DIFFERENTIAL GEOMETRY AND ITS APPLICATIONS
Proc. Conf. Prague, August 30 – September 3, 2004
Charles University, Prague (Czech Republic), 2005, 487 – 498

Geometric Foundations of Classical General Relativity Theory

G.S. Hall

ABSTRACT. This paper attempts to give a brief introduction to the geometrical ideas which led to Einstein's general theory of relativity. It starts by discussing the principles of equivalence and covariance and how, following the success of the special theory of relativity, they can be used to suggest a theory of the gravitational field based on Riemann's geometry with a Lorentz metric. Following this, Einstein's field equations are introduced and the theory is interpreted within Lorentzian geometry. Some important exact solutions of Einstein's field equations are briefly described and some of the techniques used for finding further solutions are introduced, including a geometrical approach to the study of symmetries and a brief account of the Petrov classification. An introduction to the comparative study of metric, connection and curvature is also included. Finally, a short account of causality in space-times is given.

1. Geometry and Physics

The work of Lobachevski and Bolyai showed how a geometry other than Euclid's could arise and the later development of the axiomatic system in geometry by Hilbert and others revealed rather clearly how the difference between these two structures could be reduced to the parallel axiom. Riemann's researches in the intervening period, of course, laid the foundations of differential geometry. It is, perhaps, remarkable that these developments in mathematics arose at a time when physics was going through a crisis and that they would be just what Einstein required to create his general theory of relativity which, in turn, was to resolve many of the problems threatening the foundations of physics. It is always an interesting point of discussion whether academic intercourse between mathematicians and physicists inadvertently, or otherwise, sometimes guides mathematics in the direction of physical problems or whether physicists with only the mathematics available to them at the time and which may, if unsuitable, lead them astray, have often chosen wisely. Whichever is the case, differential geometry led Einstein to the most successful theory of gravitation so far developed.

Of course, the work of Lagrange, Hamilton and others following them had shown how analytic Newtonian mechanics could be handled in a rather elegant way using Riemann's geometrical ideas. The advent of the Lagrangian formulation and the tensor calculus in arbitrary coordinate systems allowed classical theory to be described formally in a "generally covariant" way; a criterion to be emphasised later by Einstein in his general theory of relativity. However, there is a significant conceptual

This paper is in final form and no version of it will be published elswhere.

difference in this feature between classical theory and Einstein's theory. In the classical case, Euclidean geometry controlled the arena in which all physical problems were considered but itself remained unchanged and immutable by very definition. Newton's absolute space thus plays a significant role in classical physics but is itself unaffected by physics. However, the notation of the tensor calculus concealed the fact that physics had not been truly "geometrised" and that Euclidean geometry still held an absolute position in its foundations.

The asymmetrical roles played by the background space and by physics (and which has been commented on by several authors [**1**], [**2**], [**3**]) were largely resolved by Einstein since, for him, the space (and time) geometry on the one hand and the physics on the other undergo mutual interaction. In fact, one may say, in a loose way, that the geometry ***is*** the physics (and vice versa). This allowed Einstein (after much effort) to write down the "field equations" of general relativity in a covariant manner and which (excluding the gravitational source terms) contain only the metric tensor (representing the geometry of space-time) and its first and second derivatives. In this sense, Einstein's theory contains no "absolute" variables but only the ***dynamical*** (metric) variable. [It is remarked here that, since the connection in Einstein's theory is assumed to be torsion-free, it may be argued that this zero torsion constitutes an absolute variable [**3**].]

The principle of covariance, as described briefly above, was an important guideline for Einstein. Another principle, the ***principle of equivalence*** was also given prominence by him. The reliance of Newtonian theory on absolute space necessarily means that this latter theory claims the ability to distinguish between inertial and non-inertial frames and hence between "true" gravitational forces and the inertial forces which arise when the observer's frame accelerates with respect to an inertial frame. Now, there is compelling experimental evidence that the gravitational interaction results in an ***acceleration field*** which affects all bodies equally, irrespective of their composition (that is, given a body's initial location in space and time and its initial velocity, its path is then determined independently of the body) and, what is more, shares this indiscriminate behaviour with the "inertial" interaction. In fact Einstein, in his "lift" experiment, drew attention to the fact that a gravitational force on an observer at a certain event (that is, a certain point in space at a certain epoch in time) could be simulated by a particular acceleration of the observer in a region of no gravitational field. Thus, building on the success of the special theory of relativity, if one regards the latter acceleration above as an acceleration of an inertial observer in special relativity, this would result in the Minkowski metric η at the event P in question undergoing an accelerative coordinate transformation to, say, the metric g at P. But this simulation of the gravitational field will, in general, require a different accelerative balance at different events and so the accelerative coordinate transformation that changed the components of η to those of g will depend on P. It follows that, although the curvature tensor arising from η through its unique symmetric Levi-Civita connection in special relativity is zero, the corresponding curvature tensor for g constructed as above in some neighbourhood of P is ***not***, in general, zero. Thus one is led to considering a ***non-flat*** metric on a space-time as, in some sense, representative of the gravitational field and which then plays a role similar to the Newtonian gravitational potential φ.

The function φ is controlled by the equations of Laplace and Poisson. Einstein sought for some time the equivalent equations to restrict the metric and finalise his theory. The essential result is that, in general relativity, space-time is a 4-dimensional Hausdorff connected manifold M admitting a Lorentz metric g of signature $(-,+,+,+)$, the latter being restricted by Einstein's field equations (see, e.g., [**4**], [**5**]).

$$\text{Ricc} - \frac{1}{2}Rg + \Lambda g = \kappa \mathcal{T} \tag{1}$$

where Ricc is the Ricci tensor, R the Ricci scalar, κ a constant, Λ the ***cosmological constant*** and $\mathcal{T}$ the energy-momentum tensor; a symmetric type $(0,2)$ tensor which describes the gravitating matter and energy contant of space-time, together with boundary conditions and other restrictions imposed to make the solution in some sense "physically realistic" [**4**]. The tensor $\mathcal{T}$ satisfies the "conservation law" ${\mathcal{T}_a}^b{}_{;b} = 0$ where a semi-colon denotes a covariant derivative with respect to the Levi-Civita connection of g (which follows by applying the (contracted) Bianchi identity to (1)). Frequently, Λ is set equal to zero. The 4-dimensionality of M arises from experience and from the success of the 4-dimensional space-time of special relativity theory. The Hausdorff condition is also assumed to arise from experience (or mathematical expediency). The connectedness assumption arises from the fact that "communication" in general relativity results from information travelling through space-time on a certain type of path in M (see section 4). If M were not path connected (which, for a manifold, is equivalent to M not being connected) one component of M would have no communication with another and, in this sense, M would have no physics. These conditions laid down for M in fact force M to be paracompact [**6**] and so M admits a reasonable mathematical structure.

2. The Interpretation of Einstein's Theory

To understand the physical interpretation of a solution g of (1) consider, in a general space-time M, the tangent space T_mM to M at $m \in M$ and the metric $g(m)$. The pair (V, γ) where $V = T_mM$ and $\gamma = g(m)$ together result in a ***Lorentz inner product space***. If $v \in T_mM$, $v \neq 0$, v is called ***spacelike*** (respectively, ***timelike*** or ***null***) if $\gamma(v,v) > 0$ (respectively, $\gamma(v,v) < 0$ or $\gamma(v,v) = 0$) and so γ partitions V into the zero vector and the subsets S, T and N of spacelike, timelike and null vectors, respectively. Thus (V, γ) is, essentially, the Minkowski space of special relativity and the interpretation of timelike members of V being the potential space-time directions of a material particle passing through the event m is taken over into general relativity, as is the interpretation of null members of V being the potential space-time directions of photons at m.

A 1-dimensional subspace (a ***direction***) of T_mM is called ***spacelike*** (respectively, ***timelike***, ***null***) if it is spanned by a spacelike (respectively, timelike, null) vector. A 2-dimensional subspace of T_mM is called ***spacelike*** if all its non-zero members are spacelike, ***timelike*** if it contains two independent null directions and ***null*** if it contains a unique null direction. A 3-dimensional subspace of T_mM is called ***spacelike*** (respectively, ***timelike*** or ***null***) if its normal direction is timelike (respectively, spacelike or null). This classification is exclusive and exhaustive in each case. A vector field on M is then called ***spacelike*** if it is spacelike at each $m \in M$ and a submanifold of M of dimension 1, 2 or 3 is called ***spacelike*** if its tangent space is everywhere ***spacelike*** (and similarly for timelike and null).

Recalling the decomposition $T_mM = \{0\} \cup S \cup T \cup N$ $(= \mathbb{R}^4)$ one sees that in the usual topology on $\mathbb{R}^4$, S is open and connected and T is open and disconnected whilst N is a disconnected subspace. Now T can be written as $T = T_f \cup T_p$ where T_f and T_p are disjoint connected open subsets of T_mM which arise as the equivalence classes under the equivalence relation $\sim$ on T given by $u, v \in T$ and $u \sim v \Leftrightarrow \gamma(u, v) < 0$. One of these, T_f, is called the set of ***future pointing timelike vectors*** at m and the other, T_p, the set of ***past pointing timelike vectors*** at m. The set N splits up in a similar way. From thermodynamics, it is usual to impose on M the condition that it is ***time-orientable***. Intuitively, this means that one may make a smooth choice of T_f at each $m \in M$. More precisely, M is time-orientable if the association $m \to T_f(m)$ may be made at each $m \in M$ such that for any $m' \in M$ there exists $v \in T_f(m')$, an open neighbourhood U of m' and a vector field X on U such that $X(m') = v$ and $X(m'') \in T_f(m'')$ for each $m'' \in U$. If such is the case then for any $m' \in M$ and ***any*** $v \in T_f(m')$ such a local future pointing extension X of v exists. In fact, M is time orientable if and only if M admits a global nowhere zero timelike vector field and if M is simply connected it is necessarily time orientable [**7**].

A special case of (1) is a ***vacuum space-time***, when $\mathcal{T}$ is chosen to be zero. In this case (and with Λ set equal to zero) (1) is equivalent to $\mathrm{Ricc} = 0$. In such a space-time, Einstein assumed that a spherically symmetric, chargeless and non-spinning test particle followed a path in spacetime which was a geodesic of the Levi-Civita connection associated with g whose tangent vector was everywhere timelike (or zero). There have been several attempts to ***prove*** this under reasonable conditions from the remainder of the physical and mathematical "axioms" of general relativity (including attempts by Einstein and his collaborators) and with some success (see [**4**]). It is perhaps interesting to see a traditional solution of (1) and to see that the solution actually involves finding g ***and*** M. That is, although one might have an intuitive feel for what M should be in certain situations, the actual solution g may not allow such an M. Thus the solution is ***for the whole space-time geometry***. A comparison with classical electromagnetism may be helpful. If one solves Maxwell's equations for a single stationary point charge e and selects the origin O of polar coordinates at the charge, one finds a solution where the electric and magnetic fields $\mathbf{E}$ and $\mathbf{H}$ satisfy $\mathbf{E} = \frac{ke}{r^2}$ (k constant) and $\mathbf{H} = 0$. This solution is interpreted by saying that $\mathbf{E}$ is not defined at O but that there are no problems with (the Euclidean) space or the time coordinate at O, that is, there is no problem with the ***geometry*** at O.

In the first and still very important exact solution of Einstein's vacuum equation (and with $\Lambda = 0$), Schwarzschild attempted to find the time-independent gravitational field due to a single spherically symmetric gravitating body of mass m. His solution, in a coordinate system chosen to be reminiscent of polar coordinates r, θ, φ, t, is written in the traditional relativistic form $ds^2 = g_{ab}dx^a dx^b$ (and using Einstein's summation convention) as

$$ds^2 = -\left(1 - \frac{2m}{r}\right) dt^2 + \left(1 - \frac{2m}{r}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta \, d\varphi^2). \tag{2}$$

Here one must carry the restriction $r > 2m$. However this is a defect of the coordinate system (just as the polar coordinate chart in the Euclidean plane does not cover the non-positive x-axis). The coordinate transformation of (2) to coordinates r', θ', φ',

t', where $t' = r + t + 2m\log(r - 2m)$, $r' = r$, $\theta' = \theta$, $\varphi' = \varphi$ gives

$$ds^2 = -\left(1 - \frac{2m}{r}\right) dt'^2 - 2dt'dr + r^2(d\theta^2 + \sin^2\theta\, d\varphi^2) \tag{3}$$

and the metric is defined for $r > 0$. However the solution cannot be extended to $r = 0$ since it is easily checked that the curvature tensor components get arbitrarily large as $r \to 0$. In this sense solving (1) involves finding g ***and*** M and the space-time manifold is not defined at $r = 0$, in contrast to the example from classical electromagnetism in the previous paragraph.

If one regards the solution (2) or (3) as the gravitational field of a sun, the problem of finding planetary orbits involves finding the appropriate timelike geodesics of this metric. This can be done, using an approximation technique, in a straightforward manner and one consequence of the solution was the immediate resolution of a long standing problem concerning the orbit of the planet Mercury. In a similar way the null geodesics were calculated and the phenomenon of light bending was theoretically predicted and ultimately experimentally confirmed. This solution also forms the basis for a simple model of a black hole.

Another rather important exact solution of the field equations (1) occurs when $\mathcal{T} \neq 0$. It is intended to represent a general smoothed out approximation to the whole universe and evolved through the work of Einstein, Friedmann, Robertson, Walker and Lemaître. In this model the gravitating matter is regarded as a fluid, the particles of which are galaxies or "gravitationally bound" clusters of galaxies. This is represented by writing the energy-momentum tensor $\mathcal{T}$ in (1) as

$$\mathcal{T}_{ab} = (p + \rho)u_a u_b + p g_{ab} \tag{4}$$

where u is the tangent vector field (4-velocity) of the particles (so that $g(u, u) = -1$), ρ is the energy-density and p the isotropic pressure associated with the fluid. The situation is simplified by assuming the cosmological hypothesis, that is, the solution is spatially homogeneous and isotropic (but not "time independent" so as not to conflict with the expansion of the universe observed by astronomers). In mathematical language this is usually understood to mean that M admits a 6-dimensional Lie algebra of Killing vector fields representing this symmetry and which lie in 3-dimensional spacelike orbits (see section 3). It then turns out that u is a normal geodesic vector field which is everywhere orthogonal to these orbits and that these orbits with their induced geometry are spaces of constant curvature. One then chooses local coordinates t, χ, θ, φ with t an affine parameter along the geodesic integral curves of u (***cosmic time***) and χ, θ, φ coordinates in the hypersurfaces (orbits) of constant t. Thus the coordinates χ, θ and φ are constant along an integral curve of u (a "galactic path") and p and ρ depend at most on t. The metric is

$$ds^2 = -dt^2 + R^2(t)\left[d\chi^2 + f^2(\chi)(d\theta^2 + \sin^2\theta\, d\varphi^2)\right] \tag{5}$$

where $f(x)$ is $\sin\chi$, χ or $\sinh\chi$ when the (normalised) constant curvature K of the hypersurface is $+1$, 0 or -1, respectively. The field equations (1) then reduce to differential equations for $R(t)$ and the behaviour of the metric (5) then depends on the choice of K and Λ. Many of these models have a "beginning" in time (a "big bang") and once again reveal how (1) restricts M as well as g.

Further details of these and other exact solutions of Einstein's field equations may be found in [**4**], [**5**].

3. Some Techniques for Finding Exact Solutions

The difficulties involved in solving (1) are clear. To solve them one must make some simplifying assumptions. Two of these can be briefly discussed here.

Historically, the first (and, perhaps, most natural) is to assume the space-time possesses some form of symmetry. Usually, this means assuming the metric tensor g admits a symmetry and is taken as the existence of a global smooth Killing vector field X on M so that, with $\mathcal{L}$ denoting a Lie derivative, $\mathcal{L}_X g = 0$. In components, this is equivalent to Killing's equations

$$X_{a;b} + X_{b;a} = 0 \quad (\Leftrightarrow X_{a;b} = F_{ab} = -F_{ba}) \tag{6}$$

where F is the ***Killing bivector***. The collection $K(M)$ of all Killing vector fields on M is a finite-dimensional Lie algebra under the Lie bracket operation called the ***Killing algebra***. The appeal of Killing symmetries in physics and geometry arises from the fact that given $X \in K(M)$ then for each $m' \in M$ there exists a neighbourhood U of m' and a real number $\varepsilon > 0$ such that for each $m \in U$ there is an integral curve c of X starting from m (that is, $c(0) = m$) and defined on $(-\varepsilon, \varepsilon)$. Then for $|t| < \varepsilon$ the bijection $\varphi_t : U \to \varphi_t(U)$ which moves each point of U a parameter distance t along the integral curve of X through that point is a smooth ***local isometry*** on M ($\varphi_t^* g = g$ where φ_t^* is the ***pullback*** of φ_t). Thus, in the component description loved by relativists, if U is chosen as a coordinate domain with coordinate functions x^a, the metric components g_{ab} at $m \in U$ in this coordinate system equal those at $\varphi_t(m)$ in the coordinate system $y^a = x^a \circ \varphi_t^{-1}$. The Schwarzschild solution (2) admits a 4-dimensional Killing algebra containing a 3-dimensional subalgebra isomorphic to $so(3)$ representing spherical symmetry and another independent Killing vector field $\frac{\partial}{\partial t}$ representing the time independence. The Killing algebra for the metric (5) has already been described.

A major feature of this metric symmetry concept in the finding of exact solutions (in addition to the intuitive value) is the use of the associated orbits of $K(M)$ in the construction of coordinate systems. Let $X_1, \cdots, X_k \in K(M)$ and let $\varphi_t^1, \cdots, \varphi_t^k$ be the associated local isometries. Consider all local maps f (where defined) of the form

$$m \to f(m) \equiv \varphi_{t_1}^1 \left(\varphi_{t_2}^2 (\cdots \varphi_{t_k}^k (m) \cdots) \right) \quad (m \in M) \tag{7}$$

for appropriate $t_1, \cdots, t_k$. Define an equivalence relation $\sim$ on M by $m_1 \sim m_2 \Leftrightarrow f(m_1) = m_2$ for some f of the form (7). The associated equivalence classes of $\sim$ are the ***orbits*** of $K(M)$. One can also define the ***generalised distribution*** Δ associated with $K(M)$ as the map $m \to \Delta(m) \subseteq T_m M$ where

$$\Delta(m) = \{X(m) : X \in K(M)\}. \tag{8}$$

If $\dim \Delta(m)$ is constant on M, Δ is a distribution of the Fröbenius type and ***each orbit is known to admit the structure of a submanifold of*** M ***and is then an integral manifold of*** Δ. This means that if O is such an orbit and $m \in O$, then the tangent space to O at m equals $\Delta(m)$. But it may be that $\dim \Delta(m)$ is not constant on M. However, the more recent work of Hermann [**8**] and Sussmann [**9**] shows that if $K(M)$ is finite-dimensional then ***the above italicised result is still true***. It is also remarked that, in spite of statements in the literature to the contrary, the maps (7) above may not result from a Lie group action on M. This arises if and only if each member of $K(M)$ is a ***complete*** vector field [**10**].

Two technical points arise here. Suppose M' is any manifold and f a smooth map $M' \to M$ into the space-time M and whose range lies in an orbit O of $K(M)$. Then is the obvious associated map $\tilde{f} : M' \to O$ smooth? If O is a ***regular submanifold*** (that is, the manifold topology of O equals its induced topology from the manifold topology of M) then $\tilde{f}$ is easily shown to be smooth [**11**]. Otherwise, this result is not obviously true (and orbits associated with Lie algebras of vector fields need not necessarily be regular). The smoothness of $\tilde{f}$ is rather crucial for the resulting calculus required. Fortunately the work of Stefan [**12**] shows that whilst orbits are not necessarily regular, they have the special property of being ***leaves*** and the above map $\tilde{f}$ ***is*** smooth. Another point concerns certain desirable features of orbits which are often assumed in the relativistic literature but which may be false. For example, if $\{X_1, \cdots, X_k\}$ is an independent set of members of $K(M)$ and O an orbit of $K(M)$, with inclusion $i : O \to M$, then each member of this set is tangent to O and induces vector fields $\{\tilde{X}_1, \cdots, \tilde{X}_k\}$ in O. In fact, it is easily checked that, since O is an orbit of $K(M)$, it has the same nature (spacelike, timelike or null) at each point and if O is not null these latter vector fields on O are Killing with respect to the induced metric on O from g. ***Is this latter set an independent set of vector fields on*** O? In fact it is not necessarily the case. So call an orbit O ***dimensionally stable*** if for each $m \in O$ there is an open neighbourhood U of m such that the orbit through any $m' \in U$ has the same dimension as O. For dimensionally stable orbits the answer to the above italicised question is "yes" and from it follows a number of convenient results relating orbit dimension and the dimension of $K(M)$ [**13**],[**14**].

Space does not permit discussion of the other symmetries considered in Einstein's theory, for example, homothetic, general affine, conformal and projective symmetry together with the more complicated curvature symmetries, nor of the associated isotropies. Details can be found in [**14**]. In fact the study of affine and projective symmetry can, with profit, be undertaken in connection with the ***holonomy group*** structure of space-time [**14**],[**15**].

The second technique invoked in facilitating the solution of (1) is the ***Petrov classification*** of the Weyl tensor [**16**]. The curvature tensor components R_{abcd} may be decomposed in the following way

$$R_{abcd} = C_{abcd} + E_{abcd} + \frac{R}{12}(g_{ac}g_{bd} - g_{ad}g_{bc}) \tag{9}$$

where C is the Weyl tensor, E is the tensor with components

$$2E_{abcd} = \tilde{R}_{ac}g_{bd} - \tilde{R}_{ad}g_{bc} + \tilde{R}_{bd}g_{ac} - \tilde{R}_{bc}g_{ad} \quad (\tilde{R}_{ab} = R_{ab} - \frac{1}{4}Rg_{ab}) \tag{10}$$

and $\tilde{R}$ is the tracefree Ricci tensor. The tensor E is algebraically determined by $\tilde{R}$ and, in fact, algebraically determines it since, from (10), $\tilde{R}_{ab} = E^c{}_{acb}$. Thus, in a very simplistic way, one can think of the last two terms on the right hand side of (9) as arising from the "gravitating matter" contribution to the space-time curvature and the Weyl tensor as a "vacuum" contribution to this curvature. Petrov's classification is an algebraic classification of C and can be described in several different ways of which the following is quite straightforward. Let Ω_m denote the vector space of type $(2,0)$ skew-symmetric tensors (2-***forms***) at m (so that $\Omega_m \approx \mathbb{R}^6$). Then let V ($\approx \mathbb{C}^6$) be the complexification of Ω_m. Thus V is the 6-dimensional complex vector space of "complex 2-forms" at m. Finally, define the linear map $f : V \to V$ given in components by

$$(F \in V) \quad f : F^{ab} \to C^{ab}{}_{cd} F^{cd}. \tag{11}$$

The idea is to study the eigenvector-eigenvalue structure of the linear map f taken as a representative of C at m. To do this one notes the ***self-dual*** property of the Weyl tensor given by ${}^*C = C^*$ where $*$ denotes the ***Hodge duality operator*** over a skew pair of indices. (This follows immediately from the trace-free property of C and the well-known Ruse-Lanczos identity - see, e.g. [**14**].) Extending the operator $*$ in an obvious way to V by linearity, define a member F of V to be ***self dual*** (respectively, ***anti-self dual***) if $\overset{*}{F} = -iF$ (respectively, $\overset{*}{F} = iF$). Let S^+ (respectively, S^-) denote the subset of self dual (respectively, anti-self dual) members of V. Then it is easily checked that S^+ and S^- are subspaces of V and that $V = S^+ \oplus S^-$. It can then be seen that f maps S^+ to itself and S^- to itself. The obvious restrictions of f yield linear maps $S^+ \to S^+$ and $S^- \to S^-$ whose Jordan-Segre types, including degeneracies, are identical and whose eigenvalues differ only by conjugation. This common Jordan-Segre type is then called the ***algebraic type of*** C. The possibilities for this, in Segre notation, are $\{111\}$, $\{21\}$ and $\{3\}$ and these are the Petrov types **I**, **II** and **III**, respectively. Type **I** has a degeneracy $\{1(11)\}$ (Petrov type **D**) as does type **II** ($\{(21)\}$ - Petrov type **N**). These, together with the trivial case when $C = 0$ at m, constitute the Petrov types for C. It should be noted that the tracefree condition $C^a{}_{bad} = 0$ on C results in all eigenvalues in the Petrov types **N** and **III** cases being zero.

Many powerful techniques have evolved for finding exact solutions of Einstein's equations where one restricts a solution to having a particular Petrov type. This is particularly convenient since these types often have rather elegant physical interpretations. Further details of the application of these techniques can be found in [**5**].

The energy-momentum tensor T may also be algebraically classified as a type $(0, 2)$ symmetric tensor by its Segre-Jordan type. The details, physical interpretation and some history of this can be found in [**14**].

4. Causal Structure

Let M be a time orientable space-time with metric g. This section closely follows the work of Penrose [**17**]. A smooth path is called ***timelike*** if its tangent vector is timelike everywhere and ***causal*** if its tangent vector is timelike or null everywhere. A timelike or causal path is called ***future oriented*** (respectively ***past oriented***) if its tangent vector is everywhere future pointing (respectively, past pointing). If $c : I \to M$ is a path in a space-time M and with I some open or closed interval of $\mathbb{R}$ with $a = \inf I$, $b = \sup I$ (and possibly $a = -\infty$ or $b = \infty$) then x is called an ***endpoint*** of c if, for each sequence $x_n \in I$ with $x_n \to a$, then $c(x_n) \to x$ or, for each sequence $x_n \in I$ with $x_n \to b$, then $c(x_n) \to x$. All timelike and causal curves will be required to contain their endpoints and thus a timelike (respectively, causal) curve must be (extendible as a) timelike (respectively, causal) curve at its endpoints. In the above notation, if c is timelike or causal and future oriented, the first of the above limits is called a ***past endpoint*** of c and the second a ***future endpoint*** of c.

Now define two relations $<<$ and $<$ on M by [**17**] $x, y \in M$, $x << y \Leftrightarrow$ there exists a piecewise smooth future oriented affinely parametrised timelike geodesic with past endpoint x and future endpoint y (and say x ***chronologically precedes*** y) and

$x < y \Leftrightarrow$ either $x = y$ or there exists a piecewise smooth future oriented affinely parametrised causal geodesic with past endpoint x and future endpoint y (and say x ***causally precedes*** y). Note that, by definition, $x < x$, $\forall\, x \in M$ but $x << x$ means there is a ***closed*** piecewise future oriented affinely parametrised timelike geodesic in M whose past and future endpoints are equal. It can, in fact, be shown [**17**] that $x << y \Leftrightarrow$ there exists a timelike curve from x to y and that $x < y \Leftrightarrow$ either $x = y$ or there exists a causal curve from x to y.

Now define the following subsets of M.

$I^+(x) = \{y \in M : x << y\}$ — the ***chronological future of*** x,
$I^-(x) = \{y \in M : y << x\}$ — the ***chronological past of*** x,
$J^+(x) = \{y \in M : x < y\}$ — the ***causal future of*** x,
$J^-(x) = \{y \in M : y < x\}$ — the ***causal past of*** x.

If $x, y \in M$, define the subset $\langle x, y\rangle$ of M by

$$\langle x, y\rangle = I^+(x) \cap I^-(y).$$

It can be shown that $I^+(x)$, $I^-(x)$ and hence $\langle x, y\rangle$ are open subsets of M. It can also be shown that if $x, y, a, b \in M$ and if $z \in \langle x, y\rangle \cap \langle a, b\rangle$ then there exist $p, q \in M$ such that $z \in \langle p, q\rangle \subseteq \langle x, y\rangle \cap \langle a, b\rangle$. Thus the collection of subsets of M of the form $\langle x, y\rangle$, $x, y \in M$, is a basis for a topology on M called the ***Alexandrov topology*** [**18**] for M and which is contained in the manifold topology for M. These topologies clearly agree on Minkowski space $(\mathbb{R}^4, \eta)$ but if one takes the portion $0 \leqslant t \leqslant 1$ of Minkowski space with $(x, y, z, 0)$ identified with $(x, y, z, 1)$ for each $x, y, z \in \mathbb{R}$ then the Alexandrov topology is indiscrete. There is a sense in which the failure of the Alexandrov topology to equal the manifold topology on M reflects some kind of physically unreasonable behaviour in the causal structure of M. For example, if M admits a closed timelike curve and a and b are distinct points of M lying on that curve then any Alexandrov basis member containing either a or b necessarily contains both a and b. Thus the Alexandrov topology fails the separation axiom T_0 and hence differs from the manifold topology on M.

For philosophical and other reasons one would not want M to admit closed causal curves (as, for example, occurs in the above (identified) version of Minkowski space) or even "nearly" closed causal curves. Various conditions have been imposed on M to make it causally well behaved. To investigate one of these let an open subset U of M be called ***causally convex*** if no piecewise smooth future oriented affinely parametrised timelike geodesic intersects U in a disconnected set. Then M is called ***strongly causal*** if for each $m \in M$ and each open neighbourhood V of m there exists a causally convex set U such that $m \in U \subseteq V$ [**17**]. It is clear that if M is strongly causal it admits no closed causal curves but the converse is false (the left-hand diagram in fig.1 gives an example of a space-time with no closed causal curves but which is not strongly causal - in fact it possesses an "almost" closed causal curve). One pleasing feature of strong causality is its relationship to the Alexandrov topology as seen in the following theorem due to Penrose [**17**].

THEOREM 1. *The following three conditions on a simply connected space-time are equivalent:*

(i) M is strongly causal
(ii) the Alexandrov topology coincides with the manifold topology
(iii) the Alexandrov topology is Hausdorff.

Thus in a strongly causal space-time the (manifold) space-time topology follows from the causal structure.

However a hierarchy of examples due to Carter [**19**] (see, e.g. the right-hand diagram in fig.1) shows that strong causality does not forbid "almost" closed causal curves (and so, in some sense, a small deformation of such a metric could admit a closed such curve). Hawking [**20**] got round this problem by introducing the concept of ***stable causality***. Using a Whitney topology on the bundle of metrics over M, a space-time M with metric g is called ***stably causal*** if g and no metric "close to" g admits a closed causal curve. One consequence of stable causality is that M necessarily admits a "cosmic time function" [**20**].

One rather elegant application of the Alexandrov topology for M is in proving the following theorem [**17**] (cf. [**21**]) and which reveals why one would not want a space-time M to be compact.

THEOREM 2. *If a space-time M is compact (in its manifold topology) it contains a closed timelike curve.*

5. Curvature Structure

From the space-time metric g on M one can construct the associated Levi-Civita connection Γ and its type $(1,3)$ curvature tensor $\mathcal{R}$. Each of the quantities g, Γ and $\mathcal{R}$ in some way "represent" the gravitational field on M and it is interesting to ask how tightly any one of these geometrical quantities determines the others. Clearly, g uniquely determines Γ and $\mathcal{R}$ and Γ uniquely determines $\mathcal{R}$. Using the Whitney topology on the bundle of metrics on M mentioned in the last section, one can also show that, generically, $\mathcal{R}$ or Γ determine g up to a constant conformal factor (that is, up to units of measurement) and hence that, generically, $\mathcal{R}$ determines Γ uniquely. These results can be found in [**22**]-[**25**] and are collected together in [**14**]. A useful technique for part of this work is the study of the holonomy group of a Lorentz manifold [**26**],[**27**],[**14**].

From the physical viewpoint, and assuming that test particles and photons follow timelike and null geodesics, respectively, one can also ask how tightly the geodesic structure of M (which can, in principle, be found experimentally) determines g. The answer is that g is determined up to a constant conformal factor.

One final remark may be made regarding the sectional curvature function. Let $m \in M$, let G_m denote the Grassmann manifold of 2-dimensional subspaces of T_mM and let $\bar{G}_m$ denote the open submanifold of G_m consisting of ***non-null*** members of G_m. If $\tilde{F} \in \bar{G}_m$ is spanned by $p, q \in T_mM$, define the sectional curvature function $\sigma_m : \bar{G}_m \to \mathbb{R}$ at m by

$$\sigma_m(\tilde{F}) = \frac{R_{abcd}p^a q^b p^c q^d}{(g_{ac}g_{bd} - g_{ad}g_{bc})p^a q^b p^c q^d} \tag{12}$$

It is clear that (12) depends only on $\tilde{F}$ and not on the spanning pair p and q and the restriction of σ_m to $\bar{G}_m$ is because the denominator in (12) vanishes if and only if $\tilde{F}$ is null. It turns out that, provided σ_m is not a constant function for any $m \in M$ then, generically, the prescription of σ_m at each $m \in M$ uniquely determines g [**28**],[**29**],[**14**],[**25**]. A crucial step in this argument arises from the fact that the continuous extension of σ_m to at least one member of $G_m \backslash \bar{G}_m$ forces σ_m to be defined, and a constant function, on the whole of G_m [**28**].

In summary, it is seen that g, Γ, R and the sectional curvature function are, generically, each equally representative of the gravitational field on M.

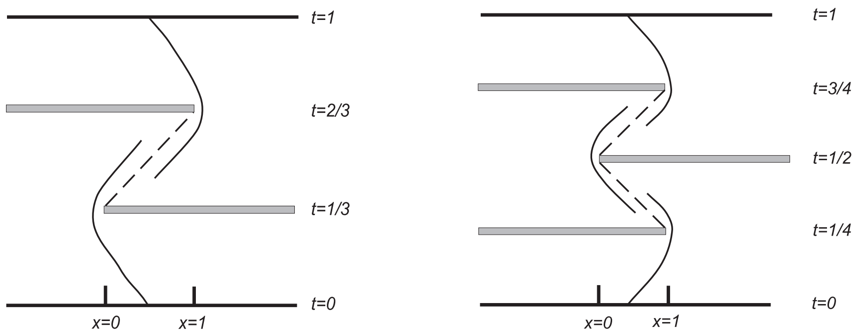

FIGURE 1. Each of the diagrams represents 2-dimensional Minkowski space with the points (x, t) and $(x, t+1)$ identified for each x and t. The lines $t = 1/3$, $x \geqslant 0$ and $t = 2/3$, $x \leqslant 1$ are removed in the left-hand diagram and the lines $t = 1/4$, $x \leqslant 1$, $t = 1/2$, $x \geqslant 0$ and $t = 3/4$, $x \leqslant 1$ are removed in the right-hand diagram. The null cone is always at 45° and the dotted lines are all null. "Almost" closed causal curves are indicated in each diagram. In the left-hand diagram strong causality fails at any point on the dotted line whereas, in the right-hand diagram, strong causality is satisfied. For each example, extension to 4-dimensions is clear.

Acknowledgements. The author wishes to thank the organisers of DGA9 for their kindness, hospitality and many valuable discussions during his stay in Prague.

References

[1] E. Kretschmann. ***Ann. Physik*** **55**, (1917), 575.
[2] J.L. Anderson. ***Gen. Rel. Grav.*** **2**, (1971), 161.
[3] A. Trautman in "Cosmology and Gravitation", ed. P.G. Bergmann and V. de Sabatta, Plenum Press, New York and London, 1980.
[4] S.W. Hawking and G.F.R. Ellis. "The Large Scale Structure of Space-Time", Cambridge, 1973.
[5] H. Stephani, D. Kramer, M.A.H. MacCallum, C. Hoenselears and E. Herlt. "Exact Solutions of Einstein's Field Equations", Cambridge, 2003.
[6] R.P. Geroch. ***J. Math. Phys.*** **9**, (1968), 1739.
[7] L. Markus. ***Ann. Maths.*** **62**, (1955), 411.
[8] R. Hermann. International Symposium on Nonlinear Differential Equations and Nonlinear Mechanics, Academic Press, New York, 1963, 325.
[9] H.J. Sussmann. ***Trans. Am. Math. Soc.*** **180**, (1973), 171.
[10] R.S. Palais. ***Mem. Am. Math. Soc.*** No.22, 1957.
[11] F. Brickell and R.S. Clark. "Differentiable Manifolds", Van Nostrand, 1970.
[12] P. Stefan. ***Proc. London Math. Soc.*** **29**, (1974), 699.
[13] G.S Hall. ***Class. Quant. Grav.*** **20**, (2003), 3745.
[14] G.S. Hall. "Symmetries and Curvature Structure in General Relativity", ***World Scientific***, 2004.
[15] G.S. Hall and D.P. Lonie. ***Class. Quant. Grav.*** **21**, (2004), 4549.
[16] A.Z. Petrov. "Einstein Spaces", Pergamon, 1969.
[17] R. Penrose. "Techniques of Differential Topology", Siam, (1972), No.7.
[18] A.D. Alexandrov. ***Vaprosy Filosofii*** No.1, **67**, 1959.
[19] B. Carter. ***Gen. Rel. Grav.*** **1**, (1971), 349.

[20] S.W. Hawking. ***Proc. Roy. Soc.*** A **308**, (1968), 433.
[21] R.W. Bass and L. Witten. ***Rev. Mod. Phys.*** **29**, (1957), 452.
[22] O. Kowalski. ***Math. Z.*** **125**, (1972), 129.
[23] G.S. Hall. ***Gen. Rel. Grav.*** **15**, (1983), 581.
[24] G.S. Hall. ***Gen. Rel. Grav.*** **20**, (1988), 399.
[25] A.D. Rendall. ***J. Math. Phys.*** **29**, (1988), 1569.
[26] H. Wu. ***Illinois J. Math.*** **8**, (1964), 291.
[27] G.S. Hall and D.P. Lonie. ***Class. Quant. Grav.*** **17**, (2000), 1369.
[28] G.S. Hall. ***Gen. Rel. Grav.*** **16**, (1984), 79.
[29] B. Ruh. ***Math. Z.*** **184**, (1985), 371.

Department of Mathematical Sciences
University of Aberdeen
Aberdeen AB24 3UE
Scotland, UK

e-mail: g.hall@maths.abdn.ac.uk