Vol. 29 (1998) ACTA PHYSICA POLONICA B No 4

GAUGE THEORIES OF GRAVITY*

Y. Ne'eman †

Raymond and Beverly Sackler Faculty of Exact Sciences
Tel-Aviv University, Tel-Aviv, Israel 69978
and Center for Particle Physics, University of Texas
Austin, Texas 78712, USA

(Received November 7, 1997)

Dedicated to Andrzej Trautman in honour of his 64^{th} birthday

The relatively simple *Fibre-Bundle* geometry of a Yang–Mills gauge theory — mainly the clear distinction between base and fibre — made it possible, between 1953 and 1971, to construct a fully *quantized* version and prove that theory's *renormalizability*; moreover, *nonperturbative* (topological) solutions were subsequently found in both the fully symmetric and the spontaneously broken modes *(instantons, monopoles)*. Though originally constructed as a model formalism, it became in 1974 the mathematical mold holding the entire *Standard Model* (*i.e.* QCD and the Electroweak theory). On the other hand, between 1974 and 1984, Einstein's theory was shown to be perturbatively nonrenormalizable. Since 1974, the *search for Quantum Gravity* has therefore provided the main motivation for the construction of Gauge Theories of Gravity. Earlier, however, in 1958-76 several such attempts were initiated, for aesthetic or heuristic reasons, to provide a better understanding of the algebraic structure of GR. A third motivation has come from the interest in *Unification*, making it necessary to bring GR into a form compatible with an enlargement of the Standard Model. Models can be classified according to the relevant structure group in the fibre. Within the Poincaré group, this has been either the R^4 translations, or the Lorentz group $\mathrm{SL}(2,C)$ — or the entire Poincaré $\mathrm{SL}(2,C) \times R^4$. Enlarging the group has involved the use of the *Conformal* $\mathrm{SU}(2,2)$, the *special Affine* $\overline{\mathrm{SA}}(4,R) = \overline{\mathrm{SL}}(4,R) \times R^4$ or *Affine* $\overline{A}(4,R)$ groups. Supergroups have included *supersymmetry*, *i.e.* the graded-Poincaré group ($n = 1 \ldots 8$ in its extensions) or the *superconformal* $\mathrm{SU}(2,2/n)$. These *supergravity* theories have exploited the lessons of the aesthetic-heuristic models — Einstein–Cartan *etc.* — and also achieved the Unification target. Although perturbative renormalizability has been achieved in some models, whether they satisfy *unitarity* is not known. The nonperturbative Ashtekar program has exploited the understanding of instantons and self-dual solutions in QCD, in the complexification and in the selection of new variables. Note that supergravity involves Lie Derivatives as supertranlations, and several models have treated local spacetime translations similarly. The reduction of the larger groups, down to Poincaré, has involved spontaneous fibration and spontaneous symmetry breakdown. In this context, *noncommutative geometry* may allow for further geometrization.

PACS numbers: 11.15. –q, 04.50. +h

---

* Presented at the Workshop on Gauge Theories of Gravitation, Jadwisin, Poland, September 4–10, 1997.

† Wolfson Distinguished Chair in Theoretical Physics.

1. Historical background

This article constituted the opening "introductory" lecture at a workshop held in Jadwisin in September 1997 and honoring the 64th birthday of our distinguished colleague and friend, Andrzej Trautman. I noted in my lecture that 64 is 100 in numerical base 8, and that the choice of that base, motivated by Trautman's characterization of *The Spinorial Chessboard* (a spacetime feature) happened to resonate in my case, even though my *eightfold* way counted *internal* degrees of freedom. In some ways this similarity yet difference, this dichotomy, is also reflected in this title and in the subject matter of this meeting — the 'authentic' gauge theories having their *gauge group* acting on *internal degrees of freedom*, whereas in the case of gravity, the group acts on *spacetime*.

The first gauge field theory was introduced — in an attempt to merge Electrodynamics with General Relativity — by Hermann Weyl in 1919 (gauge group R^1, for scale [1], a noncompact spacetime feature), then withdrawn and reformulated in 1929 with the compact gauge group U(1) acting on the complex phase of the electron wave-function [2] (the model for all future *internal* degrees of freesom), after F. London had identified that feature. It played an important role in the construction of QED and particularly (in the form of the derived Ward–Takahashi identities) in the renormalization procedure. Note that Emmy Noether had meanwhile also published her two theorems, establishing the algebraic linkage between the gauge group and the conserved current in the first — and the actual coupling of the gauge potential to this current in the second. In 1953, Yang and Mills [3] generalized the gauge mechanism to SU(2), and thereby to any compact non-Abelian gauge group.

In this case — as in many others — physics and geometry developed (independently) along related lines and the physical gauge theory paralleled the emergence of *fiber bundles* as geometrical constructs, a fact which was only realized in the Sixties — Andrzej being one of the pioneers who made the connection. Weyl was both mathematician and physicist and it is not surprising that his gospel spanned both disciplines: Cartan, Chern, Eckmann, Ehresmann, Hirsch, Hopf, Lichnerowicz, Pontrjagin, Steenrod, Whitney, W.T. Wu are some of the names on the mathematical side. Physics has fully repaid that debt, first in 1984 when Sam Donaldson and Michael Friedman used the exact solutions of physical gauge theories to make serious advances in the classification of 4-manifolds, and again in 1994 when Nathan Seiberg and Edward Witten's solutions for supersymmetruc gauge theories were applied to the mathematical program.

It was only in the Sixties that some particle physicists — E. Lubkin, L. Susskind and others became aware of the mathematicians' efforts and results. Yang himself, with T.T. Wu, explored the relationship and wrote a "dictionary": gauge theory = fiber bundle, gauge group = fiber group, field potential = connection, field strength = curvature, *etc.* [4]. To any student interested in learning about fiber bundles and yet preserve a physicist's view, I particularly recommend Yang's contribution to the 1977 Marshak Festschrift "Fifty Years of Weak Interactions" [5].

In physics, though the Yang–Mills theory was originally a physical model with no direct application, by 1975 it had become the mold for the entire Standard Model — the gauge theory of $\mathrm{SU(3)}_{\mathrm{color}}$ for Quantum Chromodynamics (the "Strong Interactions") and the *spontaneously broken* gauge theory of $\mathrm{SU(2)} \times \mathrm{U(1)}$ for the Electroweak interactions. This was the consequence of a successful renormalization program, started by Feynman in 1958 [6], continued by B. DeWitt, A.A. Slavnov, J.C. Taylor, L.D. Faddeev and V.N. Popov, B.W. Lee and J. Zinn–Justin and others, completed by 't Hooft in 1971 [7], with added final touches by Becchi, Rouet and Stora ("BRS") [8]. Thierry–Mieg [9] provided in 1979-80 an elegant geometrical interpretation to the unitarity-guaranteeing BRS construction. Note that the reason Feynman started this program was the difficulty he was experiencing in his attempt to quantize and renormalize General Relativity. Feynman took up the Yang–Mills model as an easier pilot program for gravity... We see that Gauge theories and General Relativity were very close from the start and throughout their evolution.

The progression from Feynman to 't Hooft describes the acquisition of a *perturbative* solution. This was practically all that had been needed in QED. In non-Abelian QCD, however, this turned out to be good for the high energy (UV) sector (due to *asymptotic freedom*) but useless for the low-energy (IR) region. One answer to this problem was the discovery of *exact solutions*, both for the fully symmetric case (instantons [10]) and for the broken symmetry case (monopoles [11]).

In the first part of this article, I shall review the mathematical and physical characteristics of Gauge Theories in general. I shall then analyze the various motivations for the construction of Gauge Theories of Gravity and the possible algebraic routes, after which I shall discuss results.

2. Geometrical structure

A Principal Fiber Bundle is a manifold $\mathcal{P}(\mathcal{M}, \mathcal{G}, \pi, \bullet)$. $\mathcal{M}$ is the base manifold (generally flat spacetime), $\mathcal{G}$ is the gauge group, π the projection $\pi : \mathcal{P} \to \mathcal{M}$; $\bullet$ is the right-multiplication of $\mathcal{P}$ by $\mathcal{G}$. For $p, p' \in \mathcal{P}, g, g' \in \mathcal{G}$

we have a *verticality* condition, $\pi(p \bullet g) = \pi(p)$ (the group acts only on the fiber, staying above the same point in $\mathcal{M}$, *i.e.* there is a well-defined *vertical* direction) and *equivariance* $(p \bullet g) \bullet g' = p \bullet (gg')$., *i.e.* the group product is faithfully mapped. The latter (right-multiplication) is realized as a map from *the abstract infinitesimal Lie algebra* Λ of $\mathcal{G}$ onto the tangent manifold $\mathcal{P}_*$: $\forall \lambda \in \Lambda$, with $[\lambda_a, \lambda_b] = if_{a\ b}^{\ c}\lambda_c$, we have $t : \lambda \to \tilde{\lambda} \in \mathcal{P}_*$, the abstract Lie bracket becoming mapped into a differentiation bracket $[\tilde{\lambda_a}, \tilde{\lambda_b}]f = \tilde{\lambda_a}(\tilde{\lambda_b}f) - \tilde{\lambda_b}(\tilde{\lambda_a}f)$. Thus $[\lambda_a, \tilde{\lambda_b}] = [\tilde{\lambda}_a, \tilde{\lambda}_b]$ Note that if the dimensionalities are $dim\mathcal{M} = m, dim\mathcal{G} = k, dim\mathcal{P} = m + k$, the map t is from k onto $k + m$. To have an inverse, this map needs a structure which is aware of which is the correctly parallel-transported vertical direction at any point of $\mathcal{P}$. This is the *connection* $\omega : \mathcal{P}_* \to \Lambda, \forall \lambda \in \Lambda, \omega(\tilde{\lambda}) = \lambda$. The gauge-potential or connection is thus a one-form, acting by contraction, *i.e.* at a point p, $\omega^a(\tilde{\lambda}_b) = \omega^a \rfloor \tilde{\lambda}_b = \delta^a_b$.

Defining the field-strength (or curvature two-form) as $\Omega = d\omega + \frac{1}{2}[\omega, \omega]$, the BRS equations reexpress [12] the Cartan–Maurer structure equations stating the *horizontality* of the curvature, $\tilde{\lambda} \rfloor \Omega = 0$.

3. Physical characteristics

Given a "free" Quantum Field Theory Lagrangian for a matter field ψ, namely, $\mathcal{L}_\psi(\psi, d\psi)$, we obtain the full Lagrangian as $(\mathcal{L}_\psi)_{d\to D} + \mathcal{L}_{YM}(\Omega)$, with $D = d + \omega^a\lambda_a$ and $\mathcal{L}_{YM} = \Omega \wedge {}^*\Omega$. The Bianchi identity is $D\Omega = 0$ and the equation of motion is $D^*\Omega = {}^*j$. Infinitesimally, the gauge group acts on the fields via $\delta\psi^j = i\alpha^a(\lambda_a)^j_k\psi^k$. Note that the other invariant bilinear in the curvatures, $\Omega \wedge \Omega$ is a topological invariant and corresponds (up to a numerical factor) to the exact solutions (instantons) of the symmetric theory [10]. They turn out to be 4-divergences — here of the axial vector unitary singlet current, an important point in the physical hadron theory. One can also thus freely add them (with some factor θ) to the Lagrangian — except that they may violate P or CP ("the θ problem" in QCD).

This entails special features, all relating to the couplings:

(a) The couplings are *universal*, *i.e.* they are given (in ${}^*j(d \to D)$) for the connection ω^a by the matrix element of the corresponding algebraic generator, in the ψ^i representation of Λ, $c^j_{ai} =< \psi_i \mid \lambda_a \mid \psi^j >$. If for any symmetry the coupling is a Clebsch–Gordan coefficient, here it is a specific one, the a index specifying the adjoint representation of the algebra. For Abelian $\mathcal{G}$ this is a constant number, representing the large-distance value (Ze for electric charges, whatever the measured system); for non-Abelian $\mathcal{G}$ we have an extra factor, making it into a *running* coupling $g^j_{ai} = c^j_{ai}f(q^2)$ and with $f(q^2)$ possibly diverging at some "large" distance (around 1 fm, for

QCD). In gravity, the coupling is to the matrix-elements of the density of the generator of *translations* in an Abelian invariant subgroup of the Poincaré group, a matter related to the Equivalence Principle.

(b) The coupling is also dimensionless, due to the structure of the Lagrangian, quadratic in the curvatures — thus a four-form, whose integration over four-dimensional spacetime yields a dimensionless action, fitting the Quantum Postulate, with no need for the coupling to contribute dimensionalities. This feature is essential for a program of perturbative renormalization — otherwise one would need new counter terms at every order of the perturbative expansion.

(c) These couplings are to *conserved* Noether currents. *Renormalization effects are forbidden (for dimensionless couplings) if the relevant interactions obey the same symmetry* . This is how the CVC — conserved vector current nature of the Weak Interactions was first identified by Gershtein and Zeldovich, and later by Feynman and Gell–Mann, through the equality of the Weak vector couplings in neutron beta-decay and in muon decay. The Strong Interactions, which should have renormalized the coupling in neutron decay, respect the isospin symmetry generating the Weak current (up to the Cabibbo term). Similar algorithms were found for all Lie-group-generated symmetries of the hadrons (such as the Goldberger–Treiman relation, etc). They were especially important in 1958–1975 (the era when QFT was taboo) because they could be given the form of a dispersion relation [12]. Again, there is a similarity with the Equivalence Principle — in Gravity, the coupling is to a conserved current of the Poincaré group — a symmetry respected by all known interactions (and expressing itself in the Eotvos experiments' precision nonrenormalization results).

(d) The potential (or connection) can be gauged away by a local *active* gauge transformation (of the relevant internal degree of freedom). This feature (corresponding to an acceleration replacing a gravitational potential for gravity, *i.e.* the Equivalence Principle) entails local phase effects in QED or QCD, but becomes nominal for broken symmetries like the electroweak gauge.

Using the path-integral formulation for the quantization (v is the group volume, $\Delta\omega$ the measure in the space of connections),

$$Z = (v(\mathcal{G}))^{-1} \int [\Delta\omega] e^{i\int \mathcal{L} d^4x}$$

one may define canonical variables

$$[\omega^a_\mu(x), \tilde{E}^\nu_b(y)] = i\hbar \delta^a_b \delta^\nu_\mu \delta^3(\vec{x} - \vec{y}) \,,$$

where (r, s, t are space indices, 0 is time) the momentum $\tilde{E}^r_b = \frac{\delta \mathcal{L}}{\delta \Omega^b_{st}} = \varepsilon_r{}^{st} \Omega^{0r}_b$.

Weyl already introduced what we now know as the "Wilson loop" (a holonomy) $T_\omega := \mathrm{tr}\, Pe^{\oint \omega}$, a gauge invariant quantity, therefore a possible observable. Anticipating on the next sections, we note that it is also invariant under the diffeomorphisms of the integration loop, which is why it is useful in Ashtekar's canonical treatment of gravity. In QCD, the Wilson loop is the basic tool for computation, in what is now known as *lattice gauge theory* — the loop being selected around a rectangle. Using two spacelike and two timelike sides and making the timelike ones tend to infinity, we have Wilson's lattice proof of color confinement.

Note that in Hamiltonian quantization, the Bianchi identity becomes a constraint.

4. Gravity seen as a twisted and deformed gauge theory

The story here sounds like a Freudian Oedipus' (or Electra's) complex . Gravity was the mother of all gauge-like theories, with an interplay between *two* "gauge groups" (part of the mystery) — the *diffeomorphisms Diff*(R^4) ("the Principle of *Covariance*") and in addition, the *Lorentz* group $\mathrm{SL}(2, C)$ on the *local frames* (a factor which became more explicit after Dirac's equation for the electron and its inclusion in GR through *local frames* in 1928). Also, it is to the local Lorentz group we turn when we want to implement the Equivalence Principle and replace a potential by an acceleration.

Remember that it was because of Gravity's GR that the Weyl and the Yang–Mills gauge theories were born and that it was also because of gravity that Feynman launched the quantization program for the YM model; yet when we now go into details, we shall draw a picture showing gravity to be like a caricature of a gauge theory — thus also motivating the search for a different presentation (yet preserving the macroscopic predictions).

First — *covariance*. Is this really a gauge group? For one thing, *it does not have an active mode*. Example: a change of scale is a diffeomorphism, and GR is indeed *passively* invariant under such a transformation (*i.e.* changing the unit from centimeters to inches), but it is not invariant under an *active* physical invariance, such as a doubling of all distances. The forces would really weaken, whereas in Weyl's scale-invariant 1919 theory (or in Englert's modern version), they would not. One reason is that Newton's constant has dimensions. In Englert's theory [13], there is no such constant, it is replaced by a scalar field (whose vacuun expectation value happens to have that value, but could take any other).

Secondly, mathematically, *diffeomorphisms appear equivalent to "gauging the translations"*. Again, although this route has been explored by Cho and others, I do not consider this as a valid mode because the translations ∂_μ

are not covariant and we would not be able to perform *active* displacements with them. The covariant operators are the *covariant derivatives with frame indices* (here ω becomes Γ and Ω is R) $D_a = e^\mu_a(\partial_\mu - \Gamma^\rho_{\mu\sigma})$. However, the D_a, *under commutation* do not make a Lie algebra. This kind of translation is algebraically known as a *Lie derivative*. It was discussed in Ref. [14]. We called it an *anholonomized general coordinate transformation* (AGCT) in Ref. [15]. *In supergravity, the spinorial displacements consist in such Lie derivatives, as we showed in that work.* As a matter of fact, gauging the "modified Poincaré algebra", with translations replaced by the AGCT would be a conceptually clean answer, but this also means that the group we are "gauging" is not a Lie group with a Lie algebra. Its translations subalgebra has four generators — but *structure functions* instead of structure constants. As a result, even the variations of the gauge potentials are not the usual $\delta\omega^a = D\varepsilon^a$; instead, one has an additional piece $\varepsilon\rfloor R^a$. We shall return to this approach in the sequel [16].

The Principle of Covariance is thus not really a physical gauge principle, but it is certainly mathematically useful. Equivalence, on the other hand, has many of the attributes of a gauge theory (e.g. universality, a potential that can be gauged away) but no mathematical derivation. Our third point, indeed, is that the Lorentz subgroup $\mathrm{SL}(2, C) \subset \overline{\mathrm{Diff}}(R^4)$ (the overline denotes the double covering group) *is indeed actively implementable.* And yet the dynamical theory, as expressed — our points (a, b) — by the Noether content of the coupled conserved current *is not that of the Lorentz group.* On the contrary, the relevant current is the energy-momentum tensor, *i.e.* the density of the generators of *translations*, the quotient of the Poincaré group by that same Lorentz group! And yet in the implementation of our point (d), *i.e.* gauging away the potential, we do have to use the local Lorentz group!

The fourth point relates to the dimensionality of Newton's constant, the theory's coupling. Again, with a Lagrangian linear in the curvature, *i.e.* a two-form, we need to assign to the coupling a dimensionality of the inverse of a squared length. This will impact heavily on the perturbative renormalization program.

All of this is due to the Einstein–Hilbert Lagrangian, linear in the curvature. Whereas in the YM field equation $D^*\Omega = {}^*j$, the Ω and j relate to the same algebraic generators, in GR the Ω (now R) is the field-strength of rotations and the current j is that of the translations. In the Einstein–Cartan version, we have yet another cross-eyed equation relating the *torsion* — the field strength of translations — to the spin, *i.e.* the Lorentz group current.

We may gain some consolation from a 1977 demonstration by MacDowell and Mansouri [17]. They showed that if you undo the Wigner–Inönü contraction, replacing translations by rotations into the fifth dimension (the contraction consisting in having taken the radius to infinity) and write a "somewhat YM-like" 5th component $\varepsilon_{abcd5}R^{ab}\wedge R^{cd}$ of a topological $R\wedge R$, you will have an "aesthetic" understanding of our problem. The curvatures into the 'old' spacetime dimensions now have an extra term $R^{ab\prime} = R^{ab} - (1/2)\Gamma^{a5}\wedge\Gamma^{5b}$. When you now reimplementate the contraction, $\Gamma^{a5} \to e^a$ and the additional term is $\eta_{ab} = \varepsilon_{abcd}e^c \wedge e^d$. The quadratic Lagrangian yields a quadratic topological (instanton-like) invariant, a cosmological term coming from squaring the η_{ab} — and the Einstein Lagrangian $\eta_{ab}\wedge R^{ab}$ — which has thus acquired a 'respectable' progenitor (gauging the de Sitter group in an almost YM fashion..). The advantage of this presentation is that it can directly be extended to Supergravity — the contraction holding for OSp(1/4) (as $\overline{\mathrm{SO}}(3,2) = \mathrm{Sp}(4,R)$; alternatively, for the other de Sitter group $\overline{\mathrm{SO}}(4,1) = \mathrm{Sp}(2,2)$, we would thus have OSp(1/2,2)).

5. Motivations for a Gauge Theory of Gravity — conceptual simplicity

The first motivation to look for a more YM-like alternative was explorative, a search for *conceptual simplicity*, for a YM-like interpretation of GR. It started almost immediately after the YM paper, mainly in the contributions of Utiyama [18], Kibble [19] and Sciama [20]. One result was the renewed interest in *torsion*, appearing in all treatments based on the Poincaré group, since it is the field-strength of the translations. Here there was an encounter with the veterans of the 1920–1950 Einstein-stirred search for a unification betwen GR and Electromagnetism. Other than Weyl and his gauge approach, there had been Eddington, Cartan, Mme Tonnelat, Stueckelberg, Finkelstein, Rodichev, Ivanenko and collaborators, Pellegrini and Plebanski, *etc.* Torsion with its antisymmetric indices plays an important role in all this. Before we leave this subject, it is interesting to note that Einstein's quest found its tightest solution in N=2 Supergravity: between the $J=2$ graviton and the $J=1$ photon, two $J=3/2$ gravitinos (*i.e.* one charged complex field) are what was needed to construct an irreducible theory of GR plus EM [21]. However, when this was found, the stir was minimal. After Einstein's death, there were not many left who would disregard the Strong and the Weak interactions in an effort for unification.

This phase ended up in (1) a rebirth of the *Einstein–Cartan* theory, as a mild gauge-like facelift to GR, described in [15] — and (2) in the *Poincaré Gauge Theory*, quadratic in the curvatures and torsions [22]. This includes the option of a more drastic piece of surgery, namely *teleparallelism* [23], exploiting mostly the possibility of replacing the Einstein–Hilbert linear term by a quadratic squared-torsion term — the particular irreducible component known as the Weitzenböck invariant. Another conceptually valid version is the use of the “modified” Poincaré group — with Lie derivatives (AGCT) replacing the translations — providing an elegant *physical* interpretation — including an understanding of what in gravity is not simply gauge-like. This approach is particularly appropriate for treatments which include supergravity, since that is the nature of the relevant supersymmetric transformation [24]. Note that in supersymmetry a way was also found to return to an orthodox Lie superalgebra — by adding *auxiliary fields*.

Another interpretative result was the method of *gauging on the group manifold*. With Regge [15] and Thierry–Mieg [24], we showed that starting from two copies of the group manifold, taken both as base space and as fiber (the copy which becomes the base manifold is allowed to curve, the one in the fiber is rigid — as usual in a fiber bundle), but using a Lagrangian breaking the group symmetry (in the gravity case the group is the Poincaré group and the Lagrangian is locally only Lorentz-invariant) there occurs a process of *spontaneous compactification and factorization*. Thus, on mass shell, the fiber reduces to the Lorentz subgroup and the base manifold reduces to the space of translations, *i.e.* spacetime. This methodology works also for supergravity.

An alternative approach is to have both the Einstein linear term and terms quadratic in the curvatures — with different interpretations, including the possibility of new contact interactions [14] or a contribution to the Strong Interactions with confinement [25]. Two schools, Hehl's group in Cologne and Trautman, Kopczyński, Tafel *etc.*, in Warsaw, have led these movements, with Trautman [26] providing the most thorough analysis of a U(4) geometry as the Riemannian V_4 geometry is indeed replaced here by a U(4), *i.e.* with the inclusion of torsion.

The equations of motion now involve asymmetric tensors, closer to the canonical derivations (first Noether theorem) for the relevant currents. $G^{\mu\nu}$ is an asymmetric Einstein tensor, $T^{\mu\nu\rho}$ is the torsion, $\Sigma^{\mu\nu}$ the canonical energy-momentum current tensor, $\tau^{\mu\nu\rho}$ the spin angu;ar momentum current tensor. The equations of motion are,

$$\begin{aligned} G^{\mu\nu} &= \kappa\Sigma^{\mu\nu}\,, \\ T^{\mu\nu\rho} &= \kappa\tau^{\mu\nu\rho}\,, \end{aligned} \tag{1}$$

with the Noether currents, now as defined by their couplings (2nd Noether theorem),

$$\begin{aligned}
e\tau_\rho^{\mu\nu} &:= \frac{\delta\mathcal{L}}{\delta K_{\mu\nu}{}^\rho}\,, \\
\Sigma^{\mu\nu} &:= \sigma^{\mu\nu} - \tilde{\nabla}_\rho(\tau^{\mu\nu\rho} - \tau^{\nu\rho\mu} + \tau^{\rho\mu\nu})\,, \\
\sigma^{\mu\nu} &:= \frac{\delta\mathcal{L}}{\delta g_{\mu\nu}}\,, \\
\tilde{\nabla}_\mu &:= \nabla_\mu + 2S_{\mu\nu}^\nu\,, \\
S_{\mu\nu}{}^\rho &:= (1/2)(\Gamma_{\mu\nu}{}^\rho - \Gamma_{\nu\mu}{}^\rho)\,, \\
T_{\mu\nu}{}^\rho &:= S_{\mu\nu}{}^\rho + 2\delta^\rho_{[\mu}S_{\nu]\sigma}{}^\sigma\,, \\
K_{\mu\nu}^\rho &:= -S_{\mu\nu}{}^\rho + S_\nu{}^\rho{}_\mu - S^\rho{}_{\mu\nu}\,,
\end{aligned} \tag{2}$$

where we have also defined the symmetric energy current σ and the various torsion tensors T and S and the contortion K.

There are two possible direct routes to extend this formalism (still without invoking supersymmetry), according to whether one embeds the Poincaré group in the homothetic [1] and conformal [13] groups $\mathrm{SL}(2,C) \times R^{3,1} \subset \mathrm{SU}(2,2)$ (the homothetic is the middle step inwhich one only adds the R^1 of dilations) or in the affine [30] groups, $\mathrm{SL}(2,C) \times R^{3,1} \subset \mathrm{GL}(4,R) \times R^4$, with the special affine $\mathrm{SL}(4,R) \times R^4$ as middle step. Either route has been used and we shall return to affine geometry when we look at perturbative quantization, where it has been applied. The interest in these possibilities simply followed the search for a more general outlook [28]. There appeared to be an inherent difficulty in the affine case, in which the Lorentz SO(1,3) is replaced by the linear $\mathrm{SL}(4,R)$, because *it was (wrongly) assumed in the GRG community that, while* $\overline{SO}(1,3) = SL(2,C)$*i.e. there is a double covering to the Lorentz group — hence the existence of spinors — there is no double covering group for* $SL(4,R)$. I broke this superstition in 1978 [29] and constructed (infinite-component) linear and affine spinors [30] and even *world spinors* [31] on which the diffeomorphisms are represented nonlinearly over their linear subgroup, just as for tensors. Note that in Metric–Affine Gravity, we have new components to the curvature, deriving from the nonmetricity $Dg_{\mu\nu} = Q_{\mu\nu} \neq 0$. Starting from gauge-like considerations, Yang constructed such a model (coinciding with Stephenson's and Kilmister's, the SKY model [32]), but this model, taken macroscopically, does not approximate to Newtonian gravity.

6. Motivations for a Gauge Theory of Gravity: perturbative quantization

When the renormalization program for the Yang–Mills field achieved its goal in 1971, it was natural that Veltman and 't Hooft should turn to gravity [33]. The first answer was a nice surprise — considering the various reasons which predicted failure, mainly the dimensionality of the coupling.

It turned out that the one-loop vacuum contribution (gravitons interacting with gravitons) is *accidentally* finite, due to a topological identity. However, adding matter — scalar, spinor or vector — leads to infinities [34]. This line of theoretical experimentation was not pursued — had it, supergravity might have been discovered somewhat earlier, since N=1 supergravity can be regarded as gravity plus $J = 3/2$ "matter". Due to the supersymmetric algebra, as was later proved by Kallosh, anything true of gravitons can be generalized to the entire supersymmetric multiplet.

Hope for another set of miracles was finally dashed when Goroff and Sagnotti [35] using a supercomputer, managed to evaluate the two-loop vacuum diagram and found it to be infinite. Meanwhile, Stelle had shown [36] that a curvature-squared term in the Lagrangian would make it finite — its dimension frees it from the need for a dimensional coupling; moreover, it would also ensure that it dominate the linear term in the high energy regime. Such terms would certainly be generated in the renormalization procedure, even if the original Lagrangian were the linear one. However, Stelle also showed the theory to be non-unitary, due to its p^{-4} propagators. These propagators are created because of the Riemannian nature of the theory, *i.e.* the dependence of the connection on the metric (resulting from $Dg_{\mu\nu} = 0$). As a result, $\Gamma \sim \partial g$, $R \sim \partial^2 g + (\partial g)^2$ and $\mathcal{L} \sim (\partial^2 g)^2$, $(\partial g)^4$, $\partial^2 g(\partial g)^2$, all producing p^{-4} terms in the inverse Fourier transform. Such propagators can be simulated by a difference between two poles $p^{-4} \sim (1/p^2) - [1/(p^2 - m^2)]$, one of which has to be a ghost. Tomboulis [37] has recently provided a proof of the breakdown of unitarity at the nonperturbative level.

To cope with this issue, we have assumed [38] that the fundamental (high energy) theory is an affine or metric-affine gauge like model, such as the SKY Lagrangian. An appropriate Higgs field causes the local $\overline{\mathrm{SL}}(4, R)$ symmetry to break spontaneously (presumably at Planck energies), reducing to $\mathrm{SL}(2, C)$ and the low-energy theory is then Riemannian (and Einsteinian, for instance if we prepare a $\phi^2 R$ term — which is dimensionally OK — in the original Lagrangian). This theory was proven to be renormalizable [39,40] a la YM, but we nevertheless lack a proof of unitarity because (only) the gauge-fixing term still involves p^{-4}.

At this point, I am opening a parenthesis. In 1979, David Fairlie and I [41] indpendently conceived of an *internal supersymmetry* gauge model, the simple supergroup SU(2/1), whose even subgroup is SU(2) × U(1), which we identified with the Electroweak gauge group. The odd quotient of SU(2/1) by its even subgroup has the quantum numbers of the (complex) Higgs field, with $I_{\mathrm{weak}} = 1/2, Y_{\mathrm{weak}} = 1$. The supergroup constrains the 15–20 free parameters of the electroweak theory and predicts for the Higgs mass $m_H = 2m_W$, about 170 GeV. We had, however, problems with the interpretation, since the odd generators of supersymmetry are assumed to relate bosons to fermions and vice versa, whereas in this construction, the matter fields are $[\nu_L, e_L], e_R$ so that the odd generators only connect different chiralities. However, when taking the boson multiplet and working with forms, the action fits the statistics ansatz, since the W, Z, A vector mesons in the even part make one-forms, while the Higgs in the odd part are zero-forms. With J. Thierry–Mieg and S. Sternberg we worked hard throughout 1980–1990 to overcome the interpretative issue. Meanwhile, however, the mathematician D. Quillen published his *"theory of the superconnection"* [42] and we understood that this was what I had constructed, a superconnection. One should count together both parities, that of the Grassmann elements over which the supermatrix is valued, and that of the generating superalgebra. The even parts are the Grassmann-odd Yang–Mills connections and are superalgebraically even, the odd parts are Grassmann-even Higgs fields but they are superalgebraically odd. The overall parity is thus odd, as befits a (super) connection, still a connection. We reformulated the theory, therefore as a Quillen superconnection [43].

Meanwhile, another advance had happened in Mathematics, namely Connes' Noncommutative Geometry [44]. Connes and Lott applied it to the electroweak theory and reproduced the Weinberg–Salam model geometrically [45]. Soon afterwards, Coquereaux, Scheck and collaborators [46] showed that by modifying some steps in NCG, the same geometric derivation yields our superconnection!

Already in 1980, we had shown [47] that such supergroups (and their superconnections) reproduce the Higgs–Kibble model of spontaneous symmetry breakdown in other examples. At the recent Marcel Grossmann VIII in Jerusalem I presented [48] such a superconnection for the model we discussed here, namely a SKY affine Lagrangian whose $\overline{\mathrm{SL}}(4,R)$ is spontaneously broken to $\mathrm{SL}(2,C)$, *i.e.* Einsteinian gravity as the low energy theory of a fundamentally "post-Riemannian" affine high energy theory. The relevant supergroup is the double-covering of the simple (rank 3) $\mathrm{P}(4,R)$, whose even subgroup is $\mathrm{SL}(4,R)$

I recommend this extension of gauge theories over noncommutative geometry in any case, for its aesthetic characteristics [49].

7. Nonperturbative (canonical) quantization

In Section 3 we noted the existence of an exact (*i.e.* nonperturbtive) solution to the YM theory, namely the *instanton.*This is the Chern–Pontrjagin topological invariant, characterizing a bundle manifold, $\nu = \frac{1}{64\pi^2}\int \Omega \wedge \Omega$. This is a four-divergence $\nu = dG$, that of a Chern–Simons three-form $G \sim \mathrm{tr}\,\{\omega \wedge \Omega - (1/3)\omega \wedge \omega \wedge \omega)\}$. With matter fields, this is the term generating the chiral anomaly. As an invariant and a constant, it can be added to the

QCD Lagrangian, with some coefficient θ. It does, however, generate violations of parity P, because it is a pseudoscalar — due to the $\varepsilon^{\mu\nu\rho\sigma}$ of the four-form in Minkowski spacetime. In the physical term, this is compensated by the second epsilon tensor, that of the Hodge duality. Violating P, it also violates CP (and T). These problems can be resolved by assigning a small value to θ (experimentally, $\theta \leq 10^{-9}$) or inserting the imaginary unit i, but QCD would appear more natural if one would not have to make arbitrary assignments.

Once we have both terms in the Lagrangian, minimizing the action implies putting either one of $\Omega \pm i^*\Omega$ to zero, *i.e.* selecting self-dual or anti-self-dual fields and connections. The imaginary unit here is related to $^{**}f = -f$ and the overall result amounts also to complexification.

We now turn to gravity. In striving to achieve a nonperturbative quantization, Ashtekar [50] has emulated the YM model, using self-duality eigenfields as canonical variables, thus also enacting a complexification of the model. The Chern–Simons terms here are [51] $C_{rr} = -(\Gamma \wedge R + (1.3)\Gamma \wedge \Gamma \wedge \Gamma)$ for Lorentz curvatures and $C_{tt} = (1/k^2)(\vartheta_a \wedge d\vartheta^a - \vartheta_a \wedge \vartheta^b \wedge \Gamma^a_b)$. The "instanton term" here is the divergence dC, to be added to Einstein's Lagrangian. As a result one gets the constraints, the vector $X^a = D_i\tilde{E}^{ai}$, $(i, j = 1..3)$, corresponding to the Gauss constraint in the YM case, the three $X_i = F^a_{ij}\tilde{E}^j_a$ guaranteeing invariance under 3-dimensional diffeomorphisms — and the scalar (time-evolution) Hamiltonian constraint $X = \varepsilon^{abc}F_{ija}\tilde{E}^i_b\tilde{E}^j_c$.

Ashtekar's program has succeeded in generating quantum gravitational states, realized through the *loop quantization* as an observables' representation These are the Wilson loops we discussed in Section 3, as applied by Rovelli [52] and Smolin [53]. The program's difficulties are partly in our inexperience with the interpretation of nonperturbative states, where we cannot count quanta.

It is interesting that the success in gravity has already produced new applications in YM theory, in the search for a proof of color confinement in QCD. Some ten years ago, we conjectured [54] that QCD itself produces something resembling gravity, in its IR region, and that it is this gravity-like component which generates color confinement, a geometrical feature. We were indeed able to prove that such a component does exist [55,56], using semi-perturbative and algebraic methods. Independently, D.Z. Freedman, K. Johnson and several collaborators have been attacking the same problem nonperturbatively, applying variables inspired by Ashtekar's [57]. Let me mention that Seiberg and Witten have recently supplied a mathematical proof of confinement [58], but this depends crucially on the presence of supersymmetry, beyond $SU(3)_{color}$.

8. Unification as a motivation

With all other interactions quantized and sitting in a gauge theory mold, it seems obvious that gravity should be reformulated in a similar mold, if we are to end up with a unified theory. The drive for unification goes both ways — physicists working on internal degrees of freedom have always looked for a spacetime origin — with the Kaluza–Klein approach as the simplest unifying mechanism. Right after the advent of (flavor) SU(3) we already tried this route [59]. With SU(6), combining spacetime spin with "internal" unitary spin, the prospects appeared good at first for an algebraic unification. Such hopes were dashed by the various no-go theorems of the Sixties, culminating in the Coleman–Mandula (very negative) formulation. Hope was born again after the discovery of supersymmetry and the options left open by the Haag–Łopuszański–Sohnius theorems.

The theorems, however, did not restrict gauge groups, and the result was Supergravity. There was the additional hope that the new algebraic constraints would tame at least some of the divergences, as had happened in the Wess–Zumino model. Of the two original formulations of supergravity [60,61], the Freedman et al. dealt with gravity in the classical manner, whereas the Deser–Zumino presentation was influenced by the algebraically more elegant Einstein–Cartan formulation, as discussed by Kibble, for example (including the "first order" approach). Unification is maximal with the $N = 8$ model, which we were hoping, Gell-Mann and I, would also enjoy improved renormalizability, as indeed later happened with the related $N = 4$ Yang–Mills theory (which is simply finite, no radiative corrections at all). $N = 8$ supergravity was constructed in 1978 by Cremmer and Julia [62], as a 4-dimensional reduction of $N = 1$ in eleven dimensions. This approach calls for a Kaluza–Klein interpretation and there have been interesting leads for *spontaneous compactification*. Yet the answer has to await the verdict as to the theory's renormalizability — which is still not known. For some years (1984–1995) this model was abandoned, due to pessimistic evaluations of its chances.

Supergravity itself (but not the 11-dimensional model) is just the QFT (in 10 dimensions) obtained when truncating Superstring Theory beneath Planck energies — this was the view throughout the above period. Recently, this picture has veered again in the direction of the 11-dimensional model, which was shown to emerge from the truncation of a supermembrane in that dimensionality [63]. In the last two years, interest has grown enormously, with the discovery of *dualities* which relate all Superstring theories to this "M Theory" [64]. Should this be the answer, we could rest from our search for a gravitational gauge theory ... My own suggestion is to restrain the excitement at this stage and continue in our quest.

REFERENCES

[1] H. Weyl, *Sitz. Preuss. Akad. Wiss.* 465 (1918).

[2] H. Weyl, *Z. Phys.* **56**, 330 (1929).

[3] C.N. Yang, R.L. Mills, *Phys. Rev.* **95**, 631 (1954) and **96**, 191 (1954).

[4] T.T. Wu, C.N. Yang, *Phys. Rev.* **D12**, 3845 (1975).

[5] C.N.Yang, in *Ann. N. Y. Acad. Sci.* (R.E. Marshak Festschrift "Fifty Years of Weak Interactions") **274** (1977).

[6] R.P. Feynman, *Acta Phys. Pol.* **24**, 647 (1963); B.S. DeWitt, *Phys. Rev.* **162**, 1195 (1967).

[7] G 't Hooft, *Nucl. Phys.* **B33**, 173 (1971); **B35**, 167 (1971).

[8] C. Becchi, A. Rouet, R. Stora, *Commun. Math. Phys.* 127 (1975).

[9] J. Thierry–Mieg, *J. Math. Phys.* **21**, 2834 (1980).

[10] A.M. Polyakov, *Phys. Lett.* **B59**, 82 (1975).

[11] G. 't Hooft, *Nucl. Phys.* **B79**, 276 (1974); A.M. Polyakov, *JETP Lett.* **20**, 194 (1974).

[12] Y. Ne'eman, Proc. Uppsala 1973 5th Int. Conf. High En. and Nucl. Struc., J. Tibell, ed., Almquist and Wiksell Pub., Stockholm 1974, p. 10-21.

[13] F. Englert *et al.*, *Phys. Lett.* **B57**, 73 (1975).

[14] F.W. Hehl, P.v.d. Heyde, G.D. Kerlick, D. Nester, *Rev. Mod. Phys.* **48**, 393 (1976).

[15] Y. Ne'eman, T. Regge, *Phys. Lett.* **B74**, 54 (1978); *Riv. Nuovo Cim.* (series 3) **1** #5, 1 (1978).

[16] Y. Ne'eman, in *Differential Geometrical Methods in Mathematical Physics*, Springer series *Lecture Notes in Mathematics* **676** (1979), p. 189.

[17] S.W. MacDowell, F. Mansouri, *Phys. Rev. Lett.* **38**, 739, 1376 (E)(1977).

[18] R. Utiyama, *Phys. Rev.* **101**, 1597 (1956).

[19] T.W.B. Kibble, *J. Math. Phys.* **2**, 212 (1961).

[20] D.W. Sciama, in *Recent Developments in General Relativity*, Pergamon Press and Oxford, p. 415 ff.

[21] S. Ferrara, J. Scherk, P.v. Nieuwenhuizen, *Phys. Rev. Lett.* **37**, 1035 (1976); S. Ferrara, P.v. Nieuwenhuizen, *Phys. Rev. Lett.* **37**, 1669 (1976).

[22] F.W. Hehl, in *Cosmology and Gravitation: Spin , Torsion, Rotation and Supergravity*, Erice 1979, P. Bergmann and V. de Sabbata eds., Plenum P., New York 1980, p. 5.

[23] A. Einstein, *Sitz. Preuss. Akad. Wiss. (Berlin), Phys. -Math. Kl.* 217, 224 (1928); R. Weitzenböck, *Sitz. Preuss. Akad. Wiss. (Berlin), Phys. -Math. Kl.* 466 (1928).

[24] J. Thierry–Mieg, Y. Ne'eman, *Ann. Phys.* (NY) **123**, 247 (1979).

[25] F.W. Hehl, Y. Ne'eman, J. Nitsch, P. v.d. Heyde, *Phys. Lett.* **B78**, 102 (1978).

[26] A. Trautman, *Bull. Acad. Pol. Sci., ser. Sci. Math. Astr. Phys.* **20**, 185, 503, 895 (1972); **21**, 345 (1973).

[27] F.W. Hehl, J.D. McCrea, E.W. Mielke, Y. Ne'eman, *Phys. Rep.* **258**, 1 (1995).

[28] F.W. Hehl, G.D. Kerlick, P. v.d. Heyde, *Phys. Lett.* **B63**, 446 (1976).

[29] Y. Ne'eman, *Ann. Inst. H. Poincaré* **28**, 369 (1978).

[30] Y. Ne'eman, Dj. Šijački, *Int. J. Mod. Phys.* **A2**, 1665 (1987).

[31] Y. Ne'eman, Dj. Šijački, *Phys. Lett.* **B157**, 275 (1985); **B170**, 431 (E) (1985).

[32] G. Stephenson, *Nuovo Cim.* **9**, 263 (1958); C.W. Kilmister, D.J. Newman, *Proc. Cam. Phil. Soc.* **57**, 851 (1961); C.N. Yang, *Phys. Rev. Lett.* **33**, 445 (1974).

[33] G. t' Hooft, M. Veltman, *Ann. I.H. Poincaré* **20**, 69 (1974).

[34] S. Deser, P.v. Nieuwenhuizen, *Phys. Rev.* **D10**, 401, 410 (1974); same authors and H.S. Tsao, *Phys. Rev.* **D10**, 3337 (1974).

[35] M.H. Goroff, A. Sagnotti, *Phys. Lett.* **B160**, 81 (1985).

[36] K.S. Stelle, *Phys. Rev.* **D16**, 953 (1977); *Gen. Relativ. Gravitation* 353 (1978).

[37] E.T. Tomboulis, *Phys. Lett.* **B389**, 225 (1996).

[38] Y. Ne'eman, Dj. Šijački, *Phys. Lett.* **B200**, 489 (1988).

[39] Y. Ne'eman, C.Y. Lee, *Phys. Lett.* **B242**, 59 (1990).

[40] C.Y. Lee, *Classical Quantum Gravity* **9**, 2001 (1992).

[41] Y. Ne'eman, *Phys. Lett.* **B81**, 190 (1979); D.B. Fairlie, *Phys. Lett.* **B82**, 97 (1979).

[42] D. Quillen, *Topology* **24**, 89 (1985).

[43] S. Sternberg, Y. Ne'eman, *Proc. Nat. Acad. Sci. USA* **87**, 7875 (1990).

[44] A. Connes, *Pub. Math. IHES* **62**, 44 (1983).

[45] A. Connes, J. Lott, *Nucl. Phys. (Proc. Suppl.)* **B18**, 29 (1990).

[46] R. Coquereaux, G. Esposito-Farese, F. Scheck, *Int. J. Mod. Phys.* **A7**, 6555 (1992).

[47] Y. Ne'eman, J. Thierry–Mieg, *Proc. Nat. Acad. Sci. USA* **77**, 720 (1980).

[48] Y. Ne'eman, Proc. VIIIth Marcel Grossmann Conf. (Jerusalem, 1997), to be published.

[49] Y. Ne'eman, in *Group Theory and its Applications (30th Latin-Am. Sch. of Phys., Mexico C., 1995)*, O. Castaños et al. eds., AIP Proc. **365**, Woodbury, NY (1996), pp. 311.

[50] A. Ashtekar, *Phys. Rev. Lett.* **57**, 2244 (1986); *Phys. Rev.* **D36**, 1587 (1987).

[51] E.W. Mielke, *Ann. Phys.* (NY) **219**, 78 (1992).

[52] C. Rovelli, L. Smolin, *Phys. Rev. Lett.* **61**, 1155 (1988).

[53] Y. Ne'eman, Dj. Šijački, *Phys. Lett.* **B247**, 571 (1990).

[54] Y. Ne'eman, Dj. Šijački, *Phys. Lett.* **B270**, 173 (1992).

[55] Y. Ne'eman, Dj. Šijački, *Int. J. Mod. Phys.* **A10**, 4399 (1995).

[56] Y. Ne'eman, Dj. Šijački, *Mod. Phys. Lett.* **A11**, 217 (1996).

[57] D.Z. Feedman, P.E. Haagensen, K Johnson, J.I. Latorre, CERN-TH 7010/93, unpub. See also D.Z. Freedman, R.R. Khuri, *Phys. Lett.* **B329**, 263 (1994); P.E. Haagensen, K. Johnson, *Nucl. Phys.* **B439**, 597 (1995).

[58] N. Seiberg, E. Witten, *Nucl. Phys.* **B426**, 19 (1994) and **B431**, 484 (1994).

[59] Y. Ne'eman, *Rev. Mod. Phys.* **37**, 227 (1965).

[60] D.Z. Freedman, P. v. Nieuwenhuizen, S. Ferrara, *Phys. Rev.* **D13**, 3214 (1976).

[61] S. Deser, B. Zumino, *Phys. Lett.* **B62**, 335 (1976).

[62] E. Cremmer, B. Julia, J. Scherk, *Phys. Lett.* **B76**, 409 (1978).

[63] E. Bergshoeff, E. Sezgin, P.K. Townsend, *Phys. Lett.* **B189**, 75 (1987).

[64] See for example J.H. Schwarz, *Phys. Lett.* **B367**, 97 (1996).