

Einstein memorial lecture.

Delivered at the Israel Academy
of Sciences and Humanities,
Jerusalem, Israel
March 21, 2006

**General covariance and the passive
equations of physics.**

Shlomo Sternberg

By the “**passive equations**” of physics I mean those equations which describe the motion of a small object in the presence of a force field where we ignore the effect produced by this small object. For example, Newton’s laws say that any two objects attract one another. But if we study the motion of a ball or a rocket in the gravitational field of the earth, we ignore the tiny effect that the ball or rocket has on the motion of the earth.

If we have a small charged particle in an electromagnetic field, the Lorentz equations describe the motion of the particle when we ignore the field produced by the motion of the particle itself.

To explain what I mean by “general covariance” will take the whole lecture.

The source of today's lecture is a late (1938) paper by Einstein, Infeld and Hoffman.

ANNALS OF MATHEMATICS
Vol. 39, No. 1, January, 1938

THE GRAVITATIONAL EQUATIONS AND THE PROBLEM OF MOTION

BY A. EINSTEIN, L. INFELD, AND B. HOFFMANN

(Received June 16, 1937)

Introduction. In this paper we investigate the fundamentally simple question of the extent to which the relativistic equations of gravitation determine the motion of ponderable bodies.

I was unable to find on the web a picture of E., I., &H. but here is a photo of Einstein, Infeld, and Bergmann from 1938.



theory of generalized functions) was developed. The person who extracted the key idea from this paper in the modern mathematical language was J. M. Souriau in 1974 who applied the EIH method to determine the equations of motion of a spinning charged particle in an electromagnetic field.

My purpose today is to explain how the E I H method as formulated for spinning particles by Souriau can be viewed as

The Souriau paper.

Ann. Inst. Henri Poincaré,

Vol. XX, n° 4, 1974, p. 315-364.

Section A :

Physique théorique.

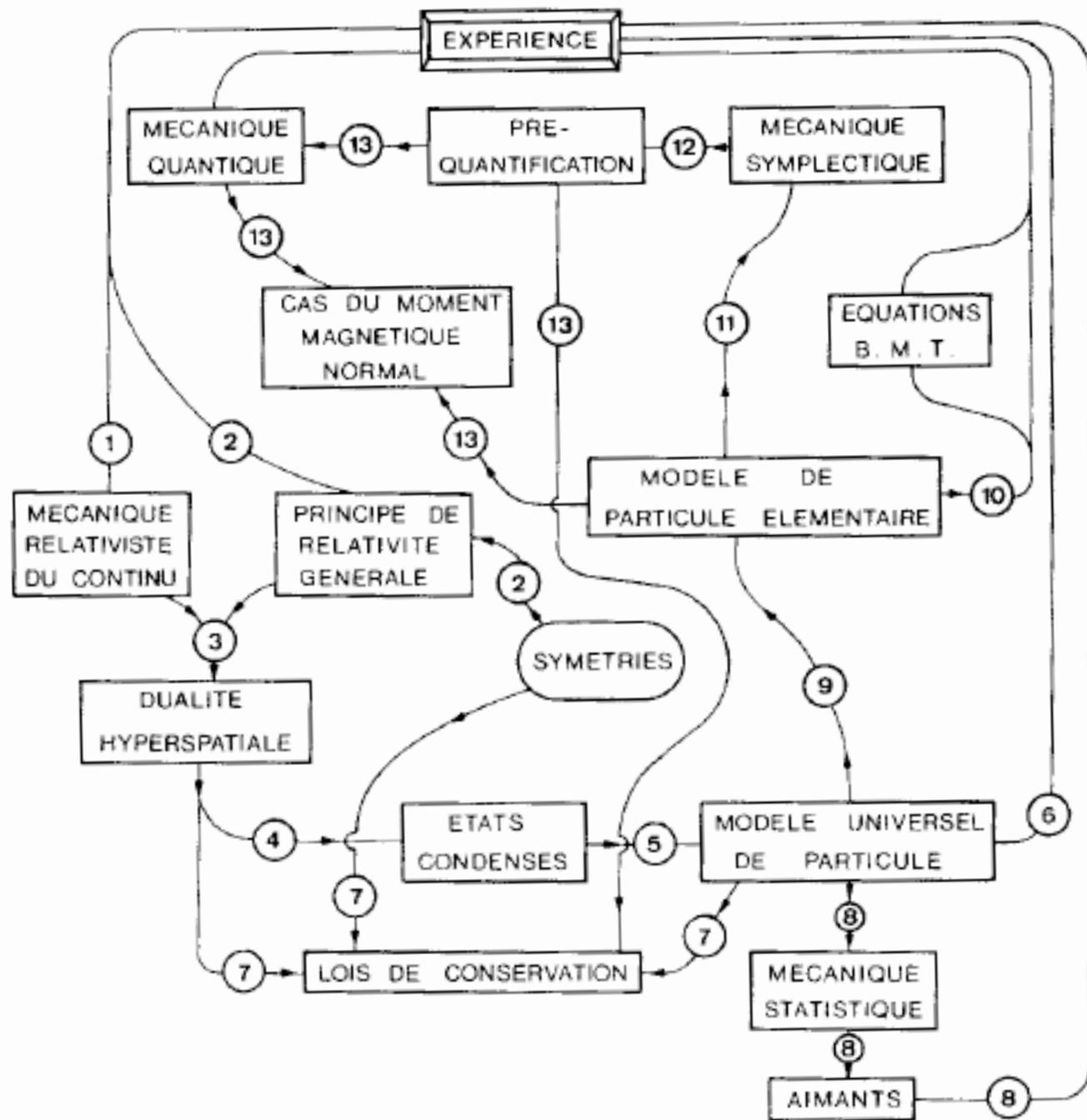
Modèle de particule à spin dans le champ électromagnétique et gravitationnel (*)

par

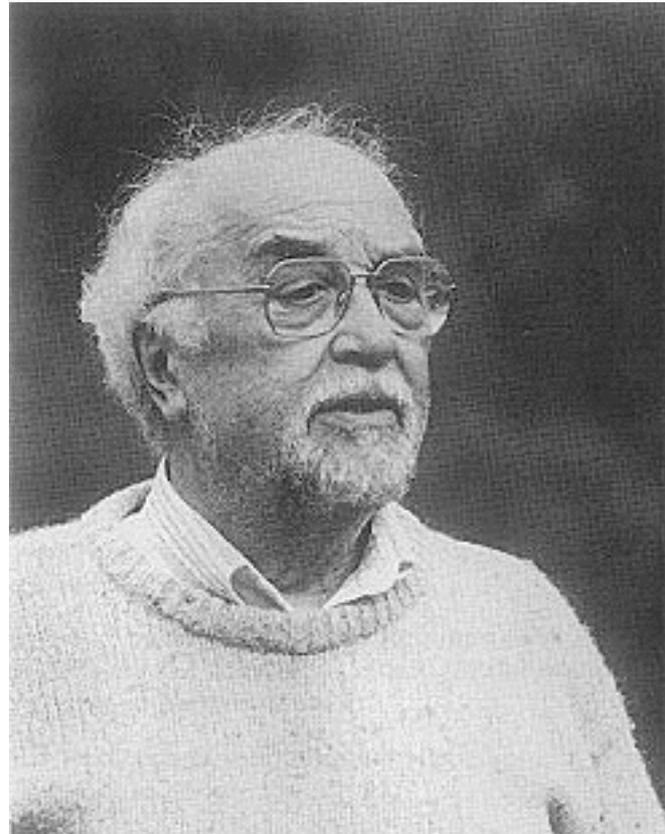
Jean-Marie SOURIAU ()**

Centre de Physique Théorique. C. N. R. S.,
31, chemin J.-Aiguier, 13274 Marseille Cedex 2 (France)

... paper is not an
easy read. He has a wonderful but
idiosyncratic mode of exposition.
For example, here is the flow
chart for the paper presented on



Jean Marie Souriau



Back to the E I H paper:

Introduction. In this paper we investigate the fundamentally simple question of the extent to which the relativistic equations of gravitation determine the motion of ponderable bodies.

What is this “fundamentally simple question” ?

The two principles of general relativity:

- The distribution of energy-matter determines the geometry of space time.
- A “small” piece of ponderable matter moves along a “geodesic” in the geometry determined as above. I will spend some time in today’s lecture explaining the meanings of the word “geodesic”.

THE QUESTION, THERE, HOWEVER
question is - what is the relation
(if any) between these two
principles? Many distinguished
physicists thought that these were

principle:

**People slowly accustomed
themselves to the idea that the
physical states of space itself**





Drawing by Rea Irvin; © 1929, 1957 The New Yorker Magazine

“People slowly accustomed themselves to the idea that the physical states of space itself were the final physical reality.”

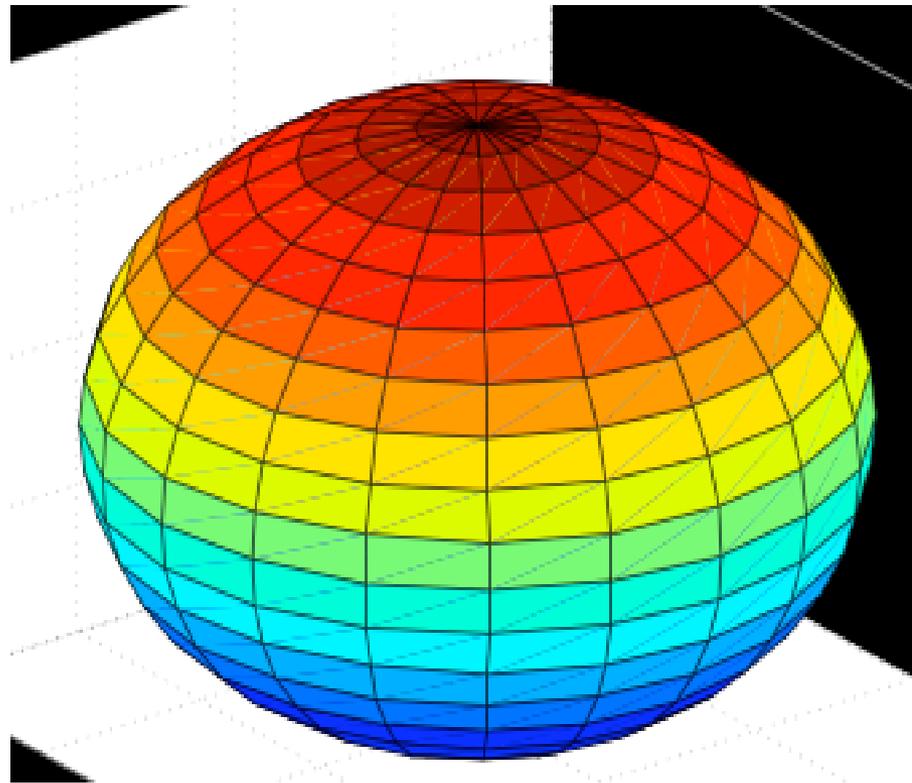
What is a geodesic?

Before the EIH paper and the Souriau paper there were several (equivalent) definitions of what a geodesic is. They all try to extend to more general geometries a characteristic property that straight lines have in Euclidean geometry:

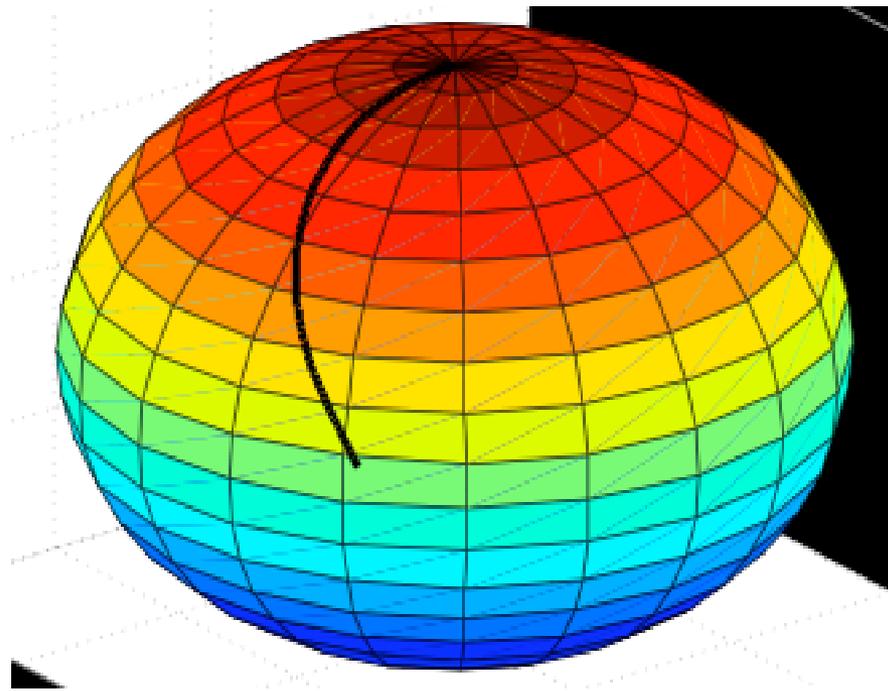
- A straight line is the “shortest distance between two points”.
- A straight line is “self-parallel” in the sense that it always points in the same direction at all its points. A curved line will (in general) be pointing in different directions at different points.

On a sphere, the shortest distance is a
piece of a great circle.

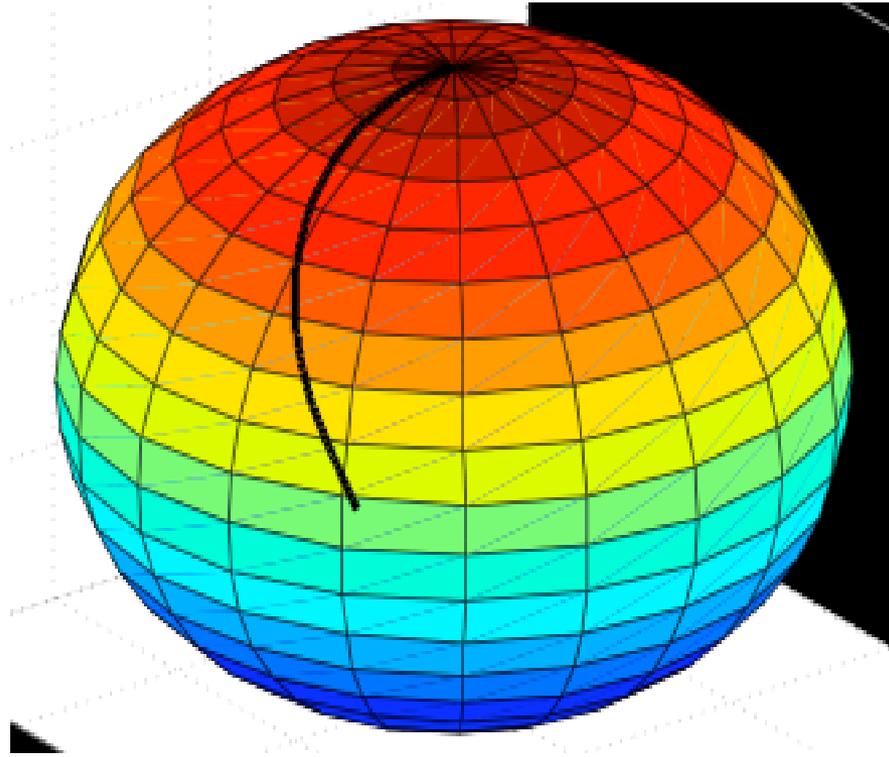
Here is a sphere drawn with
Matlab:



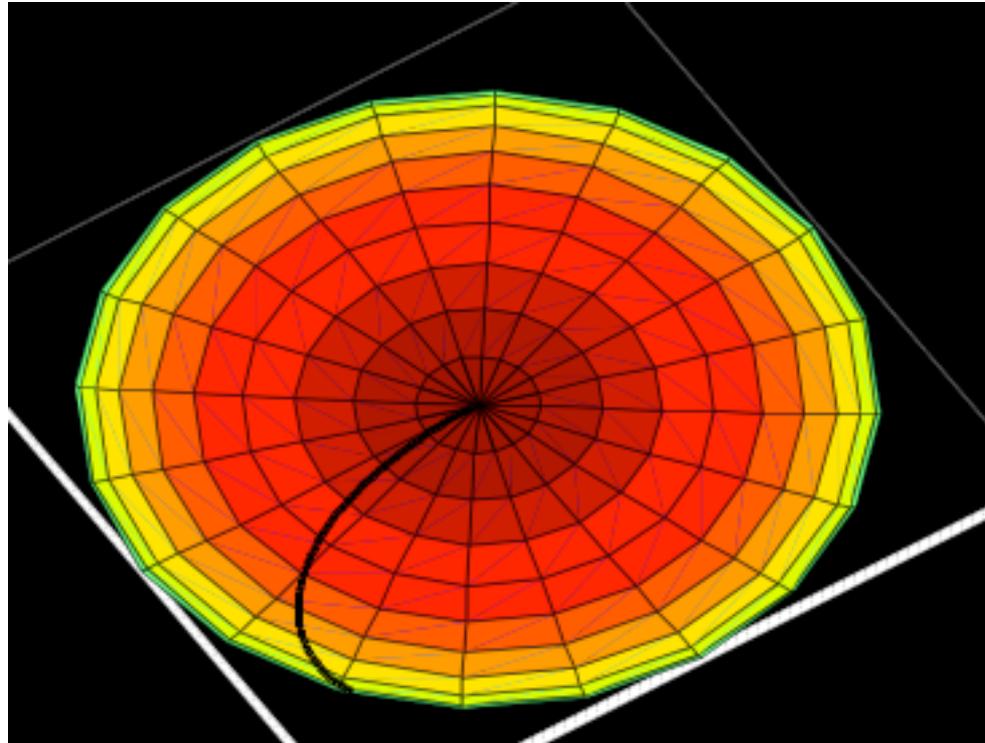
Here is a curve on the sphere
starting at the north pole.



Notice that the great circles emanating from the north pole (the circles of longitude) are consistently shorter than the corresponding piece of the curve.



View from the top:



Notice that from this point of view, the circles of longitude look almost like straight lines, and these lines are perpendicular to the circles of latitude. This is an illustration of a special case of what is known as **Gauss' lemma** although in a sense this was anticipated by al Biruni.

Abu Arrayhan Muhammad ibn Ahmad al-Biruni



Born: 15 Sept 973 in Kath, Khwarazm (now Kara-Kalpakskaya, Uzbekistan)
Died: 13 Dec 1048 in Ghazna (now Ghazni, Afghanistan)

The book [The history of cartography](#) details the mathematical contributions of al-Biruni. These include: theoretical and practical arithmetic, summation of series, combinatorial analysis, the rule of three, [irrational](#) numbers, ratio theory, algebraic definitions, method of solving algebraic equations, geometry, [Archimedes'](#) theorems, [trisection of the angle](#) and other problems which cannot be solved with [ruler and compass](#) alone, [conic sections](#), stereometry, [stereographic projection](#), trigonometry, the sine theorem in the plane, and solving spherical triangles.

Important contributions to geodesy and geography were also made by al-Biruni. He introduced techniques to measure the earth and distances on it using triangulation. He found the radius of the earth to be 6339.6 km, a value not obtained in the West until the 16th century. His Masudic canon contains a table giving the coordinates of six hundred places, almost all of which he had direct knowledge. Not all, however, were measured by al-Biruni himself, some being taken from a similar table given by [al-Khwarizmi](#). al-Biruni seemed to realise that for places given by both [al-Khwarizmi](#) and [Ptolemy](#), the value obtained by [al-Khwarizmi](#) is the more accurate. Al-Biruni also wrote a treatise on time-keeping, wrote several treatises on the astrolabe and describes a mechanical calendar. He makes interesting observations on the velocity of light, stating that its velocity is immense compared with that of sound. He also describes the [Milky Way](#) as

... a collection of countless fragments of the nature of nebulous stars.

Gauss and Riemann.

The geometry of surfaces, especially the “intrinsic” geometry of surfaces, those properties of surfaces which are independent of how they are embedded in Euclidean space, was developed by Gauss. But the full higher dimensional notion of intrinsic geometry of a possibly curved space was developed by his student Riemann. The equations for geodesics as curves which locally minimize arc length plays a key role in this theory. It was Riemann’s theory of the curvature of such spaces which played a key role in Einstein’s theory of general relativity.

Johann Carl Friedrich Gauss



Born: 30 April 1777 in Brunswick, Duchy of Brunswick (now Germany)

Died: 23 Feb 1855 in Göttingen, Hanover (now Germany)

Georg Friedrich Bernhard Riemann



Born: 17 Sept 1826 in Breselenz, Hanover (now Germany)

Died: 20 July 1866 in Selasca, Italy

Parallelism along curves.

Can we attach a meaning to the assertion that two vectors tangent to the sphere at two different points \mathbf{p} and \mathbf{q} are parallel? The answer to this question is **no**. However it does make sense if we join \mathbf{p} to \mathbf{q} by a curve: Let \mathbf{c} be a curve on the sphere which starts at \mathbf{p} and ends at \mathbf{q} . Place the sphere on a plane so that it just touches the plane at \mathbf{p} . If \mathbf{u} is a vector tangent to the sphere at \mathbf{p} we can also think of \mathbf{u} as being a vector \mathbf{U} in the plane, since this plane is tangent to the sphere at \mathbf{p} . Now roll the sphere on the plane along the curve \mathbf{c} . This will give us a curve \mathbf{C} in the plane, and at the end of this process we end up with the point \mathbf{q} touching the plane. A tangent vector \mathbf{v} at \mathbf{q} can be thought of as being a vector \mathbf{V} in the plane.

We say that \mathbf{u} and \mathbf{v} are *parallel along* \mathbf{c} if the vectors \mathbf{U} and \mathbf{V} are parallel in the plane. This notion of parallelism depends on the choice of the curve. A different curve joining \mathbf{p} to \mathbf{q} will give a different criterion for when vectors at \mathbf{p} and \mathbf{q} are parallel.

Geodesics as self-parallel curves.

We now can define geodesics to be self-parallel curves - curves \mathbf{c} which have the property that when you perform the rolling process the curve \mathbf{C} that you get in the plane is a (piece of) a straight line. For the sphere, the curves \mathbf{c} which roll out to straight lines in the plane are exactly the great circles. But we can make this definition for any curve on any surface.

It is then a mathematical theorem that this definition of geodesics, as curves which roll out to straight lines, coincides with the earlier definition of geodesics as curves which locally minimize arc length.

What about more general spaces such as those considered by Riemann?
Here the key result is due to Levi-Civita who introduced a general concept of parallelism of vectors along curves and showed that for a Riemannian manifold there is a unique such notion with certain desirable properties, and that the self-parallel curves are exactly the geodesics in Riemann's sense.

Tullio Levi-Civita



Born: 29 March 1873 in Padua, Veneto, Italy

Died: 29 Dec 1941 in Rome, Italy

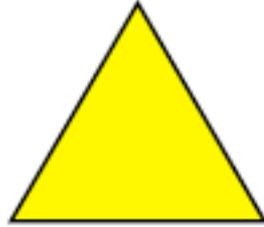
Back to the EIH paper again.

Introduction. In this paper we investigate the fundamentally simple question of the extent to which the relativistic equations of gravitation determine the motion of ponderable bodies.

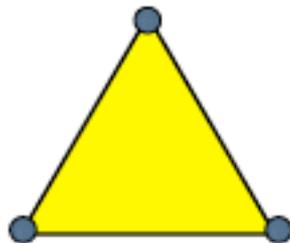
The question is: what do the “relativistic equations of gravitation” have to do with the equations which determine geodesics? In order to understand the EIH-Souriau answer to this question, we really do not need to know in detail what the “relativistic equations of gravitation” are. (This would require a whole course in general relativity.) All that we need to know is something very general about the form of these equations, in particular the symmetry which is built in to these equations. It is an amazing fact that these **symmetry conditions alone** determine the equations for geodesics.

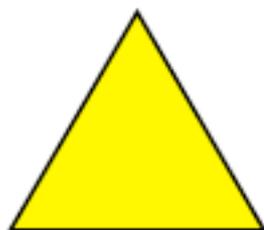
For this we need to state some elementary facts about constraints imposed by symmetry.

Constraints imposed by symmetry.

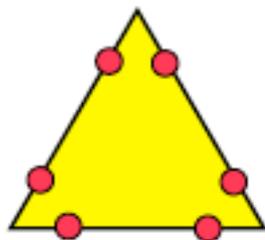


Here is an equilateral triangle. If I want to attach an object, say a little disk ● to one of the corners of this triangle, and still preserve the symmetry, then I must attach an identical object to all three corners.

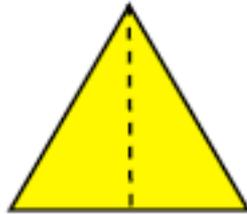




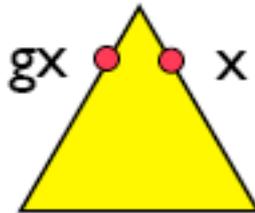
If I want to place a little red disk ● on one of the sides of the triangle and still maintain the complete symmetry, I must also place the same disk at all points (in general six of them) which can be obtained from this point by a symmetry transformation of the triangle.



x and gx .



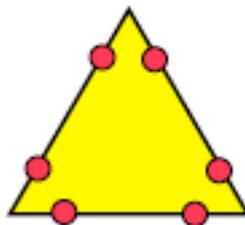
Let g denote the symmetry transformation consisting of flipping the triangle around the vertical axis. If x is a point of the triangle, then gx denotes the point obtained from x by applying the symmetry operation g to x :



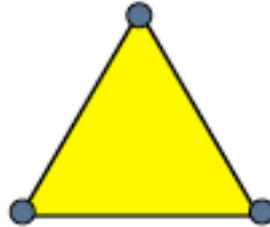
Orbits.

There are six symmetry operations of the triangle: flipping around each of the three perpendicular bisectors and rotations about the center through 0 degrees, 120 degrees and 240 degrees.

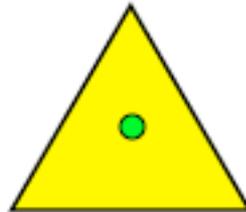
The set of all points that I obtain from a single point x by applying all the symmetry operations to x is called the **orbit** through x and is denoted by Gx . Typically there will be six points in an orbit.



Exceptional orbits.



But there will be some orbits with three elements, for example the orbits of the vertices, or more generally orbits of points lying on one of the perpendicular bisectors, and there is an orbit with only one point - the center of the triangle.



General formulation.

Suppose that X is a set (or object) and G is a group of symmetries of X . If x is a point of X and g in G is a symmetry, then we let gx denote the point of X obtained from x by applying the symmetry g . We let Gx denote the collection of all such points gx and call Gx the **orbit** of x under the symmetries G .

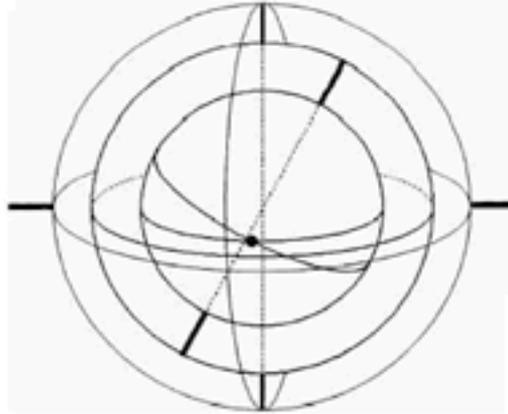
Then if F is a (say) numerical function on X which is invariant under the action of G , then F must take on a constant value on each orbit.

Example: Rotations.

Suppose that X is ordinary three dimensional space with a preferred point \mathbf{O} as origin, and G consists of all rotations about \mathbf{O} . If x is a point different from \mathbf{O} then the orbit Gx is the sphere of radius r where r is the distance from \mathbf{O} to x . If $x = \mathbf{O}$ then the orbit Gx consists of the single point \mathbf{O} . So the orbits are spheres centered about \mathbf{O} with the exception of the single orbit consisting of one point \mathbf{O} . Notice that in this example the (sphere) orbits each form a continuous manifold of points rather than a discrete collection of points as in the preceding examples.

Our symmetry conserving condition says that if F is a function which is invariant under G then F must be constant on each of these spheres.

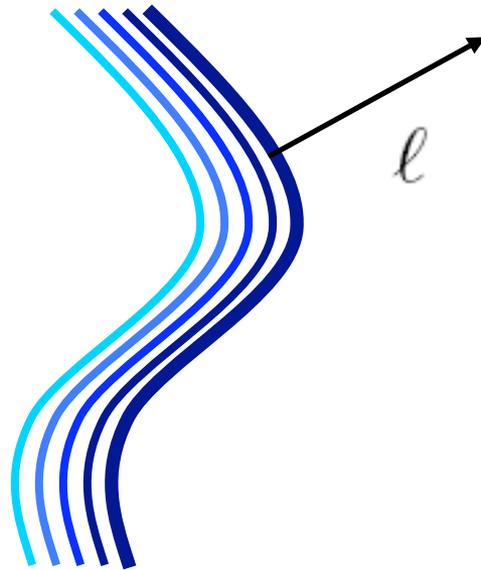
Orbits of the rotation group are
concentric spheres.



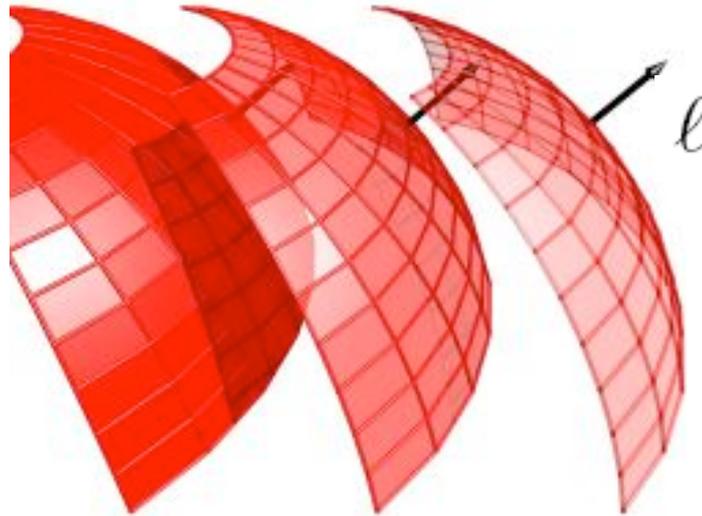
Here is a picture of a function F (the intensity of the blue) which is constant along each curve in a family. We wish to examine the infinitesimal change in F (or as we say the differential (change) of F) at any point.



The infinitesimal change ℓ of F vanishes on tangents to the orbits.



Another picture



Repeat of statement:

Suppose that G is a group of symmetries of a space X and suppose that F is a function on X which is invariant under the action of G . Let x be a point of X , and let $\mathcal{O} = G \cdot x$ be the G orbit through x . Let $T_x\mathcal{O}$ denote the tangent space to \mathcal{O} at x . Then if

$$dF_x = \ell$$

denotes the differential of F at x then

$$\ell \text{ vanishes on } T_x\mathcal{O}.$$

The punch line: The EIHS equations for a geodesic.

We now come to the punch line of today's lecture: For an appropriate choice of X and G , we can associate to certain data along a curve \mathbf{c} (in technical language - a contravariant symmetric tensor field along the curve) an ℓ , that is an object which measures infinitesimal change of a function, and the condition

$$\ell \text{ vanishes on } T_x \mathcal{O}$$

implies that the curve must be a geodesic!.

The punch line continued: the form of the field equations

Furthermore, the distribution of matter energy in space time can be thought of itself as an ℓ and for appropriate choice of the function F the equation

$$dF_x = \ell$$

thought of as an equation for x given ℓ is in fact the Einstein field equations for determining the geometry of space time from the distribution of matter energy.

Some technical details.

At this point I have to get a little technical. Let M be a manifold. Let X denote the set of all semi-Riemannian metrics on M of a given signature. (For example M could be space time and X the set of all Lorentzian metrics on M .) We will let G be the group of diffeomorphisms of compact support on M . (So an element of G will be a diffeomorphism of M which equals the identity outside some compact subset of M .) An element $\phi \in G$ acts on a metric x by sending it into $(\phi^{-1})^*x$. This gives an action of G on X .

The full tangent space.

The full tangent space $T_x X$ can be identified with the space $S_2(X)$ of smooth symmetric covariant tensor fields of degree 2. Notice that we have identified all tangent spaces with the same fixed vector space. We have *trivialized* the tangent bundle to X . We want to consider the subspace $S_2^0(X) \subset S_2(X)$ consisting of those smooth tensor fields of compact support. The corresponding subspace of $T_x X$ will be denoted by $T_x^0 X$. It is to be thought of as those infinitesimal variations of the metric x which vanish outside some compact subset.

The tangent space to the orbit.

How should we think of $T_x\mathcal{O}$? If u is a vector field on M , then differential geometry attaches a meaning to $D_u x$, the Lie derivative of the metric x with respect to the vector field u . It is the symmetric covariant two tensor whose value $D_u x(v, w)$ on two other vector fields v and w is defined as follows:

$$D_u x(v, w) = u\langle v, w \rangle_x - \langle [u, v], w \rangle_x - \langle v, [u, w] \rangle_x,$$

Here $\langle v, w \rangle_x$ denotes the scalar product of v and w (a function on M) determined by the metric x .

Since G consists of diffeomorphisms of compact support, we let $T_x\mathcal{O}$ consist of those $D_u x$ where u is a vector field of compact support. Clearly

$$T_x\mathcal{O} \subset T_x^0 X.$$

Possible ℓ 's .

What are the possible ℓ 's? We want an ℓ to be a (continuous) linear function on $S_2^0(M)$. (This is where the theory of generalized functions comes in.) For example, suppose that τ is smooth contravariant symmetric 2-tensor. Then if σ is a symmetric covariant 2 tensor of compact support, then the double contraction

$$\sigma \bullet \tau$$

is a smooth function of compact support on M . The metric x determines a volume vol_x and we can integrate the function $\sigma \bullet \tau$ with respect to this volume.

$$\sigma \bullet \tau$$

is a smooth function of compact support on M . The metric x determines a volume vol_x and we can integrate the function $\sigma \bullet \tau$ with respect to this volume. That is we can form the integral

$$\int_M \sigma \bullet \tau \text{ vol}_x .$$

In this way, which depends on the metric x , we have associated to τ a continuous linear function ℓ_τ on $T_x^0 X$:

$$\ell_\tau(\sigma) = \int_M \sigma \bullet \tau \text{ vol}_x .$$

Here is a different kind of ℓ , one associated to a curve: Let I be a (compact) interval on the real line and $\mathbf{c} : I \rightarrow M$ a smooth non-degenerate curve on M . Let τ be a smooth tensor field along \mathbf{c} . This means that $\tau(s)$ is a contravariant two tensor at the point $\mathbf{c}(s)$. (We will assume that $\tau(s) \neq 0$ for any s .) If σ is a symmetric covariant 2 tensor then

$$\sigma(\mathbf{c}(s)) \bullet \tau(s)$$

is a smooth function of s and we can form the integral

$$\int_I \sigma(\mathbf{c}(s)) \bullet \tau(s) ds.$$

An $\ell_{\mathbf{c},\tau}$ associated to a curve \mathbf{c} .

$$\sigma(\mathbf{c}(s)) \bullet \tau(s)$$

is a smooth function of s and we can form the integral

$$\int_I \sigma(\mathbf{c}(s)) \bullet \tau(s) ds.$$

So the pair consisting of the curve \mathbf{c} and the tensor field τ along \mathbf{c} gives rise to a continuous linear function $\ell_{\mathbf{c},\tau}$ on $T_x^0(X)$ by

$$\ell_{\mathbf{c},\tau}(\sigma) := \int_I \sigma(\mathbf{c}(s)) \bullet \tau(s) ds.$$

The main result.

We can now state the main result of EIH as reformulated by Souriau: If $\ell_{\mathbf{c},\tau}(\sigma)$ satisfies the condition

$$\ell_{\mathbf{c},\tau}(\sigma) \text{ vanishes on } T_x\mathcal{O}$$

then up to a suitable reparametrization of \mathbf{c} , the curve \mathbf{c} is a geodesic and

$$\tau(s) = \pm \mathbf{c}'(s) \otimes \mathbf{c}'(s)$$

where \mathbf{c}' denotes the tangent vector to \mathbf{c} .

The proof of this result is by a certain amount of integration by parts which I will omit.

The Hilbert “function”.

For any metric x , let $S(x)$ denote the scalar curvature of x . Try to define the “function” $F(x)$ by

$$F(x) = - \int_M S(x) \operatorname{vol}_x .$$

The trouble is that this integral may not be defined since M , in general, is not compact.

The variation *is* defined.

Nevertheless, the variation

$$dF_x(\sigma)$$

for $\sigma \in T_x^0 X$ is well defined since we may replace integration over M by integration over any compact subset K containing the support of σ and then define

$$dF_x(\sigma) = -\frac{d}{dt} \int_K S(x + t\sigma) \operatorname{vol}_{x+t\sigma} \Big|_{t=0} .$$

This clearly does not depend of the choice of K .

The Einstein-Hilbert field equations.

Suppose that ℓ is a linear function on $S_2^0(M)$ corresponding to a smooth tensor field in some (and hence every) metric. The Einstein-Hilbert field equations for the metric x are the equations

$$dF_x = \ell.$$

We know that a necessary condition for the solvability of these equations is

$$\ell \text{ vanishes on } T_x \mathcal{O}.$$

Passivity.

Suppose we replace ℓ by $\ell + \ell'$ where (say) ℓ' is a smooth approximation to $\ell_{\mathbf{c},\mu}$. Then we get, in principle, a different x' as a solution to

$$dF_{x'} = \ell + \ell'$$

and hence also a different orbit \mathcal{O}' . The “passivity approximation” that I stated in the first slide says that we will ignore this change in x and hence assume the necessary condition

$$\ell + \ell' \text{ vanishes on } T_x \mathcal{O}.$$

$\ell + \ell'$ vanishes on $T_x\mathcal{O}$.

Since this condition is linear, and we know that

ℓ vanishes on $T_x\mathcal{O}$

we conclude that

ℓ' vanishes on $T_x\mathcal{O}$

and hence (in the limit) that

$\ell_{\mathbf{c},\mu}$ vanishes on $T_x\mathcal{O}$

and so \mathbf{c} is a geodesic. This is the E-I-H solution to their “fundamentally simple question”.

The Schrödinger equation.

I will illustrate the “integration by parts” argument that I omitted, by studying an analogue of this procedure in a finite dimensional model of “quantum mechanics”. Let V be a finite dimensional vector space and let $G = Gl(V)$ be the group of all invertible linear transformations of V . Let X be the space of all linear transformations of V with G acting on X by conjugation:

$$g \cdot x := gxg^{-1}.$$

$$g \cdot x := gxg^{-1}.$$

The tangent space to the orbit \mathcal{O} through x consists of all $[x, y]$ as y varies over X . Since X is a vector space, we can identify $T_x X$ with X for every x . We can also identify the space of linear functions on X with X using the trace: If $z \in X$ define

$$\ell_z(w) = \operatorname{tr} zw.$$

Every linear function on X is of this form. The condition

$$\ell_z \text{ vanishes on } T_x \mathcal{O}$$

becomes

$$\operatorname{tr} z[x, y] = 0 \quad \forall y \in X.$$

The “integration by parts” argument.

$$\operatorname{tr} z[x, y] = 0 \quad \forall y \in X.$$

But

$$\operatorname{tr} zyx = \operatorname{tr} xzy$$

so

$$\operatorname{tr} z[x, y] = \operatorname{tr} z(xy - yx) = \operatorname{tr}(zx - xz)y = \operatorname{tr}[z, x]y.$$

(This was the “integration by parts”.)

The condition $\text{tr}[z, x]y = 0$ for all y implies that $[x, z] = 0$. So the condition

$$\ell_z \text{ vanishes on } T_x\mathcal{O}$$

is

$$[x, z] = 0$$

in the current example.

Suppose we look at a special kind of z (like we considered the special ℓ associated to a curve in the general relativity case). Suppose that z is of rank one, so maps the entire space V onto the line through a vector v . Then the above condition implies that v is an eigenvector of X :

$$xv = \lambda v$$

for some λ .

To write this in more familiar notation replace the letter x by the letter H . We obtain

$$H\phi = \lambda\phi.$$

To write this in more familiar notation replace the letter x by the letter H . We obtain

$$H\phi = \lambda\phi.$$

So the condition

$$\ell \text{ vanishes on } T_x\mathcal{O}$$

together with the assumption that we are looking at a special kind of ℓ , one corresponding to a rank one operator gives Schrödinger's equation. We have derived Schrödinger's from the same principle which gave us the equation for geodesics!