Gauge and optical aspects of gravitation

Andrzej Trautman
Instytut Fizyki Teoretycznej, Uniwersytet Warszawski, Hoża 69, 00681 Warszawa, Poland
E-mail: `amt@fuw.edu.pl`

Received 14 April 1999

Abstract. This is a rather personal review of a few fields of research in which the author has been involved and that he believes to be of relevance in the future. A brief description of the idea of *gauge* invariance and symmetry breaking is followed by a review of theories of the Kaluza-Klein type and of the Einstein-Cartan theory gravitation with spin and torsion. It is shown that the early work of Bateman can be considered to provide a basis for an *optical* geometry in Lorentzian manifolds with shear-free congruences of null geodesics, a notion introduced by Robinson and closely connected to that of algebraically special gravitational fields. There is a natural, one-to-one correspondence between the set of such optical geometries and that of Cauchy-Riemann spaces. A few odd remarks are devoted to the problem of 'large numbers', an EIH problem, variational principles, and elementary links between gravitation and quantum physics.

PACS numbers: 04.20.q, 04.50.+h, 01.30.Rr

Reprinted from *Classical and Quantum Gravity* **16** (1999) A157–A175.

1. Introduction; personal reminiscences

Instead of providing a formal introduction, I present a few recollections of my first ten years of contacts with General Relativity and its People. These remarks may explain the choice of topics in this review. In 1955, I started postgraduate work in Leopold Infeld's group in Warsaw; I learned much from Jerzy Plebański and by reading the books by Pauli and Landau&Lifshitz. In 1957, Felix Pirani came to Warsaw (he asked: 'What is the physical significance of Petrov type III?') and invited me to Hermann Bondi's group at King's College. During that first visit to London, I gave a series of lectures (Trautman 1958) and went for a couple of days to Cambridge, at the invitation of Dennis Sciama ('What is your opinion of Mach's Principle?... So you will be put up in the Judge's Room at Trinity'). At tea, I met there P. A. M. Dirac ('Professor Dirac, do you think that harmonic coordinate systems have a physical significance?' A long silence. 'I think they have no such significance'. End of conversation). Dennis introduced me also to Roger Penrose. One day, Ivor Robinson visited King's, explained to me the idea of shear-free congruences of null geodesics and proposed that we search for new, radiative solutions of Einstein's equations (Robinson and Trautman 1960). Royaumont 1959 (GRG2) was the first major conference that I attended; I met there many people who also exerted a great influence on my research: Peter Bergmann (his work on Kaluza-Klein and generally-covariant theories), Jürgen Ehlers (contributions to the work of the Jordan group at Hamburg; Jordan *et al* 1960 and 1961), V. A. Fock [Fo], Josh Goldberg (at that time: a proper treatment of gauge invariance in the EIH approximation scheme, Goldberg 1955; later: the Goldberg-Sachs theorem), André Lichnerowicz [L] (he advised me to read the work of Élie Cartan), Ted Newman (his subsequent work on exact solutions and the NP formalism: Newman and Penrose 1962, Newman *et al* 1963 and Newman 1974), Engelbert Schücking (work on cosmology, Heckmann and Schücking 1959, Ozsváth and Schücking in [In]), John Synge (general relativity as a geometric theory [S]), and John Wheeler (general relativity as a physical theory [MTW]). The academic year 1959-60 I spent at Imperial College in the group of Abdus Salam ('in physics, the Poincaré group is more important than the Lorentz group'). T. W. B. Kibble was then completing his work on gravitation with torsion as a gauge theory (Kibble 1961); during that year, I interacted also with Alfred Schild, Ray Sachs and Michel Cahen at King's College. In the spring of 1961, in Peter Bergmann's group at Syracuse University, I watched with fascination how Newman and Penrose developed their spin coefficient method; Ivor and I completed our work on spherical gravitational waves and I wrote a review on conservation laws in general relativity (article in [W]). GRG3 took place in 1962 in Jabłonna near Warsaw, with lectures and seminars given by Bergmann, Bondi, B. S. DeWitt, Dirac, Feynman, Fock, Ginzburg, Lichnerowicz, Mandelstam, Møller, Newman, Penrose, Robinson&Trautman, Schiff, Synge, Wheeler and others [J]. Subrahmanyan Chandrasekhar who attended the meeting, but did not lecture there, later became a friend. After the conference, John Stachel stayed for several weeks in Warsaw to help in editing the proceedings. Feynman criticized, in strong terms, that conference and research on gravitation at the time [Fe]. In a letter to his wife he wrote: 'Because there are no experiments this field is not an active one, so few of the best men are doing work on it... Remind me not to come to any more gravity conferences!' Hawking (Lecture 1 in [HaP]) quotes Feynman's remarks to emphasize the essential progress made, since the early 60s, in the development of general relativity as part of physics. In 1965 I gave a talk at the GRG4 conference in London and met there, for the first time, Stephen Hawking, István Ozsváth, and Kip Thorne. Some of the main lectures at GRG4 were given by Chandrasekhar, Ehlers, Fierz, Novikov&Zel'dovich and Taub; Khalatnikov and Penrose conducted a historic dispute on the genericness of singularities in cosmology.

1.1. Notation and terminology

The gravitational constant and the Planck length are denoted by $\mathtt{G}$ and ℓ, respectively. Occasionally, 'general-relativistic units' are used; they are such that both $\mathtt{G}$ and c are 1 and $\hbar = \ell^2$. Most of the time, I follow the notation and terminology standard in differential geometry and general relativity [MTW]. Greek indices range from 0 to 3 and refer to space-time. Given a co-frame (a linear basis in the space of 1-forms) θ^μ, one writes the metric tensor as $g = g_{\mu\nu}\theta^\mu\theta^\nu$. The exterior product of forms is denoted with a wedge. The Levi-Civita symbol is $\epsilon_{\mu\nu\rho\sigma}$ and $\eta_{\mu\nu\rho\sigma} = \sqrt{-\det(g_{\alpha\beta})}\,\epsilon_{\mu\nu\rho\sigma}$. One defines the forms

$$\eta_{\mu\nu\rho} = \theta^\sigma\eta_{\mu\nu\rho\sigma}, \quad \eta_{\mu\nu} = \tfrac{1}{2}\theta^\rho \wedge \eta_{\mu\nu\rho}, \quad \eta_\mu = \tfrac{1}{3}\theta^\nu \wedge \eta_{\mu\nu}, \quad \eta = \tfrac{1}{4}\theta^\mu \wedge \eta_\mu. \tag{1}$$

The Hodge dual is denoted with a star.

The Killing form k on a Lie algebra $\mathfrak{g}$ is defined by $k(a,b) = \mathrm{tr}(\mathrm{ad}\, a \circ \mathrm{ad}\, b)$, where $(\mathrm{ad}\, a)b = [a,b]$ and $a, b \in \mathfrak{g}$. Given a linear basis (e_i) in $\mathfrak{g}$, one introduces the structure constants c^k_{ij} by $[e_i, e_j] = c^k_{ij}e_k$, where $i, j, k = 1, \ldots, \dim G$; putting $k_{ij} = k(e_i, e_j)$, one has $k_{ij} = c^l_{ik}c^k_{jl}$. The dual of $\mathfrak{g}$ is denoted by $\mathfrak{g}^*$.

All manifolds and maps are assumed to be smooth. If $f : M \to N$ is a map of manifolds, then $Tf : TM \to TN$ is the tangent (derived) map of the corresponding tangent bundles. If φ is a differential n-form on N, then $f^*\varphi$ is its pull-back to M; in particular, for $n = 0$ one has $f^*\varphi = \varphi \circ f$.

2. Gauge ideas connected with gravitation

The idea of gauge invariance has its roots in Weyl's attempt at unification of gravitation and electromagnetism; it later underwent profound changes and generalizations, leading to the present notion of theories of the Yang-Mills type. The question of whether, and in what sense, relativistic gravitation can be regarded as a gauge theory has been the subject of some interest and controversy. In particular, the following issues have been often discussed: what is the 'gauge group' of gravitation (Lorentz, Poincaré, affine or perhaps the group of all diffeomorphisms?); the Maxwell and Yang-Mills theories are based on Lagrangians quadratic in the field strengths: should not the same be expected of a theory gravitation? (see Hehl *et al* 1995 and Ne'eman 1998 for recent reviews and further references).

2.1. Generalities on gauge theories

The kinematic framework of a classical gauge theory is rather simple: it is well described by the mathematicians' notion of a principal bundle with a connection and some additional structure that depends on, and defines the physics of, the particular theory. Wu and Yang (1975) gave a dictionary to translate between the mathematical and physical languages. One can supplement it by a remark on two meanings of the expression *gauge group* in physics and by a geometric interpretation of symmetry breaking; this is described below.

2.1.1. A principal bundle $G \to P \xrightarrow{\pi} M$ has a Lie *structure group* G that acts freely and transitively on the fibres of π: there is a map $P \times G \to P$, $(p, a) \mapsto \delta(a)p$ such that $\delta(a) \circ \delta(b) = \delta(ba)$, $\delta(1_G) = \mathrm{id}_P$, and $\pi \circ \delta(a) = \pi$. Let $\exp : \mathfrak{g} \to G$ be the exponential map from the Lie algebra $\mathfrak{g}$ of G into the group. For every $u \in \mathfrak{g}$, the vector $\bar{u}(p)$ tangent to the curve $t \mapsto \delta(\exp tu)p$ at $p \in P$ is *vertical*, $T_p\pi(\bar{u}(p)) = 0$ and the map $u \mapsto \bar{u}(p)$ is an isomorphism of $\mathfrak{g}$ onto the vertical vector space at $p \in P$. Given a representation ρ of G in a (real or complex) vector space V, one says that a V-valued n-form φ on P is of *type* ρ if it satisfies the *transformation law* $\delta(a)^*\varphi = \rho(a^{-1})\varphi$ for every $a \in G$. This form is said to be *horizontal* if the contraction of every vertical vector with φ is 0. In most practical computations, one considers local sections of π and uses them to pull-back forms from P to the base M.

2.1.2. Let $\aleph$ be the 'absolute' (non-dynamical) elements underlying the theory. The *group of gauge transformations* $\mathcal{G}$ consists of all bundle automorphisms of π that preserve $\aleph$; e.g. in special-relativistic electrodynamics one has $P = \mathbb{R}^4 \times \mathsf{U}_1$ whereas $\aleph$ is the Minkowski metric tensor and $\mathcal{G}$ is the semi-direct product of the Poincaré group by the group $\mathcal{G}_0 = \{f : \mathbb{R}^4 \to \mathsf{U}_1\}$ of 'gauge transformations of the second kind'. If a *general-relativistic theory of gravitation* is understood to be a theory of a connection on a principal $\mathsf{GL}_4(\mathbb{R})$-bundle P over a 4-manifold M, supplemented by additional dynamical fields (such as the metric tensor) and suitable invariant field equations, then the absolute is the *canonical* ('soldering') 1-form $\theta : TP \to \mathbb{R}^4$, of type id, such that the kernel of $\theta(p)$ is the vertical vector space at $p \in P$. The bundle P with such a θ is isomorphic to the bundle LM of all linear frames of M. The group $\mathcal{G}$ consists of all diffeomorphisms of the 4-dimensional space-time manifold M.

2.1.3. A *connection* on the principal bundle is described by a 1-form $\omega : TP \to \mathfrak{g}$ of type given by the adjoint representation Ad of G in $\mathfrak{g}$ and such that $\omega(\bar{u}(p)) = u$ for every $p \in P$ and $u \in \mathfrak{g}$. Given a local section s of π, the pull-back $s^*\omega$ is a $\mathfrak{g}$-valued 1-form on M; depending on the context, it is given the name of *potential in gauge* s, or referred to as the *Christoffel symbol* (when s is a 'natural' or holonomic frame) or the *Ricci rotation coefficient* (when s is a field of orthonormal frames). Given a representation $\rho : G \to \mathsf{GL}(V)$, let $\rho' : \mathfrak{g} \to \mathsf{End}(V)$ be the corresponding derived representation, $\rho'(u) = \frac{\mathrm{d}}{\mathrm{d}t}\rho(\exp tu)|_{t=0}$. If φ is a horizontal n-form of type ρ, then its covariant derivative $D\varphi = \mathrm{d}\varphi + \rho'(\omega) \wedge \varphi$ is a horizontal $(n+1)$-form of the same type. The *curvature* $\Omega = \mathrm{d}\omega + \frac{1}{2}[\omega, \omega]$ is a horizontal 2-form of type Ad; it satisfies the *Bianchi identity* $D\Omega = 0$. Given a basis (e_k) in $\mathfrak{g}$, one writes $\omega = \omega^k e_k$ and $\Omega = \Omega^k e_k$ so that $\Omega^k = \mathrm{d}\omega^k + \frac{1}{2}c^k_{lm}\omega^l \wedge \omega^m$.

2.1.4. Consider now a 0-form (scalar) φ of type ρ with values in an *orbit* W of the action of G in V. Given any point $\varphi_0 \in W$, one defines its stability ('little') group $H = \{a \in G | \rho(a)\varphi_0 = \varphi_0\}$ and the set $Q = \{p \in P | \varphi(p) = \varphi_0\}$ which is a principal H-bundle over M; it is called, by the mathematicians, a *reduction* of P to the subgroup H of G; in physics one says that φ *breaks the symmetry* from G to H. The equation $D\varphi = 0$ provides a necessary and sufficient condition for the restriction of ω to Q to have values in the Lie algebra $\mathfrak{h}$ of H and, therefore, to define a connection on this reduced bundle. A more general case of interest is when the Lie subalgebra $\mathfrak{h}$ admits a complement $\mathfrak{k}$ in $\mathfrak{g}$ such that, for every $a \in H$ and $u \in \mathfrak{k}$, one has $\mathrm{Ad}(a)u \in \mathfrak{k}$: the restriction ρ of the adjoint representation to H defines a representation of that subgroup in the vector space $\mathfrak{k}$. For example, if the group G is semi-simple, i.e. when its Killing form k is non-degenerate and, moreover, if the restriction of k to $\mathfrak{h}$ is also non-degenerate, then one can take $\mathfrak{k}$ to be the vector subspace of $\mathfrak{g}$, orthogonal to $\mathfrak{h}$ with respect to k.

When the restriction of ω to Q is split into its components γ and χ with values in $\mathfrak{h}$ and $\mathfrak{k}$, respectively, then one finds γ to be a connection on the reduced bundle and χ to be a horizontal 1-form of type ρ. In theories of the Yang-Mills type, such as the Weinberg-Salam theory of electro-weak forces, γ corresponds to massless vector bosons and χ describes spin 1 particles that acquire mass through the process of spontaneous symmetry breaking induced by the *Higgs field* φ, a 0-form of type Ad related to a (yet to-be-discovered) massive particle. There is a description, in this language, of the generation of mass of the χ field (Kerbrat *et al* 1989).

2.1.5. In the case of a general-relativistic theory of gravitation, the connection form has values in the Lie algebra $\mathfrak{g} = \mathsf{End}\,\mathbb{R}^4$, $\omega = (\omega^\mu{}_\nu)$, and the curvature is

$$\Omega^\mu{}_\nu = \mathrm{d}\omega^\mu{}_\nu + \omega^\mu{}_\rho \wedge \omega^\rho{}_\nu = \tfrac{1}{2} R^\mu{}_{\nu\rho\sigma}\theta^\rho \wedge \theta^\sigma.$$

The covariant derivative of the canonical form is the *torsion* 2-form,

$$\Theta^\mu = \mathrm{d}\theta^\mu + \omega^\mu{}_\nu \wedge \theta^\nu = \tfrac{1}{2} Q^\mu{}_{\nu\rho}\theta^\nu \wedge \theta^\rho,$$

and the Bianchi identity for torsion is $D\Theta^\mu = \Omega^\mu{}_\nu \wedge \theta^\nu$. Let V be now the vector space of all quadratic forms over $\mathbb{R}^4$; the metric tensor can be considered as a 0-form $g : LM \to V$ of type given by the transformation law $g_{\mu\nu}(ea) = g_{\rho\sigma}(e)a^\rho{}_\mu a^\sigma{}_\nu$, where $e \in LM$ and $a = (a^\mu{}_\nu) \in \mathsf{GL}(\mathbb{R}^4)$. The subset W of V consisting of all quadratic forms of a given signature is, by the theorem on inertia of quadratic forms, an orbit of $\mathsf{GL}(\mathbb{R}^4)$ for the action defined by the transformation law. In particular, if the signature is $(1, 3)$, then the (Lorentzian) metric tensor g breaks the symmetry from $G = \mathsf{GL}(\mathbb{R}^4)$ to the Lorentz group $H = \mathsf{O}_{1,3}$; the reduced bundle Q consists of all linear frames on M orthonormal with respect to g. The Lie algebra $\mathfrak{g} = \mathsf{End}(\mathbb{R}^4)$ admits a decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{k}$, where $\mathfrak{h} = \bigwedge^2 \mathbb{R}^4$ is the Lie algebra of the Lorentz group and $\mathfrak{k}$ is the vector space of all real, symmetric 4 by 4 matrices. In the decomposition of a linear connection ω restricted to Q, the form γ is a metric, but not necessarily symmetric, connection, and χ corresponds to a tensor field, $\chi_{\mu\nu\rho} = \chi_{\nu\mu\rho}$. This field appears in the 'metric-affine' theories of gravitation considered by Hehl *et al* (1995).

Two modifications of Einstein's relativistic theory of gravitation appeared already in the 1920s; they have been the object of some interest and development also during the second half of the XXth century. They both are related to gauge ideas and may present some interest in the future; for these reasons they are briefly reviewed here.

2.2. The Kaluza-Klein theory

In the 1920s Kaluza and Klein proposed to unify gravitation and electromagnetism by using the geometry of a five-dimensional Riemannian manifold with a one-parameter group of isometries. After the advent of Yang-Mills theory, it became clear that the Kaluza-Klein idea can be extended to any gauge theory of this type (DeWitt 1964). The geometry of such a (generalized) *Kaluza-Klein theory* can be briefly described as follows (Kopczyński 1980):

2.2.1. One considers a principal G-bundle P over a four-dimensional space-time (M, g) so that $\dim P = \dim G + 4$. The bundle $\pi : P \to M$ has a connection given by a 1-form ω with values in the Lie algebra $\mathfrak{g}$ of G, as described in section 2.1.3. Let $\rho : G \to \mathsf{U}(N)$ be a *unitary* representation; the matrices of the derived representation ρ' are antihermitian; given a basis (e_k) in $\mathfrak{g}$, define the *hermitian* matrix $\tau_k = \mathrm{i}\rho'(e_k)$. Let $\varphi : P \to \mathbb{C}^N$ be a 0-form of type ρ and put $\psi = s^*\varphi$ for a (local) section s of π. Let q be a coupling constant so that the potential in gauge s is $A = (\hbar c/q)s^*\omega$ and the gauge derivative of ψ is

$$s^* D\varphi = \mathrm{d}\psi - \frac{\mathrm{i}q}{\hbar c} A^k \tau_k \psi. \tag{2}$$

The *principle of minimal coupling* requires that A should appear only through the gauge derivative (2) in the Lagrangian determining the field equations of ψ.

2.2.2. Assume that $\mathfrak{g}$ admits a non-degenerate scalar product h, invariant with respect to the adjoint action of G on $\mathfrak{g}$; if G is semi-simple—and only in this case—one can take $h \propto k$. Let $\phi : P \to \mathbb{R}$ be a scalar field constant on the fibres of π. These data define a (pseudo) Riemannian metric $\boldsymbol{g}$ on P, namely $\boldsymbol{g} = \phi.h \circ \omega + g \circ T\pi$. This 'global' formula is translated into a local, but explicit, expression by introducing a field of frames (e_μ) on M, and lifting by π the dual field of coframes to P so that

$$\boldsymbol{g} = \phi.h_{ij}\omega^i\omega^j + g_{\mu\nu}e^\mu e^\nu.$$

The metric $\boldsymbol{g}$ admits G as a group of isometries. Therefore, the Ricci scalar of $\boldsymbol{g}$ descends to a function $R(\boldsymbol{g})$ on M. Put $s^*\Omega = (q/2\hbar c)e_k F^k_{\mu\nu} e^\mu \wedge e^\nu$. For $\phi = \text{const.}$, a computation (Jensen 1973, Cho 1975) gives

$$R(\boldsymbol{g}) = R(g) + (q/2\hbar c)^2 \phi h_{ij} g^{\mu\nu} g^{\rho\sigma} F^i_{\mu\rho} F^j_{\nu\sigma} + \tfrac{1}{4}\phi^{-1} h^{ij} k_{ij}. \qquad (3)$$

Choosing ϕ so that $(q/2\hbar c)^2\phi = \mathtt{G}/c^4$ and varying $\int R(\boldsymbol{g})\eta$ with respect to $A^i = s^*\omega^i$ and $g_{\mu\nu}$ one obtains, respectively, the Yang-Mills equation $D * F = 0$ and an Einstein equation with the energy-momentum tensor T of F and a term $\Lambda g_{\mu\nu}$, cosmological in form, on the right side. If $\mathfrak{g}$ is compact, then the invariant form h can be chosen to be (positive) definite; this implies $T_{00} \geq 0$. If the group G is Abelian, then $\Lambda = 0$; if the group is semi-simple and $h = k$, then the constant

$$\Lambda = \tfrac{1}{8}\ell^{-2}\alpha \dim G, \quad \text{where} \quad \alpha = q^2/\hbar c,$$

is seen to be of microscopic, rather than cosmological, nature. It is undesirably large.

2.2.3. For every $u \in \mathfrak{g}$, the vertical field $\bar{u}$ is a Killing vector for $\boldsymbol{g}$; therefore, if the curve $\mathbb{R} \ni t \mapsto \xi(t) \in P$ is an affinely parametrized geodesic of $\boldsymbol{g}$, then $\boldsymbol{g}(\dot{\xi}, \bar{u}) = \mu_\xi(u)$ is constant along the geodesic and defines a *charge* per unit mass $\mu_\xi \in \mathfrak{g}^*$. A horizontal geodesic characterized by $\mu_\xi = 0$ projects to a geodesic in M. Non-horizontal geodesics project to world-lines of particles with Yang-Mills charges, satisfying an equation of motion with a generalized Lorentz force on the right side.

2.2.4. In the original theories of Kaluza and Klein, the group G was assumed to be one-dimensional and ϕ to be constant; Jordan (1947) introduced a field ϕ to account for the possibility of a variable $\mathtt{G}$ conjectured by Dirac (1938).

It is intriguing that P plays here a double role: it is at the same time the total space of the bundle underlying a (first) quantized theory of particles interacting with a gauge field, in accordance with (2), and the Riemannian space whose geodesics determine the motion of 'classical', gauge-charged point particles. Formula (3) 'explains' why the gravitational Lagrangian is linear and the Yang-Mills one is quadratic in curvature.

2.3. The Einstein-Cartan theory

Since the early work of Weyl (1918) and Cartan (1922a) it has been known that there are two basic structures underlying local differential geometry and relevant to physics: the metric tensor and the linear connection. The metric determines a (weaker) conformal structure defining, in space-time, the propagation of light. A projective connection—a weakening of the linear one—gives the world-lines of 'material', freely-falling particles. Ehlers, Pirani and Schild (article in [O'R]) have shown how to reconstruct a Lorentzian geometry from compatible conformal and projective structures. Weyl assumed the connection to be symmetric; Cartan introduced the notion of *torsion* and related it to the density of *intrinsic angular momentum*; he also derived (Cartan 1924 p 22), from a variational principle, a set of gravitational field equations. Cartan required, without justification, that the covariant divergence of the energy-momentum tensor be zero; this led to an algebraic equation, bilinear in curvature and torsion, severely restricting the geometry. This misguided observation has probably discouraged Cartan from pursuing his theory. We now know that conservation laws in relativistic theories of gravitation follow from Bianchi identities and, in the presence of torsion, the divergence of the energy-momentum tensor need not vanish.

2.3.1. To present the Einstein-Cartan equations, it is convenient to use the differential forms (1) and refer all geometric and physical quantities to sections of the bundle of frames LM. In particular, θ now denotes the pull-back to M of the canonical form: it is a field of co-frames on M; similarly ω is the pull-back, by the same section, of the connection form, assumed for the moment to be completely general. A change of the section induced by $a : M \to \mathsf{GL}(\mathbb{R}^4)$ gives $\theta'^{\mu} = a^{-1\mu}{}_{\nu}\theta^{\nu}$, $\omega'^{\mu}{}_{\nu} = a^{-1\mu}{}_{\rho}\omega^{\rho}{}_{\sigma}a^{\sigma}{}_{\nu} + a^{-1\mu}{}_{\rho}\mathrm{d}a^{\rho}{}_{\nu}$, etc. Let (e^A), $A = 1, \dots, \dim V$, be a linear basis in V and, given a representation ρ of $\mathsf{GL}(\mathbb{R}^4)$ in V and $u = (u^{\mu}{}_{\nu}) \in \mathsf{End}(\mathbb{R}^4)$, put $\rho'(u)e^A = \rho^{A\rho}_{B\sigma}e^B u^{\sigma}{}_{\rho}$, so that the covariant derivative of an n-form $\varphi = e^A\varphi_A$ of type ρ is $D\varphi_A = \mathrm{d}\varphi_A + \rho^{B\mu}_{A\nu}\omega^{\nu}{}_{\mu} \wedge \varphi_B$. Consider a Lagrangian L which is an invariant 4-form on M, depends on the metric tensor, and is a polynomial function of φ, θ and ω, and of their first derivatives. The general variation of the Lagrangian is

$$\delta L = L^A \wedge \delta\varphi_A + \tfrac{1}{2}\tau^{\mu\nu}\delta g_{\mu\nu} + \delta\theta^{\mu} \wedge t_{\mu} - \tfrac{1}{2}\delta\omega^{\mu}{}_{\nu} \wedge s_{\mu}{}^{\nu} + \text{ an exact form} \tag{4}$$

so that $L^A = 0$ is the Euler-Lagrange equation for φ. If the changes of the functions θ, ω, g and φ are induced by an infinitesimal change of the frames,

$$\delta\theta^{\mu} = -v^{\mu}{}_{\nu}\theta^{\nu}, \quad \delta\omega^{\mu}{}_{\nu} = Dv^{\mu}{}_{\nu}, \quad \text{etc.,} \quad \text{where} \quad (v^{\mu}{}_{\nu}) : M \to \mathsf{End}(\mathbb{R}^4),$$

then $\delta L = 0$ and (4) gives the identity

$$\tau_{\mu}{}^{\nu} - \theta^{\nu} \wedge t_{\mu} + \tfrac{1}{2}Ds_{\mu}{}^{\nu} - \rho^{B\nu}_{A\mu}L^A \wedge \varphi_B = 0. \tag{5}$$

It follows from the identity that the two sets of equations obtained by varying L with respect to the triples $(\varphi, \theta, \omega)$ and (φ, g, ω) are equivalent. In the sequel, the first triple is chosen to derive the field equations.

2.3.2. Following Cartan, I consider, as the Lagrangian for the gravitational field, the 4-form

$$\eta_{\mu}{}^{\nu} \wedge \Omega^{\mu}{}_{\nu} = R\eta, \tag{6}$$

where $\eta_{\mu}{}^{\nu} = g^{\nu\rho}\eta_{\mu\rho}$ and $R = g^{\mu\nu}R_{\mu\nu}$ is the Ricci scalar; the Ricci tensor $R_{\mu\nu} = R^{\rho}{}_{\mu\rho\nu}$ is, in general, asymmetric. The form (6) is invariant with respect to 'projective transformations' of the connection,

$$\omega^{\mu}{}_{\nu} \mapsto \omega^{\mu}{}_{\nu} + \delta^{\mu}_{\nu}\lambda, \tag{7}$$

where λ is an arbitrary 1-form. Projectively related connections have the same (unparametrized) geodesics. The variations of the factors in (6) are $\delta\eta_{\mu}{}^{\nu} = \delta\theta^{\rho} \wedge \eta_{\mu}{}^{\nu}{}_{\rho}$ and $\delta\Omega^{\mu}{}_{\nu} = D\delta\omega^{\mu}{}_{\nu}$. If the total Lagrangian for gravitation interacting with the 'matter' field φ is $(16\pi)^{-1}\eta_{\mu}{}^{\nu} \wedge \Omega^{\mu}{}_{\nu} + L$, then the field equations, obtained by varying it with respect to φ, θ and ω are: $L^A = 0$,

$$\tfrac{1}{2}\eta_{\mu\nu}{}^{\rho} \wedge \Omega^{\nu}{}_{\rho} = -8\pi t_{\mu}, \tag{8}$$

and

$$D\eta_{\nu}{}^{\mu} = 8\pi s_{\nu}{}^{\mu}, \tag{9}$$

respectively. If

$$s_{\mu\nu} + s_{\nu\mu} = 0, \tag{10}$$

then $s_{\nu}{}^{\nu} = 0$ and L is also invariant under (7). One shows that, if (10) holds, then, among the projectively related connections satisfying (9), there is precisely one that is metric (Trautman 1973a). If $Dg = 0$ is assumed—this is done from now on—then (10) holds and the Cartan field equation (9) becomes

$$\eta_{\mu\nu\rho} \wedge \Theta^{\rho} = 8\pi s_{\mu\nu}. \tag{11}$$

Introducing the 'canonical' energy-momentum tensor $t_{\mu\nu}$ and the spin density tensor $s_{\mu\nu\rho}$ by

$$t_{\mu} = t_{\mu\nu}\eta^{\nu} \quad \text{and} \quad s_{\mu\nu} = s_{\mu\nu\rho}\eta^{\rho}. \tag{12}$$

one can write the Einstein-Cartan equations (8) and (11) in the form, given by Sciama (article in [In]) and Kibble (1961),

$$R_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} R \;\; = 8\pi t_{\mu\nu} \tag{13}$$

$$Q^{\rho}{}_{\mu\nu} + \delta^{\rho}_{\mu} Q^{\sigma}{}_{\nu\sigma} - \delta^{\rho}_{\nu} Q^{\sigma}{}_{\mu\sigma} = 8\pi s_{\mu\nu}{}^{\rho}. \tag{14}$$

In the limit of special relativity, in Cartesian coordinates (x^{μ}), one has $\theta^{\mu} = \mathrm{d}x^{\mu}$ and $D = \mathrm{d}$. If $L^{A} = 0$ holds, then $\mathrm{d}t_{\mu} = 0$ and the antisymmetrization of (5) gives the conservation of total angular momentum in differential form,

$$\mathrm{d}(x^{\mu} t^{\nu} - x^{\nu} t^{\mu} + s^{\mu\nu}) = 0.$$

As in classical general relativity, the right sides of the Einstein-Cartan equations need not necessarily be derived from a variational principle; they may be determined by phenomenological considerations. For example, consider a 'spinning dust' characterized by

$$t_{\mu\nu} = P_{\mu} u_{\nu} \quad \text{and} \quad s_{\mu\nu}{}^{\rho} = S_{\mu\nu} u^{\rho} \quad \text{with} \quad S_{\mu\nu} + S_{\nu\mu} = 0 \tag{15}$$

(Weyssenhoff and Raabe 1947). Indicating with a dot a 'particle' derivative in the direction of the unit vector field u, from the Bianchi identities one obtains (Trautman in [Hl])

$$P^{\mu} = \rho u^{\mu} - u_{\nu} \dot{S}^{\nu\mu}, \qquad \text{where} \qquad \rho = g_{\mu\nu} P^{\mu} u^{\nu},$$

an equation of motion of spin,

$$\dot{S}^{\mu\nu} = u^{\mu} u_{\rho} \dot{S}^{\rho\nu} - u^{\nu} u_{\rho} \dot{S}^{\rho\mu},$$

and an equation of translatory motion,

$$\dot{P}_{\mu} = (Q^{\rho}{}_{\mu\nu} P_{\rho} - \tfrac{1}{2} R^{\rho\sigma}{}_{\mu\nu} S_{\rho\sigma}) u^{\nu},$$

which is a generalization to the Einstein-Cartan theory of the Mathisson-Papapetrou equation for point particles with an intrinsic angular momentum (Mathisson 1937 and Papapetrou 1951).

2.3.3. There are heuristic, fairly convincing arguments in favour of the Einstein-Cartan theory: the integrability condition for the equation

$$D r^{\mu} + \theta^{\mu} = 0 \tag{16}$$

defining a 'radius' vector field r is $\Omega^{\mu}{}_{\nu} r^{\nu} + \Theta^{\mu} = 0$. Integration of (16) along a curve defines the 'Cartan displacement' of r; if this is done along a small loop spanned by the bivector Δf, then the radius vector changes by

$$\Delta r^{\mu} = \tfrac{1}{2} (R^{\mu}{}_{\nu\rho\sigma} r^{\nu} + Q^{\mu}{}_{\rho\sigma}) \Delta f^{\rho\sigma}.$$

This holonomy theorem (somewhat imprecisely formulated here) shows that torsion bears to translations a relation similar to that of curvature to rotations. The rest of the heuristic argument can be read off from the diagram:

rotations	······	curvature	······	energy-momentum	······	translations
	holonomy		field eqs		Noether	
translations	······	torsion	······	spin	······	rotations

It follows from (14) that torsion vanishes in the absence of spin and then (13) is the classical Einstein field equation. In particular, there is no difference between the Einstein and Einstein-Cartan theories in empty space. Since practically all tests of relativistic gravity are based on consideration of Einstein's equations in empty space, there is no difference, in this respect, between the Einstein and the Einstein-Cartan theories: the latter is as viable as the former.

2.3.4. Inside spinning matter, one can solve (14) for Q and express the full linear connection $\mathit{\Gamma}$ as a sum of the Levi-Civita connection and a 'contortion' tensor linear in Q. Substituting $\mathit{\Gamma}$ in (13), one obtains the standard Einstein equation with an effective, symmetric, energy-momentum tensor T^{eff} on the right side [Hh]. Neglecting the indices, one can write symbolically

$$T^{\text{eff}} = T + s^2, \tag{17}$$

where T is the usual, symmetric energy-momentum tensor. In special relativity, it is obtained from the canonical t by the Belinfante-Rosenfeld symmetrization procedure (symbolically: $T = t + \operatorname{div} s$); here it appears as a consequence of the field equations. It is clear from (17) that for the gravitational effects of torsion to be comparable to those of curvature, the latter should be of the same order of magnitude as the square of the former. For example, to achieve this, a nucleon of mass m should be squeezed so that its radius r_{Cart} be such that

$$\left(\frac{\mathtt{G}\hbar}{c^3 r^3_{\text{Cart}}}\right)^2 \approx \frac{\mathtt{G}m}{c^2 r^3_{\text{Cart}}}.$$

Introducing the Planck length $\ell \approx 1.6 \times 10^{-33}$ cm and the Compton wavelength $r_{\text{Compt}} = \hbar/mc \approx 10^{-13}$ cm, one can write

$$r_{\text{Cart}} \approx (\ell^2 r_{\text{Compt}})^{1/3}. \tag{18}$$

The 'Cartan radius' of the nucleon, $r_{\text{Cart}} \approx 10^{-26}$ cm, so small when compared to its physical radius under normal conditions, is much larger than the Planck length. Curiously enough, the energy $\hbar c/r_{\text{Cart}}$ is comparable to the energy at which, according to some estimates, the 'grand unification' of interactions is presumed to occur.

2.3.5. In the presence of spinning matter, T^{eff} need not satisfy the (positive) energy conditions [HaE] even if T does. Therefore, the classical theorems of Penrose and Hawking can be here 'overcome' and, in this theory, there are simple cosmological solutions without a singularity (Kopczyński 1972, Tafel 1973, Trautman 1973b and in [Hl]). For example, consider a Universe filled with a spinning dust (15) such that $u^\mu = \delta^\mu_0$, $S_{23} = \sigma$, $S_{\mu\nu} = 0$ for $\mu + \nu \neq 5$ and both ρ and σ are functions of $t = x^0$ alone. These assumptions are compatible with the Robertson-Walker line element $\mathrm{d}t^2 - a(t)^2(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2)$, where $x = x^1$, etc., and torsion is determined from (14). The Einstein equations (13) reduce to the modified Friedmann equation,

$$\tfrac{1}{2}\dot{a}^2 - Ma^{-1} + \tfrac{3}{2}S^2 a^{-4} = 0, \tag{19}$$

supplemented by the conservation laws of mass and spin,

$$M = \tfrac{4}{3}\pi\rho a^3 = \text{const.}, \quad S = \tfrac{4}{3}\pi\sigma a^3 = \text{const.}$$

The last term on the left side of (19) plays the role of a 'repulsive potential', effective at small values of a; it prevents the solution from vanishing (Trautman 1973b). It should be noted, however, that even a very small amount of shear in u will result in a term counteracting the repulsive potential due to spin.

2.3.6. In any case, the consideration of torsion amounts to a slight change of the energy-momentum tensor that can be also obtained by the introduction of a new term in the Lagrangian. This observation was made already by Weyl (1950) in the context of the Dirac equation. Supergravity theories also require torsion to implement the supersymmetries between ω and the gravitino field (Van Nieuwenhuizen 1981).

In Einstein's theory one can also satisfactorily describe spinning matter without introducing torsion; see, e.g., the article by Israel in [CB]. To derive the field equations along the lines presented above, one can use the 'Palatini trick' and express ω in terms of the orthonormal frames θ and their derivatives. Introducing the corresponding expression for $\delta\omega$ in (4) and 'integrating by parts', one obtains the symmetric T as the source in (13).

To summarize: the Einstein-Cartan theory is a viable theory of gravitation that differs very slightly from the Einstein theory; the effects of spin and torsion can be significant only at densities of matter that are very high, but nevertheless much smaller than the Planck density at which quantum gravitational effects are believed to dominate. It is likely that the Einstein-Cartan theory will prove to be a better classical limit of a future quantum theory of gravitation than the theory without torsion.

3. Optical and Cauchy-Riemann structures

Throughout the 1960s there was much research on algebraically special gravitational fields and the associated geometrical structures. It culminated with the discovery of important solutions of Einstein's equations, such as the Kerr metric, and special kinds of gravitational waves [K]. Some of this research was criticized as being dominated by an obsessive striving to discover new solutions in closed form, with little regard to their physical significance. It is certainly true that algebraically special fields are *exceptional* and, in some sense, constitute a set of measure zero in the set of all empty space-times. It is worth remembering, however, that generic stationary metrics with axial symmetry are of type I (non-degenerate). On the other hand, the Kerr metric—which is of type D—has been shown to describe the most *general* black hole resulting from a collapsing star. The dominant part—of order r^{-1}—of the curvature tensor of a generic, isolated gravitating system is algebraically special at large distances r.

The ideas associated with shear-free congruences of null geodesics, which underlie algebraically special fields, influenced the emergence of the fundamental notion of a Penrose twistor.

This section is based, in part, on (Trautman 1984, 1985 and in [Har], Robinson and Trautman 1986 and 1989); see also the references given there and (Nurowski 1996, 1997).

3.1. The differential-geometric aspect

Bateman (1910) was a precursor of the idea of an 'optical geometry' associated with a null electromagnetic field in Minkowski space. He proved a result that, in the light of later developments and in the present terminology, can be summarized as follows. Consider a Lorentzian, space- and time-oriented, 4-manifold (M, g). (Bateman assumed g to be the Minkowski metric.) Let $K \subset TM$ be a line bundle of null directions on M; a section k of $K \to M$ is a null vector field; let $\lambda = g(k)$. For any positive function ρ and any 1-form ξ, the metric

$$g' = \rho g + \lambda\xi \tag{20}$$

is also Lorentzian, k is null also with respect to g' and $g'(k) \wedge \lambda = 0$. Given K, equation (20) defines an equivalence relation in the set of Lorentzian metrics and an *optical geometry* is defined to be the triple $(M, K, [g])$, where $[g] = [g']$ iff (20) holds. The bundle $K^\perp \subset TM$ is well-defined by the optical geometry and $[g]$ induces a conformal, positive-definite metric in the fibres of the plane (*screen*) bundle $K^\perp/K \to M$. Together with the orientation of M, this defines a complex structure J in the fibres of the screen bundle. Conversely, given (M, K, L, J), where K is a line subbundle of a bundle $L \subset TM$ of fibre dimension 3 and J is a complex structure in the fibres of L/K, one can construct an optical geometry on M such that $L = K^\perp$. If Φ is a complex 2-form on M such that

$$\lambda \wedge \Phi = 0, \tag{21}$$

then self-duality of Φ,

$$*\Phi = \mathrm{i}\Phi, \tag{22}$$

is an optical property: it holds in g iff it holds in g', $[g] = [g']$. Therefore, a 2-form Φ satisfying (21) and (22) can be said to be *optical*. An *optical automorphism* of $(M, K, [g])$ is defined as an orientation-preserving diffeomorphism φ of M leaving K invariant and such that $[\varphi^* g] = [g]$. Bateman discovered that an optical automorphism transforms an optical 2-form into an optical 2-form. Moreover, since $\mathrm{d}\varphi^*\Phi = \varphi^*\mathrm{d}\Phi$, it transforms a real null solution $F = \mathrm{Re}\,\Phi$ of Maxwell's equations, $\mathrm{d}F = 0$, $\mathrm{d}{*F} = 0$, into another such solution. It was also observed (Mariot 1954) that, given such a non-zero solution, the congruence of curves tangent to the vector field k consists of *null geodesics*. An essential progress was made by Robinson (1961) who proved that if there is a nowhere vanishing closed 2-form Φ which is optical for $(M, K, [g])$, then the null vector field k generates a one-parameter group (φ_t) of optical automorphisms; this group leaves Φ invariant. The *shear-free condition* is obtained by differentiating $[\varphi_t^* g] = [g]$ with respect to t; if it is satisfied, then the congruence defined by k is *null geodetic* and *shear-free*; it is also called a shear-free congruence of rays (SFR). Robinson also sketched a proof of the converse: given such a congruence, there is an associated with it, non-zero and null, solution of Maxwell's equations. An optical geometry $(M, K, [g])$ satisfying the shear-free condition is said to be *shear-free*. Equivalently, the optical geometry given by (M, K, L, J) is shear-free if the group (φ_t) preserves L and J.

3.2. The algebraic aspect and the Goldberg-Sachs theorem

Cartan (1922b) described the four null (he called them: *optical*) directions defined by a non-zero conformal curvature tensor C in a Lorentzian space. He recognized that in the case of the Schwarzschild solution these directions degenerate into two coinciding pairs. This implicit classification of gravitational fields was independently discovered, in a coarse form, by Petrov (1954): he found only three algebraic types of conformal curvature. In its final, spinorial form, the classification was given by Penrose (1960). A link between algebraically special fields, for which at least two of those directions coincide, and the existence of an SFR is given by the well-known Goldberg-Sachs theorem and its generalizations; see §7.3 in [K] and [PR]. The physical interpretation of the five Cartan-Petrov-Penrose types of C follows from the results on peeling of curvature at large distances; see the article by Sachs in [In] and §9.7 in [PR].

3.3. The complex and analytic aspects

3.3.1. Lewy (1957) considered the differential operator on $\mathbb{R}^3$,

$$Z = \frac{\partial}{\partial x} + \mathrm{i}\frac{\partial}{\partial y} - \mathrm{i}(x + \mathrm{i}y)\frac{\partial}{\partial u}, \tag{23}$$

and showed, to the surprise of many mathematicians, that there are smooth (of class C^∞) functions $h : \mathbb{R}^3 \to \mathbb{C}$ such that the differential equation

$$Z(f) = h \tag{24}$$

has no solution, even locally, for $f : \mathbb{R}^3 \to \mathbb{C}$. (By the Cauchy-Kowalewski theorem, there are such solutions if h is analytic.) At about the same time, Robinson constructed a remarkable congruence of straight null lines in Minkowski space: he took the line element $g = 2\mathrm{d}U\mathrm{d}r - \mathrm{d}X^2 - \mathrm{d}Y^2$ and performed the coordinate transformation,

$$U = u + \tfrac{1}{2}r(x^2 + y^2), \quad X + \mathrm{i}Y = (r + \mathrm{i})(x + \mathrm{i}y),$$

to obtain

$$g = 2\lambda\mathrm{d}r - (r^2 + 1)(\mathrm{d}x^2 + \mathrm{d}y^2), \quad \text{where} \quad \lambda = \mathrm{d}u + x\mathrm{d}y - y\mathrm{d}x$$

is null and twisting, $\lambda \wedge \mathrm{d}\lambda = 2\mathrm{d}u \wedge \mathrm{d}x \wedge \mathrm{d}y$. The complex 2-form

$$\varPhi = (\exp f(x, y, u, r))\lambda \wedge (\mathrm{d}x + \mathrm{i}\mathrm{d}y)$$

is self-dual and the Maxwell equation $\mathrm{d}\varPhi = 0$ reduces to $\partial f/\partial r = 0$ and the homogeneous Lewy equation

$$Z(f) = 0. \tag{25}$$

The twisting, shear-free *Robinson congruence* generated by the vector field $k = \partial/\partial r$ played a role in the development of twistors (Penrose in [RT]).

3.3.2. The connection between the Lewy equation and the Robinson congruence is now understood in terms of the associated Cauchy-Riemann (CR) structure, a mathematical notion that appeared for the first time in physics as the underlying geometry of the five-dimensional space of null, projective twistors (Penrose 1967, 1978 and 1983). There is a one-to-one local correspondence between CR spaces and shear-free optical geometries. To see this, consider such a geometry (M, K, L, J) and assume that the set of null geodesics tangent to the fibres of K is a (three-dimensional) manifold N. Since both L and J are invariant with respect to (φ_t), these structures descend to N and define a complex line bundle $H \subset \mathbb{C} \otimes TN$ such that $H \cap \bar{H}$ is the zero bundle. By definition, a 3-manifold N with such an H is a CR space. Conversely, given a CR space (N, H), one can define its *lift* as an optical geometry on

$M = N\times\mathbb{R}$ or $N\times\mathbb{S}_1$ described as follows. The complex line bundle H defines locally the direction of a real 1-form κ such that $H = \ker\kappa$ and the complex direction of a vector field Z, section of $H \to N$. Let ν be a complex 1-form on N such that $\nu(Z) = 0$ and $\nu(\bar{Z}) = 1$. The pair (κ, ν) is defined up to transformations $(\kappa, \nu) \mapsto (a\kappa, b\nu + c\kappa)$, where $a \neq 0$ is a real function and $b \neq 0$ and c are complex functions. Denoting by π the projection $M \to N$, introducing r as the coordinate along the fibres of π and putting $\lambda = \pi^*\kappa$, $\mu = \pi^*\nu$, one can write a Lorentzian metric g, representative for an optical geometry on M, as

$$g = 2\lambda \mathrm{d}r - \mu\bar{\mu} \quad \text{and} \quad k = \partial/\partial r. \tag{26}$$

A closed optical 2-form $\varPhi$ on M descends to N: there is a closed 2-form $\varPsi$ such that $\varPhi = \pi^*\varPsi$ and $\varPsi(Z, V) = 0$ for every vector field V on N. If $\varPhi$ is nowhere zero, then there is a function $f : N \to \mathbb{C}$ such that $\varPsi = (\exp f)\kappa \wedge \nu$ and the problem of finding a null solution of Maxwell's equations associated with the SFR generated by k reduces to the CR problem of solving the equation $\mathrm{d}((\exp f)\kappa \wedge \nu) = 0$, equivalent in form to (24), where now h is a complex function on N *determined* by the equation $h\kappa \wedge \bar{\nu} \wedge \nu = \mathrm{d}(\kappa \wedge \nu)$. If the 'tangential CR equation' (25) has two solutions f_1 and f_2 such that the map $(f_1, f_2) : N \to \mathbb{C}^2$ is an immersion, then the CR structure is said to be *realizable* (or: *embeddable*) and N can be (locally) identified with a hypersurface in $\mathbb{C}^2$. The 2-form $\varPsi = \mathrm{d}f_1 \wedge \mathrm{d}f_2$ lifts to a closed form optical in the geometry given by (26). Such is the case of the CR structure underlying the Robinson congruence: if Z is as in (23), then one can take $f_1 = u + \frac{1}{2}\mathrm{i}z\bar{z}$, $f_2 = z$ and N is identified with the *hyperquadric* in $\mathbb{C}^2$ of equation $f_1 - \overline{f_1} = \mathrm{i}f_2\overline{f_2}$. Cartan (1932) has shown this CR space to be the most symmetric among those with a non-vanishing 'Levi form' ϖ defined by $\varpi\kappa \wedge \bar{\nu} \wedge \nu = \mathrm{i}\kappa \wedge \mathrm{d}\kappa$; such CR spaces correspond to twisting SFRs. The group of automorphisms of the hyperquadric is locally isomorphic to $\mathsf{SU}_{1,2}$; there are many CR spaces with a maximal group of automorphisms of dimension $n = 3$, but none with $3 < n < 8$. As a CR space, the hyperquadric is locally isomorphic to $\mathbb{S}_3 \subset \mathbb{C}^2$. Penrose (1967) has shown how the Robinson congruence is related to the Hopf fibration $\mathbb{S}_3 \to \mathbb{S}_2$. The Taub-NUT space, the Hauser solution of type N and the Gödel universe have an underlying optical geometry lifting the CR structure of the hyperquadric.

3.3.3. Nirenberg (1974) has constructed a smooth complex vector field Z on $\mathbb{R}^3$ such that (25) implies $f = \text{const.}$: the CR structure defined by this Z is not embeddable. Under the influence of the results of Lewy and Nirenberg, Tafel (1985) pointed out that the proof of the Robinson theorem is valid only in the real-analytic case and there may be smooth Lorentzian spaces with an SFR not admitting an associated, non-zero null solution of Maxwell's equations. He has also observed that if a CR space N admits a non-zero solution of the tangential CR equation (25) and a non-zero closed 2-form $\varPsi$ such that $\varPsi(Z, V) = 0$ for every vector field V, then N is embeddable. According to Rosay (1989) there are CR spaces that admit one solution of (25), but are non-embeddable; therefore they could be used to construct Lorentzian, smooth manifolds with an SFR, not admitting a non-zero null solution of Maxwell's equations. It has been conjectured that a CR space is embeddable if the associated optical geometry does admit such a solution of Maxwell's equations (Trautman in [Har]). Another result along these lines is due to Lewandowski *et al* (1990): these authors have shown that if a Lorentzian space with an SFR is Einstein, then the corresponding CR space is embeddable.

3.3.4. According to the Penrose (1968) formulation of the Kerr theorem, every analytic SFR in the conformally compactified Minkowski space is obtained as the lift of a CR space arising from the intersection of the CR 5-manifold of projective null twistors with a holomorphic hypersurface in the ambient $\mathbb{CP}_3$. Penrose (1983) points out that the set of all analytic CR spaces is (essentially) labeled by functions of three variables, whereas its subset corresponding to SFRs in (compactified) Minkowski space is labeled by functions of two variables only. An open problem is to *characterize the CR spaces that lift to an optical geometry admitting a representative which is an Einstein manifold.* Lewandowski *et al* (1991) have shown that every embeddable CR space with $\varpi \neq 0$ lifts to an optical geometry admitting a representative g satisfying the Einstein equation with pure radiation as the source, i.e. such that $R_{\mu\nu}$ is proportional to $k_\mu k_\nu$.

Cartan (1932) has shown that, in the case of an analytic CR space with a non-vanishing Levi form, one can choose the forms κ and ν defining the CR structure so that

$$\nu = \mathrm{d}z \quad \text{and} \quad \mathrm{d}\kappa = \mathrm{i}\nu \wedge \bar{\nu} + \kappa \wedge (w\nu + \overline{w\nu}),$$

where z and w are complex functions on N. He presented a method of constructing (relative) invariants of the CR structure in terms of w and its derivatives, $\mathrm{d}w = w_0\kappa + w_1\nu + w_2\bar{\nu}$. The simplest among them,

$$w_{211} - w_1 w_2 - 3ww_{21} - 2\mathrm{i}w_{01} + 2w^2 w_2 + 4\mathrm{i}ww_0,$$

vanishes iff the CR space is isomorphic to the hyperquadric. Problem: *characterize the SFRs of Minkowski space by means of the Cartan invariants*.

4. General remarks and speculations

4.1. Large numbers

The large numbers that can be formed from cosmological and atomic quantities, and the coincidences between them, have received much attention since they were noticed for the first time by Weyl, Eddington and Dirac [B]. One such large number is the inverse of the *gravitational fine structure constant* $m^2/\ell^2 \approx 6 \times 10^{-39}$, where m is the mass of the proton. Chandrasekhar (1937) has suggested to use the numbers $N_r = (\ell/m)^r$, where r is rational, to keep a record of the large and small numbers. The fundamental coincidence between the age of the Universe—determined by the Hubble constant—and $\ell^2 m^{-1} N_2$ has led Dirac to conjecture that the gravitational constant is inversely proportional to the present age of the Universe (Dirac 1938, 1974). Some of the definitions (D), theoretical results (T) and observations (O) involving the numbers N_r are summarized in the Table. Since, in gravitational units, mass is of the dimension of length, one can choose $m^\alpha \ell^\beta$ with $\alpha + \beta = 1$ as the unit of mass or length; similarly, the choice $\alpha + \beta = -2$ gives a unit of density. In the Table, these units have been chosen so that the row with $r = 0$ corresponds to the 'nuclear' quantities m, $\hbar/mc$ and $m^4 c^3/\hbar^3$.

Sometimes the values of the fundamental constants and the associated coincidences are explained by reference to the *anthropic principle*: had they been different, intelligent life could not have evolved and there would be no one to puzzle over them. I am now inclined to consider this to be an *explanation without explanation*, to use a phrase coined by Wheeler.

4.2. Gravitation provides arguments for the unity of physics

The *Chandrasekhar limit* on the mass of white dwarfs is the most significant entry in the Table; it turns out to be, essentially, also the limiting mass of neutron stars. In a form given by Landau (1932), neglecting a numerical factor of order 1, the Chandrasekhar limit is

$$M_{\max} = m \left(\frac{\hbar c}{\mathtt{G} m^2} \right)^{3/2}. \tag{27}$$

Its derivation is based on Newtonian gravitation and quantum statistical mechanics of fermions; it takes into account effects of special relativity. The result pertains to large, macroscopic systems and is confirmed by astrophysical observations; it is unique in combining the macroscopic $\mathtt{G}$ with microscopic constants such as $\hbar$ and the proton mass m. I consider (27) to be *one of the most striking results in theoretical physics of the XXth century*: it demonstrates the unity and universal applicability of fundamental physics.

4.3. An EIH problem

Recall that the Einstein, Infeld and Hoffmann (Einstein *et al* 1938) method of finding the post-Newtonian corrections to the equations of motion of gravitating bodies postulates the existence of coordinates such that the solutions of the Einstein field equations are of the form $g_{\mu\nu} = g_{\mu\nu}^{\mathrm{Mink}} + h_{\mu\nu}$ and the functions $h_{\mu\nu}$ can be expanded in power series in $1/c$. These authors assumed that the solutions were of the *standing*

Table. The Dirac-Chandrasekhar numbers.

r	mass mN_r	length $\ell^2 m^{-1} N_r$	density $m^4 \ell^{-6} N_r$
-2		gravitational radius of proton (D)	Universe (O)
-1		Planck (D)	
$-\frac{2}{3}$		r_{Cart} defined in (18) (D)	
0	nucleon (D)	nucleon radius (D)	nuclear (D)
$\frac{2}{3}$		$R_{\min}$ of the observable Universe with spin and torsion (T)	
1	Planck (D)	radius of neutron star (T& O)	
2		Hubble radius (O), gravitational Bohr radius (D)	$\rho_{\max}$ in a model of the Universe with spin and torsion (T)
3	Chandrasekhar limit (T& O)		
$\frac{7}{2}$	galaxy (O)		
4	observable Universe (O)		Planck (D)

wave type: in their work the expansions of the components h_{00} and h_{ij}, $i, j = 1, 2, 3$, contained only even and those of h_{0j} only odd powers of $1/c$. Therefore, these solutions were assumed to be invariant with respect to the time reversal and did not take into account the possibility of gravitational radiation and its reaction on the motion of the bodies. Later it was recognized that gravitational radiation can be introduced into the EIH method provided more general expansions are used; see (Trautman 1978) for a historical survey.

In the Newtonian theory, the simplest, non-trivial gravitating system consists of two bodies of equal masses moving on a circular orbit. Tacitly, Einstein, Infeld and Hoffmann assumed that there is an analogous situation in general relativity, but, to my knowledge, this has never been demonstrated. In particular, it is not clear whether the corresponding space-time is—in any sense—asymptotically flat. The doubt arises from the observation that an analogous solution of Maxwell's equations in special relativity has an infinite amount of energy stored in the electromagnetic field. One is tempted to speculate that a double-star system in periodic motion would result in a closed space. In any case, it would be interesting to know, not only for historical reasons, whether there exists a strictly periodic general-relativistic configuration for which the EIH computations provide the first approximation. Essentially the same problem. but in a different context, has been formulated by Schmidt (article in [Har]).

4.4. On variational principles

Hilbert (1915) introduced the Ricci scalar density as the basis for the variational principle of Einstein's theory. Einstein himself, in the main paper on general relativity (Einstein 1916), derived the equations from a Lagrangian quadratic in the Christoffel symbols. It differs from the Hilbert Lagrangian by a total divergence. There are several examples in physics of Lagrangians that change by a divergence under the action of the symmetries of the theory. The simplest is the Newtonian kinetic energy: it changes by a total time derivative under Galilean transformations. In relativistic mechanics, the corresponding action integral, $-m\int ds$, is invariant with respect to Lorentz transformations. On the basis of these and similar examples one can argue that the appearance of such a divergence is an indication that one is dealing with some approximation or a limiting case of a'better' theory, in which the corresponding, possibly modified symmetries fully preserve the action integral. In particular, I feel that supersymmetric theories would gain much if they were reformulated so as to make supersymmetries into proper invariant transformations; see (Trautman 1996) for further examples.

4.5. Gravitation and quantum theory

The problem of constructing a quantum theory of gravitation is considered to be one of the most important and difficult in fundamental physics; it is reviewed elsewhere in this issue. Not being a specialist in this field, I restrict myself to a few short, tentative comments. It is my impression that some of the difficulties in the development of contemporary theoretical physics arise from unjustified generalizations to all of physics of notions suitable only in parts of it. Pre-Maxwellian physics was dominated by mechanics. Many attempts were then made to explain electromagnetic phenomena in terms of elastic forces. It took us a long time to realize—and this became possible only after the advent of quantum theory and general relativity—that all the forces that can be legitimately put on the right side of Newton's equation are of electromagnetic origin. The principle of equivalence implies that there is really no such thing as gravitational force. The understanding of interactions in the XXth century has been dominated by electrodynamics. We are now not so naïve as to try to reduce all phenomena to electromagnetism, but we attempt to model all theories after electrodynamics, classical or quantum. When doing this, we rely on many notions, such as that of energy, which may be traced back to that of force.

Much of the work on quantization of the gravitational field has been based on analogies with electromagnetism, in an attempt to build a theory of gravitons, similar to that of photons. Ultimately, something along these lines will probably be achieved. One should keep in mind, however, that there are some deep links between gravitation and quantum theory at a rather elementary—and, therefore, fundamental—level. They may be interpreted to indicate that a synthesis of general-relativistic and quantum ideas should occur at a 'low level', before the construction of a theory of gravitons.

4.5.1. The beautiful arguments due to Bondi (pp 418–9 in [BPT]) and Schild (pp 27–37 in [Eh]) show that elementary quantum notions applied to (Newtonian) gravitation lead to the conclusion that space-time is curved. There is a similarly simple way of 'deriving $E = mc^2$ from general relativity and $\Delta E = h\nu$' (pp 164–5 in [R]). It is a small miracle that these arguments work.

4.5.2. The (algebraically special) Kerr-Newman solution describes the external gravitational and electromagnetic fields of a black hole with angular momentum and charge. It is remarkable that its gyromagnetic ratio is the same as that of an electron, derived in quantum mechanics from the Dirac equation (p 883 in [MTW]).

4.5.3. Of even greater significance is the mysterious miracle consisting in similarities between the laws of black holes and those of thermodynamics; for recent accounts see Wald in [Wa] and his article in this issue. In this context it is worth recalling Einstein's view on classical thermodynamics, expressed in his *Autobiographical Notes*: 'It is the only physical theory of universal content concerning which I am convinced that, within the framework of the applicability of its basic concepts, it will never be overthrown (for the special attention of those who are skeptics on principle)' (p 33 in [Sch]). Einstein probably did not appreciate the significance of the 1939 paper by Oppenheimer and Snyder portending the discovery of black holes, but—had he lived long enough—he would have been impressed by the Bekenstein analogies and the prediction of the Hawking radiation.

Acknowledgments

I am grateful to W. Kopczyński, P. Nurowski and H. Urbantke for their helpful remarks. Work on this paper was supported in part by the Polish Committee for Scientific Research (KBN).

Books and reviews

[B] Bondi H 1960 *Cosmology* (Cambridge: Cambridge University Press)
[BPT] Bondi H, Pirani F A E and Trautman A 1965 *Lectures on general relativity: Brandeis Summer Institute 1964* vol 1 (Englewood Cliffs: Prentice-Hall)
[C] Cartan É 1986 *On manifolds with an affine connection and the theory of general relativity* transl. from French by A Magnon and A Ashtekar (Napoli: Bibliopolis)
[CB] Choquet-Bruhat Y ed 1974 *Ondes et radiations gravitationnelles: Coll. intern. CNRS 220* (Paris: Éditions du CNRS)
[Eh] Ehlers J ed 1967 *Relativity theory and astrophysics 1. Relativity and Cosmology* Lectures in Appl. Math. vol 8 (Providence: American Mathematical Society)
[Fe] Feynman R P 1988 *"What Do You Care What Other People Think?"* (New York: Norton) pp 88–93
[Fo] Fock V A 1964 *The Theory of Space, Time, and Gravitation* transl. from Russian by N N Kemmer 2nd ed (New York: Macmillan)
[Har] Harvey A ed 1999 *On Einstein's Path: Essays in Honor of Engelbert Schucking* (New York: Springer)
[HaE] Hawking S W and Ellis G F R 1973 *The large scale structure of space-time* (Cambridge: Cambridge University Press)
[HaP] Hawking S W and Penrose R 1996 *The nature of space and time* (Princeton: Princeton University Press)
[Hh] Hehl F W, von der Heyde P, Kerlick G D and Nester J M 1976 General relativity with spin and torsion: foundations and prospects *Rev. Mod. Phys.* **48** 393–416
[Hl] Held A ed 1980 *General relativity and gravitation* (New York: Plenum)
[In] Infeld L (volume dedicated to) 1962 *Recent developments in general relativity* (Oxford and Warszawa: Pergamon Press and PWN)
[J] Infeld L ed 1964 *Conf. intern. sur les théories relativistes de la gravitation (Warszawa and Jabłonna 1962)* (Paris and Warszawa: Gauthier-Villars and PWN)
[K] Kramer D, Stephani H, MacCallum M and Herlt E 1980 *Exact solutions of Einstein's equations* (Berlin: VEB Deutscher Verlag der Wissenschaften)
[L] Lichnerowicz A 1955 *Théories relativistes de la gravitation et de l'électromagnétisme* (Paris: Masson)
[MTW] Misner C W, Thorne K S and Wheeler J A 1973 *Gravitation* (San Francisco: Freeman)
[O'R] O'Raifeartaigh L ed 1972 *General relativity: papers in honour of J L Synge* (Oxford: Clarendon Press)
[PR] Penrose R and Rindler W 1986 *Spinors and space-time, vol 2: Spinor and twistor methods in space-time geometry* (Cambridge: Cambridge University Press)
[R] Rindler W 1977 *Essential relativity* 2nd ed (New York: Springer)
[RT] ——and Trautman A eds 1987 *Gravitation and Geometry: A volume in honour of Ivor Robinson* (Napoli: Bibliopolis)
[Sch] Schilpp P A ed 1969 *Albert Einstein: Philosopher, Scientist* 3rd ed vol 1 (La Salle and London: Open Court and Cambridge University Press)
[S] Synge J L 1960 *Relativity: the general theory* (Amsterdam: North-Holland)
[Wa] Wald R M ed 1998 *Black holes and relativistic stars* (Chicago: University of Chicago Press)
[W] Witten L ed 1962 *Gravitation: an introduction to current research* (New York: Wiley)

Other references

Bateman H 1910 The transformation of coordinates which can be used to transform one physical problem into another *Proc. London Math. Soc.* **8** 469–88
Cartan É 1922a Sur une généralisation de la notion de courbure de Riemann et les espaces à torsion *C. R. Acad. Sci. (Paris)* **174** 593–5
——1922b Sur les espaces conformes généralisées et l'univers optique *ibid.* **174** 857–9 (English translation in [C])
——1924 Sur les variétés à connexion affine et la théorie de la relativité généralisée (première partie, suite) *Ann. Éc. Norm.* **41** 1–25 (English translation in [C])
——1932 Sur la géométrie pseudo-conforme des hypersurfaces de l'espace de deux variables complexes I *Annali Mat. Pura Appl.* **11** 17–90
Chandrasekhar S 1937 The cosmological constants *Nature* **139** 757–8
Cho Y M 1975 Higher-dimensional unifications of gravitation and gauge theories *J. Math. Phys.* **16** 2029–35

DeWitt B S 1964 Dynamical theory of groups and fields *Relativity, groups, and topology* ed C and B DeWitt (New York: Gordon and Breach) pp 587–820

Dirac P A M 1938 A new basis for cosmology *Proc. R. Soc.* A **165** 199–208

——1974 Cosmological models and the Large Numbers hypothesis *Proc. R. Soc.* A **338** 439–46

Einstein A 1916 Die Grundlage der allgemeinen Relativitätstheorie *Ann. Phys., Lpz.* **49** 769–822

——, Infeld L and Hoffmann B 1938 The gravitational equations and the problem of motion *Ann. of Math.* **39** 65–100

Goldberg J N 1955 Gravitational radiation *Phys. Rev.* **99** 1873–83

Heckmann O H L and Schücking E 1959 Newtonsche und Einsteinsche Kosmologie *Handbuch der Physik* **53** 489–519

Hehl F W, McCrea J D, Mielke E W and Ne'eman Y 1995 Metric-affine theory of gravity: field equations, Noether identities, world spinors, and breaking of dilation invariance *Physics Reports* **258** 1–171

Hilbert D 1915 Die Grundlagen der Physik *K. Gesell. Wiss. Göttingen, Nachr. Math.-Phys. Kl.* pp 395–407

Jensen G R 1973 Einstein metrics on principal fibre bundles *J. Differential Geometry* **8** 599–614

Jordan P 1947 Erweiterung der projektiven Relativitätstheorie *Ann. Phys., Lpz.* 6 Folge **1** 219–228

——, Ehlers J and Kundt W 1960 Strenge Lösungen der Feldgleichungen der Allgemeinen Relativitätstheorie *Akad. Wiss. Lit. Mainz Abhandl. Math.-Nat. Kl.* Nr 2

————and Sachs R K 1961 Beiträge zur Theorie der reinen Gravitationsstrahlung *ibid* Nr 1

Kerbrat Y, Kerbrat-Lunc H and Śniatycki J 1989 How to get masses from Kaluza-Klein theory *J. Geom. Phys.* **6** 311–30

Kibble T W B 1961 Lorentz invariance and the gravitational field *J. Math. Phys.* **2** 212–21

Kopczyński W 1972 A non-singular universe with torsion *Phys. Lett.* **39**A 219–20

——1980 A fibre bundle description of coupled gravitational and gauge fields *Differential Geometrical Methods in Mathematical Physics* ed P L Garcıa *et al* Lecture Notes in Math. 836 (Berlin: Springer) pp 462–483

Landau L 1932 On the theory of stars *Phys. Z. Sowjetunion* **1** 285–8

Lewandowski J, Nurowski P and Tafel J 1990 Einstein's equations and realizability of CR manifolds *Class. Quantum Grav.* **7** L241–6

————1991 Algebraically special solutions of the Einstein equations with pure radiation fields *Class. Quantum Grav.* **8** 493–501

Lewy H 1957 An example of a smooth linear partial differential equation without solution *Ann. of Math.* **66** 514–22

Mariot L 1954 Le champ électromagnétique singulier *C. R. Acad. Sci. Paris* **238** 2055–6

Mathisson M 1937 Neue Mechanik materieller Systeme *Acta Phys. Polon.* **6** 163–200

Ne'eman Y 1998 Gauge theories of gravity *Acta Phys. Polon.* **B29** 827–43

Newman E T 1974 Lienard-Wiechert fields and general relativity *J. Math. Phys.* **15** 44–5

——and Penrose R 1962 An approach to gravitational radiation by a method of spin coefficients *J. Math. Phys.* **3** 896–902

——, Tamburino L and Unti T 1963 Empty space generalization of the Schwarzschild metric *J. Math. Phys.* **4** 915–23

Nirenberg L 1974 On a question of Hans Lewy (in Russian) *Uspekhi Mat. Nauk* **29** 241–51; English transl. in *Russian Math. Surveys* **29** 251–62

Nurowski P 1996 Optical geometries and related structures *J. Geom. Phys.* **18** 335–48

——1997 Twistor bundles, Einstein equations and real structures *Class. Quantum Grav.* **14** A261–90

Papapetrou A 1951 Spinning test-particles in general relativity I *Proc. R. Soc.* A **209** 248–58

Penrose R 1960 A spinor approach to general relativity *Ann. Phys., NY* **10** 171–201

——1967 Twistor algebra *J. Math. Phys.* **8** 345–66

——1968 Twistor quantisation and curved space-time *Int. J. Theor. Phys.* **1** 61–99

——1978 The complex geometry of the natural world *Proc. Intern Congress Math. Helsinki* pp 189–94

——1983 Physical space-time and nonrealizable CR-structures *Proc. Symp. Pure Math.* **39** 401–22

Petrov A Z 1954 Classification of spaces defined by gravitational fields (in Russian) *Uch. Zapiski Kazan Gos. Univ.* (Sci. Not. Kazan State Univ.) **114** 55–69

Robinson I 1961 Null electromagnetic fields *J. Math. Phys.* **2** 290–1

——and Trautman A 1960 Spherical gravitational waves *Phys. Rev. Lett.* **4** 431–2

————1986 Cauchy-Riemann structures in optical geometry *Proc. of the Fourth Marcel Grossmann Meeting on General Relativity (Rome)* ed R Ruffini (Amsterdam: Elsevier) pp 317–324

————1989 Optical geometry *New theories in physics: Proc. XI Warsaw Symp. on elem. particle phys. (Kazimierz)* ed Z Ajduk *et al* (Singapore: World Scientific) pp 454–97

Rosay J P 1989 New examples of non-locally embeddable CR structures (with no non-constant distributions) *Ann. Institut Fourier (Grenoble)* **39** 811-23

Tafel J 1973 A non-singular homogeneous universe with torsion *Phys. Lett.* **45**A 341–2

——1985 On the Robinson theorem and shear free geodesic null congruences *Lett. Math. Phys.* **10** 33–9

Trautman A 1958 *Lectures on general relativity* mimeographed notes (London: King's College)

——1973a On the structure of the Einstein-Cartan equations *Symp. Math.* **12** 139–62

——1973b Spin and torsion may avert gravitational singularities *Nature* **242** 7–8

——1978 Motion and radiation according to the theory of general relativity *Leopold Infeld: His Life and Scientific Work* ed E Infeld (Warszawa: Polish Scientific Publishers) pp 17–31

——1984 Deformations of the Hodge map and optical geometry *J. Geom. Phys.* **1** 85–95

——1985 Optical structures in relativistic theories, *Astérisque* numéro hors série 401–20

——1996 A metaphysical remark on variational principles *Acta Phys. Polon. B* **27** 839–84

Van Nieuwenhuizen P1981 Supergravity *Physics Reports* **68** 189–398

Weyl H 1918 Reine Infinitesimalgeometrie *Math. Zeitschrift* **2** 384–411

——1950 A remark on the coupling of gravitation and electron *Phys. Rev.* **77** 699–701

Weyssenhoff J and Raabe A 1947 Relativistic dynamics of spin-fluids and spin-particles *Acta Phys. Polon.* **9** 7–18

Wu T T and Yang C N 1975 Concept of nonintegrable phase factors and global formulation of gauge fields *Phys. Rev.* D **12** 3845–57