K_{\rho\sigma}|}\,(K^{-1})^{\mu\nu}\,(K^{-1})^{\alpha\beta}\,F_{\mu\alpha}\,F_{\nu\beta}\,, \tag{85}$$

  where $F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field (Faraday tensor) given in terms of its four-potential A_μ, describes the Maxwell-Einstein theory.

For more examples see, e.g., [12], [2], [3].

5 Transition to the universal Palatini picture

In this Section we introduce the universal "Palatini picture", where the configuration fields are: metric tensor g and the symetric connection Γ. In contrast to the "naive" Palatini picture, discussed in Section 3, no metricity condition for Γ is assumed *a priori*. The non-metricity equation (79) and the Einstein equation (80) arise here as Euler-Lagrange field equations, implied by variation with respect to Γ and to g, respectively.

This universal Palatini formulation can be obtained directly *via* a simple Legendre transformation from the affine picture. Let us begin, therefore, with a generic affine Lagrangian

$$\mathcal{L}_A = \mathcal{L}_A(\phi,\, \partial\phi,\, {\Gamma^\kappa}_{\lambda\mu},\, {\Gamma^\kappa}_{\lambda\mu,\nu}) = \mathcal{L}_A(\phi,\, \nabla\phi,\, K_{\mu\nu})\,, \tag{86}$$

where derivatives of the connection coefficients enter *via* the symmetric part (82) of the Ricci, exclusively. Let us introduce the "coordinate transformation":

$$(\Gamma,\, \partial\Gamma) \mapsto (\Gamma,\, K)$$

in space of control parameters.

Lemma 5.1. *The following identity holds:*

$$\partial_\kappa \left(\pi_\kappa{}^{\lambda\mu\nu}\, \delta\Gamma^\kappa{}_{\lambda\mu} \right) = \left(\nabla_\kappa \pi_\lambda{}^{\mu\nu\kappa} \right) \delta\Gamma^\lambda{}_{\mu\nu} + \pi^{\mu\nu}\, \delta K_{\mu\nu} \,. \tag{87}$$

The proof of this identity is given in Appendix D. It enables us to rewrite generating formula (75) as follows:

$$\begin{aligned} \delta\mathcal{L}_A &= \partial_\kappa \left(p^\kappa\, \delta\phi \right) + \left(\nabla_\kappa \pi_\lambda{}^{\mu\nu\kappa} \right) \delta\Gamma^\lambda{}_{\mu\nu} + \pi^{\mu\nu}\, \delta K_{\mu\nu} = & (88) \\ &= \partial_\kappa \left(p^\kappa\, \delta\phi \right) + \left(\nabla_\kappa \pi_\lambda{}^{\mu\nu\kappa} \right) \delta\Gamma^\lambda{}_{\mu\nu} - K_{\mu\nu}\, \delta\pi^{\mu\nu} + \delta \left(\pi^{\mu\nu}\, K_{\mu\nu} \right) . & (89) \end{aligned}$$

Putting the complete variation $\delta \left(\pi^{\mu\nu}\, K_{\mu\nu} \right)$ on the left-hand-side, we obtain the Universal Palatini Lagrangian:

$$\begin{aligned} \mathcal{L}_P(\phi,\, \partial\phi,\, \Gamma^\lambda{}_{\mu\nu}, g_{\mu\nu}) &:= \mathcal{L}_A - \pi^{\mu\nu}\, K_{\mu\nu} = \\ &= \mathcal{L}_A - \frac{\sqrt{|\det g|}}{16\pi}\, R \,, \end{aligned} \tag{90}$$

where the subtracted term is analogous to the Hilbert Lagrangian $\mathcal{L}_H(g,\, \Gamma,\, \partial\Gamma)$ for a general (possibly non-metric) connection, like in formula (31). Due to (89), this Lagrangian generates field equations according to:

$$\delta\mathcal{L}_P = \left(\nabla_\kappa \pi_\lambda{}^{\mu\nu\kappa} \right) \delta\Gamma^\lambda{}_{\mu\nu} + \frac{1}{16\pi}\, \mathcal{G}^{\mu\nu}\, \delta g_{\mu\nu} + \partial_\kappa \left(p^\kappa\, \delta\phi \right) \,, \tag{91}$$

(cf. identity (37)) or, equivalently:

$$\begin{aligned} \frac{1}{16\pi}\, \mathcal{G}^{\mu\nu} &= \frac{\partial\mathcal{L}_P}{\partial g_{\mu\nu}} \,, \\ \nabla_\kappa \pi_\lambda{}^{\mu\nu\kappa} &= \frac{\partial\mathcal{L}_P}{\partial\Gamma^\lambda{}_{\mu\nu}} \,, \end{aligned} \tag{92}$$

(cf. (79) and (80)).

In case of a special theory (6), where matter Lagrangian density does not depend upon derivatives of the metric tensor (i.e. upon the metric connection $\overset{\circ}{\Gamma}$), the non-metricity of Γ vanishes (see Lemmas 4.1 and 4.2) and we have $K_{\mu\nu} = \overset{\circ}{R}_{\mu\nu}$ (cf. Appendix E). Consequently, definition (90) implies in that case:

$$\mathcal{L}_P = \mathcal{L}_A - \pi^{\mu\nu} K_{\mu\nu} = \mathcal{L}_{matt} + \mathcal{L}_H(g,\, \partial g,\, \partial^2 g) - \pi^{\mu\nu}\, \overset{\circ}{R}_{\mu\nu} = \mathcal{L}_{matt}\,, \tag{93}$$

i.e. our "universal Palatini picture" reduces to the classical ("naive") Palatini method of variation presented in Section 3. But, in a generic case (5), the two pictures differ from each other. In particular, we have:

$$\begin{aligned} \mathcal{L}_P &= \mathcal{L}_{matt} - \pi^{\mu\nu} \left(N^\sigma{}_{\mu\nu}\, N^\kappa{}_{\kappa\sigma} - N^\sigma{}_{\kappa\mu}\, N^\kappa{}_{\nu\sigma} \right) = \\ &= \mathcal{L}_{matt} - \frac{\sqrt{|\det g|}}{16\pi}\, g^{\mu\nu} \left(N^\sigma{}_{\mu\nu}\, N^\kappa{}_{\kappa\sigma} - N^\sigma{}_{\kappa\mu}\, N^\kappa{}_{\nu\sigma} \right)\,, \end{aligned} \tag{94}$$

where formula (94) is implied by (77), (90) and the following identity, which is an easy consequence of relation between $K_{\mu\nu}$ and $\overset{\circ}{R}_{\mu\nu}$ (see Appendix E):

$$\pi^{\mu\nu}\, \overset{\circ}{R}_{\mu\nu} - \pi^{\mu\nu}\, K_{\mu\nu} - \overset{\circ}{\nabla}_\kappa \mathcal{R}_\sigma{}^{\sigma\kappa} = -\pi^{\mu\nu} \left(N^\sigma{}_{\mu\nu}\, N^\kappa{}_{\kappa\sigma} - N^\sigma{}_{\kappa\mu}\, N^\kappa{}_{\nu\sigma} \right)\,. \tag{95}$$

We see explicitly that:

$$\frac{\partial \mathcal{L}_{matt}}{\partial \Gamma^\lambda{}_{\mu\nu}} = 0 \;\Leftrightarrow\; \mathcal{P}^{\mu\nu}{}_\lambda = 0 \;\Leftrightarrow\; N^\lambda{}_{\mu\nu} = 0 \;\Rightarrow\; \mathcal{L}_P = \mathcal{L}_{matt}\,. \tag{96}$$

5.1 Fast track from the metric picture to the Palatini picture

Formula (94) provides the method of a direct transition from the metric picture to the universal Palatini picture, without transition *via* the affine picture. The formula gives directly the value of the universal Palatini Lagrangian density as the function of the following variables:

$$\mathcal{L}_P = \mathcal{L}_P(g, \phi, \nabla\phi, \Gamma, N) = \mathcal{L}_{matt}(g, \phi, \overset{\circ}{\nabla}\phi, \overset{\circ}{\Gamma}) - \pi^{\mu\nu} \left(N^\sigma{}_{\mu\nu}\, N^\kappa{}_{\kappa\sigma} - N^\sigma{}_{\kappa\mu}\, N^\kappa{}_{\nu\sigma} \right)\,, \tag{97}$$

where $\overset{\circ}{\Gamma}$ has been replaced by "$\Gamma - N$". What remains now is only elimination of the "illegal" variable N, which must be expressed in terms of the remaining ("legal") variables. For this purpose we have to solve (with respect to N) equation (53), with $\mathcal{P}$ given by (79), i.e. by the definition of non-metricity:

$$\mathcal{P}^{\mu\nu}{}_{\lambda} = \frac{\partial \mathcal{L}_{matt}}{\partial \overset{\circ}{\Gamma}{}^{\lambda}{}_{\mu\nu}} .$$

Example: As the "metric matter-Lagrangian-density" of the vector field ϕ^α, take the standard, quadratic form of its covariant derivatives:

$$\mathcal{L}_{matt}(g, \overset{\circ}{\nabla}\phi) = \frac{\sqrt{|\det g|}}{16\pi} \left(\overset{\circ}{\nabla}_{\alpha}\phi^{\beta}\right) \left(\overset{\circ}{\nabla}{}^{\alpha}\phi_{\beta}\right) .$$

We prove in Appendix G that the corresponding Palatini Lagrangian (94) is given by following formula:

$$\begin{aligned} \mathcal{L}_P(g, \nabla\phi) \;\; = \;\; & \frac{\sqrt{|\det g|}}{32\pi\, D_1\, D_2\, (\phi^2 - 1)} \Big\{ D_1\, D_2\, \left(\phi^2 - 2\right) (\nabla_\alpha\phi^\mu)\, (\nabla_\beta\phi^\nu)\, g^{\alpha\beta}\, g_{\mu\nu} + \\ & + D_1\, D_2\, \phi^2\, (\nabla_\alpha\phi^\mu)\, (\nabla_\mu\phi^\alpha) + 4\phi^2 \left(\phi^2 - 1\right) \left[\phi^\alpha\, (\nabla_\alpha\phi^\mu)\, \phi_\mu\right]^2 + \\ & + 4D_2 \left(\phi^2 - 1\right) \left[\phi^\alpha\, (\nabla_\alpha\phi^\mu)\, \phi_\mu\right] \left(\nabla_\beta\phi^\beta\right) - \phi^2\, D_2 \left(\phi^2 - 1\right) \left(\nabla_\beta\phi^\beta\right)^2 + \\ & - D_1 \left(\phi^2 - 2\right)^2 \left[(\nabla_\mu\phi^\alpha)\, \phi_\alpha\right] \left[(\nabla_\nu\phi^\beta)\, \phi_\beta\right] g^{\mu\nu} + \\ & - 2D_1 \left(\phi^2 - 2\right) \phi^2 \left[(\nabla_\mu\phi^\alpha)\, \phi_\alpha\right] \left[\phi^\beta\, (\nabla_\beta\phi^\mu)\right] + \\ & - D_1\, \phi^4 \left[\phi^\alpha\, (\nabla_\alpha\phi^\mu)\right] \left[\phi^\beta\, (\nabla_\beta\phi^\nu)\right] g_{\mu\nu} \Big\} , \end{aligned} \tag{98}$$

where D_1, D_2 are polynomials of the variable $\phi^2 := \phi_\alpha\, \phi^\alpha$ defined in Appendix G.

Example: Transformation in the opposite direction. As the "Palatini-Lagrangian-density" of the vector field X^α, take the simplest, quadratic form of its covariant (with respect to the non-metric connection Γ) derivatives:

$$\mathcal{L}_P(g,\, \nabla X) = \frac{\sqrt{|\det g|}}{16\pi} \left(\nabla_\alpha X^\mu\right) \left(\nabla_\beta X^\nu\right) g^{\alpha\beta}\, g_{\mu\nu} .$$

We prove in Appendix G that the corresponding metric matter Lagrangian density is equal to:

$$
\begin{aligned}
\mathcal{L}_{matt}(g, \overset{\circ}{\nabla} X) \quad = \quad & \frac{\sqrt{|\det g|}}{32\pi\, \widetilde{D}_1\, \widetilde{D}_2\, (X^2+1)} \Big\{ \widetilde{D}_1\, \widetilde{D}_2\, \left(X^2+2\right) \left(\overset{\circ}{\nabla}_\alpha X_\beta\right) \left(\overset{\circ}{\nabla}{}^\alpha X^\beta\right) + \\
& + \widetilde{D}_1\, \widetilde{D}_2\, X^2 \left(\overset{\circ}{\nabla}_\alpha X_\beta\right) \left(\overset{\circ}{\nabla}{}^\beta X^\alpha\right) + 4\left(X^2+1\right)\, X^2 \left[X^\alpha \left(\overset{\circ}{\nabla}_\alpha X^\mu\right) X_\mu\right]^2 + \\
& + 4\widetilde{D}_2\, \left(X^2+1\right) \left[X^\alpha \left(\overset{\circ}{\nabla}_\alpha X^\mu\right) X_\mu\right] \left(\overset{\circ}{\nabla}_\beta X^\beta\right) + \\
& - X^2\, \widetilde{D}_2\, \left(X^2+1\right) \left(\overset{\circ}{\nabla}_\alpha X^\alpha\right)^2 + \\
& - \widetilde{D}_1\, \left(X^2-2\right)^2 \left[(\overset{\circ}{\nabla}_\mu X^\alpha)\, X_\alpha\right] \left[(\overset{\circ}{\nabla}{}^\mu X^\beta)\, X_\beta\right] + \\
& - 2\widetilde{D}_1\, \left(X^2+2\right)\, X^2 \left[(\overset{\circ}{\nabla}_\mu X^\alpha)\, X_\alpha\right] \left[X^\beta\, (\overset{\circ}{\nabla}_\beta X^\mu)\right] + \\
& - \widetilde{D}_1\, X^4 \left[X^\alpha\, (\overset{\circ}{\nabla}_\alpha X^\mu)\right] \left[X^\beta\, (\overset{\circ}{\nabla}_\beta X_\mu)\right] \Big\} ,
\end{aligned}
\tag{99}
$$

where $\widetilde{D}_1$ i $\widetilde{D}_2$ are polynomials of the variable $X^2 := X_\alpha\, X^\alpha$ defined in Appendix H.

The two examples illustrate nicely our conclusion that "what is simple in the metric picture can be computationally difficult in the affine picture and *vice versa*". Being conceptually much simpler (as far as the canonical structure of the theory is concerned), the affine formulation (together with its consequence: the universal Palatini picture) are – in our opinion – better suited to become the starting point for any attempt to extrapolate validity of the present Gravity Theory by many orders of magnitude, namely from our Solar System scale to the cosmological scale (scale of the Universe).

6 From affine to metric picture

The equivalence proof of the three different formulations of the Gravity Theory (purely metric, metric-affine which we call "Palatini" and – finally – purely affine) would not be complete without a brief discussion of what happens if we begin with a purely affine Lagrangian density. Even if these results are contained *implicite* in our previous Sections, we have decided, for the convenience of the reader, to give below the short equivalence proof starting from the affine picture.

Consider, therefore, a generic affine variational principle, where the field configuration is described by the first jet of a symmetric connection (Γ, $\partial\Gamma$) and the first jet of a matter field $(\phi, \partial\phi)$:

$$\begin{aligned} \delta\mathcal{L}_A(\Gamma,\,\partial\Gamma,\,\phi,\partial\phi) &= \partial_\nu\left(\mathcal{P}_\kappa{}^{\lambda\mu\nu}\,\delta\Gamma^\kappa{}_{\lambda\mu} + p^\nu\,\delta\phi\right) = \\ &= \left(\partial_\nu\mathcal{P}_\kappa{}^{\lambda\mu\nu}\right)\delta\Gamma^\kappa{}_{\lambda\mu} + \mathcal{P}_\kappa{}^{\lambda\mu\nu}\,\delta\Gamma^\kappa{}_{\lambda\mu,\nu} + \\ &\quad + \left(\partial_\nu p^\nu\right)\delta\phi + p^\nu\delta\phi_{,\nu}\,, \end{aligned} \tag{100}$$

or, equivalently,

$$\partial_\nu\mathcal{P}_\kappa{}^{\lambda\mu\nu} = \frac{\partial\mathcal{L}_A}{\partial\Gamma^\kappa{}_{\lambda\mu}}\,,\ \mathcal{P}_\kappa{}^{\lambda\mu\nu} = \frac{\partial\mathcal{L}_A}{\partial\Gamma^\kappa{}_{\lambda\mu,\nu}} \quad\Rightarrow\quad \frac{\delta\mathcal{L}_A}{\delta\Gamma^\kappa{}_{\lambda\mu,\nu}} = 0\,, \tag{101}$$

$$\partial_\nu p^\nu = \frac{\partial\mathcal{L}_A}{\partial\phi}\,,\ p^\nu = \frac{\partial\mathcal{L}_A}{\partial\phi_{,\nu}} \quad\Rightarrow\quad \frac{\delta\mathcal{L}_A}{\delta\phi} = 0\,. \tag{102}$$

As was explained in Section 2, the formula (100) is equivalent to Euler-Lagrange equations, together with definition of momenta: $\mathcal{P}_\kappa{}^{\lambda\mu\nu}$ and p^ν, canonically conjugate to $\Gamma^\kappa{}_{\lambda\mu}$ and ϕ, respectively.

To manufacture an invariant scalar-density $\mathcal{L}_A$ from derivatives of Γ, the only method is to use the unique tensorial object which can by constructed from them, namely the Riemann tensor (32):

$$R^\kappa{}_{\lambda\mu\nu} = \Gamma^\kappa{}_{\lambda\nu,\mu} - \Gamma^\kappa{}_{\lambda\mu,\nu} + \Gamma^\kappa{}_{\sigma\mu}\,\Gamma^\sigma{}_{\lambda\nu} - \Gamma^\kappa{}_{\sigma\nu}\,\Gamma^\sigma{}_{\lambda\mu}\,.$$

Moreover, the connection Γ can only enter into the game either *via* the curvature $R^\kappa{}_{\lambda\mu\nu}$ itself, or *via* some "covariant derivatives" of the matter field, which can be symbolically written as:

$$\nabla\phi := \partial\phi + \text{“}\Gamma\cdot\phi\text{''}\,. \tag{103}$$

This means that, in fact, we have $\mathcal{L}_A = \mathcal{L}_A(R^\kappa{}_{\lambda\mu\nu}, \phi, \nabla\phi)$. The Riemann tensor satisfies identities:

$$\begin{aligned} R^\kappa{}_{\lambda\mu\nu} &= -R^\kappa{}_{\lambda\nu\mu}\,(\text{skew-symmetry})\,, \\ R^\kappa{}_{[\lambda\mu\nu]} &= 0\ \ (\text{Bianchi I identity})\,, \end{aligned}$$

and, consequently, splits into three irreducible parts. First, its splits into its trace – the Ricci tensor (33) $R_{\lambda\nu} := R^\kappa{}_{\lambda\kappa\nu}$ – and the remaining traceless part $W^\kappa{}_{\lambda\mu\nu}$. The latter is an analogue of the Weyl tensor in the metric case. It is defined by the following identities:

$$W^\kappa{}_{\kappa\mu\nu} = W^\kappa{}_{\mu\kappa\nu} = W^\kappa{}_{\mu\nu\kappa} = 0\,,$$

but other identities fulfilled by the Weyl tensor cannot even be formulated here, because there is **no** metric tensor to lower the upper index! Furthermore, the Ricci tensor splits into its symmetric – $K_{\mu\nu}$ – and its skew-symmetric – $F_{\mu\nu}$ – parts, according to $R_{\mu\nu} = K_{\mu\nu} + F_{\mu\nu}$. Finally, decomposition of the Riemann tensor into three irreducible parts (i.e.: K, F and W) enables us to rewrite the gravitational component (100) of the generating formula in a way similar to Lemma 5.1 in the purely metric theory (cf. [20] and see Appendix C):

Theorem 6.1. *The following identity holds:*

$$\partial_\nu \left(\mathcal{P}_\kappa{}^{\lambda\mu\nu}\, \delta\Gamma^\kappa{}_{\lambda\mu} \right) = \left(\nabla_\nu \mathcal{P}_\kappa{}^{\lambda\mu\nu} \right) \delta\Gamma^\kappa{}_{\lambda\mu} + \pi^{\mu\nu}\, \delta K_{\mu\nu} + \chi^{\mu\nu}\, \delta F_{\mu\nu} - 2\, \Omega_\kappa{}^{\lambda[\mu\nu]}\, \delta W^\kappa{}_{\lambda\mu\nu} \,, \tag{104}$$

where

$$\begin{aligned} \pi^{\mu\nu} &= -\frac{2}{3} \mathcal{P}_\kappa{}^{\kappa(\mu\nu)} \,, \\ \chi^{\mu\nu} &= -\frac{2}{5} \mathcal{P}_\kappa{}^{\kappa[\mu\nu]} \,, \end{aligned}$$

and Ω *is the remaining, traceless part of* $\mathcal{P}$.

Applying above decomposition to the variational formula (100), we obtain:

$$\begin{aligned} \delta\mathcal{L}_A(K_{\mu\nu}, F_{\mu\nu}, W^\kappa{}_{\lambda\mu\nu}, \phi, \nabla\phi) &= \partial_\nu \left(p^\nu\, \delta\phi \right) + \left(\nabla_\nu \mathcal{P}_\kappa{}^{\lambda\mu\nu} \right) \delta\Gamma^\kappa{}_{\lambda\mu} + \pi^{\mu\nu}\, \delta K_{\mu\nu} + \\ &\quad + \chi^{\mu\nu}\, \delta F_{\mu\nu} - 2\, \Omega_\kappa{}^{\lambda[\mu\nu]}\, \delta W^\kappa{}_{\lambda\mu\nu} \,. \end{aligned} \tag{105}$$

Because Γ enters here through "covariant derivatives" of matter field, the gravitational part of the Euler-Lagrange equation reads:

$$\nabla_\nu \mathcal{P}_\kappa{}^{\lambda\mu\nu} = \frac{\partial \mathcal{L}_A}{\partial \Gamma^\kappa{}_{\lambda\mu}} = \frac{\partial \mathcal{L}_A}{\partial (\nabla_\alpha \phi)} \frac{\partial (\nabla_\alpha \phi)}{\partial \Gamma^\kappa{}_{\lambda\mu}} = \text{"} p \cdot \phi \text{"} \,. \tag{106}$$

Mathematical structure and the physical interpretation of a general theory (105) will be discussed in another paper. Here, we limit ourselves to the conventional General Relativity Theory. This means, as proved in Section 4, that we consider Lagrangians which depend upon curvature *via* the symmetric part of the Ricci tensor, exclusively: $\mathcal{L}_A = \mathcal{L}_A(K_{\mu\nu}, \phi, \nabla\phi)$. Consequently, we have:

$$\chi^{\mu\nu} = \frac{\partial \mathcal{L}_A}{\partial F_{\mu\nu}} = 0 \,, \tag{107}$$

$$-2\, \Omega_\kappa{}^{\lambda[\mu\nu]} = \frac{\partial \mathcal{L}_A}{\partial W^\kappa{}_{\lambda\mu\nu}} = 0 \,, \tag{108}$$

$$\pi^{\mu\nu} = \frac{\partial \mathcal{L}_A}{\partial K_{\mu\nu}} \,, \tag{109}$$

$$\mathcal{P}_\kappa{}^{\lambda\mu\nu} = \pi_\kappa{}^{\lambda\mu\nu} = \pi^{\lambda\mu}\, \delta^\nu_\kappa - \delta^{(\lambda}_\kappa\, \pi^{\mu)\nu} \,. \tag{110}$$

This means, that we obtain formula (88), used already in Section 5:

$$\delta\mathcal{L}_A = \partial_\nu \left(p^\nu \, \delta\phi\right) + \left(\nabla_\nu \pi_\kappa{}^{\lambda\mu\nu}\right) \, \delta\Gamma^\kappa{}_{\lambda\mu} + \pi^{\mu\nu} \, \delta K_{\mu\nu} \, , \tag{111}$$

or equivalently:

$$\delta\mathcal{L}_A = \partial_\nu \left(\pi_\kappa{}^{\lambda\mu\nu} \, \delta\Gamma^\kappa{}_{\lambda\mu} + p^\nu \, \delta\phi\right) \, , \tag{112}$$

cf. formula (75). The gravitational field equations are:

$$\nabla_\nu \pi_\kappa{}^{\lambda\mu\nu} = \frac{\partial\mathcal{L}_A}{\partial\Gamma^\kappa{}_{\lambda\mu}} = \frac{\partial\mathcal{L}_A}{\partial\nabla_\alpha\phi} \, \frac{\partial\nabla_\alpha\phi}{\partial\Gamma^\kappa{}_{\lambda\mu}} = p^\alpha \, \frac{\partial\nabla_\alpha\phi}{\partial\Gamma^\kappa{}_{\lambda\mu}} = \text{"}p\cdot\phi\text{"} \, , \tag{113}$$

$$\pi^{\mu\nu} = \frac{\partial\mathcal{L}_A}{\partial K_{\mu\nu}} \, . \tag{114}$$

The first equation enables us to find the non-metricity tensor N as a difference between our affine connection Γ and the Levi-Civita connection $\overset{\circ}{\Gamma}$ – see formula (52). Then, equation (112) reads:

$$\delta\mathcal{L}_A = \partial_\nu \left(\pi_\kappa{}^{\lambda\mu\nu} \, \delta \, \overset{\circ}{\Gamma}{}^\kappa{}_{\lambda\mu} + \pi_\kappa{}^{\lambda\mu\nu} \, \delta N^\kappa{}_{\lambda\mu} + p^\nu \, \delta\phi\right) \, . \tag{115}$$

The first term can be calculated from identity (40) which was already presented in Section 2:

$$\partial_\nu \left(\pi_\kappa{}^{\lambda\mu\nu} \, \delta \, \overset{\circ}{\Gamma}{}^\kappa{}_{\lambda\mu}\right) = \delta\mathcal{L}_H + \frac{1}{16\pi} \, \overset{\circ}{\mathcal{G}}{}^{\mu\nu} \, \delta g_{\mu\nu} \, . \tag{116}$$

For the second term we use the "inverted" Lemma 4.2 and obtain:

$$\partial_\nu \left(\pi_\kappa{}^{\lambda\mu\nu} \, \delta N^\kappa{}_{\lambda\mu}\right) = \partial_\kappa \left(\mathcal{R}^{\mu\nu\kappa} \, \delta g_{\mu\nu}\right) - \delta \left[\overset{\circ}{\nabla}_\kappa \mathcal{R}_\sigma{}^{\sigma\kappa}\right] \, , \tag{117}$$

where

$$\mathcal{R}^{\mu\nu\kappa} = \frac{\sqrt{|\det g|}}{16\pi} \left[N^{\kappa\mu\nu} - g^{\kappa(\mu} \, N_\sigma{}^{\nu)\sigma} + \frac{1}{2} \left(N^{\kappa\sigma}{}_\sigma - N_\sigma{}^{\sigma\kappa}\right) \, g^{\mu\nu}\right] \, . \tag{118}$$

Analogously, the "inverted" Lemma 4.1 gives us:

$$\partial_\kappa \left(\mathcal{R}^{\mu\nu\kappa} \, \delta g_{\mu\nu}\right) = \mathcal{P}^{\mu\nu}{}_\lambda \, \delta \, \overset{\circ}{\Gamma}{}^\lambda{}_{\mu\nu} + \left(\overset{\circ}{\nabla}_\kappa \mathcal{R}^{\mu\nu\kappa}\right) \, \delta g_{\mu\nu} \, ,$$

where:

$$\mathcal{P}^{\mu\nu}{}_\lambda := \mathcal{R}_\lambda{}^{\mu\nu} + \mathcal{R}_\lambda{}^{\nu\mu} \, . \tag{119}$$

This way, variational formula (115) can be rewritten as follows:

$$\delta\mathcal{L}_A = \partial_\nu\left(p^\nu\,\delta\phi\right) + \delta\mathcal{L}_H + \mathcal{P}^{\mu\nu}{}_\lambda\,\delta\,\overset{\circ}{\Gamma}{}^\lambda{}_{\mu\nu} + \left(\frac{1}{16\pi}\,\overset{\circ}{\mathcal{G}}{}^{\mu\nu} + \overset{\circ}{\nabla}{}_\kappa\mathcal{R}^{\mu\nu\kappa}\right)\delta g_{\mu\nu} + \\ -\delta\left[\overset{\circ}{\nabla}{}_\kappa\mathcal{R}_\sigma{}^{\sigma\kappa}\right]. \qquad (120)$$

Putting the complete variation $\delta\left[\overset{\circ}{\nabla} R\right]$ on the left hand side, we finally obtain the metric Lagrangian (cf. (74)):

$$\mathcal{L}_g := \mathcal{L}_A + \overset{\circ}{\nabla}{}_\kappa\mathcal{R}_\sigma{}^{\sigma\kappa}$$

together with its variation:

$$\delta\mathcal{L}_g = \partial_\nu\left(p^\nu\,\delta\phi\right) + \delta\mathcal{L}_H + \mathcal{P}^{\mu\nu}{}_\lambda\,\delta\,\overset{\circ}{\Gamma}{}^\lambda{}_{\mu\nu} + \left(\frac{1}{16\pi}\,\overset{\circ}{\mathcal{G}}{}^{\mu\nu} + \overset{\circ}{\nabla}{}_\kappa\mathcal{R}^{\mu\nu\kappa}\right)\delta g_{\mu\nu}\,.$$

Define the matter Lagrangian density $\mathcal{L}_{matt} := \mathcal{L}_g - \mathcal{L}_H$. It satisfies equation:

$$\delta\left(\mathcal{L}_g - \mathcal{L}_H\right) = \delta\mathcal{L}_{matt} = \partial_\nu\left(p^\nu\,\delta\phi\right) + \mathcal{P}^{\mu\nu}{}_\lambda\,\delta\,\overset{\circ}{\Gamma}{}^\lambda{}_{\mu\nu} + \left(\frac{1}{16\pi}\,\overset{\circ}{\mathcal{G}}{}^{\mu\nu} + \overset{\circ}{\nabla}{}_\kappa\mathcal{R}^{\mu\nu\kappa}\right)\delta g_{\mu\nu}\,. \qquad (121)$$

This equation implies the gravitational field equations (79) – (80) in the purely metric formulation.

7 Conclusions

The "Palatini method of variation" leads to the standard, metric formulation of General Relativity Theory only for a special class of matter fields: where the matter Lagrangian density $\mathcal{L}_{matt} = \mathcal{L}_{matt}(\phi, \partial\phi, g)$ does not depend upon connection coefficients (i.e. does not depend upon derivatives of the metric tensor). But, for a generic matter field we have

$$\mathcal{L}_{matt} = \mathcal{L}_{matt}(\phi, \overset{\circ}{\nabla}\phi, g) = \mathcal{L}_{matt}(\phi, \partial\phi, g, \partial g)$$

and, whence, the "Palatini method" leads *a priori* to the non-metric connection $\Gamma \neq \overset{\circ}{\Gamma}$. In this paper we propose to accept this non-metricity of the connection and show that the theory, rewritten in terms of this connection – although entirely equivalent to the conventional GR theory – acquires a much simpler mathematical structure.

Physically, this means that similarly as "mass generates curvature" (i.e. departure from flattness), the matter sensitivity towards the connection generates the departure of the connection from metricity. The two spacetime geometric structures, namely: metric and connection, acquire an independent role. The relation between them is not assumed *a priori* but is governed by field equations.

A Proof of Lemma 4.1

Proof. Using explicit formula for the metric connection coefficients, we obtain:

$$\begin{aligned}
\mathcal{P}^{\mu\nu}{}_{\lambda}\,\delta\,\overset{\circ}{\Gamma}{}^{\lambda}{}_{\mu\nu} &= \mathcal{P}^{\mu\nu}{}_{\lambda}\,\delta\left[\frac{1}{2}\,g^{\lambda\kappa}\,\left(g_{\kappa\mu,\nu}+g_{\kappa\nu,\mu}-g_{\mu\nu,\kappa}\right)\right]= \\
&= \mathcal{P}^{\mu\nu}{}_{\lambda}\,\delta\left[g^{\lambda\kappa}\,\left(g_{\kappa\mu,\nu}-\frac{1}{2}\,g_{\mu\nu,\kappa}\right)\right]= \\
&= \mathcal{P}^{\mu\nu}{}_{\lambda}\,\left(g_{\kappa\mu,\nu}-\frac{1}{2}\,g_{\mu\nu,\kappa}\right)\,\delta g^{\lambda\kappa}+\mathcal{P}^{\mu\nu\kappa}\,\delta\left(g_{\kappa\mu,\nu}-\frac{1}{2}\,g_{\mu\nu,\kappa}\right)= \\
&= -\mathcal{P}^{\mu\nu}{}_{\lambda}\,\left(g_{\kappa\mu,\nu}-\frac{1}{2}\,g_{\mu\nu,\kappa}\right)\,g^{\lambda\alpha}\,g^{\kappa\beta}\,\delta g_{\alpha\beta}+\mathcal{P}^{\mu\nu\kappa}\,\delta\left(g_{\kappa\mu,\nu}-\frac{1}{2}\,g_{\mu\nu,\kappa}\right)= \\
&= -\mathcal{P}^{\mu\nu\alpha}\,\overset{\circ}{\Gamma}{}^{\beta}{}_{\mu\nu}\,\delta g_{\alpha\beta}+\partial_{\kappa}\left[\left(\mathcal{P}^{\nu\kappa\mu}-\frac{1}{2}\,\mathcal{P}^{\mu\nu\kappa}\right)\,\delta g_{\mu\nu}\right]+ \\
&\quad-\left[\partial_{\kappa}\left(\mathcal{P}^{\nu\kappa\mu}-\frac{1}{2}\,\mathcal{P}^{\mu\nu\kappa}\right)\right]\,\delta g_{\mu\nu}= \\
&= -\left(\mathcal{R}^{\mu\nu\kappa}{}_{,\kappa}+\mathcal{P}^{\alpha\beta\mu}\,\overset{\circ}{\Gamma}{}^{\nu}{}_{\alpha\beta}\right)\,\delta g_{\mu\nu}+\partial_{\kappa}\left(\mathcal{R}^{\mu\nu\kappa}\,\delta g_{\mu\nu}\right)\,.
\end{aligned}$$

Taking into account that both $\mathcal{R}$ and $\mathcal{P}$ are tensor densities, we have:

$$\begin{aligned}
\overset{\circ}{\nabla}{}_{\kappa}\mathcal{R}^{\mu\nu\kappa} &= \mathcal{R}^{\mu\nu\kappa}{}_{,\kappa}-\overset{\circ}{\Gamma}{}^{\sigma}{}_{\sigma\kappa}\,\mathcal{R}^{\mu\nu\kappa}+\overset{\circ}{\Gamma}{}^{\mu}{}_{\sigma\kappa}\,\mathcal{R}^{\sigma\nu\kappa}+\overset{\circ}{\Gamma}{}^{\nu}{}_{\sigma\kappa}\,\mathcal{R}^{\mu\sigma\kappa}+\overset{\circ}{\Gamma}{}^{\kappa}{}_{\sigma\kappa}\,\mathcal{R}^{\mu\nu\sigma}= \\
&= \mathcal{R}^{\mu\nu\kappa}{}_{,\kappa}+\frac{1}{2}\,\overset{\circ}{\Gamma}{}^{\mu}{}_{\sigma\kappa}\,\mathcal{P}^{\sigma\kappa\nu}+\frac{1}{2}\,\overset{\circ}{\Gamma}{}^{\nu}{}_{\sigma\kappa}\,\mathcal{P}^{\sigma\kappa\mu}= \\
&= \mathcal{R}^{\mu\nu\kappa}{}_{,\kappa}+\mathcal{P}^{\sigma\kappa(\mu}\,\overset{\circ}{\Gamma}{}^{\nu)}{}_{\sigma\kappa}\,,
\end{aligned}$$

which implies the thesis of the Lemma. ∎

B Proof of Lemma 4.2

Proof. Using formula (37) we obtain:

$$
\begin{aligned}
\mathcal{R}^{\mu\nu\kappa}\,\delta g_{\mu\nu} \;&=\; \frac{16\pi}{\sqrt{|\det g|}}\left(\frac{1}{2}\,\mathcal{R}_{\lambda}{}^{\lambda\kappa}\,g_{\alpha\beta}-\mathcal{R}_{\alpha\beta}{}^{\kappa}\right)\delta\pi^{\alpha\beta} = \\
&= 16\pi\left(\frac{1}{2}\,R_{\lambda}{}^{\lambda\kappa}\,g_{\alpha\beta}-R_{\alpha\beta}{}^{\kappa}\right)\delta\pi^{\alpha\beta} = \\
&= \delta\left\{16\pi\left(\frac{1}{2}\,R_{\lambda}{}^{\lambda\kappa}\,g_{\alpha\beta}-R_{\alpha\beta}{}^{\kappa}\right)\pi^{\alpha\beta}\right\} + \\
&\quad -\pi^{\alpha\beta}\delta\left[16\pi\left(\frac{1}{2}\,R_{\lambda}{}^{\lambda\kappa}\,g_{\alpha\beta}-R_{\alpha\beta}{}^{\kappa}\right)\right] = \\
&= \delta\mathcal{R}_{\sigma}{}^{\sigma\kappa}-\pi^{\alpha\beta}\,\delta\left[16\pi\left(\frac{1}{2}\,R_{\sigma}{}^{\sigma\kappa}\,g_{\alpha\beta}-R_{\alpha\beta}{}^{\kappa}\right)\right] ,
\end{aligned}
$$

and, consequently:

$$
\partial_{\kappa}\left(\mathcal{R}^{\mu\nu\kappa}\delta g_{\mu\nu}\right)=\delta\left[\mathcal{R}_{\sigma}{}^{\sigma\kappa}{}_{,\kappa}\right]+\partial_{\kappa}\left\{\pi^{\alpha\beta}\,\delta\left[16\pi\left(R_{\alpha\beta}{}^{\kappa}-\frac{1}{2}\,R_{\sigma}{}^{\sigma\kappa}\,g_{\alpha\beta}\right)\right]\right\} .
$$

Since $\mathcal{R}^{\mu\nu\kappa}\,g_{\mu\nu}$ is a vector density, we have:

$$
\mathcal{R}_{\sigma}{}^{\sigma\kappa}{}_{,\kappa}=\overset{\circ}{\nabla}_{\kappa}\mathcal{R}_{\sigma}{}^{\sigma\kappa} ,
$$

and, whence:

$$
\partial_{\kappa}\left(\mathcal{R}^{\mu\nu\kappa}\delta g_{\mu\nu}\right)=\delta\left[\overset{\circ}{\nabla}_{\kappa}\mathcal{R}_{\sigma}{}^{\sigma\kappa}\right]+\partial_{\kappa}\left\{\pi^{\alpha\beta}\,\delta\left[16\pi\left(R_{\alpha\beta}{}^{\kappa}-\frac{1}{2}\,R_{\sigma}{}^{\sigma\kappa}\,g_{\alpha\beta}\right)\right]\right\} .
$$

Now, it remains to check that the quantity N defined by (71) satisfies identity

$$
\pi_{\kappa}{}^{\lambda\mu\nu}\,\delta N^{\kappa}{}_{\lambda\mu}:=\pi^{\alpha\beta}\,\delta\left[16\pi\left(R_{\alpha\beta}{}^{\nu}-\frac{1}{2}\,R_{\sigma}{}^{\sigma\nu}\,g_{\alpha\beta}\right)\right] .
$$

□

C Different representations of the curvature and its decomposition into irreducible components

The Riemann tensor of a symmetric connection, defined by (32), splits into the three irreducible conmponents: the symmetric ($K_{\mu\nu}$) and the antisymmetric ($F_{\mu\nu}$) parts of the Ricci tensor $R_{\mu\nu} := R^{\kappa}{}_{\mu\kappa\nu} = K_{\mu\nu} + F_{\mu\nu}$, and the remaining traceless part $W^{\kappa}{}_{\lambda\mu\nu}$, according to the following decomposition formula:

$$R^{\kappa}{}_{\lambda\mu\nu} = \frac{1}{3}\left(\delta^{\kappa}_{\mu}\, K_{\lambda\nu} - \delta^{\kappa}_{\nu}\, K_{\lambda\mu}\right) + \frac{1}{5}\left(2\,\delta^{\kappa}_{\lambda}\, F_{\mu\nu} + \delta^{\kappa}_{\mu}\, F_{\lambda\nu} - \delta^{\kappa}_{\nu}\, F_{\lambda\mu}\right) + W^{\kappa}{}_{\lambda\mu\nu}\,. \tag{122}$$

The Riemann tensor satisfies the following two algebraic identities:

$$R^{\kappa}{}_{\lambda\mu\nu} = -R^{\kappa}{}_{\lambda\nu\mu}\,, \quad R^{\kappa}{}_{[\lambda\mu\nu]} = 0\,, \tag{123}$$

where the square bracket denotes antisymmetrization.

As will be seen in the next Appendix (see also [20], [19]), for many calculational purposes it is useful to represent the same geometric object (the curvature) by a different, but totally equivalent, *curvature tensor*, defined in terms of the first jet of the connection by the following formula:

$$K^{\kappa}{}_{\lambda\mu\nu} := \Gamma^{\kappa}{}_{\lambda\mu\nu} - \Gamma^{\kappa}{}_{(\lambda\mu\nu)} + \Gamma^{\sigma}{}_{\lambda\mu}\,\Gamma^{\kappa}{}_{\nu\sigma} - \Gamma^{\sigma}{}_{(\lambda\mu}\,\Gamma^{\kappa}{}_{\nu)\sigma}\,. \tag{124}$$

The equivalence between the two objects is assured by the following identities:

$$K^{\kappa}{}_{\lambda\mu\nu} = -\frac{2}{3}\, R^{\kappa}{}_{(\lambda\mu)\nu}\,, \quad R^{\kappa}{}_{\lambda\mu\nu} = -2\, K^{\kappa}{}_{\lambda[\mu\nu]}\,, \tag{125}$$

where the round bracket denotes symmetrization. The curvature tensor K satisfies the following two algebraic identities:

$$K^{\kappa}{}_{\lambda\mu\nu} = K^{\kappa}{}_{\mu\lambda\nu}\,, \qquad K^{\kappa}{}_{(\lambda\mu\nu)} = 0\,, \tag{126}$$

analogous to (123), and can be decomposed into its irreducible components in a way analogous to (122), namely:

$$K^{\kappa}{}_{\lambda\mu\nu} = -\frac{1}{9}\left(\delta^{\kappa}_{\lambda}\, K_{\mu\nu} + \delta^{\kappa}_{\mu}\, K_{\lambda\nu} - 2\delta^{\kappa}_{\nu}\, K_{\lambda\mu}\right) - \frac{1}{5}\left(\delta^{\kappa}_{\lambda}\, F_{\mu\nu} + \delta^{\kappa}_{\mu}\, F_{\lambda\nu}\right) + U^{\kappa}{}_{\lambda\mu\nu}\,, \tag{127}$$

where

$$U^{\kappa}{}_{\lambda\mu\nu} := -\frac{2}{3}\, W^{\kappa}{}_{(\lambda\mu)\nu}\,.$$

D Proof of Theorem 6.1

Proof. Because curvature tensor is equivalent with Riemann, we can assume that the invariant Lagrangian density $\mathcal{L}_A(\Gamma,\, \partial\Gamma,\, \phi, \partial\phi)$ depends upon derivatives of the connection coefficients *via* the tensor $K^{\kappa}{}_{\lambda\mu\nu}$, i.e. *via* the expression "$\Gamma^{\kappa}{}_{\lambda\mu\nu} - \Gamma^{\kappa}{}_{(\lambda\mu\nu)}$" which is: 1) symmetric in indices (λ,μ) and 2) its totally symmetric part $\Gamma^\kappa{}_{(\lambda\mu\nu)} - \Gamma^\kappa{}_{(\lambda\mu\nu)} = 0$ vanishes identically. Being equal to the derivative of $\mathcal{L}_A$ with respect to this quantity, the momentum $\mathcal{P}_\kappa{}^{\lambda\mu\nu}$ must fulfill *a priori* identities similar to (126), namely:

$$\mathcal{P}_\kappa{}^{\lambda\mu\nu} = \mathcal{P}_\kappa{}^{\mu\lambda\nu}\,, \quad \mathcal{P}_\kappa{}^{(\lambda\mu\nu)} = 0\,. \tag{128}$$

We have, therefore:

$$\begin{aligned}\partial_\nu\left(\mathcal{P}_\kappa{}^{\lambda\mu\nu}\,\delta\Gamma^\kappa{}_{\lambda\mu}\right) &= \mathcal{P}_\kappa{}^{\lambda\mu\nu}{}_{,\nu}\,\delta\Gamma^\kappa{}_{\lambda\mu} + \mathcal{P}_\kappa{}^{\lambda\mu\nu}\,\delta\Gamma^\kappa{}_{\lambda\mu,\nu} = \\ &= \mathcal{P}_\kappa{}^{\lambda\mu\nu}{}_{,\nu}\,\delta\Gamma^\kappa{}_{\lambda\mu} + \mathcal{P}_\kappa{}^{\lambda\mu\nu}\,\delta\left(\Gamma^\kappa{}_{\lambda\mu,\nu} - \Gamma^\kappa{}_{(\lambda\mu,\nu)}\right)\,.\end{aligned}$$

because the last term vanishes. Using definition (124) of the curvature tensor, we get:

$$\begin{aligned}\partial_\nu\left(\mathcal{P}_\kappa{}^{\lambda\mu\nu}\,\delta\Gamma^\kappa{}_{\lambda\mu}\right) &= \mathcal{P}_\kappa{}^{\lambda\mu\nu}\,\delta K^\kappa{}_{\lambda\mu,\nu} + \mathcal{P}_\kappa{}^{\lambda\mu\nu}{}_{,\nu}\,\delta\Gamma^\kappa{}_{\lambda\mu} - \mathcal{P}_\kappa{}^{\lambda\mu\nu}\,\delta\left(\Gamma^\sigma{}_{\lambda\mu}\,\Gamma^\kappa{}_{\nu\sigma} - \Gamma^\sigma{}_{(\lambda\mu}\,\Gamma^\kappa{}_{\nu)\sigma}\right) = \\ &= \mathcal{P}_\kappa{}^{\lambda\mu\nu}\,\delta K^\kappa{}_{\lambda\mu,\nu} + \mathcal{P}_\kappa{}^{\lambda\mu\nu}{}_{,\nu}\,\delta\Gamma^\kappa{}_{\lambda\mu} - \mathcal{P}_\kappa{}^{\lambda\mu\nu}\,\delta\left(\Gamma^\sigma{}_{\lambda\mu}\,\Gamma^\kappa{}_{\nu\sigma}\right)\,.\end{aligned}$$

But, we have:

Lemma D.1. *For a tensor density $\mathcal{P}_\kappa{}^{\lambda\mu\nu}$, the following identity holds:*

$$\mathcal{P}_\kappa{}^{\lambda\mu\nu}{}_{,\nu}\,\delta\Gamma^\kappa{}_{\lambda\mu} - \mathcal{P}_\kappa{}^{\lambda\mu\nu}\,\delta\left(\Gamma^\sigma{}_{\lambda\mu}\,\Gamma^\kappa{}_{\nu\sigma}\right) = \nabla_\nu\,\mathcal{P}_\kappa{}^{\lambda\mu\nu}\,\delta\Gamma^\kappa{}_{\lambda\mu}\,.$$

Proof. It is entirely computation:

$$\begin{aligned}\mathcal{P}_\kappa{}^{\lambda\mu\nu}{}_{,\nu}\,\delta\Gamma^\kappa{}_{\lambda\mu} - \mathcal{P}_\kappa{}^{\lambda\mu\nu}\,\delta\left(\Gamma^\sigma{}_{\lambda\mu}\,\Gamma^\kappa{}_{\nu\sigma}\right) &= \mathcal{P}_\kappa{}^{\lambda\mu\nu}{}_{,\nu}\,\delta\Gamma^\kappa{}_{\lambda\mu} - \mathcal{P}_\kappa{}^{\lambda\mu\nu}\,\Gamma^\sigma{}_{\lambda\mu}\,\delta\Gamma^\kappa{}_{\nu\sigma} - \mathcal{P}_\kappa{}^{\lambda\mu\nu}\,\Gamma^\kappa{}_{\nu\sigma}\,\delta\Gamma^\sigma{}_{\lambda\mu} = \\ &= \left(\mathcal{P}_\kappa{}^{\lambda\mu\nu}{}_{,\nu} - \mathcal{P}_\sigma{}^{\lambda\mu\nu}\,\Gamma^\sigma{}_{\kappa\nu} - \mathcal{P}_\kappa{}^{\sigma\nu\lambda}\,\Gamma^\mu{}_{\sigma\nu}\right)\,\delta\Gamma^\kappa{}_{\lambda\mu}\,.\end{aligned}$$

Using again (128), the last term can be rewritten as follows:

$$\begin{aligned}-\mathcal{P}_\kappa{}^{\sigma\nu\lambda}\,\Gamma^\mu{}_{\sigma\nu}\,\delta\Gamma^\kappa{}_{\lambda\mu} &= \mathcal{P}_\kappa{}^{\sigma\nu\lambda}\,\Gamma^\mu{}_{\nu\lambda}\,\delta\Gamma^\kappa{}_{\sigma\mu} + \mathcal{P}_\kappa{}^{\sigma\nu\lambda}\,\Gamma^\mu{}_{\lambda\sigma}\,\delta\Gamma^\kappa{}_{\nu\mu} = \\ &= \mathcal{P}_\kappa{}^{\lambda\nu\sigma}\,\Gamma^\mu{}_{\nu\sigma}\,\delta\Gamma^\kappa{}_{\lambda\mu} + \mathcal{P}_\kappa{}^{\sigma\lambda\nu}\,\Gamma^\mu{}_{\nu\sigma}\,\delta\Gamma^\kappa{}_{\lambda\mu} = \\ &= \mathcal{P}_\kappa{}^{\lambda\sigma\nu}\,\Gamma^\mu{}_{\sigma\nu}\,\delta\Gamma^\kappa{}_{\lambda\mu} + \mathcal{P}_\kappa{}^{\nu\lambda\sigma}\,\Gamma^\mu{}_{\sigma\nu}\,\delta\Gamma^\kappa{}_{\lambda\mu}\,,\end{aligned}$$

and, whence:

$$\mathcal{P}_\kappa{}^{\lambda\mu\nu}{}_{,\nu}\,\delta\Gamma^\kappa{}_{\lambda\mu} - \mathcal{P}_\kappa{}^{\lambda\mu\nu}\,\delta\left(\Gamma^\sigma{}_{\lambda\mu}\,\Gamma^\kappa{}_{\nu\sigma}\right) = \left(\mathcal{P}_\kappa{}^{\lambda\mu\nu}{}_{,\nu} - \mathcal{P}_\sigma{}^{\lambda\mu\nu}\,\Gamma^\sigma{}_{\kappa\nu} + \mathcal{P}_\kappa{}^{\lambda\sigma\nu}\,\Gamma^\mu{}_{\sigma\nu} + \mathcal{P}_\kappa{}^{\nu\lambda\sigma}\,\Gamma^\mu{}_{\sigma\nu}\right)\,\delta\Gamma^\kappa{}_{\lambda\mu}\,.$$

On the other hand, we have:

$$\left(\nabla_\nu\,\mathcal{P}_\kappa{}^{\lambda\mu\nu}\right)\,\delta\Gamma^\kappa{}_{\lambda\mu} = \left(\mathcal{P}_\kappa{}^{\lambda\mu\nu}{}_{,\nu} - \mathcal{P}_\sigma{}^{\lambda\mu\nu}\,\Gamma^\sigma{}_{\kappa\nu} + \mathcal{P}_\kappa{}^{\sigma\mu\nu}\,\Gamma^\mu{}_{\sigma\nu} + \mathcal{P}_\kappa{}^{\lambda\sigma\nu}\,\Gamma^\mu{}_{\sigma\nu}\right)\,\delta\Gamma^\kappa{}_{\lambda\mu}\,,$$

which ends the proof of the Lemma. □

Inserting this formula:

$$\partial_\nu \left(\mathcal{P}_\kappa{}^{\lambda\mu\nu}\, \delta\Gamma^\kappa{}_{\lambda\mu} \right) = \mathcal{P}_\kappa{}^{\lambda\mu\nu}\, \delta K^\kappa{}_{\lambda\mu\nu} + \nabla_\nu\, \mathcal{P}_\kappa{}^{\lambda\mu\nu}\, \delta\Gamma^\kappa{}_{\lambda\mu}\,, \tag{129}$$

into (127) we obtain:

$$\partial_\nu \left(\mathcal{P}_\kappa{}^{\lambda\mu\nu}\, \delta\Gamma^\kappa{}_{\lambda\mu} \right) = \pi^{\mu\nu}\, \delta K_{\mu\nu} + \Omega_\kappa{}^{\lambda\mu\nu}\, \delta U^\kappa{}_{\lambda\mu\nu} + \chi^{\mu\nu}\, \delta F_{\mu\nu} + \left(\nabla_\nu\, \mathcal{P}_\kappa{}^{\lambda\mu\nu} \right)\, \delta\Gamma^\kappa{}_{\lambda\mu}\,,$$

where we have defined:

$$\pi^{\mu\nu} = -\frac{2}{3}\mathcal{P}_\kappa{}^{\kappa(\mu\nu)}\,, \tag{130}$$

$$\chi^{\mu\nu} = -\frac{2}{5}\mathcal{P}_\kappa{}^{\kappa[\mu\nu]}\,, \tag{131}$$

whereas Ω represents the remaining, traceless part of $\mathcal{P}$. This finally gives rise to the following decomposition of the momentum $\mathcal{P}$, analogous to (123) and (126):

$$\mathcal{P}_\kappa{}^{\lambda\mu\nu} = -\frac{1}{2}\left(\delta^\lambda_\kappa\, \pi^{\mu\nu} + \delta^\mu_\kappa\, \pi^{\lambda\nu} - 2\delta^\nu_\kappa\, \pi^{\lambda\mu} \right) - \frac{1}{2}\left(\delta^\lambda_\kappa\, \chi^{\mu\nu} + \delta^\mu_\kappa\, \chi^{\lambda\nu} \right) + \Omega_\kappa{}^{\lambda\mu\nu}\,. \tag{132}$$

The techniques used above prove that the use of the curvature tensor $K^\kappa{}_{\lambda\mu\nu}$, instead of the Riemann tensor $R^\kappa{}_{\lambda\mu\nu}$ simplifies considerably description of the canonical structure of the theory. Indeed, the canonical momentum (101) is directly equal to the derivative of the Lagrangian density $\mathcal{L}$ with respect to the curvature:

$$\mathcal{P}_\kappa{}^{\lambda\mu\nu} := \frac{\partial \mathcal{L}_A}{\partial \Gamma^\kappa{}_{\lambda\mu,\nu}} = \frac{\partial \mathcal{L}_A}{\partial K^\kappa{}_{\lambda\mu\nu}}\,, \tag{133}$$

whereas derivative with respect to the Riemann tensor would need further symmetrization, which obscures considerably this, relatively simple and transparent, structure. □

E Relation between the metric and non-metric Ricci tensors

Using definition (81) of the Ricci tensor $R_{\mu\nu}$ of the connection $\Gamma^\kappa{}_{\mu\nu} = \overset{\circ}{\Gamma}{}^\kappa{}_{\mu\nu} + N^\kappa{}_{\mu\nu}$, we can easily decompose the complete Ricci into the metric Ricci $\overset{\circ}{R}$ and the remaining part, which depends upon non-metricity and its derivatives:

$$\begin{aligned} R_{\mu\nu} &= -\Gamma^{\kappa}{}_{\kappa\mu,\nu} + \Gamma^{\kappa}{}_{\mu\nu,\kappa} - \Gamma^{\sigma}{}_{\kappa\mu}\,\Gamma^{\kappa}{}_{\nu\sigma} + \Gamma^{\sigma}{}_{\mu\nu}\,\Gamma^{\kappa}{}_{\kappa\sigma} = \\ &= \overset{\circ}{R}_{\mu\nu} - \underline{N^{\kappa}{}_{\kappa\mu,\nu}} + \underbrace{N^{\kappa}{}_{\mu\nu,\kappa}} - \underbrace{\overset{\circ}{\Gamma}{}^{\sigma}{}_{\kappa\mu}\,N^{\kappa}{}_{\nu\sigma}} + \underline{\overset{\circ}{\Gamma}{}^{\sigma}{}_{\mu\nu}\,N^{\kappa}{}_{\kappa\sigma}} - \underbrace{N^{\sigma}{}_{\kappa\mu}\,\overset{\circ}{\Gamma}{}^{\kappa}{}_{\nu\sigma}} + \underbrace{N^{\sigma}{}_{\mu\nu}\,\overset{\circ}{\Gamma}{}^{\kappa}{}_{\kappa\sigma}} + \\ &\quad - N^{\sigma}{}_{\kappa\mu}\,N^{\kappa}{}_{\nu\sigma} + N^{\sigma}{}_{\mu\nu}\,N^{\kappa}{}_{\kappa\sigma}\,. \end{aligned}$$

The terms, which have been marked, gather to the covariant derivatives of the non-metricity N:

$$R_{\mu\nu} = \overset{\circ}{R}_{\mu\nu} - \underline{\overset{\circ}{\nabla}_{\nu} N^{\kappa}{}_{\kappa\mu}} + \underbrace{\overset{\circ}{\nabla}_{\kappa} N^{\kappa}{}_{\mu\nu}} - N^{\sigma}{}_{\kappa\mu}\,N^{\kappa}{}_{\nu\sigma} + N^{\sigma}{}_{\mu\nu}\,N^{\kappa}{}_{\kappa\sigma}\,. \tag{134}$$

F Non-metricity tensor N

To find the non-metricity tensor N we have to solve equation (53):

$$\mathcal{P}^{\mu\nu}{}_{\lambda} = \pi^{\mu\alpha} N^{\nu}{}_{\lambda\alpha} + \pi^{\nu\alpha} N^{\mu}{}_{\lambda\alpha} - \pi^{\mu\nu} N^{\alpha}{}_{\lambda\alpha} - \frac{1}{2}\left(\delta^{\mu}_{\lambda}\,N^{\nu}{}_{\alpha\beta} + \delta^{\nu}_{\lambda}\,N^{\mu}{}_{\alpha\beta}\right)\pi^{\alpha\beta}\,.$$

First, we calculate the trace, putting $\nu = \lambda$:

$$\mathcal{P}^{\mu\lambda}{}_{\lambda} = -\frac{3}{2}\,\frac{\sqrt{|\det g|}}{16\pi}\,N^{\mu\alpha}{}_{\alpha}\,,$$

and contracting the equation (53) with metric tensor $g_{\mu\nu}$:

$$\mathcal{P}^{\sigma}{}_{\sigma\lambda} = \frac{\sqrt{|\det g|}}{16\pi}\left[-2N^{\alpha}{}_{\lambda\alpha} - N_{\lambda\alpha}{}^{\alpha}\right]\,.$$

Then, a simple algebra leads to the final result:

$$\begin{aligned} N^{\kappa}{}_{\lambda\mu} &= \frac{8\pi}{\sqrt{|\det g|}}\,g^{\kappa\sigma}\left[\mathcal{P}_{\sigma\lambda\mu} + \mathcal{P}_{\sigma\mu\lambda} - \mathcal{P}_{\lambda\mu\sigma} + g_{\lambda\mu}\left(\mathcal{P}_{\sigma\nu}{}^{\nu} - \frac{1}{2}\,\mathcal{P}^{\nu}{}_{\nu\sigma}\right) + \right. \\ &\quad \left. + \mathcal{P}^{\nu}{}_{\nu(\lambda}\,g_{\mu)\sigma} - \frac{2}{3}\,g_{\sigma(\lambda}\,\mathcal{P}_{\mu)\nu}{}^{\nu}\right]\,. \end{aligned}$$

G Transition from metric to Palatini picture – example

Take the following matter Lagrangian density for the vector field ϕ^{μ}:

$$\mathcal{L}_{matt}(g, \overset{\circ}{\nabla}\phi) = \frac{\sqrt{|\det g|}}{16\pi}\left(\overset{\circ}{\nabla}_{\alpha}\phi^{\beta}\right)\left(\overset{\circ}{\nabla}{}^{\alpha}\phi_{\beta}\right)\,.$$

To find the corresponding Palatini Lagrangian density (94), we have to calculate the non-metricity tensor N as a function of $g, \phi, \nabla\phi$ from equation (53). The idea of solving is quite easy (see Appendix F), but, unfortunately, calculations are difficult and cumbersome. Using the symbolic calculation package[4], we obtain the final result. In terms of the following auxiliary quantities:

$$
\begin{aligned}
\phi^2 &= \phi^\alpha \phi_\alpha \,, \\
D_1 &= 3 + 2\,\phi^2 \,, \\
D_2 &= -\phi^4 - \phi^2 + 3 \,, \\
B_\alpha{}^\beta &= \nabla_\alpha \phi^\beta \,, \\
E_\mu &= B_\mu{}^\alpha \, \phi_\alpha \,, \\
F^\mu &= \phi^\alpha \, B_\alpha{}^\mu \,, \\
A &= \phi^\alpha \, B_\alpha{}^\beta \, \phi_\beta = E_\mu \, \phi^\mu = \phi_\mu \, F^\mu \,, \\
B &= B_\alpha{}^\alpha \,,
\end{aligned}
$$

the Palatini Lagrangian is equal:

$$
\begin{aligned}
\mathcal{L}_P(g,\, \nabla\phi) \;=\; & \frac{\sqrt{|\det g|}}{32\pi\, D_1\, D_2\, (\phi^2 - 1)} \Big[D_1\, D_2\, \left(\phi^2 - 2\right)\, B_{\kappa\lambda}\, B^{\kappa\lambda} + D_1\, D_2\, \phi^2\, B_{\kappa\lambda}\, B^{\lambda\kappa} + \\
& + 4A^2 \left(\phi^2 - 1\right) \phi^2 + 4D_2\, \left(\phi^2 - 1\right)\, A\, B - \phi^2\, D_2\, \left(\phi^2 - 1\right)\, B^2 + \\
& - D_1\, \phi^4\, F_\mu\, F^\mu - 2D_1\, \left(\phi^2 - 2\right)\, \phi^2\, E_\mu\, F^\mu - D_1\, \left(\phi^2 - 2\right)^2\, E_\mu\, E^\mu \Big] \,.
\end{aligned}
$$

H Transition from Palatini to metric picture – example

Take as the Palatini Lagrangian for the vector field X^μ the quadratic expression:

$$
\mathcal{L}_P(g,\, \nabla X) = \frac{\sqrt{|\det g|}}{16\pi}\, \left(\nabla_\alpha X^\mu\right)\, \left(\nabla_\beta X^\nu\right)\, g^{\alpha\beta}\, g_{\mu\nu} \,.
$$

In a way analogous to the one used in Appendix G, using the symbolic computations provided by Mathematica, we obtain the final result in terms of the auxiliary quantities:

[4] We used the suite of free packages xAct version 1.2.0 (especially xTensor package) prepared by José M. Martín-García (see www.xAct.es) for Mathematica version 12.2 environment.

$$
\begin{aligned}
X^2 &= X^\alpha X_\alpha\,,\\
\widetilde{D}_1 &= 2X^2-3\,,\\
\widetilde{D}_2 &= X^4-X^2-3\,,\\
\widetilde{B}_\mu{}^\alpha &= \overset{\circ}{\nabla}_\mu X^\alpha\,,\\
\widetilde{E}_\mu &= \widetilde{B}_\mu{}^\alpha X_\alpha\,,\\
\widetilde{F}^\alpha &= X^\mu\,\widetilde{B}_\mu{}^\alpha\,,\\
\widetilde{A} &= X_\alpha\,\widetilde{F}^\alpha = \widetilde{E}_\alpha X^\alpha\,,\\
\widetilde{B} &= \widetilde{B}_\alpha{}^\alpha\,.
\end{aligned}
$$

The corresponding metric matter Lagrangian defined by equation (97) has the following form:

$$
\begin{aligned}
\mathcal{L}_{matt}(g,\overset{\circ}{\nabla}X) \;=\; & \frac{\sqrt{|\det g|}}{32\pi\,\widetilde{D}_1\,\widetilde{D}_2\,(X^2+1)}\Big[\widetilde{D}_1\,\widetilde{D}_2\,\left(X^2+2\right)\,\widetilde{B}_{\kappa\lambda}\,\widetilde{B}^{\kappa\lambda}+\widetilde{D}_1\,\widetilde{D}_2\,X^2\,\widetilde{B}_{\lambda\kappa}\,\widetilde{B}^{\kappa\lambda}+\\
& +4\left(X^2+1\right)\,X^2\,\widetilde{A}^2+4\widetilde{D}_2\,\left(X^2+1\right)\,\widetilde{A}\,\widetilde{B}+\\
& -X^2\,\widetilde{D}_2\,\left(X^2+1\right)\,\widetilde{B}^2-\widetilde{D}_1\,\left(X^2-2\right)^2\,\widetilde{E}_\mu\,\widetilde{E}^\mu+\\
& -2\widetilde{D}_1\,\left(X^2+2\right)\,X^2\,\widetilde{E}_\mu\,\widetilde{F}^\mu-\widetilde{D}_1\,X^4\,\widetilde{F}_\mu\,\widetilde{F}^\mu\Big]\,.
\end{aligned}
$$

References

[1] D. Hilbert, *Die Grundlagen der Physik*, Königliche Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse, S. 395–407, 1915;

[2] J. Kijowski and G. Magli, *Unconstrained variational principle for relativistic elasticity theory,*Rep. Math. Phys. 39 p. 99 - 112, 1997;

[3] J. Kijowski and G. Magli, *Unconstrained hamiltonian formulation of General Relativity with thermo-elastic sources*, Class. Quantum Grav. 15 p. 3891 - 3916, 1998;

[4] A. Palatini, *Deduzione invariantiva delle equazioni gravitazionali dal principio di Hamilton*, Rendiconti del Circolo Matematico di Palermo, 43, 203-212, 1919;

[5] F.J. Belinfante, *On the spin angular momentum of mesons*, Physica 6 (**1939**) p.887,

[6] F.J. Belinfante, *On the current and the density of the electric charge, the energy, the linear momentum and the angular momentum of arbitrary fields*, Physica 7, (**1940**) p.449

[7] Rosenfeld L.: *A theorem on stress–energy tensors*, Acad. Roy. Belg. 18 (**1940**) p.l

[8] W. Clifford, *On the Space-Theory of Matter*, Proceedings of the Cambridge philosophical society 2, 1876;

[9] A. Einstein *Autobiographical Notes* in: P. A. Schild (ed.) *Albert Einstein: Philosopher-Scientist*, Northwestern University Press, Evanstone Illinois (1949)

[10] J. Kijowski, *On a new variational principle in general relativity and the energy of gravitational field*, Gen. Relat. Grav. Journal **9** (1978), p. 857 – 881.

[11] J. Kijowski, W.M. Tulczyjew, *A symplectic framework for field theories*, Springer Lecture Notes in Physics, vol. 107, 1979;

[12] J. Kijowski, R. Werpachowski, *Universality of affine formulation in General Relativity*, Rep. Math. Phys. **59** (2007) pp. 1 – 31

[13] B.. Bąk, *Variational formulations of General Relativity*, Acta Phys. Pol. B Proc. Suppl. 15, 2022;

[14] C.W. Misner, K.S. Thorne, J.A. Wheeler, *Gravitation*, W.H. Freeman and Company, San Francisco, 1973;

[15] J. Kijowski, *A simple Derivation of Canonical Structure and quasi-local Hamiltonians in General Relativity*, Gen. Relat. Grav. Journal 29, p. 307-343, 1997;

[16] P. T. Chruściel, J. Jezierski and J. Kijowski, *Hamiltonian Field Theory in the Radiating Regime*, monograph (174 pages), volume **70** of the series: Springer Lecture Notes in Physics, Monographs (2001).

[17] J. Kijowski, G. Moreno, *Symplectic structures related with higher order variational problems*, International Journal of Geometric Methods in Modern Physics, **12**, 2015;

[18] V.A. Fock, *The theory of Space, Time and Gravitation*, Pergamon, 1964;

[19] J. Kijowski, K. Senger, *Covariant jets of a connection and higher order curvature tensors*, Journal of Geometry and Physics **163** (2021) 104092

[20] J. Kijowski, *Universality of the Einstein theory of gravitation*, Int. J. of Geometric Methods in Modern Physics, **13** (2016), p. 1640008