CHAPTER 3: RIEMANN GEOMETRY

3.1 Affine connection

According to the definition, a vector field $X \in D^1(M)$ determines a derivation of the algebra of smooth real valued functions on M. This action is linear in X such that $(fX)g = f(Xg)$ for any pair f, g of smooth functions. Next, we want to define an action of X on $D^1(M)$ itself, which has similar properties. Let $\nabla_X : D^1(M) \to D^1(M)$ for any $X \in D^1(M)$ be an operator satisfying the following conditions:

(1) The map $Y \mapsto \nabla_X Y$ is real linear in Y for any fixed X,
(2) $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$ for any vector fields X, Y, Z and any smooth real valued functions f, g, and
(3) $\nabla_X(fY) = f\nabla_X Y + Y(X \cdot f)$ for any vector fields X, Y and any smooth real valued function f.

An operator ∇ satisfying these conditions is called an *affine connection* on the manifold M.

Example 1 Let $M = \mathbb{R}^n$ and define

$$\nabla_X Y = (X \cdot Y^j)\frac{\partial}{\partial x^j}.$$

Then, ∇ is an affine connection.

Warning! The above example needs a modification when applied to an arbitrary manifold M. The difficulty is that the right-hand side depends on the choice of local coordinates and it does not transform like a true vector. If we transform to coordinates $y^j = y^j(x^1, \dots, x^n)$, then in the new coordinates

$$Y'^j(y) = \frac{\partial y^j}{\partial x^i} Y^i(x) \tag{1}$$

and therefore

$$(X \cdot Y'^j)\partial'_j = (X \cdot Y^i)\frac{\partial y^j}{\partial x^i}\partial'_j + Y_i\left(X \cdot \frac{\partial yj}{\partial x^i}\right)\partial'_j. \tag{2}$$

The coordinates of the first term on the right-hand side are equal to $\frac{\partial y^j}{\partial x^i}(\nabla_X Y)^i$, but for any non-linear coordinate transformation we also have a second inhomogeneous term.

Choosing local coordinates, the difference

$$H^i(X,Y) = (\nabla_X Y)^i - X \cdot Y^i \tag{3}$$

is linear in both arguments in the extended sense

$$H^i(fX, gY) = fgH^i(X,Y), \tag{4}$$

for any smooth functions f and g. For this reason, we can write

$$H^i(X,Y) = \Gamma^i_{jk} X^j Y^k. \tag{5}$$

Here $\Gamma^i_{jk} = \Gamma^i_{jk}(x)$ are smooth (local) functions on M. Once again,

$$(\nabla_X Y)^i = X \cdot Y^i + \Gamma^i_{jk} X^j Y^k. \tag{6}$$

The functions Γ^i_{jk} are called the *Christoffel symbols* of the affine connection ∇. Let us look what happens to the Christoffel symbols in a coordinate transformation $y = y(x)$. Let us denote by ∇_i the *covariant derivative* $\nabla_{\frac{\partial}{\partial x^i}}$. Then,

$$\nabla_i \partial_j = \Gamma^k_{ij} \partial_k. \tag{7}$$

Denoting $\partial'_i = \frac{\partial}{\partial y^i}$, we get

$$\begin{aligned} \nabla'_i \partial'_j &= \Gamma'^k_{ij} \partial'_k = \frac{\partial x^a}{\partial y^i} \nabla_a \left(\frac{\partial x^b}{\partial y^j} \partial_b \right) \\ &= \frac{\partial x^a}{\partial y^i} \left[\frac{\partial x^b}{\partial y^j} \nabla_a \partial_b + \partial_a \left(\frac{\partial x^b}{\partial y^j} \right) \partial_b \right] \\ &= \frac{\partial x^a}{\partial y^i} \frac{\partial x^b}{\partial y^j} \Gamma^c_{ab} \partial_c + \frac{\partial^2 x^b}{\partial y^i \partial y^j} \partial_b. \end{aligned} \tag{8}$$

Transforming back to the x coordinates on the left-hand side, we finally get

$$\Gamma'^k_{ij}(y) = \frac{\partial x^a}{\partial y^i} \frac{\partial x^b}{\partial y^j} \frac{\partial y^k}{\partial x^c} \Gamma^c_{ab}(x) + \frac{\partial y^k}{\partial x^c} \frac{\partial^2 x^c}{\partial y^i \partial y^j}. \tag{9}$$

Note that in linear coordinate transformations, the inhomogeneous term containing second derivatives vanishes and the Christoffel symbols transform like components of a third rank tensor.

Exercise 1 We define the Christoffel symbols on the unit sphere, using spherical coordinates (θ, ϕ). When $\theta \neq 0, \pi$, we set

$$\Gamma^\theta_{\phi\phi} = -\frac{1}{2} \sin 2\theta, \ \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = \cot\theta,$$

and all the other Γ's are equal to zero. Show that the apparent singularity at $\theta = 0, \pi$ can be removed by a better choice of coordinates at the poles of the sphere. Thus, the above affine connection extends to the whole S^2.

3.2 Parallel Transport

The tangent vectors at a point $p \in M$ form a vector space T_pM. Thus, tangent vectors at the *same point* can be added. However, at different points p and q, there is in general no way to compare the tangent vectors $u \in T_pM$ and $v \in T_qM$. In particular, the sum $u + v$ is ill-defined. An affine connection gives a method to relate tangent vectors at p to tangent vectors at q, provided that we have fixed some smooth curve $\gamma(t)$ starting from p and ending at q.

A curve γ defines a distribution of tangent vectors along the curve by

$$X(s) = \dot{x}^i(s)\partial_i. \tag{10}$$

We have chosen a local coordinate system x^i. Thus, $X(s) \in T_{\gamma(s)}M$. Consider the system of first order ordinary differential equations given by

$$\dot{Y}^i(s) + \Gamma^i_{kj}(x(s))\dot{x}^k(s)Y^j(s) = 0,\ i = 1, 2, \dots, n. \tag{11}$$

Here $Y(s)$ is an unknown vector field along the curve $x(s)$.

Exercise 2 Show that the set of equations (11) is coordinate independent in the sense that if the equations are valid in one coordinate system, then they are also valid in any other coordinate system.

A vector field Y along the curve $x(s)$ satisfying Eq. (11) is called a *parallel vector field.* The existence and uniqueness theorem in the theory of first order differential equations gives the following fundamental theorem in geometry:

Theorem 3.1. *Given a tangent vector $v \in T_pM$ at the initial point $p = \gamma(s_0)$ of a smooth curve $\gamma(s)$ then there is a unique parallel vector field $Y(s)$ along $\gamma(s)$ satisfying the initial condition $Y(s_0) = v$.*

Definition 3.2. A curve $\gamma(s)$ is a *geodesic* if its tangent vectors $\dot{\gamma}(s)$ at each point are parallel.

Thus, the statement $\gamma(s)$ is a geodesic means that the coordinate functions $x_i(s)$ satisfy

$$\ddot{x}^i(s) + \Gamma^i_{jk}(x(s))\dot{x}^j(s)\dot{x}^k(s) = 0. \tag{12}$$

This condition is a second order ordinary differential equation for the coordinate functions. We can use existence and uniqueness results from the theory of differential equations:

Theorem 3.3. *For given point $p \in M$ and a tangent vector $u \in T_pM$ there is, in some open neighborhood of p, a unique geodesic $\gamma(s)$ such that $\gamma(0) = p$ and $\dot{\gamma}(0) = u$.*

Example 2 Let $M = S^2$ and let Γ be the affine connection in Exercise 1. Then, the coordinates $\theta(s)$ and $\phi(s)$ of a geodesic satisfy

$$\begin{aligned}
\ddot{\theta}(s) - \frac{1}{2}\sin 2\theta(s)\,\dot{\phi}(s)\dot{\phi}(s) &= 0, \\
\ddot{\phi}(s) + 2\cot\theta(s)\,\dot{\phi}(s)\dot{\theta}(s) &= 0.
\end{aligned}$$

Find the general solution to the geodesic equations. The solutions are great circles on the sphere. For example, $\theta = \alpha s + \beta$ and $\phi = \text{const}$.

Let ∇ be a connection on M and $\gamma(s)$ a curve connecting points $p = \gamma(s_1)$ and $q = \gamma(s_2)$. We define the *parallel transport* from the point p to the point q along the curve γ as a linear map

$$\hat{\gamma} : T_pM \to T_qM.$$

The map is given as follows: Let $u \in T_pM$ and let $X(s)$ be a parallel vector field along γ such that $X(s_1) = u$. We set $\hat{\gamma}(u) = X(s_2)$. The map is linear, because the differential equation

$$\dot{X}^i(s) + \Gamma^i_{kj}\dot{x}_k(s)X^j(s) = 0 \tag{13}$$

is linear in X^i and therefore the solution depends linearly on the initial condition u.

Example 3 If $M = \mathbb{R}^n$ and $\Gamma^i_{jk} = 0$, then the parallel transport $\hat{\gamma}$ is the identity map $u \mapsto u$ for any curve γ.

Example 4 Let M and Γ be as in Example 2. Let $(\theta, \phi) = (\alpha s + \beta, \phi_0)$. Now, the parallel transport is determined by the equations

$$\begin{aligned}
&\dot{X}_\theta = 0,\\
&\dot{X}_\phi + \cot\theta \cdot \dot{\theta}X_\phi = \dot{X}_\phi + X_\phi \alpha \cot(\alpha s + \beta) = 0.
\end{aligned}$$

This set has the solution $X_\theta = \text{const.}$ and $X_\phi = \text{const.} \cdot (\sin(\alpha s + \beta))^{-1}$. If u is the tangent vector $(1,1)$ at the point $(\theta,\phi) = (\pi/4, 0)$, then the parallel transported vector v at $(\theta, \phi) = (\pi/2, 0)$ is $(1, 1/\sqrt{2})$.

3.3 Torsion and Curvature

Given an affine connection ∇ on a manifold M we can define a third rank tensor field $T = (T^k_{ij})$ as follows. Any pair of vector fields X and Y gives another vector field

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]. \tag{14}$$

The dependence on X and Y is linear, after choosing local coordinates, we may write

$$T(X,Y)^i = X^j Y^k T^i_{jk}, \tag{15}$$

which defines the components T^i_{jk} of the tensor. Note that $T(X,Y)$ is linear in the extended sense,

$$T(fX,Y) = T(X,fY) = fT(X,Y),\ T(X,Y+Z) = T(X,Y) + T(X,Z)$$

for any real function f. Note further that $T(X,Y) = -T(Y,X)$. Since

$$T(\partial_i, \partial_j)^k = \Gamma^k_{ij} - \Gamma^k_{ji}, \tag{16}$$

we see that T is precisely the antisymmetric part (in the lower indices) of the Christoffel symbols.

From Eq. (16) and the transformation formula (11) for the Christoffel symbols follows that the components of the *torsion* T really transform like tensor components in coordinate transformations,

$$T'^i_{jk}(y) = \frac{\partial y^i}{\partial x^p}\frac{\partial x^\ell}{\partial y^j}\frac{\partial x^m}{\partial y^k}T^p_{\ell m}(x). \tag{17}$$

Next, we define the *curvature tensor* R. For a triple X, Y, Z of vector fields, we can define a vector field

$$R(X,Y)Z = [\nabla_X, \nabla_Y]Z - \nabla_{[X,Y]}Z. \tag{18}$$

In local coordinates,

$$R(\partial_i, \partial_j)\partial_k = R^m_{kij}\partial_m. \tag{19}$$

From the definition (18), we get

$$\begin{aligned} R^m_{kij}\partial_m &= \nabla_i\nabla_j\partial_k - \nabla_j\nabla_i\partial_k \\ &= \nabla_i(\Gamma^m_{jk}\partial_m) - \nabla_j(\Gamma^m_{ik}\partial_m) \\ &= \partial_i\Gamma^m_{jk}\partial_m + \Gamma^m_{jk}\Gamma^p_{im}\partial_p - \partial_j\Gamma^m_{ik}\partial_m - \Gamma^m_{ik}\Gamma^p_{jm}\partial_p, \end{aligned} \tag{20}$$

i.e.,

$$R^m_{kij} = \partial_i\Gamma^m_{jk} - \partial_j\Gamma^m_{ik} + \Gamma^p_{jk}\Gamma^m_{ip} - \Gamma^p_{ik}\Gamma^m_{jp}. \tag{21}$$

For fixed i and j, we may think of $R^\bullet_{\bullet ij}$ as a real $n \times n$ matrix. With this notation,

$$R^\bullet_{\bullet ij} = \partial_i\Gamma^\bullet_{j\bullet} - \partial_j\Gamma^\bullet_{i\bullet} + [\Gamma^\bullet_{i\bullet}, \Gamma^\bullet_{j\bullet}] = \left[\partial_i + \Gamma^\bullet_{i\bullet}, \partial_j + \Gamma^\bullet_{j\bullet}\right]. \tag{22}$$

The curvature is antisymmetric in i and j,

$$R^m_{kij} = -R^m_{kji}. \tag{23}$$

Using Eq. (21), one checks by a direct computation that in a coordinate transformation $y = y(x)$,

$$R'^m_{kij}(y) = \frac{\partial y^m}{\partial x^q}\frac{\partial x^r}{\partial y^k}\frac{\partial x^s}{\partial y^i}\frac{\partial x^p}{\partial y^j}R^q_{rsp}(x). \tag{24}$$

Thus, R^m_{kij} is really a 4th rank tensor in contrast to the Christoffel symbols Γ^k_{ij}, which transform inhomogeneously in coordinate transformations.

Exercise 3 Check that

$$T'^k_{ij}(y) = \frac{\partial y^k}{\partial x^m}\frac{\partial x^r}{\partial y^i}\frac{\partial x^s}{\partial y^j}T^m_{rs}(x)$$

in a coordinate transformation $y = y(x)$.

Assume that the torsion T vanishes. From Eq. (21), we deduce the *first Bianchi identity*

$$R^m_{kij} + R^m_{jki} + R^m_{ijk} = 0 \tag{25}$$

for all indices. This can also be written as

$$R(X,Y)Z + R(Z,X)Y + R(Y,Z)X = 0 \tag{26}$$

for all vector fields X, Y, Z. This identity is in general not true when $T \neq 0$.

Another important tensor in general relativity is the *Ricci tensor*

$$R_{ij} = R^k_{ikj}. \tag{27}$$

Exercise 4 Show that R_{ij} transforms like a second rank tensor in coordinate transformations.

The curvature is related to the parallel transport in the following way. Consider a *very small* parallelogram with edges at $x, x+\delta x, x+\delta x+\delta y, x+\delta y$. According to the differential equation (11), determining a parallel transport, a tangent vector Y at x when parallel transported to the point $x+\delta x$ becomes approximately (in given local coordinates)

$$Y^i(x+\delta x) = Y^i(x) - \Gamma^i_{jk}(x)Y^k(x)\delta x^j. \tag{28}$$

At the next point $x+\delta x+\delta y$, we get

$$\begin{aligned}
Y^i(x+\delta x+\delta y) &= Y^i(x) - \Gamma^i_{jk}(x)Y^k(x)\delta xj \\
&- \Gamma^i_{jk}(x+\delta x)[Y^k(x) - \Gamma^k_{\ell m}(x)Y^m(x)\delta x^j]\delta y^\ell \\
&= Y^i(x) - \Gamma^i_{jk}(x)Y^k(x)\delta x^j - \Gamma^i_{jk}(x)Y^k(x)\delta y^j \\
&- \partial_m\Gamma^i_{jk}(x)\delta x^m\delta y^j Y^k(x) \\
&+ \Gamma^i_{jk}(x)\Gamma^k_{\ell m}(x)Y^m(x)\delta x^\ell\delta y^j.
\end{aligned} \tag{29}$$

In the same way, we can compute the parallel transport of Y from x to $x+\delta y$ and further to $x+\delta y+\delta x$. The parallel transport around the parallelogram is then obtained as a combination of the right-hand side of the above formula and the latter transport (note the direction of motion!); the result is

$$\begin{aligned}
\delta Y^i &= R^i_{kmj}(x)Y^k(x)\delta x^m\delta y^j \\
&= \frac{1}{2}R^i_{kmj}(x)Y^k(x)(\delta x^m\delta y^j - \delta x^j\delta y^m).
\end{aligned} \tag{30}$$

Thus, the parallel transport around the small parallelogram is proportional to the curvature at x and the area of the parallelogram.

Example 5 We compute the curvature tensor of the unit sphere S^2. Since there are only two independent coordinates, all the non-zero components of R are given by the tensor $R^i_j = R^i_{j\theta\phi} = -R^i_{j\phi\theta}$, where $i, j = \theta, \phi$. Looking at the table (Exercise) of the Christoffel symbols, we get

$$R^\theta_\phi = \sin^2\theta,\ R^\phi_\theta = -1,$$

and the other components $= 0$.

The *second Bianchi identity*

$$\partial_i R^{\bullet}_{\bullet jk} + [\Gamma^{\bullet}_{i\bullet}, R^{\bullet}_{\bullet jk}] + \partial_j R^{\bullet}_{\bullet ki} + [\Gamma^{\bullet}_{j\bullet}, R^{\bullet}_{\bullet ki}] + \partial_k R^{\bullet}_{\bullet ij} + [\Gamma^{\bullet}_{k\bullet}, R^{\bullet}_{\bullet ij}] = 0 \tag{31}$$

follows from Eq. (22) and the Jacobi identity for matrices,

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0. \tag{32}$$

3.4 Metric and Pseudo-Metric

In order to define distances and inner products between tangent vectors on a manifold M, we have to define a *metric*. A *Riemannian metric* is an inner product defined in each of the tangent spaces. That is, for each $p \in M$, we have a non-degenerate bilinear mapping

$$g_p : T_pM \times T_pM \to \mathbb{R},$$

which is symmetric, $g_p(u, v) = g_p(v, u)$ for all tangent vectors $u, v \in T_pM$, and $g_p(u, u) > 0$ for all $u \neq 0$, and it depends smoothly on the coordinates of the point p. Choosing local coordinates x_i and writing the tangent vectors in the coordinate basis, $u = u^i\partial_i$, we can write a symmetric bilinear mapping as a second rank symmetric tensor,

$$g_p(u, v) = g_{ij}u^iv^j. \tag{33}$$

Non-degenerate means that $\det(g_{ij}) \neq 0$. Since (g_{ij}) is symmetric, it can be diagonalized. Positivity of the inner product means then that all eigenvalues of g are positive.

In relativity, we need a generalization of the Riemann metric to a *pseudo-Riemannian metric*. In this generalization, we shall drop the requirement that the inner product is positive. In particular, we want to include the Minkowski space metric $(\eta_{\mu\nu})$, which has *signature* $(1, 3)$, it has one positive eigenvalue $(= 1)$ and three negative eigenvalues $(= -1)$.

A metric (or a pseudo-metric) can be used to define distances. If $\gamma(s)$ is a parametrized curve such that its tangent vector at each point on the curve has non-negative length, then we define the length of the curve (between the parameter values a and b) as

$$\ell(\gamma) = \int_a^b \sqrt{g_{\gamma(s)}(\dot\gamma(s), \dot\gamma(s))}\, ds.$$

The extremal curves $\gamma(t)$ for the functional $\ell(\gamma)$ are the geodetic curves for a certain connection (the Levi-Civita connection, see the discussion below and the exercise 2.15). Recall the *Euler-Lagrange* variational equations: Let $x(t) = (x^1(t), x^2(t), \ldots, x^n(t))$ be a vector valued function of a real variable t and

$$S(x(\cdot)) = \int_a^b L(x(t), x'(t), x''(t), \ldots)dt$$

where L is some (differentiable) function of the derivatives $x, x', x'', \ldots$. Then the derivative of S in the direction $\delta x(t)$ of a variation of the curve $x(t)$ is

$$\delta S = \sum_i \int_a^b \delta x^i(t) \left\{ \frac{\partial L}{\partial x^i} - \frac{d}{dt}\frac{\partial L}{\partial x^{i\prime}} + (\frac{d}{dt})^2 \frac{\partial L}{\partial x^{i\prime\prime}} - \ldots \right\} dt,$$

where we have used partial integration in the variable t in order to factor out δx under the integral sign. The requirement that the variation δS vanishes in arbitrary directions δx in the path space is then equivalent to the Euler-Lagrange equations

$$\frac{\partial L}{\partial x^i} - \frac{d}{dt}\frac{\partial L}{\partial x^{i\prime}} + (\frac{d}{dt})^2 \frac{\partial L}{\partial x^{i\prime\prime}} - \cdots = 0,$$

where $i = 1, 2, \ldots n$.

Example 6 If $M = \mathbb{R}^n$, then we can define a constant metric $g_{ij} = \delta_{ij}$. This is the standard Euclidean metric. In general in $\mathbb{R}^n$, a Riemannian metric is given by smooth real functions $g_{ij}(x) = g_{ji}(x)$ such that the matrix $(g_{ij}(x))$ is strictly positive for all $x \in \mathbb{R}^n$.

Example 7 If $M \subset \mathbb{R}^n$ is any smooth surface in the Euclidean space, then we can define a metric g as follows. Let $u, v \in T_pM$ be a pair of tangent vectors to the surface at the point p. The tangent vectors are also vectors in $\mathbb{R}^n$, thus we may compute the scalar product $u \cdot v$. We set $g_p(u, v) = u \cdot v$. From the fact that the Euclidean metric is positive definite follows at once that g is a positive symmetric form.

Example 8 Let $M = S^2 \subset \mathbb{R}^3$. We compute the metric g on M, as defined in Example 7, in terms of the spherical coordinates θ and ϕ. The spherical coordinates are related to the standard coordinates by

$$\begin{aligned} \partial_\theta &= \cos\theta \cos\phi \frac{\partial}{\partial x} + \cos\theta \sin\phi \frac{\partial}{\partial y} - \sin\theta \frac{\partial}{\partial z}, \\ \partial_\phi &= -\sin\theta \sin\phi \frac{\partial}{\partial x} + \sin\theta \cos\phi \frac{\partial}{\partial y}. \end{aligned}$$

¿From this we obtain the inner products

$$\begin{aligned} g_{\theta\theta} &= g(\partial_\theta, \partial_\theta) = 1, \\ g_{\phi\phi} &= g(\partial_\phi, \partial_\phi) = \sin^2\theta, \\ g_{\theta\phi} &= g_{\phi\theta} = 0. \end{aligned}$$

For example, the inner product of vectors $(1,2)$ and $(2,-1)$ (in the θ and ϕ coordinates) is $1\cdot 2\cdot g_{\theta\theta} + 2\cdot(-1)\cdot g_{\phi\phi} = 2 - 2\sin^2\theta$, at the point (θ,ϕ). Note that the spherical coordinates are *orthogonal*, the off-diagonal matrix elements of g are equal to zero.

According to the last example, the distance between to points on a sphere along a curve $\gamma(t) = (\theta(t), \phi(t))$ is given by

$$\ell(\gamma) = \int_a^b [\theta'(t)^2 + \sin^2\theta(t)\phi'(t)^2]^{1/2} dt.$$

The Euler-Lagrange equations give then (check this!)

$$\theta''(t) - \frac{1}{2}\phi'^2 \sin\theta(t) = 0 \tag{34}$$

$$\frac{d}{dt}[\phi'(t)\sin^2\theta(t)] = 0 \tag{35}$$

which agrees with the equations in example 2.

Suppose a (pseudo) metric g is given on a manifold M. From the metric, we can construct a preferred affine connection, called the *Levi-Civita connection.* Its Christoffel symbols (in given local coordinates) are given by the formula

$$\Gamma^k_{ij} = \frac{1}{2} g^{k\ell}(\partial_i g_{j\ell} + \partial_j g_{i\ell} - \partial_\ell g_{ij}), \tag{36}$$

where g^{ij} are the matrix elements of the inverse matrix g^{-1}.

One should always be extremely careful when trying to define something with the help of local coordinates. It is not *a priori* clear that the locally defined Christoffel symbols in various coordinate systems match together to define a connection on whole manifold M. To investigate the patching problem, we compute what happens in a coordinate transformation $y = y(x)$. Since

$$\frac{\partial}{\partial y^i} = \frac{\partial x^k}{\partial y^i}\frac{\partial}{\partial x^k}, \tag{37}$$

we get

$$\begin{aligned} g'_{ij}(y) &= g_y\left(\frac{\partial}{\partial y^i}, \frac{\partial}{\partial y^j}\right) = \frac{\partial x^k}{\partial y^i}\frac{\partial x^\ell}{\partial y^j} g_x\left(\frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^\ell}\right) \\ &= g_{k\ell}(x)\frac{\partial x^k}{\partial y^i}\frac{\partial x^\ell}{\partial y^j}. \end{aligned} \tag{38}$$

Inserting this transformation law into the definition (36) of Christoffel symbols, we get

$$\Gamma'^k_{ij}(y) = \frac{\partial y^k}{\partial x^c}\frac{\partial x^a}{\partial y^i}\frac{\partial x^b}{\partial y^j}\Gamma^c_{ab} + \frac{\partial y^k}{\partial x^c}\frac{\partial^2 x^c}{\partial y^i \partial y^j}, \tag{39}$$

as expected. Thus, the Christoffel symbols defined in different coordinate systems are compatible and define indeed an affine connection.

Example 9 Since the standard Euclidean metric is constant in the standard coordinates, the Christoffel symbols of the Levi-Civita connection vanish.

Example 10 The Christoffel symbols computed from the metric defined in Example agree with the Christoffel symbols of Exercise .

The Levi-Civita connection has two characteristic properties. The first property is that its torsion $T = 0$, since $\Gamma^k_{ij} - \Gamma^k_{ji} = 0$ according to Eq. (36). The second property is that the parallel transport defined by the Levi-Civita connection is *metric compatible* in the following sense: Let $X(s)$ and $Y(s)$ be a pair of parallel vector fields along a curve $\gamma(s)$. Then,

$$\frac{d}{ds} g_{\gamma(s)}(X(s), Y(s)) = 0, \tag{40}$$

i.e., the inner products of parallel vector fields are constant along the curve. This means that the parallel transport $\hat{\gamma} : T_pM \to T_qM$ between the end points of the curve is an *isometry.*

Theorem 3.4. *An affine connection ∇ is compatible with a metric g if and only if*

$$Z \cdot g(X, Y) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y)$$

for all vector fields X, Y, Z.

A word about the notation: We write $g(X, Y)$ for the real valued smooth function $p \mapsto g_p(X(p), Y(p))$. Remember that a vector field acts on functions as derivations, so the left-hand side is a well-defined smooth function, too.

Proof 1) Assume that the condition for g in the theorem is satisfied. Let $X(s)$ and $Y(s)$ be a pair of parallel vector fields along a curve $\gamma(s)$. We shall extend X and Y to vector fields defined in an open neighborhood of the curve. Let Z be some vector field defined in a neighborhood of the curve such that along the curve $Z(\gamma(s)) = \dot{\gamma}(s)$. Since X and Y are parallel along γ, we have

$$\nabla_Z X = \nabla_Z Y = 0 \text{ on the curve } \gamma.$$

Thus,

$$\frac{d}{ds} g_{\gamma(s)}(X(s), Y(s)) = Z \cdot g(X, Y) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y) = 0 \text{ on } \gamma.$$

2) Assume that ∇ is compatible with g. Let X, Y, Z be a triple of vector fields. Let $p \in M$ and γ any curve through p such that at p, $\dot{\gamma}(s_1) = Z(p)$. Define vector fields along γ by $X(s) = X(\gamma(s))$ and $Y(s) = Y(\gamma(s))$.

Let $X_1, \ldots, X_n$ be an orthonormal basis of tangent vectors at p. We define a set of parallel vector fields $X_i(s)$ along γ such that at $p = \gamma(s_1)$, we have $X_i(s_1) = X_i$. Any pair of vector fields along γ can then be written as

$$X(s) = \alpha^i(s) X_i(s),\ Y(s) = \beta^i(s) X_i(s).$$

Now, we have

$$\begin{aligned}
\frac{d}{ds}g_{\gamma(s)}(X(s),Y(s)) &= \frac{d}{ds}\alpha^i(s)\beta^j(s)g_{\gamma(s)}(X_i(s),X_j(s)) \\
&= \frac{d}{ds}\alpha^i(s)\beta^i(s) = \dot{\alpha}^i(s)\beta^i(s) + \alpha^i(s)\dot{\beta}^i(s) \\
&= g_{\gamma(s)}(\dot{\alpha}^i(s)X_i(s), \beta^j(s)X_j(s)) \\
&+ g_{\gamma(s)}(\alpha^i(s)X_i(s), \dot{\beta}^j(s)X_j(s)) \\
&= g_{\gamma(s)}(\nabla_{\dot{\gamma}}X(s), Y(s)) + g_{\gamma(s)}(X(s), \nabla_{\dot{\gamma}}Y(s)).
\end{aligned}$$

Applying this formula to the vector field Z at p, $Z(p) = \dot{\gamma}(s_1)$, we get the condition of the theorem at (the arbitrary point) p.

Theorem 3.5. *For a given metric, the Levi-Civita connection is the unique torsion free metric compatible connection.*

Proof. Use the equation in Theorem 3.4 for coordinate vector fields $X, Y, Z = \partial_i, \partial_j, \partial_k$ and the symmetry $\Gamma^k_{ij} = \Gamma^k_{ji}$ of a torsion free connection.

Theorem 3.6. *A geodesic of the Levi-Civita connection gives an extremal for the path length between two points. If the points are close enough, then the extremal gives the minimum length.*

Proof. Compare the differential equations obtained from the Euler-Lagrange variational principle, applied to curve length, with the differential equations of a geodesic, for the Levi-Civita connection. Note that the Euler-Lagrange equations obtained from the variation of the curve length are the same as obtained from variation of the integral (without square root!)

$$\int_a^b g_{x(t)}(\dot{x}(t), \dot{x}(t))dt.$$

APPENDIX: The Einstein Field Equations

The *Einstein tensor* is defined as

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R, \tag{41}$$

where $R = g^{\mu\nu}R_{\mu\nu}$ is the *Ricci scalar.* We assume that the metric $g_{\mu\nu}$ is pseudo-Riemannian of signature $(1,3)$ (one positive direction and three negative directions). The connection is the Levi-Civita connection computed from the metric and $R_{\mu\nu} = R^{\lambda}_{\mu\lambda\nu}$ is the Ricci tensor.

Exercise Writing $R_{\alpha\beta\mu\nu} = g_{\alpha\lambda}R^{\lambda}_{\beta\mu\nu}$, show that

$$R_{\alpha\beta\mu\nu} = -R_{\beta\alpha\mu\nu} = -R_{\alpha\beta\nu\mu} = R_{\mu\nu\alpha\beta}.$$

Show that this implies that $R_{\mu\nu}$ is symmetric.

The Einstein tensor is symmetric. Furthermore, its *covariant divergence* vanishes,

$$\nabla_\mu G^{\mu\nu} = \partial_\mu G^{\mu\nu} + \Gamma^{\mu}_{\mu\alpha}G^{\alpha\nu} + \Gamma^{\nu}_{\mu\alpha}G^{\mu\alpha} = 0. \tag{42}$$

This is seen as follows. First, taking $Z = \partial_\alpha, X = \partial_\mu, Y = \partial_\nu$ in Theorem 3.4, we obtain

$$\partial_\alpha g_{\mu\nu} = \Gamma^{\beta}_{\alpha\mu}g_{\beta\nu} + \Gamma^{\beta}_{\alpha\nu}g_{\mu\beta} = \Gamma_{\alpha\nu\mu} + \Gamma_{\alpha\mu\nu}. \tag{43}$$

This can be also written as

$$(\nabla_\alpha g)_{\mu\nu} = 0. \tag{44}$$

For the inverse tensor $g^{\mu\nu} = (g^{-1})_{\mu\nu}$, one gets

$$\partial_\alpha g^{\mu\nu} + \Gamma^{\nu}_{\alpha\beta}g^{\mu\beta} + \Gamma^{\mu}_{\alpha\beta}g^{\beta\nu} = 0. \tag{45}$$

Note the difference in sign for the covariant derivative of the metric tensor and its inverse.

Exercise For any vector field $X = X^\mu\partial_\mu$ the components of the covariant derivatives are $(\nabla_\nu X)^\mu = \partial_\nu X^\mu + \Gamma^{\mu}_{\nu\alpha}X^\alpha$. Show that the *covariant divergence* is given by

$$(\nabla_\mu X)^\mu = (-\det g)^{-1/2}\partial_\mu((-\det g)^{1/2}X^\mu).$$

In relativity theory literature, it is a custom to use the abbreviation $X_{\mu;\nu} = (\nabla_\nu X)_\mu$ for the covariant differentiation of vector (and higher order tensor) indices. With this notation, we can write the second Bianchi identity as

$$R_{\alpha\beta\mu\nu;\lambda} + R_{\alpha\beta\nu\lambda;\mu} + R_{\alpha\beta\lambda\mu;\nu} = 0. \tag{46}$$

Contracting the α and μ indices in this identity with the metric tensor, we get

$$g^{\alpha\mu}(R_{\alpha\beta\mu\nu;\lambda} + R_{\alpha\beta\nu\lambda;\mu} + R_{\alpha\beta\lambda\mu;\nu}) = 0. \tag{47}$$

By the definition of the Ricci tensor, this can be written as

$$R_{\beta\nu;\lambda} + R^{\mu}_{\beta\nu\lambda;\mu} - R_{\beta\lambda;\nu} = 0, \tag{48}$$

where we have taken into account that the covariant derivative of $g^{\mu\nu}$ vanishes, implying that the multiplication with the components of the metric tensor commutes with covariant differentiation; in particular, index raising and lowering commutes with covariant derivatives. Using the results of Exercise , we get

$$g^{\alpha\mu}R_{\alpha\beta\lambda\mu;\nu} = -g^{\alpha\mu}R_{\alpha\beta\mu\lambda;\nu} = -R_{\beta\lambda;\nu}. \tag{49}$$

Contracting Eq. (48) once again with $g^{\beta\nu}$, we get

$$g^{\beta\nu}(R_{\beta\nu;\lambda} + R^{\mu}_{\beta\nu\lambda;\mu} - R_{\beta\lambda;\nu}) = 0, \tag{50}$$

or in other words,

$$R_{;\lambda} - R^{\mu}_{\lambda;\mu} - R^{\nu}_{\lambda;\nu} = 0. \tag{51}$$

Note that since R is a scalar, $R_{;\mu} = \partial_\mu R$. An equivalent form of the previous equation is

$$\left(2R^{\mu}_{\lambda} - \delta^{\mu}_{\lambda}R\right)_{;\mu} = 0. \tag{52}$$

Raising the index λ and dividing by 2 finally leads to

$$\left(R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R\right)_{;\mu} = 0. \tag{53}$$

Einstein's gravitational field equations are written simply as

$$G^{\mu\nu} = 8\pi\frac{G}{c^4}T^{\mu\nu}, \tag{54}$$

where G on the right-hand side (not to be confused with Einstein's tensor!) is Newton's gravitational constant and $T^{\mu\nu}$ is the *stress-energy (energy-momentum) tensor.* It describes the distribution of matter and energy in space-time. For example, the electromagnetic field gives a contribution to $T_{\mu\nu}$ defined by $T^{EM}_{\mu\nu} = \epsilon_0 F_{\mu}{}^{\lambda}F_{\lambda\nu} + \frac{\epsilon_0}{4}g_{\mu\nu}F^{\lambda\omega}F_{\lambda\omega}$.

Another example is the energy-momentum tensor of a *perfect fluid* .A perfect fluid is characterized by a 4-velocity field u, a scalar density field ρ_0 and a scalar pressure field p. The energy-momentum tensor is defines as

$$T_{\mu\nu} = (\rho_0 + p)u_\mu u_\nu - p g_{\mu\nu}.$$

A special case of this is $p = 0$ which can be viewed as the energy momentum tensor of a flow of noninteracting dust particles. Normally p and ρ_0 are not independent but they are related by the equation of state of the form $p = p(\rho_0, T)$, where T is the temperature. The requirement that the covariant divergence of the energy-momentum tensor vanishes leads to equations of motion for the perfect fluid. In fact, in case of Minkowski space-time and in a certain limit one gets the classical *Navier-Stokes equations* (from $\partial^\mu T_{\mu k} = 0$ for $k = 1, 2, 3$),

$$\rho\left[\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u}\right] = -\nabla p$$

and the continuity equation (from $\partial^\mu T_{\mu 0} = 0$),

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\mathbf{u}) = 0.$$

Here $\rho = \rho_0(1-\mathbf{u}^2)$.

Let S be some space-like surface with a time-like unit normal vector field n^μ, $n_0 > 0$. Then,

$$\int_S (-\det g)^{1/2} T^{\mu\nu} n_\nu \, d^3x$$

gives the energy and momentum contained in S. Equation (42) leads to the following conservation law of energy and momentum. Suppose that the metric $g_{\alpha\beta}$ does not depend on a particular coordinate x_μ. Then,

$$0 = \partial_\mu g_{\alpha\beta} = \Gamma_{\mu\beta\alpha} + \Gamma_{\mu\alpha\beta} = \Gamma_{\alpha\beta\mu} + \Gamma_{\beta\alpha\mu}. \tag{55}$$

Thus, $\Gamma_{\alpha\beta\mu}$ is antisymmetric in the first two indices. Now,

$$(\nabla_\nu T)^\nu_{\ \mu} = \partial_\nu T^\nu_{\ \mu} + \Gamma^\nu_{\nu\lambda} T^\lambda_{\ \mu} - \Gamma^\lambda_{\nu\mu} T^\nu_{\ \lambda}. \tag{56}$$

The third term on the right-hand side is equal to $-\Gamma_{\nu\lambda\mu}T^{\nu\lambda}$ and it vanishes because the second factor is symmetric in its indices, whereas the first factor is antisymmetric in λ and ν by the remark above. On the other hand, the sum of the first two terms is $(-g)^{-1/2}\partial_\nu[(-g)^{1/2}T^\nu_{\ \mu}]$, according to the result of Exercise . Thus, for fixed μ, $j^\nu = (-g)^{1/2}T^\nu_{\ \mu}$ is conserved in the usual sense,

$$\partial_\nu j^\nu = 0. \tag{57}$$

In order to avoid convergence problems with the infinite integrals, we assume that all energy and momentum are contained in a compact region K in space-time. Consider a surface S, consisting of two space-like components S_1 and S_2 and some surface S_3 'far away' such that T vanishes on S_3. Using Gauss' law and the current conservation, we conclude that the surface integral of $(-\det g)^{1/2}T^\nu_{\ \mu}n_\nu$ over S vanishes. In other words,

$$\int_{S_1} (-\det g)^{1/2} T^\nu_{\ \mu} n_\nu \, d^3x = \int_{S_2} (-\det g)^{1/2} T^\nu_{\ \mu} n_\nu \, d^3x. \tag{58}$$

We have taken into account that, since n is future pointing, one of the normal vector fields on S_1 and S_2 is outward directed and the second inward directed. Equation (58) tells us that the stress-energy, in the μ-direction, on S_1 is the same as the corresponding quantity on S_2; one could think of S_i as a fixed time slice at time t_i and one obtains the usual law of conservation of energy or momentum.

Often one uses units in which $G = 1$ and $c = 1$ so that one does not need to write explicitly the coefficient G/c^4 in Einstein's equations.

4. The Newtonian Limit

It is known that the Newtonian gravitational theory is valid for fields, which can produce only velocities much smaller than the velocity of light. Since the components T^{0i} and T^{ij}

are related to spatial momenta and T^{00} is related to energy, this condition says that $|T^{00}|$ is much larger than the other components. Because of Einstein's equations, the same is true for the components of the Einstein tensor. Furthermore, we expect that for weak gravitational fields the metric $g^{\mu\nu}$ differs slightly from the Minkowski metric $\eta^{\mu\nu}$,

$$g^{\mu\nu} = \eta^{\mu\nu} + h^{\mu\nu} \tag{59}$$

for a small perturbation $h^{\mu\nu}$. Next, we compute the connection, curvature, and finally the Ricci tensor to first order in the perturbation $h^{\mu\nu}$. A straight-forward computation, starting from the definitions of the various tensors, gives $G^{\mu\nu} = -\frac{1}{2}\Box(h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}h)$, where $h = \eta_{\mu\nu}h^{\mu\nu}$. Thus, Einstein's equations, in this approximation, are linear,

$$-\frac{1}{2}\Box\left(h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}h\right) = 8\pi\frac{G}{c^4}T^{\mu\nu}. \tag{60}$$

Taking into account the remark in the beginning of this section, only the 00-component is relevant,

$$\Box\left(h^{00} - \frac{1}{2}h\right) = -16\pi\frac{G}{c^2}\rho, \tag{61}$$

where $\rho = T^{00}/c^2$ is the matter density in the rest system of the source. We can also drop the time derivatives (in the system of coordinates, where the source is slowly moving, because $\partial_0 = \frac{1}{c}\partial_t$) and so the only relevant equation becomes

$$\nabla^2\left(h^{00} - \frac{1}{2}h\right) = 16\pi\frac{G}{c^2}\rho. \tag{62}$$

This means that,

$$h^{00} - \frac{1}{2}h = \frac{4}{c^2}\phi, \tag{63}$$

where ϕ is the gravitational potential for the matter distribution ρ. (Compare Eq. (62) with the Newtonian equation $\nabla^2\phi = 4\pi G\rho$, where $\phi = -GM/r$!)

Since all the other components of $h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}h$ vanish at this order of approximation, we finally get

$$h^{\mu\mu} = \frac{2}{c^2}\phi = -\frac{2GM}{c^2 r} \text{ (no summation!)} \tag{64}$$

for all $\mu = 0, 1, 2, 3$.

Next, we shall compute the geodesics for the metric $g^{\mu\nu} = \eta^{\mu\nu} + h^{\mu\nu}$ in the linear approximation (we neglect higher order terms in $h^{\mu\nu}$). For small velocities, the time component $\dot{x}_0(s)$ of the 4-velocity is much larger than the spatial components. For this reason, we can approximate the geodesic equations of motion as

$$\frac{d^2x^\mu}{ds^2} + \Gamma^\mu_{00}\left(\frac{dx^0}{ds}\right)^2 = 0. \tag{65}$$

In the linear approximation,

$$\Gamma^0_{00} = \partial_0 \phi, \ \ \Gamma^i_{00} = \partial_i \phi. \tag{66}$$

Thus, the geodesic equations become

$$\ddot{x}_0 + \partial_0 \phi (\dot{x}_0)^2 = 0, \ \ \ddot{x}_i + \partial_i \phi (\dot{x}_0)^2 = 0. \tag{67}$$

In the coordinate system, where the source is at rest, the first equation says that we can choose the time t as the geodesic parameter, $x_0(s) = s = ct$, and then the second equation becomes

$$\ddot{x}_i = -\partial_i \phi. \tag{68}$$

The right-hand side (after multiplication by the mass m of the test particle) is the gravitational force of the source on m, so this equation is just Newton's second law, $m\mathbf{a} = \mathbf{F}$, where $\mathbf{F} = -\nabla \Phi$ and $\Phi = m\phi$.

5. The Schwarzschild Metric

The basic problem in Newtonian celestial mechanics is to solve the equations of motions outside of a spherically symmetric mass distribution (orbits of the planets around the Sun, orbits of satellites around the Earth). In general relativity the first natural problem is to search for spherically symmetric solutions of Einstein's equations.

Actually, there is a unique 1-parameter family of spherically symmetric solutions, which are *asymptotically flat*, meaning that at large distances from the source the metric tends to the flat Minkowski metric $ds^2 = dx_0^2 - dx_1^2 - dx_2^2 - dx_3^2$. This is the content of Birkhoff's theorem (which we are not going to prove). The line element of the metric is given as

$$ds^2 = \left(1 - \frac{2GM}{c^2 r}\right) dx_0^2 - \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2 - r^2 d\Omega^2, \tag{69}$$

where $d\Omega^2$ is the angular part of the Euclidean metric in $\mathbb{R}^3$, $d\Omega^2 = d\theta^2 + \sin^2\theta \, d\phi^2$. It is clear from Eq. (69) that for large distances r the metric approaches the Minkowski metric. The line element (69) is called the *Schwarzschild metric.*

When $r > 2GM/c^2$ the Schwarzschild metric is supposed to describe the gravitational field outside of a spherically symmetric star. The other disconnect region $r < 2GM/c^2$ is the *Schwarzschild black hole.* The singularity at $r = 2GM/c^2$ is actually due to a bad choice of coordinates. There is a way to glue the inside solution in a smooth way to the outside solution by a suitable choice of coordinates; the complete discussion of this was first given by Kruskal and Szekeres in 1960. The Kruskal–Szekeres metric is given as follows. The coordinates are denoted by (u, v, θ, ϕ). The latter two are the ordinary spherical coordinates on a unit sphere. The coordinates (u, v) are restricted to the region $L \subset \mathbb{R}^2$ defined by

$$uv < \frac{2GM}{c^2 e}.$$

The metric is then

$$ds^2 = \frac{16\mu^2}{r} e^{(2\mu - r)/2\mu} du dv - r^2 d\Omega^2, \tag{70}$$

where $\mu = MG/c^2$ and r is a function of u, v defined by

$$uv = (2\mu - r)e^{(r-2\mu)/2\mu}. \tag{71}$$

Note that $f(x) = xe^{x/a}$ is monotonically increasing when $x > -a$ (and $f(x) > -a/e$) and therefore $y = f(x)$ has a unique solution x for any $y > -a/e$. We treat u as a kind of universal time; a time-like vector is future directed if its projection to ∂_u is positive. The orientation (needed in integration!) is defined by the ordering (v, u, θ, ϕ) of coordinates. Note that the radial null lines (radial light rays) are given by $du = 0$ or $dv = 0$.

The Kruskal–Szekeres space-time can be divided into four regions: K_1 consists of points $v > 0$, $u < 0$, region K_2 of points $u, v > 0$, in region K_3 we have $u, v < 0$, and finally region K_4 is characterized by $u > 0$, $v < 0$. The boundaries between these regions are non-singular points for the metric. The only singularities are at the boundary $uv = 2\mu/e$.

The region K_1 is equivalent with the outer region of a Schwarzschild space-time. This is seen by performing the coordinate transformation $(v, u, \theta, \phi) \mapsto (t, r, \theta, \phi)$, where $r = r(u, v)$ as above and the Schwarzschild time is $t = 2\mu \ln(-v/u)$. With a similar coordinate transformation the region K_3 is seen to be equivalent with the outer Schwarzschild solution. The region K_2 is equivalent with the Schwarzschild black hole. The equivalence is obtained through the coordinate transformation $(v, u, \theta, \phi) \mapsto (t, r, \theta, \phi)$, where $r = r(u, v)$ is the same as before but now $t = 2\mu \ln(v/u)$.

It is easy to construct smooth time-like curves which go from either K_1 or K_3 to the black hole K_2. However, we shall prove that once an observer falls to the black hole there is no way to go back to the 'normal' regions K_1 and K_3.

Let $x(t)$ be the time-like path of the observer. Then along the path

$$\frac{dr}{dt} = \frac{\partial r}{\partial u}\frac{du}{dt} + \frac{\partial r}{\partial v}\frac{dv}{dt} = \frac{r}{8\mu^2} e^{(r-2\mu)/2\mu} \left[\frac{\partial r}{\partial u} g(\partial_v, x'(t)) + \frac{\partial r}{\partial v} g(\partial_u, x'(t)) \right] < 0,$$

since $x(t)$ is time-like and in K_2 holds $r\frac{\partial r}{\partial u} = -2\mu v e^{(2\mu - r)/2\mu} < 0$ and similarly for the v-coordinate.

The boundary between K_2 and the normal regions is $r = 2\mu$ (*i.e.*, $u = 0$ or $v = 0$). The function $r(x(t))$ was seen to be decreasing, and therefore the path $x(t)$ can never hit the boundary $r = 2\mu$. But the observer entering K_2 has a deplorable future, since it will eventually hit the true singularity $r = 0$, again using the monotonicity of the function $r(x(t))$.

There is also another singularity, the outer boundary of region K_3. But this is of no great concern because it is in the past; no future directed time-like curve can enter that singularity.