A Brief Introduction to the Einstein Field Equations

Ryan Maguire

October 26, 2022

Many of the notions of curvature require only a semi-Riemannian metric and a choice of affine connection.

It is almost universal amongst physicists and mathematicians to work with the unique Levi-Civita connection that a given semi-Riemannian metric induces.

We'll discuss several types of curvatures, and their uses in describing the Einstein field equations.

Preliminaries

A manifold is a topological space (M, τ) that is:

- Hausdorff
- Second Countable
- Locally Euclidean

Some authors replace second countability with a plethora of other not-necessarily-equivalent notions. Paracompactness is a common one, others like the notion of $\sigma\text{-compactness}$. If the space is required to be *connected*, all of these ideas are the same[1]

Most authors omit the topology τ altogether.

[1] I believe there are 50+ alternatives to second countability one can use if the space is connected, but I've forgotten the reference. So don't quote me.

Preliminaries

A chart in a manifold (M, τ) is an ordered pair $(\mathcal{U}, \varphi)$ where $\mathcal{U} \in \tau$ is an open subset and $\varphi : \mathcal{U} \to \mathbb{R}^n$ is a continuous injective open mapping for some $n \in \mathbb{N}$. The *dimension* of the chart is this value n.

Theorem (Brauer's Invariance of Domain)

If $f : \mathbb{R}^n \to \mathbb{R}^n$ is a continuous injective mapping, and if $\mathcal{U} \subseteq \mathbb{R}^n$ is open, then $f[\mathcal{U}]$ is open.

This would make f a continuous injective open mapping, meaning it is a homeomorphism onto the image. This has two corollaries.

Theorem

$\mathbb{R}^n$ is homeomorphic to $\mathbb{R}^m$ if and only if $n = m$.

Preliminaries

Theorem

If (M, τ) is locally Euclidean, if $x \in M$, and if $(\mathcal{U}, \varphi)$ and $(\mathcal{V}, \psi)$ are charts in M with $x \in \mathcal{U}$ and $x \in \mathcal{V}$, then the charts have the same dimension.

The proof is performed by examining the map $\varphi \circ \psi^{-1}$ and noting it induces a homeomorphism from an open subset of $\mathbb{R}^n$ to an open subset of $\mathbb{R}^m$, meaning $n = m$.

This says that dimension is *locally* constant. If the manifold is connected, dimension is a constant. The only way to have a manifold with a 1 dimensional component and a 2 dimension component is via disjoint unions, like the disjoint union of a line and a sphere.

Preliminaries

Examining this composition map allows us to define differentiability. With very few exceptions (topological vector spaces where the Fréchet derivative is definable) functions between topological spaces have no notion of differentiability. Manifolds have the ability to define such things.

Given two charts $(\mathcal{U}, \varphi)$ and $(\mathcal{V}, \psi)$ where $\mathcal{U} \cap \mathcal{V} \neq \emptyset$, the function $\varphi \circ \psi^{-1}$ is a continuous function from an open subset of $\mathbb{R}^n$ to another open subset of $\mathbb{R}^n$. It is then perfectly valid to ask if this function has partial derivatives, or second partial derivatives, and so on. One can even ask if the function is *smooth*, having all partial derivatives of all orders.

If $\varphi \circ \psi^{-1}$ and $\psi \circ \varphi^{-1}$ are smooth, we say the charts $(\mathcal{U}, \varphi)$ and $(\mathcal{V}, \psi)$ are smoothly compatible. We'll say this is vauously true if $\mathcal{U} \cap \mathcal{V} = \emptyset$.

Preliminaries

An atlas on a manifold (M, τ) is a collection of charts $(\mathcal{U}_\alpha, \varphi_\alpha)$ such that $\bigcup_\alpha \mathcal{U}_\alpha = M$.

A smooth atlas is an atlas where all charts are smoothly compatible. A maximal smooth atlas is a smooth atlas with, intuitively, as many smoothly compatible charts possible. A smooth manifold is a topological manifold with a maximal smooth atlas.

Given two smooth manifolds M and N, a smooth function is a function $F : M \to N$ such that for every $x \in M$ there is a chart $(\mathcal{U}, \varphi)$ in M and a chart $(\mathcal{V}, \psi)$ in N such that $x \in \mathcal{U}$, $F[\mathcal{U}] \subseteq \mathcal{V}$, and the function $\psi \circ F \circ \varphi^{-1}$ is smooth.

Preliminaries

A side-note, the $F[\mathcal{U}] \subseteq \mathcal{V}$ criterion is important. Without it the function $F : \mathbb{R} \to \mathbb{R}$ defined by

$$F(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases} \tag{1}$$

would be considered smooth, but it's not even continuous!

A diffeomorphism is a function $F : M \to N$ that is bijective, smooth, and such that F^{-1} is smooth. For dimensions 0, 1, 2, and 3, every topological manifold has a maximal smooth atlas that is unique up to diffeomorphism. Things get weird in dimension 4 and higher where it is possible for the same manifold to have different smooth structures and for some manifolds to have **zero** smooth structures. Some compact manifolds need not be *smoothable*.

Preliminaries

Smooth manifolds have a notion of tangent spaces. Given a smooth manifold M and a point $x \in M$ the tangent space at x is denoted $T_x M$. This can be described via derivations, which are functions $D : C^\infty(M, \mathbb{R}) \to \mathbb{R}$, that take in smooth functions and output real numbers, such that:

$$D(af + bg) = aD(f) + bD(g) \tag{2}$$

$$D(fg) = f(x)D(g) + D(f)g(x) \tag{3}$$

That is, D is linear and Liebnizean. If M is an n dimensional manifold, the set of all derivations at $x \in M$ forms an n dimensional real vector space.

Preliminaries

If $(\mathcal{U}, \varphi)$ is a chart containing $x \in M$ then the partial derivative operators $\partial\varphi_k$, $k = 0, \ldots, n-1$, form a basis for $T_x M$:

$$\partial\varphi_k(f) = \frac{\partial}{\partial x_k}\left(f \circ \varphi^{-1}\right) \tag{4}$$

Note, since $f \in C^\infty(M, \mathbb{R})$, it is smooth, and $f \circ \varphi^{-1}$ is a smooth function from an open subset of $\mathbb{R}^n$ to $\mathbb{R}$ so it is valid to take the partial derivative in the k^{th} component.

The tangent bundle of a manifold is formed by taking all $T_x M$ for each $x \in M$ and gluing them together in a natural way. For $M = \mathbb{R}$ the tangent bundle is $\mathbb{R}^2$, and for $M = \mathbb{S}^1$ the tangent bundle is $\mathbb{S}^1 \times \mathbb{R}$. These are called *trivial* tangent bundles. Most tangent bundles are not of the form $\mathbb{R}^n \times M$, the easiest example is $\mathbb{S}^2$.

The tangent bundle is denoted TM and is a smooth manifold of dimension $2n$. A vector field on a manifold is a smooth function $V : M \to TM$ such that for all $x \in M$ the element $V(x)$ is of the form $V(x) = (x, v)$. That is, V assigns to every element $x \in M$ a tangent vector $v \in T_x M$ that starts at the point x. The image $V(x)$ is often denoted V_x.

Preliminaries

A Riemannian metric is a function g that assigns to every $x \in M$ a function $g_x : T_xM \times T_xM \to \mathbb{R}$ that mimics the dot product in Euclidean space. That is, for all $v_0, v_1, w \in T_xM$ and $a_0, a_1 \in \mathbb{R}$ we have:

$$g_x(a_0 v_0 + a_1 v_1,\ w) = a_0 g_x(v_0,\ w) + a_1 g_x(v_1,\ w) \tag{5}$$

$$g_x(w,\ a_0 v_0 + a_1 v_1) = a_0 g_x(w,\ v_0) + a_1 g_x(w,\ v_1) \tag{6}$$

$$g_x(v_0,\ v_1) = g_x(v_1,\ v_0) \tag{7}$$

$$g_x(w,\ w) > 0 \quad w \neq 0 \tag{8}$$

The first two say g_x is bilinear, the third equation makes g_x symmetric, and the last condition is called positive-definiteness.

Preliminaries

The algebraist can summarize this by stating that g_x is a symmetric positive-definite bilinear form. But if you're like me and forget these words half the time, then remember the Euclidean dot product.[2]

The metric g should also be *smooth*. That is, for all $x \in X$ and for every smooth vector field V, W on M the function $f : M \to \mathbb{R}$ defined by $f(x) = g_x(V_x, W_x)$ should be smooth.

[2] That's how I made the previous slide.

Preliminaries

Riemannian metrics give us a means of smoothly measuring angles between tangent vectors on a manifold at a given tangent space. Namely, one can define, for two tangent vectors $v, w \in T_x M$, the function:

$$\angle(v,\ w) = \cos^{-1}\left(\frac{g_x(v,\ w)}{g_x(v,\ v)\, g_x(w,\ w)}\right) \tag{9}$$

Preliminaries

Pseudo-Riemannian metrics are formed by replacing the positive-definite requirement with non-degeneracy. That is, for all non-zero tangent vectors $v \in T_x M$, there is a tangent vector $w \in T_x M$ such that $g_x(v, w) \neq 0$. Positive-definiteness implies non-degenerate, given $v \neq 0$ simply choose $w = v$. Pseudo-Riemannian metrics are thus a generalization of Riemannian metrics.

Preliminaries

Tangent vectors are derivations on $C^\infty(M, \mathbb{R})$. Cotangent vectors are functions that take in tangent vectors and return real numbers in a linear way. Much like tangent vectors can be represented explicitly with charts, so can cotangent vectors. Given $(\mathcal{U}, \varphi)$ we can define:

$$\mathrm{d}\varphi_k(\partial\varphi_\ell) = \begin{cases} 0 & k \neq \ell \\ 1 & k = \ell \end{cases} \tag{10}$$

Preliminaries

Theorem (Sylvester's Law of Inertia)

If M is a smooth manifold of dimension N, if g is a pseudo-Riemannian metric on M, and if $x \in M$, then there is a chart $(\mathcal{U}, \varphi)$ and a fixed integer $0 \leq n < N$ such that for all $v, w \in T_x M$ we have:

$$g_x(v,\, v) = \sum_{k=0}^{n-1} d\varphi_k(v)^2 - \sum_{k=n}^{N-1} d\varphi_k(v)^2 \tag{11}$$

The *signature* of the metric g is the tuple $(1,\, \dots, 1,\, -1,\, \dots,\, -1)$ where there are n positives and $N - n$ negatives.[3]

[3] Why is it called Sylvester's Law of Inertia? Excellent question.

Preliminaries

Pseudo-Riemannian metrics allow one to define affine connections, which are tools for transporting tangent vectors around a manifold in a parallel fashion, and defining things like curvature. Let $\mathfrak{X}(M)$ denote the set of all smooth vector fields on M. An affine connections is a function $\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ such that for all $X, Y, Z \in \mathfrak{X}(M)$, $a, b \in \mathbb{R}$, and $f, g \in C^{\infty}(M, \mathbb{R})$ we have:

$$\nabla_{aX+bY} Z = a\nabla_X Z + b\nabla_Y Z \tag{12}$$

$$\nabla_Z(aX + bY) = a\nabla_Z X + b\nabla_Z Y \tag{13}$$

$$\nabla_{fX} Y = f\nabla_X Y \tag{14}$$

$$\nabla_X(fY) = D_X Y + f\nabla_X Y \tag{15}$$

where $D_X f$ is the directional derivative of f in the direction of X. This is an attempt to axiomatize the contravariant derivative that occurs in multi-variable analysis.

Preliminaries

The general affine connection is rarely discussed in mathematics and physics. Two more desirable properties are usually added, *torsion-free* and *compatibility*. Torsion is defined in terms of the Lie bracket. Given two vector fields $X, Y \in \mathfrak{X}(M)$ the Lie bracket is another vector field $[X,\ Y]$ defined by:

$$[X,\ Y] = XY - YX \tag{16}$$

The composition XY need not be a vector field because of second order terms that make it not linear, but $-YX$ always kills those factors. Hence $[X,\ Y]$ is a vector field.

Preliminaries

A torsion free connection is a connection ∇ such that for all $X,\ Y \in \mathfrak{X}(M)$ we have:

$$\nabla_X Y - \nabla_Y X - [X,\ Y] = 0 \tag{17}$$

Preliminaries

Compatibility with the metric is defined in terms of parallel translation. Say you have a tangent vector v at a point $x \in M$ and you want to move it along a curve $\gamma : [0, 1] \to M$, $\gamma(0) = x$, in a manner that is *parallel*. This equates to solving for a vector field X using differential equations in coordinates, and you seek the solution to:

$$\nabla_{\gamma'(t)} X = 0 \tag{18}$$

$$X_{\gamma(0)} = v \tag{19}$$

Preliminaries

An affine connection that is compatible with g is one such that parallel transport is an isometry. That is, if $v, w \in T_x M$, if $\gamma : [0, 1] \to M$ is a smooth curve, and if v', w' are the results of parallel transport of v and w, respectively, for 1 second, then:

$$g_{\gamma(0)}(v, w) = g_{\gamma(1)}(v', w') \tag{20}$$

A Levi-Civita connection is an affine connection that is torsion free and compatible with the metric.

Preliminaries

Theorem (Fundamental Theorem of Semi-Riemannian Geometry)

If M is a smooth manifold, and if g is a pseudo-Riemannian metric on M, then there is a unique Levi-Civita connection ∇ on M.

Lastly, a Lorentzian metric is a pseudo-Riemannian metric on M with signature $(N-1,\,1)$, or $(+1,\,\ldots,\,+1,\,-1)$. The *negative* dimension is taken to be time. The Lorentzian *norm* (it's not a true norm) is given by:

$$||\cdot||^2 = \sum_{k=0}^{N-2} \mathrm{d}x_k^2 - \mathrm{d}t^2 \tag{21}$$

And that's it, no more preliminary stuff. On to physics!

General Relativity and Curvature

In physics it is common to work in a coordinate chart $(\mathcal{U}, \varphi)$ and express all physical quantities in terms of this chart. The semi-Riemannian metric g becomes a matrix $g_{\mu\nu}$ with entries $g_{\mu\nu} = g(\partial\varphi_\mu, \partial\varphi_\nu)$, which is called the *metric tensor* in general relativity. Other tensors and tensor fields will be described similarly.

General Relativity and Curvature

The first tensor to describe is the stress-energy tensor $T_{\mu\nu}$. It is the gravitational analogue of the stress tensor from Newtonian mechanics and describes the density and flux of energy in the manifold (M, g), which is always chosen to be Lorentzian.

The Einstein field equations relate the stress-energy tensor and the metric tensor to Ricci curvature and scalar curvature.

General Relativity and Curvature

The Ricci curvature is described in terms of the Riemann curvature tensor field (It's a tensor field, not a tensor). Given the affine connection ∇ on the semi-Riemannian manifold, the Riemann curvature tensor field is defined in one of two equivalent ways. It is a function $R : \mathfrak{X}(M)^3 \to \mathfrak{X}(M)$

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z \qquad (22)$$

Where $[X, Y]$ is the Lie bracket. We can also write this as:

$$R(X, Y) = [\nabla_X, \nabla_Y] - \nabla_{[X,Y]} \qquad (23)$$

again using the Lie bracket. With this we see that the Riemann curvature tensor field measures the failure of the second derivative to commute.

General Relativity and Curvature

If ∇ is a Levi-Civita connection (torsion free and compatible with the metric), then there are several identities the Riemann curvature tensor field enjoys. These identities can be combined with the Einstein field equations to prove the local conservation of energy and momentum, classical laws of Newtonian mechanics which still hold in general relativity.

- R is trilinear over $C^\infty(M, \mathbb{R})$.
- the Bianchi identity holds:

$$R(X, Y)Z + R(Y, Z)X + R(Z, X)Y = 0 \tag{24}$$

The Bianchi identity cyclicly permutes the vector fields. It is the Bianchi identity that helps one prove conservation of momentum and energy.

General Relativity and Curvature

The quadruple product relates the Riemann curvature tensor field to the semi-Riemannian metric. It is defined as:

$$(X, Y, Z, T) = g\big(R(X, Y)Z, T\big) \tag{25}$$

There are several identities for this operation, which are again useful for the proof of various theorems in the framework of general relativity.

$$(X, Y, Z, T) = -(Y, X, Z, T) \tag{26}$$

$$(X, Y, Z, T) = -(X, Y, T, Z) \tag{27}$$

$$(X, Y, Z, T) = (Z, T, X, Y) \tag{28}$$

Lastly, an analogue of the Bianchi identity:

$$(X, Y, Z, T) + (Y, Z, X, T) + (Z, X, Y, T) = 0 \tag{29}$$

These identities combine to give the following theorem.

Theorem

If $(\mathcal{U}, \varphi)$ is a chart in a spacetime (M, g), if ∇ is the unique Levi-Civita connection on M, and if T is the stress-energy tensor, then:

$$\sum_{n=0}^{N-1} \nabla_{\partial\varphi_n} T_{n,m} = 0 \tag{30}$$

This is the analogue of the conservation of momentum and energy laws that occur in Newtonian mechanics. The proof is about a page and simply uses the identities of the Riemannian curvature tensor field, the quadruple product, and the Einstein field equations which will be stated soon.

General Relativity and Curvature

The Einstein field equations relate the stress-energy tensor to the Ricci and scalar curvatures. The Ricci curvature is defined in terms of the Riemann curvature tensor field. There are two ways of doing this.

In the Riemann setting (g is positive-definite), fix $p \in M$ and $x = z_n \in T_pM$ to be unit length. Since T_pM is an n dimensional real inner product space, we may extend z_n via the Gram-Schmidt procedure to an orthonormal basis. Label these other elements $z_1, \ldots, z_{n-1}$. The Ricci curvature about p is defined as:

$$\mathrm{Ric}_p(x) = \frac{1}{n-1} \sum_{k=1}^{n} g_p\Big(R(x, z_k)x, z_k\Big) \tag{31}$$

It is a theorem that this result is independent of the choice of basis.

General Relativity and Curvature

In the semi-Riemannian setting T_pM is not an inner product space since g can, in general, fail to be positive definite. Such is the case in spacetimes with signature $(+,+,+,-)$. Fix two vector fields Y and Z. Given a vector field X, the mapping $X \mapsto R(X,Y)Z$ is linear at each tangent space. Because of this one may define the *trace* of this mapping. This is the Ricci curvature tensor.

$$\mathrm{Ric}_p(Y,Z) = \mathrm{tr}\big(X_p \mapsto R_p(X_p, Y_p)Z_p\big) \tag{32}$$

In local coordinates $(\mathcal{U}, \varphi)$ it can be given by a matrix $R_{\mu\nu}$.

General Relativity and Curvature

The Ricci curvature can be completely described by the sectional curvature, which is one of the older notions of curvature dating back to a time when differential geometry dealt solely with regular surfaces and curves. The sectional curvature of a 2-dimensional subspace δ of the tangent space T_pM is given by:

$$K_\delta = \frac{(v, w, v, w)}{A(v, w)} = \frac{g_p\big(R(v, w)v, w\big)}{\sqrt{||v||^2\,||w||^2 - g_p(v, w)^2}} \qquad (33)$$

where v and w are two tangent vectors that span δ, and $A(v, w)$ is the area of the parallelogram with sides v and w. K_δ is independent of choice of basis since a change of basis can be made by a combination of moves $(x, y) \mapsto (y, x)$, $(x, y) \mapsto (\lambda x, y)$, $\lambda \neq 0$, and $(x, y) \mapsto (x + \lambda y, y)$. These operations are reflection, scaling, and shearing, respectively. All of these are invariant under formula above showing K_δ is independent of basis.

General Relativity and Curvature

For constant curvature manifolds the Ricci curvature is given by a simple formula:

$$R_{\mu\nu} = (n-1)Kg_{\mu\nu} \tag{34}$$

where K is the constant curvature of the manifold. It is probably not the case that the spacetime we live in is constant curvature.

General Relativity and Curvature

The scalar curvature is defined directly by the Ricci curvature. Given the Riemannian definition, $\mathrm{Ric}_p(x)$, given a basis $\{z_1, \ldots, z_n\}$ of T_pM, the scalar curvature is defined by:

$$K(p) = \frac{1}{n}\sum_{k=1}^{n} \mathrm{Ric}_p(z_k) \tag{35}$$

It is independent of choice of basis. With respect to the second definition, we can define:

$$K(p) = \mathrm{tr}(R_{\mu\nu}) \tag{36}$$

General Relativity and Curvature

The Einstein tensor is defined in terms of the Ricci and scalar tensors. We have:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} K g_{\mu\nu} \tag{37}$$

Where $R_{\mu\nu}$ is the Ricci tensor, K is the scalar curvature, and $g_{\mu\nu}$ is the metric tensor. The Einstein field equations are:

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \kappa T_{\mu\nu} \tag{38}$$

Where $T_{\mu\nu}$ is the stress-energy tensor. Λ is the cosmological constant, and κ is the Einstein gravitational constant.

General Relativity and Curvature

In practice, one measures the stress-energy tensor and the Einstein tensor and wishes to solve for the metric in the Einstein field equation. A common simplification is to suppose the spacetime you are working in is a vacuum containing no mass-energy. The Einstein field equations simplify to:

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 0 \tag{39}$$

Expanding the Einstein tensor in terms of the Ricci and scalar curvature, we get:

$$R_{\mu\nu} - \frac{1}{2} K g_{\mu\nu} + \Lambda g_{\mu\nu} = 0 \tag{40}$$

General Relativity and Curvature

This is a purely geometrical problem. Depending on the value of Λ there are several known spacetimes with metrics that satisfy the Einstein field equations.

- Minkowski spacetime $\mathbb{M}^{3,1}$
- Milne spacetime
- Schwarzschild vacuum spacetime
- Kerr vacuum

The value of Λ was originally thought to be zero, and Einstein retracted it from the equation. In the late 1990's it was discovered the inflation of the universe is accelerating, indicating the constant may be positive. One possible value involves the Hubble constant, given by:

$$\Lambda = 1.1056 \times 10^{-52}\,\mathrm{m}^{-2} \tag{41}$$