Available online at www.sciencedirect.com
ScienceDirect
Journal of Geometry and Physics 57 (2007) 1089–1097
JOURNAL OF GEOMETRY AND PHYSICS
www.elsevier.com/locate/jgp

Higher-order Utiyama–Yang–Mills Lagrangians

Fernando Etayo Gordejuela[a], Pedro L. García Pérez[b], Jaime Muñoz Masqué[c], Javier Pérez Alvarez[d,*]

[a] *Departamento de Matemáticas, Estadística y Computación, Universidad de Cantabria, Spain*
[b] *Departamento de Matemáticas, Universidad de Salamanca, Plaza de La Merced, 1-4, E-37008 Salamanca, Spain*
[c] *Instituto de Física Aplicada, C/ Serrano 144, 28006 Madrid, Spain*
[d] *UNED, Mathematicas Fundamentales, Senda del Rey, 9, 28040 Madrid, Spain*

Received 30 June 2006; received in revised form 20 August 2006; accepted 1 September 2006
Available online 11 October 2006

Abstract

In this article we classify higher-order gauge invariant Lagrangian densities on the bundle of connections of a principal G-bundle $\pi : P \to M$, in the case where the structure group is abelian. Also we show the strong obstruction for an analogous classification in the noncommutative case.

MSC: 53C05; 49S05

Keywords: Utiyama–Yang–Mills Lagrangians; Gauge theory

1. Introduction

Yang and Mills [15] interpreted the exchange of mesons in the strong nuclear force supposing that the Lagrangian for the nucleon will have $SU(2)$ as a local symmetry group. This Lagrangian $\mathcal{L}$ is constructed using covariant derivatives associated with a connection for the given structure group $SU(2)$ (whose components as a function of a basis of the Lie algebra are the Yang–Mills fields and mediate the force between nucleons), to construct the field strength, which is none other than the curvature of the vector bundle. Later on, Utiyama [14] generalized this work proving that any given gauge invariant action density defined on the 1-jet space of connections of a principal fiber bundle must be an ad-invariant function of the curvature. While Utiyama's arguments are local in nature and rely on a trivial bundle, global formulations and geometric interpretations for nontrivial bundles have subsequently appeared in the literature. See: Garcia [9], Bleecker [3] and Eck [6]; Betounes [1] and Grassini [10] for the case of interaction of a gauge and a particle field; Betounes [2] for Yang–Mills Lagrangians coupled with matter fields; Grassini [11] for an interpretation of the Utiyama theory as a geometrical gauge theory and Bruzzo [4] for the invariant Lagrangians associated with global actions of supergauge transformations.

* Corresponding author. Tel.: +34 913987227; fax: +34 913987107.
E-mail addresses: pgarcia@usal.es (P.L. García Pérez), j1235a@hotmail.com (J. Pérez Alvarez).

0393-0440/$ - see front matter
doi:10.1016/j.geomphys.2006.09.004

The geometric version of the Utiyama theorem states that given a principal G-bundle $\pi : P \to M$ with $p : \mathcal{C} \to M$ as bundle of connections, a Lagrangian defined on the 1-jet level $\mathcal{L} : J^1(\mathcal{C}) \to \mathbb{R}$ is invariant under the natural representation of the gauge algebra of P, if and only if $\mathcal{L}$ factors through the curvature mapping Ω (which sends each connection w onto its curvature $\Omega_w = \mathrm{d}w + \frac{1}{2}[w, w]$) as follows: $\mathcal{L} = \overline{\mathcal{L}} \circ \Omega$, where $\overline{\mathcal{L}}$ is a function defined on the curvature bundle $\mathcal{K} \to M$, which in turn must be invariant under the gauge algebra representation on $\mathcal{K}$.

Lagrangian densities are a very special kind of differential forms on $\mathcal{C}$. They are horizontal and of degree equal to the dimension of M. Of course, their importance lies in the fact that such forms define variational problems on the fibred manifold $p : \mathcal{C} \to M$, and gauge invariance for Lagrangian density provides a natural symmetry condition for the variational problem under consideration. Moreover, probably the simplest method of obtaining Euler–Lagrangian equations which are invariant under gauge transformations is to start with a gauge invariant Lagrangian. In any case, we know that gauge invariant Lagrangians produce gauge invariant field equations. In this way due the importance of this notion in gauge theories we consider here the problem of gauge invariance for higher-order Lagrangians. So far, this classification problem has only been solved in the case of the electromagnetic field theory [13], that is, in the mathematical context, the case in which we consider the trivial $U(1)$-bundle over a spacetime M. Herein we solve this problem for any principal bundle whose structure group is an abelian Lie group, and we study the type of obstruction that prevents an analogous classification for any nonabelian Lie group.

As an open problem, we propose to study conditions for obtaining arbitrary order gauge invariant Lagrangians for particles interacting with external gauge fields.

2. Gauge algebra, bundle of connections and automorphic forms

Given a principal G-bundle $\pi : P \to M$, let us denote by $\mathrm{Gau}(P)$ the group of the G-invariant diffeomorphisms of P which stabilize every G-orbit, i.e. φ is a diffeomorphism $\varphi : P \to P$ such that $\varphi(ug) = \varphi(u)g, \forall g \in G$, in such a way that u and $\varphi(u)$ belong to the same G-orbit $\forall u \in P$.

Let us denote with $V(P) \subset T(P)$ the vertical bundle of the fibration. A description of $V(P)$ becomes from the trivialization $P \times \mathcal{G} \cong V(P)$ given by $(u, A) \longmapsto A^*_u$, where A^* is the fundamental vector field on P associated with the element A of $\mathcal{G}$, the Lie algebra of G. A vector field $X \in V(P)$ is said to be G-invariant if $(R_g)_* X = X, \forall g \in G$. If f_t is the flow of X, then X is G-invariant if and only if $f_t \in \mathrm{Gau}(P), \forall t \in \mathbb{R}$. We shall denote by $\mathrm{gau}(P)$ the Lie algebra of all π-vertical G-invariant vector fields on P, which is called *the gauge algebra of P*.

Definition 1. Let $\mathcal{K}^r(P, \mathcal{G})$ be the fibred manifold over M consisting of the $\mathcal{G}$-valued r-skew-forms on P such that
(a) For $X_1, \ldots, X_r \in T_p P$ we have $\Phi((R_g)_* X_1, \ldots, (R_g)_* X_r) = \mathrm{ad}(g^{-1}) \cdot \Phi(X_1, \ldots, X_r)$.
(b) If one of $X_1, \ldots, X_r$ is π-vertical, then $\Phi(X_1, \ldots, X_r) = 0$.
$\mathcal{K}^r(P, \mathcal{G})$ will be called the space of r-automorphic forms on P. In particular $\mathcal{K}^0(P, \mathcal{G})$ will be called the space of automorphic functions, i.e., maps $\Phi : P \to \mathcal{G}$ such that $\Phi(pg) = \mathrm{ad}(g^{-1})\Phi(p)$, and $\mathcal{K}^2(P, \mathcal{G})$, which we shall denote simply as $\mathcal{K}$, will be called the *curvature bundle* associated with the given principal fibration.

Definition 2. Given $D \in \mathrm{gau}(P)$ we define the automorphic function $\tau_D \in \mathcal{K}^0(P, \mathcal{G})$ as follows: $\tau_D(p)$ is the unique $A \in \mathcal{G}$ such that A^*_p is equal to $D_p \in T_p(P)$.

The automorphic character of τ_D is then guaranteed by the fact that for each $g \in G$, $(R_g)_* A^*$ is the fundamental vector field corresponding to $\mathrm{ad}(g^{-1})A \in \mathcal{G}$ (cf. [7]).

If we define $\exp : \mathcal{K}^0(P, \mathcal{G}) \to \mathrm{Gau}(P)$ as

$$\exp(\tau)(p) = p \cdot \exp(\tau(p)),$$

then for each $D \in \mathrm{gau}(P)$ its flow can be written as $f_t(p) = p \cdot \exp(t\boldsymbol{\tau}_D)(p)$.

We define the bundle of connections $p : \mathcal{C} \to M$ of the principal bundle $\pi : P \to M$ as the subbundle $\mathcal{C} \subset \wedge^1(TP, \mathcal{G})$ formed by the $\mathcal{G}$-valued 1-forms $w : TP \to \mathcal{G}$ such that $w(A^*) = A, \forall A \in \mathcal{G}$, and $(R_g)^* w = \mathrm{ad}(g^{-1})w, \forall g \in G$. If $w, w' \in \mathcal{C}$, then trivially $w - w' \in \mathcal{K}^1(P, \mathcal{G})$. Hence $\mathcal{C}$ is an affine bundle modeled over the vector bundle $\mathcal{K}^1(P, \mathcal{G})$. A connection Γ on P is a global section σ_Γ of $p : \mathcal{C} \to M$. Also we shall denote by w_Γ the $\mathcal{G}$-valued 1-form on P defined by Γ.

Given a connection Γ on P, the horizontal lift of a vector field X on M is the unique horizontal vector field X^* on P (that is $w_\Gamma(X^*) = 0$) which is G-invariant and such that $\pi_*(X^*) = X$.

Definition 3. If we fix a connection form w on P, we define a map between fibred spaces of automorphic forms

$$d^w : \mathcal{K}^r(P, \mathcal{G}) \to \mathcal{K}^{r+1}(P, \mathcal{G}), \quad (r \geq 0)$$

as follows: $d^w \Phi(X_1, \ldots, X_{r+1}) = \mathrm{d}\Phi(X_1^h, \ldots, X_{r+1}^h)$, where X_i^h denotes the horizontal component of X_i with respect to w.

Proposition 4. *For $\Phi \in \mathcal{K}^r(P, \mathcal{G})$ it holds that $d^w \Phi = \mathrm{d}\Phi + [w, \Phi]$.*

Proof. We shall only prove the case in which $\Phi \in \mathcal{K}^1(P, \mathcal{G})$. The general case is in essence identical to this one. Then we shall prove $d^w \Phi(X, Y) = \mathrm{d}\Phi(X, Y) + [w(X), \Phi(Y)] - [w(Y), \Phi(X)]$ for any tangent vectors $X, Y \in T_p P$. Since both sides of this equality are bilinear in X and Y, it is sufficient to consider the different situations in which X and Y are either horizontal or vertical. The only nontrivial case is the one where X is vertical and Y is horizontal. We can suppose that $X = A^*$ at p, where $A \in \mathcal{G}$, and that Y is a horizontal lift of a vector field on M. Then if $f_t(p) = p \cdot \exp(tA)$ is the 1-parameter subgroup of G generated by A, we have

$$[A^*, Y] = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (f_t^{-1})_* Y = 0$$

since Y is R_g-invariant for every $g \in G$. It is clear that $d^w \Phi(X, Y) = 0$, and we shall show that the right hand side of the equality vanishes. Now

$$\mathrm{d}\Phi(A^*, Y) = A^*\Phi(Y) - Y\Phi(A^*) - \Phi([A^*, Y]) = A^*\Phi(Y)$$

and $[w, \Phi](A^*, Y) = [w(A^*), \Phi(Y)] - [w(Y), \Phi(A^*)] = [A, \Phi(Y)]$. But $A^*\Phi(Y) = -[A, \Phi(Y)]$ by the very definition of the Lie derivative of functions and vector fields. ■

Definition 5. Given a connection Γ on P, with connection form w, we define the curvature of Γ as $\Omega_w = d^w w$. Then it holds that $\Omega_w = \mathrm{d}w + \frac{1}{2}[w, w]$ and $\Omega_w \in \mathcal{K}^2(P, \mathcal{G})$. (Note that $w \notin \mathcal{K}^1(P, \mathcal{G})$.)

Proposition 6. *If Γ is a connection on P defined by its connection form w and $D \in \mathrm{gau}(P)$ is the generator of the flow $f_t(p) = p \cdot \exp(t\tau_D(p)) = p \cdot \varphi_t(p)$ ($\varphi_t : P \to G$), then*

(i) $L_D w = d^w \tau_D = \mathrm{d}\tau_D + [w, \tau_D]$

(ii) $L_D \Omega_w = -[\tau_D, \Omega_w]$

where L_D stands for the Lie derivative with respect to D.

Proof. (i) It suffices to see that $L_D w$ and $d^w \tau_D$ coincide when applied to any vector field X on P. If X is vertical, we may suppose $X = A^*$; then since $f_t^* w$ is also a connection form we have

$$L_D w(X) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f_t^* w(A^*) = 0.$$

On the other hand, $d^w \tau_D(A^*) = 0$, since due to the Proposition 4, $d^w \tau_D \in \mathcal{K}^1(P, \mathcal{G})$.

If X is horizontal then $(d^w \tau_D)(X) = (\mathrm{d}\tau_D)(X) = X\tau_D$. On the other hand by the Leibniz rule we can write $(f_t)_* X = (\varphi_t)_* X + (R_{\varphi_t})_* X$, hence calling $\gamma(s)$ a flow of X, we have

$$\begin{aligned} L_D w(X) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f_t^* w(X) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\varphi_t)_* X \\ &= \left.\frac{\partial}{\partial t}\right|_{t=0} \left.\frac{\partial}{\partial s}\right|_{s=0} \varphi_t(\gamma(s)) = \left.\frac{\partial}{\partial t}\right|_{t=0} \left.\frac{\partial}{\partial s}\right|_{s=0} \exp(t\tau_D(\gamma(s))) \\ &= \left.\frac{\partial}{\partial s}\right|_{s=0} \tau_D(\gamma(s)) = X\tau_D. \end{aligned}$$

(ii) Taking into account that for a horizontal vector field $X \in T(P)$, the horizontal component of $(f_t)_* X$ is $(R_{\varphi_t})_* X$, the automorphic character of Ω_w implies $(f_t)^* \Omega_w = \mathrm{ad}(\varphi_t^{-1}) \cdot \Omega_w$, and hence

$$L_D \Omega_w = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (f_t)^* \Omega_w = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{ad}(\exp(t\tau_D)^{-1}) \cdot \Omega_w = -[\tau_D, \Omega_w]. \quad ■$$

Every vector field $D \in \mathrm{gau}(P)$ determines an element $D_{\mathcal{K}} \in \mathcal{X}(\mathcal{C})$ as follows: if f_t is the flow of D, then we define a flow $(f_t)_{\mathcal{C}}$ on $\mathcal{C}$ by pulling back connections forms, that is: $(f_t)_{\mathcal{C}} w = (f_t)^* w$, $w \in \mathcal{C}$; the corresponding infinitesimal generator is denoted by $D_{\mathcal{C}}$. In this way, the map $D \longmapsto D_{\mathcal{C}}$ defines an homomorphism of real Lie algebras whose kernel will be denoted by $\mathrm{gau}_0(P)$:

$$0 \to \mathrm{gau}_0(P) \to \mathrm{gau}(P) \to \mathcal{X}(\mathcal{C}).$$

3. Higher-order Garcia fibration: Gauge meaning

Let $q : \overline{P} \to \mathcal{C}$ be the bundle induced by the principal bundle $\pi : P \to M$ on its fibre bundle of connections $p : \mathcal{C} \to M$. Then $\overline{P}$ is a principal bundle over $\mathcal{C}$ with structural group G, such that the canonical morphism $\pi_{10} : \overline{P} \to P$ is a principal G-bundle morphism, i.e., one has the following commutative diagram:

$$\begin{array}{ccc} \overline{P} & \overset{\pi_{10}}{\rightarrow} & P \\ q \downarrow & & \downarrow \pi \\ \mathcal{C} & \overset{p}{\rightarrow} & M \end{array}$$

A useful construction of the bundle $\overline{P}$ is as follows. Let $p_1 : J^1 P \to M$ be the 1-jet bundle of local sections of $\pi : P \to M$. The group G acts on $J^1 P$ by $(j_x^1 s) \cdot g = j_x^1(R_g s)$. The quotient $J^1 P/G$ exists as a fibred differentiable manifold over M and can be identified with the bundle of connections (see [8]). Let us describe this fact; we define a mapping $q : J^1 P \to \mathcal{C}$ as follows: for each element $j_x^1 s$ we consider the vertical differential of s, $d^v s \in \wedge^1(TP, \mathcal{G})$: if $X \in T_{s(x)} P$ then $d^v_{s(x)} s(X)$, is the only element $A \in \mathcal{G}$, such that $A^*_{s(x)} = X - s_* \pi_* X$. From the fact that the fundamental vector field associated with $(R_g)_* A$, $(g \in G, A \in \mathcal{G})$ is $\left(\mathrm{ad}(g^{-1})A\right)^*$, it follows that $(R_g)^* d^v s = \mathrm{ad}(g^{-1}) d^v s$. On the other hand for each point $u \in \pi^{-1}(x)$, there exists a unique $g \in G$ such that $u = s(x)g$ and we set $d^v_u s : T_u P \to \mathcal{G}$ by $d^v_u s = (R_g)_* \circ (R_{g^{-1}})^* d^v_{s(x)} s$. In this way we obtain an element of $\mathcal{C}$ which depends only on $j_x^1 s$ and we define $q(j_x^1 s) = d^v s$. It is not difficult to prove that q is a surjective submersion whose fibres are the orbits of G, and we have a principal G-bundle $q : J^1 P \to \mathcal{C}$, which we shall call the *Garcia fibration* [8, 9]. Furthermore there is a G-principal bundle isomorphism $J^1 P = \mathcal{C} \times_M P$ given by $j_x^1 s \to (q(j_x^1 s), s(x))$ (see [8]).

Now we consider the r-jet bundle $(r > 1) J^r P \longrightarrow M$ of local sections of π. Our objective here will be to give a gauge interpretation of the quotient bundles $J^r(P)/G$.

Proposition 7. *With the canonical action of G on $J^r(P)$*

$$(j_x^r s) \cdot g = j_x^r(R_g \circ s)$$

where s is a section of π and $g \in G$, the quotient manifold exists and it is endowed with a natural fibred structure over M:

$$\pi_r : J^r(P)/G \to M.$$

Proof. We must check that G acts properly discontinuously on $J^r(P)$. By the Dieudonné criterion [5] this condition holds if and only if $R = \{(j_x^r s, j_x^r s \cdot g); j_x^r s \in J^r(P), g \in G\}$ is a closed submanifold of $J^r(P) \times J^r(P)$, and this is easy to see. ■

Proposition 8. *The local connection associated with a section $s : U \to P$ is flat.*

Proof. It suffices to take into account that the condition of being horizontal for a vector field X on $s(U)$ is $X = s_* \pi_* X$, which is stable by Lie bracket. ■

By taking jets in the submersion $q : J^1(P) \to \mathcal{C}$ we obtain a mapping $J^{r-1}(q) : J^{r-1}(J^1(P)) \to J^{r-1}(\mathcal{C})$ and restricting $J^{r-1}(q)$ to the injection $J^r(P) \hookrightarrow J^{r-1}(J^1(P))$, $j_x^r s \longmapsto j_x^{r-1}(j_x^1 s)$, we can define a homomorphism of fibred manifolds over M:

$$\varphi_r : J^r(P) \to J^{r-1}(\mathcal{C}).$$

Theorem 9. *For every* $r > 1$ *there exists an embedding*

$$\Phi_r : J^r(P)/G \to J^{r-1}(\mathcal{C})$$

whose image coincides with the subbundle $\mathcal{C}^{r-1}$ *of* $J^{r-1}(\mathcal{C})$ *given by*

$$\mathcal{C}^{r-1} = \{j_x^{r-1}(\sigma_\Gamma) \in J^{r-1}(\mathcal{C}) : j_x^{r-2}(\Omega_\Gamma) = 0\}.$$

Proof. It easily follows from the definition of the Garcia fibration $q : J^1(P) \to \mathcal{C}$ that the definition of Φ_r makes sense and that it is injective. Now by Proposition 8 it is clear that $\operatorname{Im}\Phi_r \subseteq \mathcal{C}^{r-1}$. For the converse we shall introduce some notation. Assume that $(U; x_1, \ldots x_n)$ is an open coordinate subset of M over which P is trivial: $\pi^{-1}(U) = U \times G$. Let $\{A_1, \ldots, A_m\}$ be a basis of the Lie algebra $\mathcal{G}$ of G. We have a coordinate system $(x_j; z_{ij})$, $1 \le i \le m$, $1 \le j \le n$, in $\mathcal{C}$ given by the formula:

$$\sigma_\Gamma(\partial/\partial x_j) = \partial/\partial x_j + \sum_i (z_{ij} \circ \sigma_\Gamma)\widetilde{A}_i$$

where $\widetilde{A}_i$ is the G-invariant vector field on $\pi^{-1}(U)$ induced by the 1-parameter group $\tau^i_{(x,\sigma)}(t) = [x, \exp(tA_i)\sigma]$. In what follows we shall refer to the coordinate systems $(x_j, y^i_{(\alpha)})$, $(x_j, z^{ij}_{(\alpha)})$ induced by (x_j, y_i), (x_j, z_{ij}) on $J^r(P)$, $J^r(\mathcal{C})$ respectively.

Given a section $s : U \to P$, we set $s_i = y_i \circ s$, $1 \le i \le m$. As a simple calculation shows, we have

$$d^v s = \sum_{h,i} a_{hi} \left(\mathrm{d}y_i - \sum_j \frac{\partial s_i}{\partial x_j} \mathrm{d}x_j \right) \otimes \widetilde{A}_h$$

where (a_{hi}) is an invertible matrix in a neighbourhood of the origin. In this way the equations of the projection $q : J^1(P) \to \mathcal{C}$ are the following:

$$z_{hj} \circ q = \sum_{i=1}^m a_{hi} y^i_{(j)}, \quad 1 \le h \le m;\ 1 \le j \le n.$$

Hence,

$$(z^{hj}_\alpha \circ \Phi_r)(j^r_x s) = \frac{\partial^{|\alpha|}}{\partial x^\alpha} \left(\sum_i (a_{hi} s) \frac{\partial s_i}{\partial x_j} \right)(x),$$

$\alpha \in \mathbb{N}$; $|\alpha| \le r-1$; $1 \le h, i \le m$; $1 \le j \le n$, and, by using the Leibniz formula, we obtain

$$(z^{hj}_\alpha \circ \Phi_r)(j^r_x s) = \sum_i \sum_{\beta \le \alpha} \binom{\alpha}{\beta} D_x^{\alpha-\beta}(a_{hi} s)(D_x^{\beta+(j)} s_i),$$

$\alpha, \beta \in \mathbb{N}$; $|\alpha| \le r-1$; $1 \le h, i \le m$; $1 \le j \le n$, where we have set $D^\alpha = \partial^{|\alpha|}/\partial x^\alpha$. Now, let us consider a point $j_x^{r-1}(\sigma_\Gamma) \in J^{r-1}(\mathcal{C})$, so that $j_x^{r-2}(\Omega_\Gamma) = 0$. We define a point $j^r_x s \in J^r(P)$ by giving its components $y^h_\alpha(j^r_x s) = D^\alpha_x s_h$, $|\alpha| \le r$; $1 \le h \le m$. To do so, we proceed by recurrence on $|\alpha|$. Set $s_h(x) = 0$, $1 \le h \le m$. Assume that $D^a_x s_h$ has been defined for every $|\alpha| < l$, $1 \le l \le r-1$, in such a way that the following formulas hold:

$$\sum_i \sum_{b \le a} \binom{\alpha}{\beta} D_x^{\alpha-\beta}(a_{hi} \circ s)(D_x^{\beta+(j)} s_i) = z^{hj}_a(j_x^{r-1}(\sigma_\Gamma))$$

for all $|\alpha| \le l-1$; $1 \le h, i, \le m$; $1 \le j \le n$. Let α be a multi-index of order l and let us consider an index j such that $\alpha_j > 0$. We define

$$D^\alpha_x s_h = z^{hj}_{\alpha-(j)}(j_x^{r-1}(\sigma_\Gamma)) - \sum_i \sum_{\beta < \alpha-(j)} \binom{\alpha-(j)}{\beta} D_x^{\alpha-(j)-\beta}(a_{hj} \circ s)(D_x^{\beta+(j)} s_i). \tag{a}$$

It is obvious that (a) also holds for the indices $\alpha - (j), j$. Let k be another index such that $\alpha_k > 0$. Thus we can write $\alpha = (jk) + s$. We only need to prove that (a) also holds for the couple $\alpha - (k), k$; or, in other words,

$$z^{hk}_{\sigma+(j)}(j^{r-1}_x(\sigma_\Gamma)) - z^{hj}_{\sigma+(k)}(j^{r-1}_x(\sigma_\Gamma)) = \sum_i \sum_{\beta<\sigma+(k)} \binom{\sigma+(k)}{\beta} D^{\sigma+(k)-\beta}_x (a_{hi}\circ\sigma)(D^{\beta+(j)}_x s_i)$$
$$- \sum_i \sum_{\beta<\sigma+(k)} \binom{\sigma+(j)}{\beta} D^{\sigma+(j)-\beta}_x (a_{hi}\circ\sigma)(D^{\beta+(k)}_x s_i). \tag{b}$$

Moreover, the condition $j^{r-2}_x(\Omega_\Gamma) = 0$ means

$$z^{hk}_{\sigma+(j)}(j^{r-1}_x(\sigma_\Gamma)) - z^{hj}_{\sigma+(k)}(j^{r-1}_x(\sigma_\Gamma)) = \sum_{\beta\le\sigma} \sum_{i,t} \binom{\sigma}{\beta} c^h_{it} \lambda^{tj}_{\sigma-\beta} \lambda^{ik}_\beta,$$

where c^h_{it} are the structure constants of $\mathcal{G}$ relative to the basis chosen. Hence, (b) is equivalent to the following:

$$\sum_{\beta\le\sigma} \sum_{i,t} \binom{\sigma}{\beta} c^h_{it} D^{\sigma-\beta}_x \left(\sum_u (a_{tu}\circ s) \frac{\partial s_u}{\partial x_j} \right) D^\beta_x \left(\sum_v (a_{iv}\circ s) \frac{\partial s_v}{\partial x_k} \right)$$
$$= \sum_i \sum_{\beta<\sigma+(k)} \binom{\sigma+(k)}{\beta} D^{\sigma+(k)-\beta}_x (a_{hi}\circ\sigma)(D^{\beta+(j)}_x s_i)$$
$$- \sum_i \sum_{\beta<\sigma+(k)} \binom{\sigma+(j)}{\beta} D^{\sigma+(j)-\beta}_x (a_{hi}\circ s)(D^{\beta+(k)}_x s_i).$$

The left hand side of the above equation is equal to

$$\sum_{i,t,u,v} c^h_{it} D^\sigma_x \left((a_{tu}\circ s)(a_{iv}\circ s) \frac{\partial s_u}{\partial x_j} \frac{\partial s_v}{\partial x_k} \right),$$

whereas its right hand side becomes

$$\sum_i D^{\sigma+(k)}_x \left((a_{hi}\circ s) \frac{\partial s_i}{\partial x_j} \right) - \sum_i D^{\sigma+(j)}_x \left((a_{hi}\circ s) \frac{\partial s_i}{\partial x_k} \right) = \sum_{u,v} D^\sigma_x \left(\frac{\partial a_{hu}}{\partial y_v}\circ s - \frac{\partial a_{hv}}{\partial y_u}\circ s \right) \frac{\partial s_u}{\partial x_j} \frac{\partial s_v}{\partial x_k}.$$

Moreover, it holds that $\partial a_{hu}/\partial y_v - \partial a_{hv}/\partial y_u = \sum_{i,t} c^h_{it} a^{tu} a^{iv}$, showing that the two sides coincide, thus completing the theorem proof. ■

4. Higher-order gauge invariant Lagrangians

Lemma 10. (1) *Every element $f \in \mathrm{Gau}(P)$ induces, by pulling back 2-forms, an isomorphism of fibred manifolds over M*

$$f^{(r)} : J^r(\mathcal{K}) \to J^r(\mathcal{K}), \quad (r \ge 0).$$

The image of $\mathrm{Gau}(P)$ *under this representation will be denoted* $\mathrm{Gau}^{(r)}_{\mathcal{K}}(P)$.

(2) *Every element $g \in G$ induces an isomorphism*

$$g^{(r)} : J^r(\mathcal{K}) \to J^r(\mathcal{K})$$

given by $g^{(r)}(j^r_x \Phi) = j^r_x(\mathrm{ad}(g^{-1}) \cdot \Phi)$, which provides a representation of the Lie group G denoted as $G^{(r)}_{\mathcal{K}}$.

(3) *Every element $f \in \mathrm{Gau}(P)$ induces, pulling back connection forms, an isomorphism of fibred manifolds over M*

$$f^{(r)} : J^r(\mathcal{C}) \to J^r(\mathcal{C}).$$

The corresponding image of $\mathrm{Gau}(P)$ *under this representation will be denoted as* $\mathrm{Gau}^{(r)}_{\mathcal{C}}(P)$.

Proof. (1) For $\Phi \in \mathcal{K}$ and $f \in \mathrm{Gau}(P)$ we have $(R_g)^* f^*(\Phi) = f^*(R_g)^*(\Phi) = f^*\left(\mathrm{ad}(g^{-1}) \cdot \Phi\right) = \mathrm{ad}(g^{-1}) \cdot f^*(\Phi)$. On the other hand, if A^* is a fundamental vector field on P, then it is easy to see that $f_*(A^*) = A^*$. Thus $f^*(\Phi)$ is, like Φ, null whenever one of its arguments is vertical. In this way we have an isomorphism $f^* : \mathcal{K} \to \mathcal{K}$. According to Goldschmidt [12] there is an rth prolongation $f^{(r)} : J^r(\mathcal{K}) \to J^r(\mathcal{K})$ which is also an isomorphism of fibred manifolds over X.

(2) and (3) are proved in a similar way. ∎

Lemma 11. *The orbits by the actions of the groups* $\mathrm{Gau}^{(r)}_{\mathcal{K}}(P)$ *and* $G^{(r)}_{\mathcal{K}}$ *on* $J^r(\mathcal{K})$ *coincide.*

Proof. If $f \in \mathrm{Gau}(P)$ is of the form $f(p) = p\varphi(p)$ with $\varphi : P \to G$, and $\Phi \in K$, by the proof of Proposition 6 we have $f^*(\Phi) = \mathrm{ad}(\varphi^{-1}) \cdot \Phi$. In this way, our statement follows since φ ranges over all G as f ranges over $\mathrm{Gau}(P)$. ∎

Definition 12. We shall define an rth-order Lagrangian on the fibred of connections $p : \mathcal{C} \to M$ as an arbitrary differentiable function $\mathcal{L} : J^r(\mathcal{C}) \to \mathbb{R}$. Such a Lagrangian is said to be gauge invariant if $\mathcal{L}(f^{(r)} j^r_x(\sigma)) = \mathcal{L}(j^r_x(\sigma))$ for all $f^{(r)} \in \mathrm{Gau}^{(r)}_{\mathcal{C}}(P)$.

In the same way an $(r-1)$th gauge invariant curvature Lagrangian is a differentiable function $\overline{\mathcal{L}} : J^{r-1}(\mathcal{K}) \to \mathbb{R}$, verifying $\overline{L}(g^{(r)} j^{r-1}_x(\Omega)) = \overline{L}(j^{r-1}_x(\Omega))$ for all $g^{(r)} \in G^{(r)}_{\mathcal{K}}$.

We define a map:

$$\Omega : J^1(\mathcal{C}) \to \mathcal{K},$$

called the *curvature mapping*, assigning to the 1-jet $j^1_x(\sigma_\Gamma)$ of a local connection Γ, its curvature $j^0_x(\Omega_\Gamma)$.

By applying the J^{r-1} functor to this map and restricting to the holonomic subbundle, we define an homomorphism of vector bundles

$$J^{r-1}_\Omega : J^r(\mathcal{C}) \to J^{r-1}(\mathcal{K})$$

as being the composition $J^r(\mathcal{C}) \hookrightarrow J^{r-1}(J^1(C)) \overset{J^{r-1}_\Omega}{\to} J^{r-1}(\mathcal{K})$.

Theorem 13. *Given a principal G-bundle $\pi : P \to M$, where G is any commutative Lie group, an rth-order Lagrangian defined on the fibred manifold of connections $\mathcal{L} : J^r(\mathcal{C}) \to \mathbb{R}$ is gauge invariant if and only there exists an $(r-1)$th-order curvature Lagrangian $\overline{\mathcal{L}}$ which is $G^{(r)}_{\mathcal{K}}$-invariant and a factorization through the $(r-1)$th-order curvature mapping,*[1]

$$\mathcal{L} = \overline{\mathcal{L}} \circ J^{r-1}_\Omega.$$

Proof. In order to prove the theorem, it suffices to see that the orbit through the gauge group of a point $j^r_x(\sigma_\Gamma) \in J^r(\mathcal{C})$ coincides with the inverse image through J^{r-1}_Ω of the gauge orbit of $j^{r-1}_x(\Omega_\Gamma)$ in $J^{r-1}(\mathcal{K})$. That is, if we denote as $\mathcal{O}(j^r_x(\sigma_\Gamma))$ and $\mathcal{O}(j^{r-1}_x(\Omega_\Gamma))$ such orbits, taking into account that $\mathcal{O}(j^{r-1}_x(\Omega_\Gamma)) = \{j^{r-1}_x(\Omega_\Gamma)\}$, we shall therefore prove

$$\mathcal{O}(j^r_x(\sigma_\Gamma)) = (J^{r-1}_\Omega)^{-1}(j^{r-1}_x(\Omega_\Gamma)). \tag{c}$$

In this way, with the equality between these two gauge spaces, we can correctly define the Lagrangian $\overline{\mathcal{L}}$: $J^{r-1}(\mathcal{K}) \to \mathbb{R}$ by the formula $\overline{\mathcal{L}}(j^{r-1}_x(\Omega_\Gamma)) = \mathcal{L}(j^r_x(\sigma_\Gamma))$. In fact, if we choose another connection Γ' such that $j^{r-1}_x(\Omega_{\Gamma'}) = j^{r-1}_x(\Omega_\Gamma)$, then

$$j^r_x(\sigma_{\Gamma'}) \in (J^{r-1}_\Omega)^{-1}(j^{r-1}_x(\Omega_\Gamma)),$$

[1] We have stated the theorem in the most general possible way, in order to show the similarity with the classical Utiyama theorem (true without the commutative hypothesis only for $r = 1$). In our case, in which the structure group G is abelian, due to the adjoint representation being trivial, there are no gauge orbits in $\mathcal{K}$, and hence the condition of gauge invariance on $\overline{\mathcal{L}}$ is superfluous. In this way the theorem simply states that the Lagrangian $\mathcal{L}$ is invariant if and only if it factors through the curvature mapping, which in this case is nothing but the exterior differential.

and by virtue of the hypothesis, there exists a gauge transformation $f \in \mathrm{Gau}(P)$, such that $j^r_x(\sigma_{\Gamma'}) = f^{(r)} j^r_x(\sigma_\Gamma)$, with which, taking into account the gauge invariance of $\mathcal{L}$, we conclude that $\overline{\mathcal{L}}(j^{r-1}_x(\Omega_\Gamma))$ does not depend on the chosen representative.

Therefore, in order to prove the equality (c), let us first see that the left hand side is included in the right hand side, as a consequence of the curvature mapping being gauge equivariant:

$$J^{r-1}_\Omega(f^{(r)}_{\mathcal{C}} j^r_x(\sigma_\Gamma)) = J^{r-1}_\Omega(j^r_x(f^*\sigma_\Gamma)) = f^{(r-1)}_K j^{r-1}_x(\Omega) = J^{r-1}_\Omega(j^r_x(\sigma_\Gamma)).$$

Now we compute $\dim(\mathcal{O}(j^r_x(\sigma_\Gamma)))$. If we take a local connection form w_Γ associated with the connection Γ and $f_t(p) = p\exp(t\cdot\tau_D(p)) \in \mathrm{Gau}(P)$, then a tangent vector of $\mathcal{O}(j^r_x(\sigma_\Gamma))$ at $j^r_x(\sigma_\Gamma)$ is given by $j^r_x\left(\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f^*_t w_\Gamma\right)$, but

$$j^r_x\left(\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} f^*_t w_\Gamma\right) = j^r_x(L_D w_\Gamma) = j^r_x(d^{w_\Gamma}\tau_D) = j^r_x(\mathrm{d}\tau_D).$$

Now, let U be an open set on which P is trivial and coordinated by $\{x_1, \ldots, x_n\}$ ($n = \dim X$) and let $\{A_1, \ldots, A_m\}$ be a basis of $\mathcal{G}$. Let us denote by $\{\tau^\alpha_D\}$ the mesonic components of τ_D in this basis. Then $j^r_x(\mathrm{d}\tau_D)$ is determined by the parameters

$$\frac{\partial^{|I|}\tau^\alpha_D}{\partial x_{i_1}\cdots\partial x_{i_p}}(x), \quad I = (i_1, \ldots, i_p),\ 0 < p \le r+1.$$

Thus we get $\dim(\mathcal{O}(j^r_x(\sigma_\Gamma))) = m\left(n + \binom{n+1}{2} + \cdots + \binom{n+r}{r+1}\right)$. We shall call this combinatorial number D_{r+1}.

To compute $\dim(J^{r-1}_\Omega)^{-1}(j^{r-1}_x(\Omega_\Gamma))$, we consider the following exact sequence, trivially deduced from Theorem 9:

$$0 \to J^{r+1}(P)/G \to J^r(\mathcal{C}) \to \mathrm{Im}(J^{r-1}_\Omega) \to 0, \quad (r \ge 2).$$

An easy computation proves that the dimension of the fiber of the bundle $J^{r+1}(P)/G$ over $x \in M$ is D_{r+1}, which easily implies $\dim(J^{r-1}_\Omega)^{-1}(j^{r-1}_x(\Omega_\Gamma)) = D_{r+1}$.

In this way, $\mathcal{O}(j^r_x(\sigma_\Gamma))$ is an open subset of $(J^{r-1}_\Omega)^{-1}(\mathcal{O}(j^{r-1}_x(\Omega_\Gamma)))$. Since the orbits are either disjoint or coincident, if we prove that $(J^{r-1}_\Omega)^{-1}(j^{r-1}_x(\Omega_\Gamma))$ is connected then the equality (c) between gauge spaces is established.

We shall proceed by induction on r. In order to describe $j^1_x(\sigma_\Gamma)$ we consider the local connection form $w_\Gamma = \sum_{\alpha=1}^m w^\alpha \otimes A_\alpha$. Then $j^1_x(\sigma_\Gamma)$ is determined by the values on x of the functions $w^\alpha_i = w^\alpha(\frac{\partial}{\partial x_i})$ and $w^\alpha_{ij} = \frac{\partial w^\alpha_i}{\partial x_j}$. Analogously, $j^0_x(\Omega_\Gamma)$ is described by the values on x of $\Omega^\alpha_{ij} = \Omega^\alpha(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j})$ where $\Omega = \sum_{\alpha=1}^m \Omega^\alpha \otimes A_\alpha$.

Thus, given the components $\{\Omega_{ij}\}$ of $j^0_x(\Omega_\Gamma)$ we must prove that the set of 1-jets $j^1_x(\sigma_\Gamma) = j^1_x(w_\Gamma) = \{w^\alpha_{ij}, w^\alpha_i\}$ which verify

$$J^0_\Omega(j^1_x(\sigma_\Gamma)) = j^0_x(\Omega_\Gamma), \quad \text{that is: } w^\alpha_{ij} - w^\alpha_{ji} = \Omega^\alpha_{ij} \tag{d}$$

is a connected subset of $\mathbb{R}^{m(n+n^2)}$. To see this, we take first the particular solution $j^1_x(\sigma_\Gamma) = (w^\alpha_{ij} = \frac{1}{2}\Omega^\alpha_{ij}, w^\alpha_i = 0)$. Given any other solution $j^1_x(\sigma'_\Gamma) = \{w'^\alpha_{ij}, w'^\alpha_i\}$ verifying the curvature equation (d), we can perform a deformation of the first in the second by writing

$$\widetilde{w}^\alpha_i(t) = (1-t)w'^\alpha_i, \qquad \widetilde{w}^\alpha_{ij}(t) = (1-t)w'^\alpha_{ij} - \frac{1}{2}t\Omega^\alpha_{ij}.$$

Thus, there is a path linking any two 1-jets of connections in $\left(J^0_\Omega\right)^{-1}(j^0_x(\Omega_\Gamma))$, proving the connectedness of this space.

The induction process is applied taking into account the following statement (which is easy to prove): Let C be a set of functions defined on a neighbourhood of a point x_0 of $\mathbb{R}^k$. If $D = \{f(x_0) : f \in C\}$ is a subset of $\mathbb{R}^N$ and $D' = \{f'(x_0) : f \in C\}$ is a subset of $\mathbb{R}^M$, then the path-connected character of D implies the connectedness of D'. ■

Additional Remark 14. If $r > 1$ and G is nonabelian, then the gauge invariant Lagrangians $\mathcal{L} : J^r(\mathcal{C}) \to \mathbb{R}$ are not classifiable by the Utiyama theorem. In fact, as regards the theorem notation, we have

$$\mathcal{O}(j_x^r(\sigma_\Gamma)) \subset (J_\Omega^{r-1})^{-1}(\mathcal{O}(j_x^{r-1}(\Omega_\Gamma))),$$

and the inclusion is always strict. To see this, we calculate first $\dim \mathcal{O}(j_x^r(\sigma_\Gamma))$. If we compute again the number of parameters on which a tangent vector at $j_x^r(\sigma_\Gamma) \in \mathcal{O}(j_x^r(\sigma_\Gamma))$ depends, we have

$$j_x^r\left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f_t^* w_\Gamma\right) = j_x^r(L_D w_\Gamma) = j_x^r(d^{w_\Gamma}\tau_D) = j_x^r(\mathrm{d}\tau_D + [\tau_D, w]),$$

for the flow f_t of $D \in \mathrm{gau}(P)$. Thus now, not only must we take into account the mesonic components of $\mathrm{d}\tau_D$ but also those of τ_D, and accordingly we get $\dim \mathcal{O}(j_x^r(\sigma_\Gamma)) = m + D_{r+1}$. On the other hand and in order to calculate the dimension of the gauge orbit of $j_x^{r-1}(\Omega_\Gamma)$ in $J^{r-1}(\mathcal{K})$, we have

$$j_x^{r-1}\left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f_t^* \Omega_\Gamma\right) = j_x^{r-1}(L_D \Omega_\Gamma) = -j_x^{r-1}([\tau_D, \Omega_\Gamma]),$$

and then $\dim\left(\mathcal{O}(j_x^{r-1}(\Omega_\Gamma))\right) = m + D_{r-1}$. Since $(J_\Omega^{r-1})^{-1}$ raises in D_{r+1} the dimension of any subspace in $\mathrm{Im}(J_D^{r-1})_x$, we conclude that

$$\dim(J_\Omega^{r-1})^{-1}(\mathcal{O}(j_x^{r-1}(\Omega_\Gamma))) - \dim \mathcal{O}(j_x^r(\sigma_\Gamma)) = D_{r-1}.$$

Additional Remark 15. For $r = 1$, due to $D_0 = 0$, we have $\dim(J_\Omega^0)^{-1}(\mathcal{O}(j_x^0(\Omega_\Gamma))) = m + D_2 = \dim(\mathcal{O}(j_x^1(\sigma_\Gamma)))$; an argument on the connectedness of $(J_\Omega^0)^{-1}(\mathcal{O}(j_x^0(\Omega_\Gamma)))$, similar to that in our theorem, proves that in fact

$$\mathcal{O}(j_x^1(\sigma_\Gamma)) = (J_\Omega^0)^{-1}(\mathcal{O}(j_x^0(\Omega_\Gamma))),$$

which provides the classical Utiyama–Yang–Mills theorem.

References

[1] D. Betounes, The geometry of gauge-particle field interaction: A generalization of Utiyama's theorem, J. Geom. Phys. 6 (1989) 107–125.
[2] D. Betounes, Special symmetries for the Utiyama Lagrangian with external Yang–Mills fields, Phys. Rev. D 31 (10) (1985) 2677–2681.
[3] D. Bleecker, Gauge Theory and Variational Principles, Addison-Wesley, 1971.
[4] U. Bruzzo, Global Utiyama's theorem in curved superspace, in: Proceedings of the XV International Conference on Differential Geometric Methods in Theoretical Physics (Clausthal, 1986), World Sci. Publishing, Teaneck, HJ, 1987, pp. 308–316.
[5] J. Dieudonné, Eléments d'Analyse, Tome III, Gauthier-Villars, 1970.
[6] D.J. Eck, Gauge natural bundles and generalized gauge theories, Mem. Amer. Math. Soc. 247 (1981).
[7] S. Kobayashi, K. Nomizu, Foundations of Differential Geometry, Wiley-Interscience Publication, 1963.
[8] P.L. Garcia, Connections and 1-jet bundles, Rend. Sem. Mat. Univ. Padova 47 (1972).
[9] P.L. Garcia, Gauge algebras, curvature and symplectic structure, J. Differential Geom. 12 (1977) 209–227.
[10] R. Grassini, On Utiyama's invariant theoretical interpretation of interaction, J. Math. Phys. 26 (6) (1985) 1345–1347.
[11] R. Grassini, The geometrical setting of Utiyama's interaction theory, J. Math. Phys. 26 (11) (1985) 2974–2976.
[12] H. Goldschmidt, Integrability criteria for systems of non-linear partial differential equations, J. Differential Geom. (1) (1967) 269–270.
[13] J. Muñoz Masqué, L.M. Pozo Coronado, Higher-order gauge invariant Lagrangians on T^*M, J. Phys. A: Math. Gen. 29 (1996) 7757–7767.
[14] R. Utiyama, Invariant theoretical interpretation of iteration, Phys. Rev. 101 (1956) 1597–1607.
[15] C.N. Yang, R.L. Mills, Conservation of isotopic spin and isotopic gauge invariance, Phys. Rev. 96 (1954) 191–195.