I. INTRODUCTION

Since the introduction by Yang and Mills of the non-Abelian gauge theories [1], attempts have been undertaken of describing all the known interactions as emerging from the localization of some fundamental symmetries of the laws of physics. It is now clear that all the non-gravitational fundamental interactions can be successfully given such an interpretation. The Yang–Mills (YM) theories constitute a formal basis for the standard model of particle physics. Although the attempts to describe gravity as a gauge theory were initiated by Utiyama [2] within a mere two years after the pioneering work of Yang and Mills, the construction of this theory seems yet not to be satisfactorily completed.

If a field theory in Minkowski space is given, this theory being symmetric under the global action of a representation of a Lie group, the natural way to introduce the corresponding interaction within the spirit of YM is to apply the minimal coupling procedure (MCP). However, trying to apply MCP in order to pass from a field theory in flat space to a Riemann–Cartan (RC) space (i.e. a manifold equipped with a metric tensor and a metric connection) results in difficulties. This is because adding a divergence to the flat space Lagrangian density, which is a symmetry transformation, leads to the non-equivalent theory in curved space after MCP is applied. Although this problem was observed already by Kibble [3], it has been largely ignored in the subsequent investigations concerning EC theory. The resulting ambiguity can be physically important for the standard Einstein–Cartan theory and its modifications [4, 5]. It seems that MCP should be somehow modified for the sake of connections with torsion, so that it gives equivalent results for equivalent flat space Lagrangians. An attempt to establish such a modification was made by Saa [6, 7]. Unfortunately, Saa’s solution results in significant departures from general relativity, which seem incompatible with observable data [8][9], unless some additional assumptions of rather artificial nature are made, such as demanding a priori that part of the torsion tensor vanish [10]. The main purpose of this paper is to introduce an alternative modification of MCP, which also eliminates the ambiguity. Unlike Saa’s proposal, our approach does not lead to radical changes in the predictions of the theory. In the case of gravity with fermions, the procedure simply justifies the earlier results of [11–15]. These results were obtained partly ‘by chance’, as the flat space Dirac Lagrangian was randomly selected from the infinity of equally good possibilities.

II. THE GAUGE APPROACH TO GRAVITY AND THE AMBIGUITY OF MINIMAL COUPLING

Let us recall the classical formalism of a YM theory of a Lie group G. Let

\[ S[\phi] = \int \mathcal{L}(\phi, \partial_\mu \phi) \, d^4x = \int \mathcal{L}(\phi, d\phi) \]  

represent the action of a field theory in Minkowski space \( M \). Here \( \mathcal{L} \) is a Lagrangian density and \( \mathcal{L} \) a Lagrangian four–form. Assume that \( \mathcal{V} \) is a (finite dimensional) linear space in which fields \( \phi \) take their values, \( \phi : M \to \mathcal{V} \), and \( \pi \) is a representation of \( Lie(G) \) on \( \mathcal{V} \). Let \( \rho \) denote the corresponding representation of the group\(^1\), \( \rho(\exp(g)) = \exp(\pi(g)) \). If the Lagrangian four–form is invariant under its global action \( \phi \to \phi' = \rho(g)\phi \), one can introduce an interaction associated to the symmetry group \( G \) by allowing the group element \( g \) to depend on space–time point and demanding the theory to be invariant under the local action of \( G \). This can be most easily achieved by performing the replacement

\[ d\phi \to D\phi = d\phi + A\phi, \]  

where \( A \) is a \( Lin(\mathcal{V}) \)–valued one–form field on \( M \) \( (Lin(\mathcal{V})) \) being the set of linear maps of \( \mathcal{V} \) into itself\) which transforms under the local action of \( G \) as

\[ A \to A' = \rho(g)A\rho^{-1}(g) - d\rho(g)\rho^{-1}(g). \]

\(^1\) More precisely, in a generic case \( \rho \) is a representation of the universal covering group of \( G \), which may not be a representation of \( G \) itself.
In the standard YM one requires that \( A \) takes values in a linear subspace \( \text{Ran}(\pi) \) of \( \text{Lin}(V) \), but this requirement is not necessary to make the action invariant under local transformations. We shall adopt a more general approach, in which \( A \) assumes the form
\[
A = A + B(A,e),
\]
where \( A \) is the usual YM connection taking values in \( \text{Ran}(\pi) \) and transforming according to (II.3), \( e \) denotes an orthonormal basis of one–form fields serving physically as a reference frame at each point of space–time\(^2\), \( B(A,e) \) is a \( \text{Ran}(\pi)^+ \)-valued one–form on \( M \). Here \( \perp \) denotes the orthogonal complement with respect to some natural scalar product on \( \text{Lin}(V) \). The simplest candidate for this scalar product is \( \langle (X,Y) \rangle = \text{trace} (X^\dagger Y) \), where \( \dagger \) stands for Hermitian conjugation of a matrix. However, if \( V \) admits a \( \rho \)-invariant scalar product \( \langle , \rangle_\rho \) such that \( \forall v,w \in V, \ g \in G, \langle (\rho(g)v, \rho(g)w)_\rho = \langle v,w \rangle_\rho \), then the use of the induced scalar product \( \langle , \rangle_\rho \) on \( \text{Lin}(V) \) satisfying \( \langle (\rho(g)X \rho^{-1}(g), \rho(g)Y \rho^{-1}(g)) \rangle_\rho = \langle (X,Y) \rangle_\rho \) may seem esthetically more appealing. This product may not be positive–definite, but if the subspace \( \text{Ran}(\pi) \subset \text{Lin}(V) \) is nondegenerate with respect to \( \langle , \rangle_\rho \), then the space of linear maps decouples into a simple sum \( \text{Lin}(V) = \text{Ran}(\pi) \oplus \text{Ran}(\pi)^+ \) and hence \( A \) and \( B(A,e) \) are uniquely determined by \( A \).

In order not to introduce additional fields, \( B \) is required to be determined by \( A \) and \( e \). In order not to destroy the transformation law (II.3), it is also required that \( B(A',e') = \rho(g)B(A,e)\rho^{-1}(g) \). Our final requirement is that the coupling procedure thus obtained be free of the ambiguity corresponding to the possibility of the addition of a divergence to the initial matter action. It is remarkable that in the case of the gravitational interaction and fermions these ideas, together with the natural requirement that the Leibniz rule holds for vector fields composed of spinors, fix the form of \( B(A,e) \) (up to terms that can be absorbed by other known fundamental interactions and do not influence the resulting connection on the base manifold), as we will see below. All the constructions of YM can be accomplished in terms of \( A \) and its curvature \( F = dA + A \wedge A \). The role of \( B \) is only to modify the coupling procedure such that it is unique.

In the case of gravity, it is not sufficient to perform the replacement (II.2) – one needs also to replace the Minkowski space (holonomic) basis of orthonormal one–forms \( dx^\mu \) by the cotetrad \( e^a \) and redefine the geometric structure of the base manifold such that the original Minkowski space \( M \) becomes the RC space \( \mathcal{M}(e,\omega) \) (here \( \omega \) is a spin–connection that can be extracted out of \( A \)). We shall use the Dirac field case as an instructive example. In particle physics, the most frequently used Lagrangian four–form for the Dirac field is
\[
\mathcal{L}_{F0} = -i (\gamma^\mu \partial_\mu \psi + m \psi) dx^\mu.
\]
(II.5)
Here \( \gamma^\mu \) are the Dirac matrices obeying \( \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2\eta^\mu\nu \), where \( \eta = \text{diag}(1,-1,-1,-1) \) is the Minkowski matrix, and \( \psi := \psi^\dagger \gamma^0 \), where \( \psi^\dagger \) is a Hermitian conjugation of a column matrix (think of \( \psi \) as a column of four complex–valued functions on space–time). This four–form is invariant under the global action of the Poincaré group
\[
x^\mu \rightarrow x'^\mu = \Lambda^\mu_{\nu} x^\nu + a^\mu, \quad \psi \rightarrow \psi' = S(\Lambda)\psi,
\]
\[
S(\Lambda(\varepsilon)) := \exp \left( -\frac{i}{4} \varepsilon_{\mu\nu} \Sigma^{\mu\nu} \right), \quad \Sigma^{\mu\nu} := \frac{i}{2} [\gamma^\mu, \gamma^\nu],
\]
(II.6)
where \( a^\mu \) and \( \varepsilon_{\mu\nu} = -\varepsilon_{\nu\mu} \) are the parameters of the transformation. In order to make the symmetry local, it is sufficient to replace the differentials by covariant differentials (thus introducing the connection \( \omega \)), to replace the basis of one–forms \( dx^\mu \) of \( M \) by the cotetrad basis \( e^a \) on the resulting RC space \( \mathcal{M}(e,\omega) \), and to use the Hodge star operator \( \star \) adapted to \( M \). The resulting Lagrangian four–form is
\[
\mathcal{L}_{F0} = -i (\star e_a) \wedge \bar{\psi} \gamma^a D\psi - m \bar{\psi} \psi e,
\]
\[
D\psi = d\psi - \frac{i}{4} \omega_{ab} \Sigma^{ab} \psi
\]
(II.6)
(the matrices \( \gamma^a \), \( a = 0, \ldots, 3 \) are just the same as \( \gamma^\mu \), \( \mu = 0, \ldots, 3 \)). Here \( e = e^0 \wedge e^1 \wedge e^2 \wedge e^3 \) is the canonical volume element on \( \mathcal{M} \). The coupling procedure of this kind will be referred to as the minimal coupling procedure (MCP) for the gravitational interaction. The one–forms \( \omega_{ab} = -\omega_{ba} \), which endow the space–time with the metric–compatible connection, may be interpreted as gauge–fields corresponding to Lorentz rotations. Although the relation of \( e^a \) to the translational gauge–fields is more subtle, the procedure can be given interpretation in the framework of gauge theory of the Poincaré group (see [16] for an exhaustive and simple treatment). In EC theory, the gauge–field–part of the Lagrangian is taken to be \( \mathcal{L}_G = -\frac{i}{4} \varepsilon_{abcd} e^a \wedge \omega^b \wedge e^c \wedge e^d \), where \( k \) is a constant and \( \Omega^a_b = \omega^a_b + \omega^a_c \wedge \omega^c_b \) the curvature two–form on \( M \). It is crucial that the first–order formulation of general relativity is much more adequate for gauge formulation than the standard second–order one. We shall now address the problem which the first–order approach entails.

Let (II.1) denote the action functional of a classical field theory in Minkowski space \( M \). It is well known that the transformation
\[
\mathcal{L} \rightarrow \mathcal{L}' = \mathcal{L} + \partial_{\mu} V^\mu
\]
(II.7)
of the Lagrangian density changes $\mathcal{L}$ by a differential. When introducing a new interaction, it seems reasonable to require that the resulting theory be independent on whether we have added a divergence to the initial Lagrangian density or not. Let us now specialize again to the Dirac field and consider the effect of the transformation (II.7) of the initial Lagrangian on the final Lagrangian four–form on $\mathcal{M}$. We shall consider the vector field of the form

$$V^\mu = a J^\mu_V + b J^\mu_A, \quad a, b \in \mathbb{C},$$

(II.8)

where $J^\mu_V = \bar{\psi} \gamma^\mu \psi$ and $J^\mu_A = \bar{\psi} \gamma^\mu \gamma^5 \psi$ are the Dirac vector and axial currents (this is the only possible form which is quadratic in $\psi$ and transforms as a vector under proper Lorentz transformations). It is straightforward to check that the following Leibniz rule applies

$$(D\bar{\psi}) C a \psi + \bar{\psi} C^a D\psi = d (\bar{\psi} C^a \psi) + \omega_{ab} (\bar{\psi} C^b \psi),$$

(II.9)

where $C^a := a\gamma^a + b\gamma^a\gamma^5$. Hence under the minimal coupling $d\psi \to D\psi$ the differential $dV^\mu$ of (II.8) will pass into $d\psi = dV^\mu + \omega_{ab} V^b$. Using the identity $\partial_\mu V^\rho d^4x = -\ast (d\mu) \wedge dV^\mu$ one can then conclude that the change in the resulting Lagrangian four–form on $\mathcal{M}$ (under the transformation (II.7) of the initial Lagrangian density) will be

$$\mathcal{L}' - \mathcal{L} = d (V, \varepsilon) - T_a V^a \varepsilon,$$

(II.10)

where $T^a = T^{ba} b$ is the torsion trace (the components of the torsion tensor in the tetrad basis are given by the equation $\tfrac{1}{2} T^a_{\mu c} e^b \wedge e^c = d e^a + \omega_{ab} \wedge e^b$) and $\bullet$ denotes the internal product. When deriving (II.10), it is necessary to use metrizity of $\omega$. Within the framework of classical general relativity, where the torsion of the connection is assumed to vanish, the result would be again a differential. In EC theory the torsion is determined by the gauge fields corresponding to the localization of the gravitational gauge group is $\mathbb{C}$ and the space $\text{ran}(\pi)$ is spanned by the matrices $\Sigma^{ab}$. The natural Lorentz invariant scalar product $\langle \phi, \psi \rangle_\rho = \phi^\dagger \gamma^0 \psi$ on $\mathbb{C}$ induces the product $\langle \gamma^0 X^1 \gamma^0 Y \rangle$ on $Lin(V)$. For any representation of the matrices $\gamma^a$ that is unitarily equivalent to the Dirac representation, the orthogonal complement is spanned by $1, \gamma^5, \gamma^a, \gamma^5 \gamma^a$. Hence we have

$$D\psi = D\psi + \mathbb{B} \psi,$$

$$D\psi = d\psi + \lambda \psi, \quad \mathbb{A} = -\frac{i}{4} \omega_{a\delta} \Sigma^{ab},$$

(III.1)

$$\mathbb{B} = \lambda 1 + \kappa \gamma^5 + \tau_a \gamma^a + \rho_a \gamma^a \gamma^5 \gamma^a,$$

where $\chi, \kappa, \tau_a, \rho_a$ are complex valued one–forms on space–time. We will require that the Leibniz rule hold for the Dirac vector and axial currents,

$$(D\bar{\psi}) \gamma^a \gamma^5 \psi + \bar{\psi} \gamma^a \gamma^5 D\psi = dJ^a_V + \omega^a b J^a_A,$$

$$(D\bar{\psi}) \gamma^a \gamma^5 \psi + \bar{\psi} \gamma^5 \gamma^a D\psi = dJ^a_A + \omega^a b J^a_A,$$

where $D\bar{\psi} := (D\psi) \gamma^5$ and $\omega^a b$ represents a modified connection on the RC space. Straightforward calculations show that these equations are satisfied if and only if

$$\omega^a b = \omega^a b + \lambda \delta^a b,$$

$$\mathbb{B} = \frac{1}{2} \lambda 1 + i \mu_1 1 + i \mu_2 \gamma^5,$$

where $\lambda := 2 \text{Re}(\chi), \mu_1 := \text{Im}(\chi), \mu_2 := \text{Im}(\kappa)$ are real–valued one–forms. Note that the one–forms $\mu_1$ and $\mu_2$ do not influence the resulting connection on the RC space. If non–gravitational interactions were included, the components of these one–forms could be hidden in the gauge fields corresponding to the localization of the global symmetry of the change of phase $\psi \to e^{i\alpha} \psi$ and the approximate symmetry under the chiral transformation $\psi \to e^{i\alpha} \gamma^5 \psi$. In order not to involve non–gravitational interactions, one needs to set $\mu_1$ and $\mu_2$ to zero.

According to the ideas presented at the beginning of this report, $\lambda$ should be determined by $\omega$ and $\epsilon$ in such a way that it is a scalar (compare (II.4) and the remarks concerning the dependence of $\mathbb{B}$ on $\Lambda$ and $\epsilon$). What is more, the procedure is expected to be free of the ambiguity. To see that all the requirements can be accomplished, note that the divergence $\partial_\mu V^\mu d^4x = -\ast (d\mu) \wedge dV^\mu$ will
pass into $\star (e_\lambda) \wedge (dV^a + \tilde{\omega}^a_b V^b)$. Hence, (II.10) imply that the procedure will yield unique solutions for generic $\nu$ if and only if $\nu = \mathbb{T}$, where $\mathbb{T} = T_e e^a$ is the torsion–trace–one–form, which is indeed a scalar under local Lorentz (or Poincaré) transformations $\omega \to \Lambda \omega \Lambda^{-1} - d\Lambda \Lambda^{-1}$, $e \to \Lambda e$. Hence, there exists precisely one coupling procedure which is free of the ambiguity and satisfies all the requirements.

From the perspective of the base manifold $\mathcal{M}$, it seems that the procedure could be stated briefly by saying that the modified connection $\tilde{\omega}^a_b = \omega^a_b + T\delta^a_b$ should be used in MCP, instead of the original metric connection $\omega^a_b$ entering $\mathcal{L}_G$. However, it would not be clear then how the new connection is to be implemented on spinors (the simple substitution $\omega \to \tilde{\omega}$ in $D\Psi$ would not work well). What is more, there are other possibilities of modifying the connection so that its application in MCP guarantees uniqueness. The simplest way to achieve this would be to subtract the contortion tensor. This would result in a Levi–Civita connection forming the formalism effectively to the second–order one. The torsion would entirely disappear from the theory. Less drastic possibility could be to retain only the antisymmetric part of the torsion tensor by adopting $\tilde{\omega}_{ab} = \omega_{ab} - \tfrac{1}{2}T_{[ab]}e^c$, where $\omega$ is the Levi–Civita part of $\omega$ and $T_{abc}$ the torsion of $\omega$. For the Dirac field, all such possibilities necessarily violate one of the assumptions supporting our approach (the two that were mentioned produce $B$ that does not take values in the orthogonal complement of $\text{Ran}(\pi)$ – this makes impossible reading out the connection $\omega$, that ought to be used in the construction of $\mathcal{L}_G$, from given $A = A + B(A, e)$. A different approach is possible, in which the corrected connection takes values in an extension of the original Lie algebra. One should specify what kind of extensions are allowed, how the original connection is to be retrieved from the extended one and to establish the dependence of Yang–Mills fields of the extension from those of the original theory. In the case discussed here, extending $so(1,3)$ by dilatations would work well. However, the details of such an abstract approach ought to be considered with care and this will not be done in this brief report.

The new connection $\tilde{\omega}$ on $\mathcal{M}$ is not metric. One could hope that $\omega$ could be obtained from $\tilde{\omega}$ as its metric part. This is however not the case. Let us recall that the coefficients $\Gamma^a_{bc}$ of any connection can be decomposed as

$$\Gamma^a_{bc} = \Gamma^a_{bc} + K^a_{bc} + L^a_{bc}, \quad (\text{III.2})$$

where $\Gamma^a_{bc}$ is the Levi–Civita part determined by the metric $g = \eta_{ab}e^a \otimes e^b$, $K_{abc} := \frac{1}{2}(T_{cab} + T_{bac} - T_{dbc})$ the contortion and $L_{abc} = -\frac{1}{2}((\nabla_a g_{bc} + \nabla_b g_{ac} - \nabla_c g_{ab})$. The metric part of $\Gamma_{abc}$ is therefore equal to $\Gamma_{abc} + \eta_{ab}T_c - \eta_{ac}T_b$, and not to $\Gamma_{abc}$.

**Acknowledgements**

I wish to thank the referee of PRD for bringing the solution based on anti–symmetric part of the torsion to my attention and for other important remarks that improved this report. I am also grateful to Wojciech Kamiński, Jerzy Lewandowski and Andrzej Trautman for helpful comments. This work was partially supported by the 2007–2010 research project N202 081 32/1844, the Foundation for Polish Science grant ”Master”.

---

[1] C. Yang and R. Mills, “Conservation of isotopic spin and isotopic gauge invariance”, Phys. Rev. 96, 191 (1954).
[2] R. Utiyama, “Invariant theoretical interpretation of interaction”, Phys. Rev. 101, 1597 (1956).
[3] T. Kibble, “Lorentz invariance and the gravitational field”, J. Math. Phys. 2, 212–221 (1961).
[4] M. Kazmierczak, “Nonuniqueness of gravity induced fermion interaction in the Einstein–Cartan theory”, Phys. Rev. D 78, 124025 (2008) [arXiv:0811.1332].
[5] M. Kazmierczak, “Einstein–Cartan gravity with Holst term and fermions”, Phys. Rev. D 79, 064029 (2009) [arXiv:0812.1298].
[6] A. Saa, “Propagating torsion from first principles”, Gen. Rel. Grav. 29, 205 (1997) [arXiv:gr-qc/9609011].
[7] A. Saa, “Volume–forms and minimal action principles in affine manifolds”, J. Geom. Phys. 15, 102 (1995) [arXiv:hep-th/9308087].
[8] T. Boyadjiev, P. Fiziev and S. Yazadjiev, “Neutron star in presence of torsion–dilaton field”, Class. Quant. Grav. 16, 2359 (1999) [arXiv:gr-qc/9803084].
[9] P. Fiziev and S. Yazadjiev, “Solar System Experiments and the Interpretation of Saa’s Model of Gravity with Propagating Torsion as a Theory with Variable Planck “Constant” ”, Mod. Phys. Lett. A14, 511 (1999) [arXiv:gr-qc/9807025].
[10] R. Mosna and A. Saa, “Volume elements and torsion”, J. Math. Phys. 46, 112502 (2005) [arXiv:gr-qc/0505146].
[11] F. Hehl and B. Data, “Nonlinear spinor equation and asymmetric connection in general relativity”, J. Math. Phys. 12, 1334 (1971).
[12] F. Hehl and P. von der Heyde, “Spin and the structure of space–time”, Ann. Inst. Henri Poincaré A19, 179 (1973).
[13] D. Kerlick, “Cosmology and particle pair production via gravitational spin–spin interaction in the Einstein–Cartan–Sciama–Kibble theory of gravity”, Phys. Rev. D 12, 3004 (1975).
[14] H. Rumpf, “Creation of Dirac Particles in General Relativity with Torsion and Electromagnetism I, II, III”, Gen. Rel. Grav. 10 509, 525, 647 (1979).
[15] A. Perez and C. Royelli, “Physical effects of the Immirzi parameter”, Phys. Rev. D 73, 044013 (2006) [arXiv:gr-qc/0505081].
[16] M. Kazmierczak, “On the choice of coupling procedure for the Poincaré gauge theory of gravity”, (2009) [arXiv:0902.4432].