Chapter 6

Curved spacetime and General Relativity

6.1 Manifolds, tangent spaces and local inertial frames

A manifold is a continuous space whose points can be assigned coordinates, the number of coordinates being the dimension of the manifold [for example a surface of a sphere is 2D, spacetime is 4D].

A manifold is differentiable if we can define a scalar field ϕ at each point which can be differentiated everywhere. This is always true in Special Relativity and General Relativity.

We can then define one-forms $\tilde{\mathbf{d}}\phi$ as having components $\{\phi_{,\alpha} \equiv \frac{\partial \phi}{\partial x^\alpha}\}$ and vectors $\mathbf{V}$ as linear functions which take $\tilde{\mathbf{d}}\phi$ into the derivative of ϕ along a curve with tangent $\mathbf{V}$:

$$\mathbf{V}\left(\tilde{\mathbf{d}}\phi\right) = \nabla_{\mathbf{V}}\phi = \phi_{,\alpha}V^\alpha = \frac{d\phi}{d\lambda}\,. \tag{6.1}$$

Tensors can then be defined as maps from one-forms and vectors into the reals [see chapter 3].

A Riemannian manifold is a differentiable manifold with a symmetric metric tensor g at each point such that

$$g\left(\mathbf{V},\mathbf{V}\right) > 0 \tag{6.2}$$

for any vector $\mathbf{V}$, for example Euclidian 3D space.

If however $g(\mathbf{V}, \mathbf{V})$ is of indefinite sign as it is in Special and General Relativity it is called Pseudo - Riemannian.

For a general spacetime with coordinates $\{x^\alpha\}$, the interval between two neighboring points is

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta \; . \tag{6.3}$$

In Special Relativity we can choose Minkowski coordinates such that $g_{\alpha\beta} = \eta_{\alpha\beta}$ everywhere. This will not be true for a general curved manifold. Since $g_{\alpha\beta}$ is a symmetric matrix, we can always choose a coordinate system at each point $\mathbf{x}_0$ in which it is transformed to the diagonal Minkowski form, i.e. there is a transformation

$$\Lambda^{\bar{\alpha}}{}_{\beta} \equiv \frac{\partial x^{\bar{\alpha}}}{\partial x^{\beta}} \tag{6.4}$$

such that

$$g_{\bar{\alpha}\bar{\beta}}(\mathbf{x}_0) = \eta_{\bar{\alpha}\bar{\beta}} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{6.5}$$

Note that the sum of the diagonal elements is conserved; this is the signature of the metric [+2].

In general $\Lambda^{\bar{\alpha}}{}_{\beta}$ will not diagonalize $g_{\alpha\beta}$ at every point since there are ten functions $g_{\alpha\beta}(\mathbf{x})$ and only four transformation functions $X^{\bar{\alpha}}(x^\beta)$.

We can also choose $\Lambda^{\bar{\alpha}}{}_{\beta}$ so that the first derivatives of the metric vanishes at $\mathbf{x}_0$ i.e.

$$\frac{\partial g_{\alpha\beta}}{\partial x^\gamma} = 0 \tag{6.6}$$

for all α, β and γ. This implies

$$g_{\alpha\beta}(x^\mu) = \eta_{\alpha\beta} + \mathcal{O}[(x^\mu)^2] \; . \tag{6.7}$$

That is, the metric near $\mathbf{x}_0$ is approximately that of Special Relativity, differences being of second order in the coordinates. This corresponds to the local inertial frame whose existence was deduced from the equivalence principle.

In summary we can define a local inertial frame to be one where

$$g_{\alpha\beta}(\mathbf{x}_0) = \eta_{\alpha\beta} \tag{6.8}$$

for all α, β, and

$$g_{\alpha\beta,\gamma}(\mathbf{x}_0) = 0 \tag{6.9}$$

for all α, β, γ. However

$$g_{\alpha\beta,\gamma\mu}(\mathbf{x}_0) \neq 0 \tag{6.10}$$

for at least some values of α. β, γ and μ.

It reflects the fact that any curved space has a flat tangent space at every point, although these tangent spaces cannot be meshed together into a global flat space. Recall that straight lines in a flat spacetime are the worldlines of free particles; the absence of first derivative terms in the metric of a curved spacetime will mean that free particles are moving on lines that are locally straight in that coordinate system. This makes such coordinates very useful for us, since the equations of physics will be nearly as simple as they are in flat spacetime, and if they are tensor equations they will be valid in every coordinate system.

6.2 Covariant derivatives and Christoffel symbols

In Minkowski spacetime with Minkowski coordinates (t, x, y, z) the derivative of a vector $\mathbf{V} = V^\alpha \mathbf{e}_\alpha$ is just

$$\frac{\partial \mathbf{V}}{\partial x^\beta} = \frac{\partial V^\alpha}{\partial x^\beta}\mathbf{e}_\alpha \ , \tag{6.11}$$

since the basis vectors do not vary. In a general spacetime with arbitrary coordinates, $\mathbf{e}_\alpha$ which vary from point to point so

$$\frac{\partial \mathbf{V}}{\partial x^\beta} = \frac{\partial V^\alpha}{\partial x^\beta}\mathbf{e}_\alpha + V^\alpha \frac{\partial \mathbf{e}_\alpha}{\partial x^\beta} \ . \tag{6.12}$$

Since $\partial \mathbf{e}_\alpha / \partial x^\beta$ is itself a vector for a given β it can be written as a linear combination of the bases vectors:

$$\frac{\partial \mathbf{e}_\alpha}{\partial x^\beta} = \Gamma^\mu{}_{\alpha\beta}\mathbf{e}_\mu \ . \tag{6.13}$$

The Γ's are called Christoffel symbols [or the metric connection]. Thus we have:

$$\frac{\partial \mathbf{V}}{\partial x^\beta} = \frac{\partial V^\alpha}{\partial x^\beta}\mathbf{e}_\alpha + V^\alpha \Gamma^\mu{}_{\alpha\beta}\mathbf{e}_\mu \ , \tag{6.14}$$

so

$$\frac{\partial \mathbf{V}}{\partial x^\beta} = \left(\frac{\partial V^\alpha}{\partial x^\beta} \mathbf{e}_\alpha + V^\mu \Gamma^\alpha{}_{\mu\beta} \right) \mathbf{e}_\alpha \ . \tag{6.15}$$

Thus we can write

$$\frac{\partial \mathbf{V}}{\partial x^\beta} = V^\alpha{}_{;\beta} \mathbf{e}_\alpha \ , \tag{6.16}$$

where

$$V^\alpha{}_{;\beta} = V^\alpha{}_{,\beta} + V^\mu \Gamma^\alpha{}_{\mu\beta} \ . \tag{6.17}$$

Let us now prove that $V^\alpha{}_{;\beta}$ are the components of a 1/1 tensor. Remember in section 3.5 we found that $V^\alpha{}_{,\beta}$ was only a tensor under Poincaré transformations in Minkowski space with Minkowski coordinates. $V^\alpha{}_{;\beta}$ is the natural generalization for a general coordinate transformation.

Writing $\Lambda^{\bar\alpha}{}_\beta \equiv \frac{\partial x^{\bar\alpha}}{\partial x^\beta}$, we have:

$$\begin{aligned} V^{\bar\alpha}{}_{;\bar\beta} &= V^{\bar\alpha}{}_{,\bar\beta} + V^{\bar\mu} \Gamma^{\bar\alpha}{}_{\bar\mu\bar\beta} \\ &= \Lambda^\mu{}_{\bar\beta} \frac{\partial}{\partial x^\mu} \left(\Lambda^{\bar\alpha}{}_\nu V^\nu \right) + V^{\bar\mu} \Gamma^{\bar\alpha}{}_{\bar\mu\bar\beta} \\ &= \Lambda^\mu{}_{\bar\beta} \Lambda^{\bar\alpha}{}_\nu \frac{\partial V^\nu}{\partial x^\mu} + V^\nu \Lambda^\mu{}_{\bar\beta} \frac{\partial \Lambda^{\bar\alpha}{}_\nu}{\partial x^\mu} + V^{\bar\mu} \Gamma^{\bar\alpha}{}_{\bar\mu\bar\beta} \ . \end{aligned} \tag{6.18}$$

Now

$$\frac{\partial \mathbf{e}_{\bar\mu}}{\partial x^{\bar\beta}} = \Gamma^{\bar\lambda}{}_{\bar\mu\bar\beta} \mathbf{e}_{\bar\lambda} \ , \tag{6.19}$$

therefore

$$\frac{\partial \mathbf{e}_{\bar\mu}}{\partial x^{\bar\beta}} \tilde{\mathbf{w}}^{\bar\alpha} = \Gamma^{\bar\lambda}{}_{\bar\mu\bar\beta} \mathbf{e}_{\bar\lambda} \tilde{\mathbf{w}}^{\bar\alpha} = \Gamma^{\bar\lambda}{}_{\bar\mu\bar\beta} \delta^{\bar\alpha}{}_{\bar\lambda} = \Gamma^{\bar\alpha}{}_{\bar\mu\bar\beta} \ , \tag{6.20}$$

so we obtain:

$$\begin{aligned} V^{\bar\alpha}{}_{;\bar\beta} &= \Lambda^\mu{}_{\bar\beta} \Lambda^{\bar\alpha}{}_\nu \frac{\partial V^\nu}{\partial x^\mu} + V^\nu \Lambda^\mu{}_{\bar\beta} \frac{\partial \Lambda^{\bar\alpha}{}_\nu}{\partial x^\mu} + V^{\bar\mu} \frac{\partial \mathbf{e}_{\bar\mu}}{\partial x^{\bar\beta}} \tilde{\mathbf{w}}^{\bar\alpha} \\ &= \Lambda^\mu{}_{\bar\beta} \Lambda^{\bar\alpha}{}_\nu \frac{\partial V^\nu}{\partial x^\mu} + V^\nu \Lambda^\mu{}_{\bar\beta} \frac{\partial \Lambda^{\bar\alpha}{}_\nu}{\partial x^\mu} + \Lambda^{\bar\mu}{}_\gamma V^\gamma \Lambda^{\bar\alpha}{}_\delta \tilde{\mathbf{w}}^\delta \frac{\partial \mathbf{e}_{\bar\mu}}{\partial x^{\bar\beta}} \\ &= \Lambda^\mu{}_{\bar\beta} \Lambda^{\bar\alpha}{}_\nu \frac{\partial V^\nu}{\partial x^\mu} + V^\nu \Lambda^\mu{}_{\bar\beta} \frac{\partial \Lambda^{\bar\alpha}{}_\nu}{\partial x^\mu} + \Lambda^{\bar\mu}{}_\gamma V^\gamma \Lambda^{\bar\alpha}{}_\delta \tilde{\mathbf{w}}^\delta \Lambda^\epsilon{}_{\bar\beta} \frac{\partial}{\partial x^\epsilon} \left(\Lambda^\nu{}_{\bar\mu} \mathbf{e}_\nu \right) \\ &= \Lambda^\mu{}_{\bar\beta} \Lambda^{\bar\alpha}{}_\nu \frac{\partial V^\nu}{\partial x^\mu} + V^\nu \Lambda^\mu{}_{\bar\beta} \frac{\partial \Lambda^{\bar\alpha}{}_\nu}{\partial x^\mu} + \Lambda^{\bar\mu}{}_\gamma V^\gamma \Lambda^{\bar\alpha}{}_\delta \tilde{\mathbf{w}}^\delta \Lambda^\epsilon{}_{\bar\beta} \Lambda^\nu{}_{\bar\mu} \frac{\partial \mathbf{e}_\nu}{\partial x^\epsilon} \\ &+ \Lambda^{\bar\mu}{}_\gamma V^\gamma \Lambda^{\bar\alpha}{}_\delta \tilde{\mathbf{w}}^\delta \Lambda^\epsilon{}_{\bar\beta} \mathbf{e}_\nu \frac{\partial \Lambda^\nu{}_{\bar\mu}}{\partial x^\epsilon} \end{aligned} \tag{6.21}$$

Now using $\Lambda^{\nu}{}_{\bar{\mu}}\Lambda^{\bar{\mu}}{}_{\gamma} = \delta^{\nu}{}_{\gamma}$, $\tilde{\mathbf{w}}^{\delta}\mathbf{e}_{\nu} = \delta^{\delta}{}_{\nu}$ and $\Lambda^{\bar{\alpha}}{}_{\nu}\frac{\partial \Lambda^{\nu}{}_{\bar{\mu}}}{\partial x^{\epsilon}} = -\Lambda^{\nu}{}_{\bar{\mu}}\frac{\partial \Lambda^{\bar{\alpha}}{}_{\nu}}{\partial x^{\epsilon}}$ we obtain:

$$\begin{aligned} V^{\bar{\alpha}}{}_{;\bar{\beta}} &= \Lambda^{\bar{\alpha}}{}_{\nu}\Lambda^{\mu}{}_{\bar{\beta}}\frac{\partial V^{\nu}}{\partial x^{\mu}} + V^{\nu}\Lambda^{\mu}{}_{\bar{\beta}}\frac{\partial \Lambda^{\bar{\alpha}}{}_{\nu}}{\partial x^{\mu}} + \Lambda^{\bar{\alpha}}{}_{\delta}\Lambda^{\epsilon}{}_{\bar{\beta}}V^{\nu}\Gamma^{\delta}{}_{\nu\epsilon} - \Lambda^{\mu}{}_{\bar{\beta}}\frac{\partial \Lambda^{\bar{\alpha}}{}_{\nu}}{\partial x^{\mu}}V^{\nu} \\ &= \Lambda^{\bar{\alpha}}{}_{\nu}\Lambda^{\mu}{}_{\bar{\beta}}\frac{\partial V^{\nu}}{\partial x^{\mu}} + \Lambda^{\bar{\alpha}}{}_{\delta}\Lambda^{\epsilon}{}_{\bar{\beta}}V^{\nu}\Gamma^{\delta}{}_{\nu\epsilon} \\ &= \Lambda^{\bar{\alpha}}{}_{\nu}\Lambda^{\mu}{}_{\bar{\beta}}\left(\frac{\partial V^{\nu}}{\partial x^{\mu}} + V^{\delta}\Gamma^{\nu}{}_{\delta\mu}\right) , \end{aligned} \tag{6.22}$$

so

$$V^{\bar{\alpha}}{}_{;\bar{\beta}} = \Lambda^{\bar{\alpha}}{}_{\nu}\Lambda^{\mu}{}_{\bar{\beta}}V^{\nu}{}_{;\mu} \ . \tag{6.23}$$

We have shown that $V^{\alpha}{}_{;\beta}$ are indeed the components of a 1/1 tensor. We write this tensor as

$$\nabla \mathbf{V} = V^{\alpha}{}_{;\beta}\mathbf{e}_{\alpha} \otimes \tilde{\mathbf{w}}^{\beta} \ . \tag{6.24}$$

It is called the <u>covariant derivative</u> of $\mathbf{V}$. Using a Cartesian basis, the components are just $V^{\alpha}{}_{,\beta}$, but this is not true in general; however for a scalar ϕ we have:

$$\nabla_{\alpha}\phi \equiv \phi_{;\alpha} = \frac{\partial \phi}{\partial x^{\alpha}} \ , \quad \nabla\phi = \tilde{\mathbf{d}}\phi \ , \tag{6.25}$$

since scalars do not depend on basis vectors.

Writing $\Gamma^{\alpha}{}_{\mu\beta} = \frac{\partial \mathbf{e}_{\mu}}{\partial x^{\beta}}\tilde{\mathbf{w}}^{\alpha}$, we can find the transformation law for the components of the Christoffel symbols.

$$\begin{aligned} \Gamma^{\alpha}{}_{\mu\beta} &= \frac{\partial \mathbf{e}_{\mu}}{\partial x^{\beta}}\tilde{\mathbf{w}}^{\alpha} = \Lambda^{\alpha}{}_{\bar{\gamma}}\tilde{\mathbf{w}}^{\bar{\gamma}}\Lambda^{\bar{\sigma}}{}_{\beta}\frac{\partial}{\partial x^{\bar{\sigma}}}\left(\Lambda^{\bar{\lambda}}{}_{\mu}\mathbf{e}_{\bar{\lambda}}\right) \\ &= \Lambda^{\alpha}{}_{\bar{\gamma}}\Lambda^{\bar{\sigma}}{}_{\beta}\Lambda^{\bar{\lambda}}{}_{\mu}\tilde{\mathbf{w}}^{\bar{\gamma}}\frac{\partial \mathbf{e}_{\bar{\lambda}}}{\partial x^{\bar{\sigma}}} + \Lambda^{\alpha}{}_{\bar{\gamma}}\Lambda^{\bar{\sigma}}{}_{\beta}\tilde{\mathbf{w}}^{\bar{\gamma}}\mathbf{e}_{\bar{\lambda}}\frac{\partial \Lambda^{\bar{\lambda}}{}_{\mu}}{\partial x^{\bar{\sigma}}} \\ &= \Lambda^{\alpha}{}_{\bar{\gamma}}\Lambda^{\bar{\sigma}}{}_{\beta}\Lambda^{\bar{\lambda}}{}_{\mu}\Gamma^{\bar{\gamma}}{}_{\bar{\lambda}\bar{\sigma}} + \Lambda^{\alpha}{}_{\bar{\lambda}}\Lambda^{\bar{\sigma}}{}_{\beta}\frac{\partial \Lambda^{\bar{\lambda}}{}_{\mu}}{\partial x^{\bar{\sigma}}} \end{aligned} \tag{6.26}$$

This is just

$$\Gamma^{\alpha}{}_{\mu\beta} = \frac{\partial x^{\alpha}}{\partial x^{\bar{\gamma}}}\frac{\partial x^{\bar{\sigma}}}{\partial x^{\beta}}\frac{\partial x^{\bar{\lambda}}}{\partial x^{\mu}}\Gamma^{\bar{\gamma}}{}_{\bar{\lambda}\bar{\sigma}} + \frac{\partial x^{\alpha}}{\partial x^{\bar{\lambda}}}\frac{\partial^2 x^{\bar{\lambda}}}{\partial x^{\beta}\partial x^{\mu}} \ . \tag{6.27}$$

We can calculate the covariant derivative of a one-form $\tilde{\mathbf{p}}$ by using the fact that $\tilde{\mathbf{p}}(\mathbf{V})$ is a scalar for any vector $\mathbf{V}$:

$$\phi = p_{\alpha}V^{\alpha} \ . \tag{6.28}$$

We have

$$\begin{aligned}\nabla_\beta\phi &= \phi_{,\beta} = \frac{\partial p_\alpha}{\partial x^\beta}V^\alpha + p_\alpha\frac{\partial V^\alpha}{\partial x^\beta} \\ &= \frac{\partial p_\alpha}{\partial x^\beta}V^\alpha + p_\alpha V^\alpha{}_{;\beta} - p_\alpha V^\mu\Gamma^\alpha{}_{\mu\beta} \\ &= \left(\frac{\partial p_\alpha}{\partial x^\beta} - p_\mu\Gamma^\mu{}_{\alpha\beta}\right)V^\alpha + p_\alpha V^\alpha{}_{;\beta}\ .\end{aligned} \tag{6.29}$$

Since $\nabla_\beta\phi$ and $V^\alpha{}_{;\beta}$ are tensors, the term in the parenthesis is a tensor with components:

$$p_{\alpha;\beta} = p_{\alpha,\beta} - p_\mu\Gamma^\mu{}_{\alpha\beta}\ . \tag{6.30}$$

We can extend this argument to show that

$$\begin{aligned}\nabla_\beta T_{\mu\nu} &\equiv T_{\mu\nu;\beta} = T_{\mu\nu,\beta} - T_{\alpha\nu}\Gamma^\alpha{}_{\mu\beta} - \Gamma_{\mu\alpha}\Gamma^\alpha{}_{\nu\beta}\ , \\ \nabla_\beta T^{\mu\nu} &\equiv T^{\mu\nu}{}_{;\beta} = T^{\mu\nu}{}_{;\beta} + T^{\alpha\nu}\Gamma^\mu{}_{\alpha\beta} + T^{\mu\alpha}\Gamma^\nu{}_{\alpha\beta}\ , \\ \nabla_\beta T^\mu{}_\nu &\equiv T^\mu{}_{\nu;\beta} = T^\mu{}_{\nu,\beta} + T^\alpha{}_\nu\Gamma^\mu{}_{\alpha\beta} - T^\mu{}_\alpha\Gamma^\alpha{}_{\nu\beta}\ .\end{aligned} \tag{6.31}$$

6.3 Calculating $\Gamma^\alpha{}_{\beta\gamma}$ from the metric

Since $V^\alpha{}_{;\beta}$ is a tensor we can lower the index α using the metric tensor:

$$V_{\alpha;\beta} = g_{\alpha\mu}V^\mu{}_{;\beta}\ . \tag{6.32}$$

But by linearity, we have:

$$V_{\alpha;\beta} = (g_{\alpha\mu}V^\mu)_{;\beta} = g_{\alpha\mu;\beta}V^\mu + g_{\alpha\mu}V^\mu{}_{;\beta}\ . \tag{6.33}$$

So consistency requires $g_{\alpha\mu;\beta}V^\mu = 0$. Since $\mathbf{V}$ is arbitrary this implies that

$$g_{\alpha\mu;\beta} = 0\ . \tag{6.34}$$

Thus the covariant derivative of the metric is zero in every frame.

We next prove that $\Gamma^\alpha{}_{\beta\gamma} = \Gamma^\alpha{}_{\gamma\beta}$ [i.e. symmetric in β and γ]. In a general frame we have for a scalar field ϕ:

$$\phi_{;\beta\alpha} = \phi_{,\beta;\alpha} = \phi_{,\beta\alpha} - \Gamma^\gamma{}_{\beta\alpha}\phi_{,\gamma}\ . \tag{6.35}$$

in a local inertial frame, this is just $\phi_{,\beta\alpha} \equiv \frac{\partial^2\phi}{\partial x^\beta \partial x^\alpha}$, which is symmetric in β and α. Thus it must also be symmetric in a general frame. Hence $\Gamma^\gamma{}_{\alpha\beta}$ is symmetric in β and α:

$$\Gamma^\gamma{}_{\alpha\beta} = \Gamma^\gamma{}_{\beta\alpha} \ . \tag{6.36}$$

We now use this to express $\Gamma^\gamma{}_{\alpha\beta}$ in terms of the metric. Since $g_{\alpha\beta;\mu} = 0$, we have:

$$g_{\alpha\beta,\mu} = \Gamma^\nu{}_{\alpha\mu} g_{\nu\beta} + \Gamma^\nu{}_{\beta\mu} g_{\alpha\nu} \ . \tag{6.37}$$

By writing different permutations of the indices and using the symmetry of $\Gamma^\gamma{}_{\alpha\beta}$, we get

$$g_{\alpha\beta,\mu} + g_{\alpha\mu,\beta} - g_{\beta\mu,\alpha} = 2 g_{\alpha\nu} \Gamma^\nu{}_{\beta\mu} \ . \tag{6.38}$$

Multiplying by $\frac{1}{2} g^{\alpha\gamma}$ and using $g^{\alpha\gamma} g_{\alpha\nu} = \delta^\gamma{}_\nu$ gives

$$\Gamma^\gamma{}_{\beta\mu} = \tfrac{1}{2} g^{\alpha\gamma} \left(g_{\alpha\beta,\mu} + g_{\alpha\mu,\beta} - g_{\beta\mu,\alpha} \right) \ . \tag{6.39}$$

Note that $\Gamma^\gamma{}_{\beta\mu}$ is not a tensor since it is defined in terms of partial derivatives.

In a local inertial frame $\Gamma^\gamma{}_{\beta\mu} = 0$ since $g_{\alpha\beta,\mu} = 0$. We will see later the significance of this result.

6.4 Tensors in polar coordinates

The covariant derivative differs from partial derivatives even in flat spacetime if one uses non - Cartesian coordinates. This corresponds to going to a non - inertial frame. To illustrate this we will focus on two dimensional Euclidian space with Cartesian coordinates ($x^1 = x$, $x^2 = y$) and polar coordinates ($x^{\bar{1}} = r$, $x^{\bar{2}} = \theta$). The coordinates are related by

$$x = r\cos\theta \ , \quad y = r\sin\theta \ , \quad \theta = \tan^{-1}\left(\tfrac{y}{x}\right) \ , \quad r = \left(x^2 + y^2\right)^{1/2} \ . \tag{6.40}$$

For neighboring points we have

$$dr = \frac{\partial r}{\partial x} dx + \frac{\partial r}{\partial y} dy = \cos\theta dx + \sin\theta dy \ , \tag{6.41}$$

and

$$d\theta = \frac{\partial \theta}{\partial x} dx + \frac{\partial \theta}{\partial y} dy = -\frac{1}{r}\sin\theta dx + \frac{1}{r}\cos\theta dy \ . \tag{6.42}$$

We can represent this by a transformation matrix $\Lambda^{\bar{\alpha}}{}_{\beta}$:

$$dx^{\bar{\alpha}} = \Lambda^{\bar{\alpha}}{}_{\beta} dx^{\beta} \ , \tag{6.43}$$

where

$$\Lambda^{\bar{\alpha}}{}_{\beta} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\frac{1}{r}\sin\theta & \frac{1}{r}\cos\theta \end{pmatrix} \ . \tag{6.44}$$

Any vector components must transform in the same way.

For any scalar field ϕ, we can define a one-form:

$$\tilde{\mathbf{d}}\phi \rightarrow \left(\frac{\partial\phi}{\partial r}, \frac{\partial\phi}{\partial\theta} \right) \ . \tag{6.45}$$

We have

$$\begin{aligned} \frac{\partial\phi}{\partial\theta} &= \frac{\partial\phi}{\partial x}\frac{\partial x}{\partial\theta} + \frac{\partial\phi}{\partial y}\frac{\partial y}{\partial\theta} \\ &= -r\sin\theta\frac{\partial\phi}{\partial x} + r\cos\theta\frac{\partial\phi}{\partial y} \ , \end{aligned} \tag{6.46}$$

and

$$\begin{aligned} \frac{\partial\phi}{\partial r} &= \frac{\partial\phi}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial\phi}{\partial y}\frac{\partial y}{\partial r} \\ &= \cos\theta\frac{\partial\phi}{\partial x} + \sin\theta\frac{\partial\phi}{\partial y} \ , \end{aligned} \tag{6.47}$$

This transformation can be represented by another matrix $\Lambda^{\alpha}{}_{\bar{\beta}}$:

$$\frac{\partial\phi}{\partial x^{\bar{\beta}}} = \Lambda^{\alpha}{}_{\bar{\beta}} \frac{\partial\phi}{\partial x^{\alpha}} \ , \tag{6.48}$$

where

$$\Lambda^{\alpha}{}_{\bar{\beta}} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \ . \tag{6.49}$$

Any one-form components must transform in the same way.

The matrices $\Lambda^{\bar{\alpha}}{}_{\beta}$ and $\Lambda^{\alpha}{}_{\bar{\beta}}$ are different but related:

$$\Lambda^{\bar{\alpha}}{}_{\gamma}\Lambda^{\gamma}{}_{\bar{\beta}} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\frac{1}{r}\sin\theta & \frac{1}{r}\cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \ . \tag{6.50}$$

This is just what you would expect since in general $\Lambda^{\bar{\alpha}}{}_{\gamma}\Lambda^{\gamma}{}_{\bar{\beta}} = \delta^{\bar{\alpha}}{}_{\bar{\beta}}$.

The basis vectors and basis one-forms are

$$\begin{aligned} \mathbf{e}_r &= \Lambda^{\alpha}{}_{r}\mathbf{e}_{\alpha} = \Lambda^{x}{}_{r}\mathbf{e}_x + \Lambda^{y}{}_{r}\mathbf{e}_y = \cos\theta\mathbf{e}_x + \sin\theta\mathbf{e}_y \ , \\ \mathbf{e}_\theta &= \Lambda^{\alpha}{}_{\theta}\mathbf{e}_{\alpha} = \Lambda^{x}{}_{\theta}\mathbf{e}_x + \Lambda^{y}{}_{\theta}\mathbf{e}_y = -r\sin\theta\mathbf{e}_x + r\cos\theta\mathbf{e}_y \ , \end{aligned} \tag{6.51}$$

and

$$\begin{aligned} \tilde{\mathbf{w}}^r &= \tilde{d}r = \Lambda^{r}{}_{\alpha}\tilde{\mathbf{w}}^{\alpha} = \Lambda^{r}{}_{x}\tilde{\mathbf{w}}^{x} + \Lambda^{r}{}_{y}\tilde{\mathbf{w}}^{y} = \cos\theta\tilde{\mathbf{w}}^x + \sin\theta\tilde{\mathbf{w}}^y \ , \\ \tilde{\mathbf{w}}^\theta &= \tilde{d}\theta = \Lambda^{\theta}{}_{\alpha}\tilde{\mathbf{w}}^{\alpha} = \Lambda^{\theta}{}_{x}\tilde{\mathbf{w}}^{x} + \Lambda^{\theta}{}_{y}\tilde{\mathbf{w}}^{y} = -\tfrac{1}{r}\sin\theta\tilde{\mathbf{w}}^x + \tfrac{1}{r}\cos\theta\tilde{\mathbf{w}}^y \ . \end{aligned} \tag{6.52}$$

Note that the basis vectors change from point to point in polar coordinates and need not have unit length so they do not form an orthonormal basis:

$$|\mathbf{e}_r| = 1 \ , \quad |\mathbf{e}_\theta| = r \ , \quad |\tilde{\mathbf{w}}^r| = 1 \ , \quad |\tilde{\mathbf{w}}^\theta| = \tfrac{1}{r} \ . \tag{6.53}$$

The inverse metric tensor is:

$$g^{\bar{\alpha}\bar{\beta}} = \begin{pmatrix} 1 & 0 \\ 0 & r^{-2} \end{pmatrix} \tag{6.54}$$

so the components of the vector gradient $\mathbf{d}\phi$ of a scalar field ϕ are:

$$\begin{aligned} (d\phi)^r &= g^{\alpha r}\phi_{,\alpha} = g^{rr}\phi_{,r} + g^{r\theta}\phi_{,\theta} = \frac{\partial\phi}{\partial r} \ , \\ (d\phi)^\theta &= g^{\alpha\theta}\phi_{,\alpha} = g^{r\theta}\phi_{,r} + g^{\theta\theta}\phi_{,\theta} = \frac{1}{r^2}\frac{\partial\phi}{\partial\theta} \ . \end{aligned} \tag{6.55}$$

This is exactly what we would expect from our understanding of normal vector calculus.

We also have:

$$\frac{\partial\mathbf{e}_r}{\partial r} = \frac{\partial}{\partial r}\left(\cos\theta\mathbf{e}_x + \sin\theta\mathbf{e}_y\right) = 0 \ , \tag{6.56}$$

and

$$\frac{\partial\mathbf{e}_r}{\partial\theta} = \frac{1}{r}\mathbf{e}_\theta \ , \quad \frac{\partial\mathbf{e}_\theta}{\partial r} = \frac{1}{r}\mathbf{e}_\theta \ , \quad \frac{\partial\mathbf{e}_\theta}{\partial\theta} = -r\mathbf{e}_r \ . \tag{6.57}$$

Since

$$\frac{\partial\mathbf{e}_{\bar{\alpha}}}{\partial x^{\bar{\beta}}} = \Gamma^{\bar{\mu}}{}_{\bar{\alpha}\bar{\beta}}\mathbf{e}_{\bar{\mu}} \ , \tag{6.58}$$

we can work out all the components of the Christoffel symbols:

$$\Gamma^{\theta}{}_{r\theta} = \frac{1}{r} , \quad \Gamma^{\theta}{}_{\theta r} = \frac{1}{r} , \quad \Gamma^{r}{}_{\theta\theta} = -r , \tag{6.59}$$

and all other components are zero.

Alternatively, we can work out these components from the metric [EXERCISE]:

$$\Gamma^{\bar{\alpha}}{}_{\bar{\beta}\bar{\gamma}} = \tfrac{1}{2} g^{\bar{\alpha}\bar{\delta}} \left[g_{\bar{\beta}\bar{\delta},\bar{\gamma}} + g_{\bar{\gamma}\bar{\delta},\bar{\beta}} - g_{\bar{\beta}\bar{\gamma},\bar{\delta}} \right] . \tag{6.60}$$

In fact this is the best way of working out the components of $\Gamma^{\bar{\alpha}}{}_{\bar{\beta}\bar{\gamma}}$, and it is the way we will adopt in General Relativity.

Finally we can check that all the components of $g_{\bar{\alpha}\bar{\beta};\bar{\gamma}} = 0$ as requited. For example

$$\begin{aligned} g_{\theta\theta;r} &= g_{\theta\theta,r} - \Gamma^{\bar{\mu}}{}_{\theta r} g_{\theta\bar{\mu}} - \Gamma^{\bar{\mu}}{}_{\theta r} g_{\bar{\mu}\theta} \\ &= (r^2)_{,r} - \frac{2}{r}(r^2) = 0 . \end{aligned} \tag{6.61}$$

6.5 Parallel transport and geodesics

A vector field $\mathbf{V}$ is parallel transported along a curve with tangent

$$\mathbf{U} = \frac{d\mathbf{x}}{d\lambda} , \tag{6.62}$$

where λ is the parameter along the curve [usually taken to be the proper time τ if the curve is timelike] if and only if

$$\frac{dV^{\alpha}}{d\lambda} = 0 \tag{6.63}$$

in an inertial frame. Since

$$\frac{dV^{\alpha}}{d\lambda} = \frac{\partial V^{\alpha}}{\partial x^{\beta}} \frac{dx^{\beta}}{d\lambda} = U^{\beta} V^{\alpha}{}_{,\beta} , \tag{6.64}$$

in a general frame the condition becomes:

$$U^{\beta} V^{\alpha}{}_{;\beta} = 0 , \tag{6.65}$$

i.e. we just replace the partial derivatives (,) with a covariant derivative (;). This is called "the comma goes to semicolon" rule, i.e. work things out in a local inertial

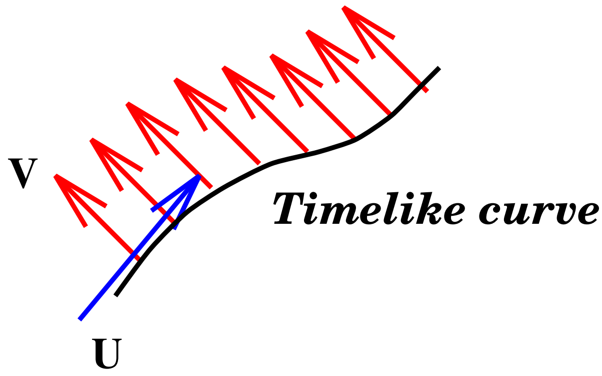

Figure 6.1: Parallel transport of a vector **V** along a timelike curve with tangent **U**.

frame and if it is a tensor equation it will be valid in all frames. The curve is a geodesic if it parallel transports its own tangent vector:

$$U^\beta U^\alpha{}_{;\beta} = 0 \; . \tag{6.66}$$

This is the closest we can get to defining a straight line in a curved space. In flat space a tangent vector is everywhere tangent only for a straight line. Now

$$U^\beta U^\alpha{}_{;\beta} = 0 \Rightarrow U^\beta U^\alpha{}_{,\beta} + \Gamma^\alpha{}_{\mu\beta} U^\mu U^\beta = 0 \; . \tag{6.67}$$

Since $U^\alpha = \frac{dx^\alpha}{d\lambda}$ and $\frac{d}{d\lambda} = U^\beta \frac{\partial}{\partial x^\beta}$ we can write this as

$$\frac{d^2 x^\alpha}{d\lambda^2} + \Gamma^\alpha{}_{\mu\beta} \frac{dx^\mu}{d\lambda}\frac{dx^\beta}{d\lambda} = 0 \; . \tag{6.68}$$

This is the geodesic equation. It is a second order differential equation for $x^\alpha(\lambda)$, so one gets a unique solution by specifying an initial position $\mathbf{x}_0$ and velocity $\mathbf{U}_0$.

6.5.1 The variational method for geodesics

We now apply the variational techniques to compute the geodesics for a given metric.

For a curved spacetime, the proper time $d\tau$ is defined to be

$$d\tau^2 = -\frac{1}{c^2} ds^2 = -\frac{1}{c^2} g_{\alpha\beta} dx^\alpha dx^\beta \; . \tag{6.69}$$

Remember in flat spacetime it was just

$$d\tau^2 = -\frac{1}{c^2}\eta_{\alpha\beta}dx^\alpha dx^\beta \ . \tag{6.70}$$

Therefore the proper time between two points A and B along an arbitrary timelike curve is

$$\begin{aligned} \tau_{AB} &= \int_A^B d\tau = \int_A^B \frac{d\tau}{d\lambda}d\lambda \\ &= \int_A^B \frac{1}{c}\left[-g_{\alpha\beta}(\mathbf{x})\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}\right]^{1/2} d\lambda \ , \end{aligned} \tag{6.71}$$

so we can write the Lagrangian as

$$\mathcal{L}(x^\alpha, \dot{x}^\alpha, \lambda) = \frac{1}{c}\left[-g_{\alpha\beta}(\mathbf{x})\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}\right]^{1/2} \ , \tag{6.72}$$

and the action becomes

$$A = \tau_{AB} = \int_A^B \mathcal{L}(x^\alpha, \dot{x}^\alpha, \lambda)d\lambda \ . \tag{6.73}$$

Extremizing the action we get the Euler - Lagrange equations:

$$\frac{\partial\mathcal{L}}{\partial x^\alpha} = \frac{d}{d\lambda}\left(\frac{\partial\mathcal{L}}{\partial\dot{x}^\alpha}\right) \ . \tag{6.74}$$

Now

$$\frac{\partial\mathcal{L}}{\partial x^\alpha} = -\frac{1}{2c}\left(-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right)^{-1/2}\frac{\partial g_{\beta\gamma}}{\partial x^\alpha}\frac{dx^\beta}{d\lambda}\frac{dx^\gamma}{d\lambda} \tag{6.75}$$

and

$$\frac{\partial\mathcal{L}}{\partial\dot{x}^\alpha} = -\frac{1}{2c}\left(-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right)^{-1/2}(2)g_{\alpha\beta}\frac{dx^\beta}{d\lambda} \ . \tag{6.76}$$

Since

$$\frac{1}{c}\left(-g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}\right)^{1/2} = \frac{d\tau}{d\lambda} \tag{6.77}$$

we get

$$\frac{\partial\mathcal{L}}{\partial x^\alpha} = -\frac{1}{2}\frac{d\lambda}{d\tau}\frac{\partial g_{\beta\gamma}}{\partial x^\alpha}\frac{dx^\beta}{d\lambda}\frac{dx^\gamma}{d\lambda} \tag{6.78}$$

and

$$\frac{\partial\mathcal{L}}{\partial\dot{x}^\alpha} = -\frac{d\lambda}{d\tau}g_{\alpha\beta}\frac{dx^\beta}{d\lambda} = -g_{\alpha\beta}\frac{dx^\beta}{d\tau} \ , \tag{6.79}$$

so the Euler - Lagrange equations become:

$$\frac{1}{2}\frac{d\lambda}{d\tau}\frac{\partial g_{\beta\gamma}}{\partial x^\alpha}\frac{dx^\beta}{d\lambda}\frac{dx^\gamma}{d\lambda} = \frac{d}{d\lambda}\left[g_{\alpha\beta}\frac{dx^\beta}{d\tau}\right] . \tag{6.80}$$

Multiplying by $\frac{d\lambda}{d\tau}$ we obtain

$$\frac{1}{2}\frac{\partial g_{\beta\gamma}}{\partial x^\alpha}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} = \frac{d}{d\tau}\left[g_{\alpha\beta}\frac{dx^\beta}{d\tau}\right] . \tag{6.81}$$

Using

$$\frac{dg_{\alpha\beta}}{d\tau} = \frac{\partial g_{\alpha\beta}}{\partial x^\gamma}\frac{dx^\gamma}{d\tau} , \tag{6.82}$$

we get

$$\frac{1}{2}\frac{\partial g_{\beta\gamma}}{\partial x^\alpha}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} = g_{\alpha\beta}\frac{d^2x^\beta}{d\tau^2} + \frac{\partial g_{\alpha\beta}}{\partial x^\gamma}\frac{dx^\gamma}{d\tau}\frac{dx^\beta}{d\tau} . \tag{6.83}$$

Multiplying by $g^{\delta\alpha}$ gives

$$\frac{d^2x^\delta}{d\tau^2} = -g^{\delta\alpha}\left[\frac{\partial g_{\alpha\beta}}{\partial x^\gamma} - \frac{1}{2}\frac{\partial g_{\beta\gamma}}{\partial x^\alpha}\right]\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} . \tag{6.84}$$

Now

$$\begin{aligned}\frac{\partial g_{\alpha\beta}}{\partial x^\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} &= \frac{1}{2}\left[\frac{\partial g_{\alpha\beta}}{\partial x^\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} + \frac{\partial g_{\alpha\beta}}{\partial x^\gamma}\frac{dx^\gamma}{d\tau}\frac{dx^\beta}{d\tau}\right] \\ &= \frac{1}{2}\left[\frac{\partial g_{\alpha\beta}}{\partial x^\gamma} + \frac{\partial g_{\alpha\gamma}}{\partial x^\beta}\right]\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} .\end{aligned} \tag{6.85}$$

Using the above result gives us

$$\begin{aligned}\frac{d^2x^\delta}{d\tau^2} &= -\frac{1}{2}g^{\delta\alpha}\left[g_{\alpha\beta,\gamma} + g_{\alpha\gamma,\beta} - g_{\beta\gamma,\alpha}\right]\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} \\ &= -\Gamma^\delta{}_{\beta\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} ,\end{aligned} \tag{6.86}$$

so we get the geodesic equation again

$$\frac{d^2x^\delta}{d\tau^2} + \Gamma^\delta{}_{\beta\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} = 0 . \tag{6.87}$$

This is the equation of motion for a particle moving on a timelike geodesic in curved spacetime. Note that in a local inertial frame i.e. where $\Gamma^\delta{}_{\beta\gamma} = 0$, the equation reduces to

$$\frac{d^2x^\delta}{d\tau^2} = 0 , \tag{6.88}$$

which is the equation of motion for a free particle.

The geodesic equation preserves its form if we parameterize the curve by any other parameter λ such that

$$\frac{d^2\lambda}{d\tau^2} = 0 \quad \Rightarrow \quad \lambda = a\tau + b \tag{6.89}$$

for constants a and b. A parameter which satisfies this condition is said to be affine.

6.5.2 The principle of equivalence again

In Special Relativity, in a coordinate system adapted for an inertial frame, namely Minkowski coordinates, the equation for a test particle is:

$$\frac{d^2x^\alpha}{d\tau^2} = 0\ . \tag{6.90}$$

If we use a non-inertial frame of reference, then this is equivalent to using a more general coordinate system $[x^{\bar{\alpha}} = x^{\bar{\alpha}}(x^\beta)]$. In this case, the equation becomes

$$\frac{d^2x^\alpha}{d\tau^2} + \Gamma^\alpha{}_{\beta\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} = 0\ , \tag{6.91}$$

where $\Gamma^\alpha{}_{\beta\gamma}$ is the metric connection of $g_{\alpha\beta}$, which is still a flat metric but not the Minkowski metric $\eta_{\alpha\beta}$. The additional terms involving $\Gamma^\alpha{}_{\beta\gamma}$ which appear, are inertial forces.

The principle of equivalence requires that gravitational forces, a well as inertial forces, should be given by an appropriate $\Gamma^\alpha{}_{\beta\gamma}$. In this case we can no longer take the spacetime to be flat. The simplest generalization is to keep $\Gamma^\alpha{}_{\beta\gamma}$ as the metric connection, but now take it to be the metric connection of a non-flat metric. If we are to interpret the $\Gamma^\alpha{}_{\beta\gamma}$ as force terms, then it follows that we should regard the $g_{\alpha\beta}$ as potentials. The field equations of Newtonian gravitation consist of second-order partial differential equations in the potential Φ. In an analogous manner, we would expect that General Relativity also to involve second order partial differential equations in the potentials $g_{\alpha\beta}$. The remaining task which will allow us to build a relativistic theory of gravitation is to construct this set of partial differential equations. We will do this shortly but first we must define a quantity that quantifies spacetime curvature.

6.6 The curvature tensor and geodesic deviation

So far we have used the local - flatness theorem to develop as much mathematics on curved manifolds as possible without having to consider curvature explicitly. In this section we will make a precise mathematical definition of curvature and discuss the remaining tools needed to derive the Einstein field equations.

It is important to distinguish two different kinds of curvature: intrinsic and extrinsic. Consider for example a cylinder. Since a cylinder is round in one direction, one thinks of it as curved. It is its extrinsic curvature i.e. the curvature it has in relation to the flat three - dimensional space it is part of. On the other hand, a cylinder can be made by rolling a flat piece of paper without tearing or crumpling it, so the intrinsic geometry is that of the original paper i.e. it is flat. This means that the distance between any two points is the same as it was in the original paper. Also parallel lines remain parallel when continued; in fact all of Euclid's axioms hold for the surface of a cylinder. A 2D ant confined to that surface would decide that it was flat; only that the global topology is funny!

It is clear therefore that when we talk of the curvature of spacetime, we talk of its intrinsic curvature since the worldlines [geodesics] of particles are confined to remain in spacetime.

The cylinder, as we have just seen is intrinsically flat; a sphere, on the other hand, has an intrinsically curved surface. To see this, consider the surface of a sphere [or balloon] in which two neighboring lines begin at A and B perpendicular to the equator, and hence are parallel. When continued as locally straight lines they follow the arc of great circles, and the two lines meet at the pole P. So parallel lines, when continued, do not remain parallel, so the space is not flat.

There is an even more striking illustration of the curvature of the surface of a sphere. Consider, first, a flat space. Let us take a closed path starting at A then going to B and C and then back to A. Let's parallel transport a vector around this loop. The vector finally drawn at A is, of course, parallel to the original one [see Figure 6.2]. A completely different thing happens on the surface of a sphere! This time the vector is rotated through 90 degrees! [see the balloon example]. We will

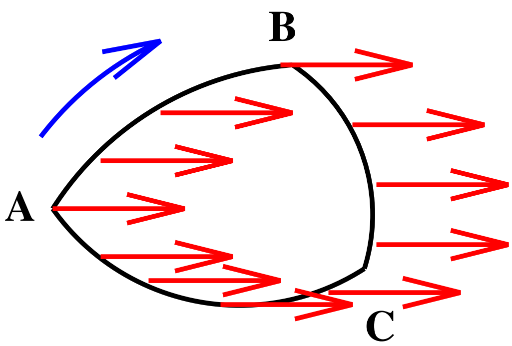

Figure 6.2: Parallel transport around a closed loop in flat space.

now use parallel transport around a closed loop in curved space to define curvature in that space.

6.6.1 The curvature tensor

Imagine in our manifold a very small closed loop whose four sides are the coordinate lines $x^1 = a$, $x^1 = a + \delta a$, $x^2 = b$, $x^2 = b + \delta b$.

A vector $\mathbf{V}$ defined at A is parallel transported to B. From the parallel transport law

$$\nabla_{\mathbf{e}} \mathbf{V} = 0 \quad \Rightarrow \quad \frac{\partial V^\alpha}{\partial x^1} = -\Gamma^\alpha{}_{\mu 1} V^\mu \ , \tag{6.92}$$

it follows that at B the vector has components

$$\begin{aligned} V^\alpha(B) &= V^\alpha(A) + \int_A^B \frac{\partial V^\alpha}{\partial x^1} dx^1 \\ \Rightarrow \quad V^\alpha(B) &= V^\alpha(A) - \int_{x^2=b} \Gamma^\alpha{}_{\mu 1} V^\mu dx^1 \ , \end{aligned} \tag{6.93}$$

where the notation "$x^2 = b$" under the integral denotes the path AB. Similar transport from B to C to D gives

$$V^\alpha(C) = V^\alpha(B) - \int_{x^1=a+\delta a} \Gamma^\alpha{}_{\mu 2} V^\mu dx^2 \ , \tag{6.94}$$

and

$$V^\alpha(D) = V^\alpha(C) + \int_{x^1=b+\delta b} \Gamma^\alpha{}_{\mu 1} V^\mu dx^1 \ . \tag{6.95}$$

The integral in the last equation has a different sign because of the direction of transport from C to D is in the negative x^1 direction.

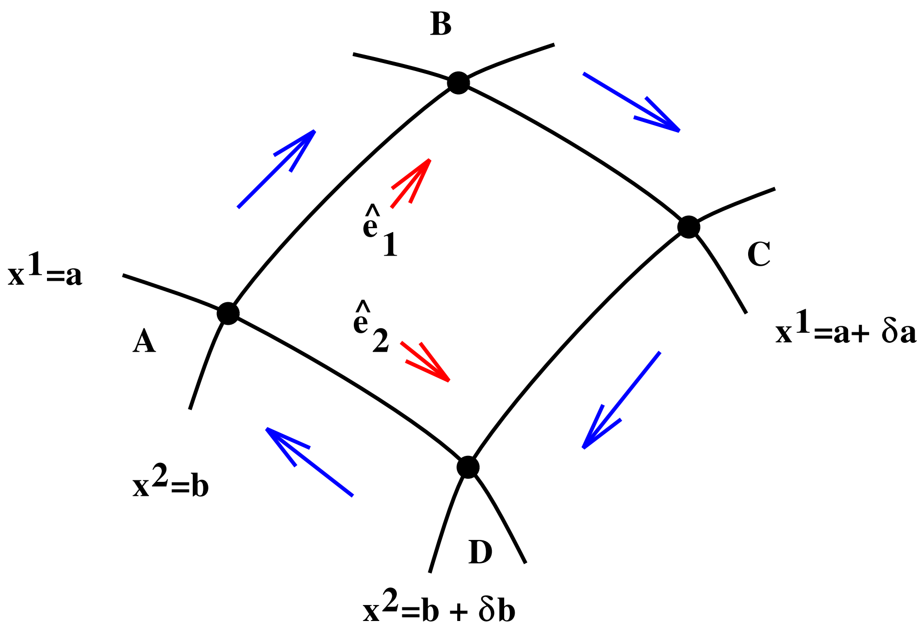

Figure 6.3: Parallel transport around a closed loop $ABCD$.

Similarly, the completion of the loop gives

$$V^\alpha(A_{final}) = V^\alpha(D) + \int_{x^1=a} \Gamma^\alpha{}_{\mu 2} V^\mu dx^2 \ . \tag{6.96}$$

The net change in $V^\alpha(A)$ is a vector δV^α, found by adding (6.93) - (6.96).

$$\begin{aligned} \delta V^\alpha &= V^\alpha(A_{final}) - V^\alpha(A_{initial}) \\ &= \int_{x^1=a} \Gamma^\alpha{}_{\mu 2} V^\mu dx^2 - \int_{x^1=a+\delta a} \Gamma^\alpha{}_{\mu 2} V^\mu dx^2 \\ &+ \int_{x^2=b+\delta b} \Gamma^\alpha{}_{\mu 1} V^\mu dx^1 - \int_{x^2=b} \Gamma^\alpha{}_{\mu 1} V^\mu dx^1 \ . \end{aligned} \tag{6.97}$$

To lowest order we get

$$\begin{aligned} \delta V^\alpha &= -\int_b^{b+\delta b} \delta a \frac{\partial}{\partial x^1} \left(\Gamma^\alpha{}_{\mu 2} V^{mu} \right) dx^2 \\ &+ \int_a^{a+\delta a} \delta b \frac{\partial}{\partial x^2} \left(\Gamma^\alpha{}_{\mu 1} V^\mu \right) dx^1 \\ &\approx \delta a \delta b \left[-\frac{\partial}{\partial x^1} \left(\Gamma^\alpha{}_{\mu 2} V^\mu \right) + \frac{\partial}{\partial x^2} \left(\Gamma^\alpha{}_{\mu 1} V^\mu \right) \right] \ . \end{aligned} \tag{6.98}$$

This involves derivatives of Γ's and of V^α. The derivatives of V^α can be eliminated using for example

$$\frac{\partial V^\mu}{\partial x^1} = -\Gamma^\mu{}_{\alpha 1} V^\alpha \ . \tag{6.99}$$

This gives

$$\delta V^\alpha = \delta a \delta b \left[\Gamma^\alpha{}_{\mu 1,2} - \Gamma^\alpha{}_{\mu 2,1} + \Gamma^\alpha{}_{\nu 2}\Gamma^\nu{}_{\mu 1} - \Gamma^\alpha{}_{\nu 1}\Gamma^\nu{}_{\mu 2}\right] V^\mu \ . \tag{6.100}$$

To obtain this, one needs to relabel dummy indices in the terms quadratic in Γ.

Notice that this just turns out to be a number times V^μ summed on μ. Now the indices 1 and 2 appear because the path was chosen to go along those coordinates. It is antisymmetric in 1 and 2 because the change δV^α would have the opposite sign if one went around the loop in the opposite direction.

If we use general coordinate lines x^σ and x^λ, we find

$$\delta V^\alpha = \delta a \delta b \left[\Gamma^\alpha{}_{\mu\sigma,\lambda} - \Gamma^\alpha{}_{\mu\lambda,\sigma} + \Gamma^\alpha{}_{\nu\lambda}\Gamma^\nu{}_{\mu\sigma} - \Gamma^\alpha{}_{\nu\sigma}\Gamma^\nu{}_{\mu\lambda}\right] V^\mu \ . \tag{6.101}$$

Defining

$$R^\alpha{}_{\mu\lambda\sigma} \equiv \Gamma^\alpha{}_{\mu\sigma,\lambda} - \Gamma^\alpha{}_{\mu\lambda,\sigma} + \Gamma^\alpha{}_{\nu\lambda}\Gamma^\nu{}_{\mu\sigma} - \Gamma^\alpha{}_{\nu\sigma}\Gamma^\nu{}_{\mu\lambda} \tag{6.102}$$

we can write

$$\delta V^\alpha = \delta a \delta b R^\alpha{}_{\mu\lambda\sigma} V^\mu \tilde{\mathbf{w}}^\lambda \tilde{\mathbf{w}}^\sigma \ . \tag{6.103}$$

$R^\alpha{}_{\beta\mu\nu}$ are the components of a 1/3 tensor. This tensor is called the Riemann curvature tensor.

6.6.2 Properties of the Riemann curvature tensor

Recall that the Riemann tensor is

$$R^\alpha{}_{\beta\mu\nu} \equiv \Gamma^\alpha{}_{\beta\nu,\mu} - \Gamma^\alpha{}_{\beta\mu,\nu} + \Gamma^\alpha{}_{\sigma\mu}\Gamma^\sigma{}_{\beta\nu} - \Gamma^\alpha{}_{\sigma\nu}\Gamma^\sigma{}_{\beta\mu} \ . \tag{6.104}$$

In a local inertial frame we have $\Gamma^\alpha{}_{\mu\nu} = 0$, so in this frame

$$R^\alpha{}_{\beta\mu\nu} = \Gamma^\alpha{}_{\beta\nu,\mu} - \Gamma^\alpha{}_{\beta\mu,\nu} \ . \tag{6.105}$$

Now

$$\Gamma^\alpha{}_{\beta\nu} = \frac{1}{2} g^{\alpha\delta} \left(g_{\delta\beta,\nu} + g_{\delta\nu,\beta} - g_{\beta\nu,\delta}\right) \tag{6.106}$$

so

$$\Gamma^{\alpha}{}_{\beta\nu,\mu} = \frac{1}{2} g^{\alpha\delta} \left(g_{\delta\beta,\nu\mu} + g_{\delta\nu,\beta\mu} - g_{\beta\nu,\delta\mu} \right) \tag{6.107}$$

since $g^{\alpha\delta}{}_{,\mu} = 0$ i.e the first derivative of the metric vanishes in a local inertial frame. Hence

$$R^{\alpha}{}_{\beta\mu\nu} = \frac{1}{2} g^{\alpha\delta} \left(g_{\delta\beta,\nu\mu} + g_{\delta\nu,\beta\mu} - g_{\beta\nu,\delta\mu} - g_{\delta\beta,\mu\nu} - g_{\delta\mu,\beta\nu} + g_{\beta\mu,\delta\nu} \right) \ . \tag{6.108}$$

Using the fact that partial derivatives always commute so that $g_{\delta\beta,\nu\mu} = g_{\delta\beta,\mu\nu}$, we get

$$R^{\alpha}{}_{\beta\mu\nu} = \frac{1}{2} g^{\alpha\delta} \left(g_{\delta\nu,\beta\mu} - g_{\delta\mu,\beta\nu} + g_{\beta\mu,\delta\nu} - g_{\beta\nu,\delta\nu} \right) \tag{6.109}$$

in a local inertial frame. Lowering the index α with the metric we get

$$\begin{aligned} R_{\alpha\beta\mu\nu} &= g_{\alpha\lambda} R^{\lambda}{}_{\beta\nu\nu} \\ &= \frac{1}{2} \delta^{\delta}{}_{\alpha} \left(g_{\delta\nu,\beta\mu} - g_{\delta\mu,\beta\nu} + g_{\beta\mu,\delta\nu} - g_{\beta\nu,\delta\nu} \right) \ . \end{aligned} \tag{6.110}$$

So in a local inertial frame the result is

$$R_{\alpha\beta\mu\nu} = \frac{1}{2} \left(g_{\alpha\nu,\beta\mu} - g_{\alpha\mu,\beta\nu} + g_{\beta\mu,\alpha\nu} - g_{\beta\nu,\alpha\mu} \right) \ . \tag{6.111}$$

We can use this result to discover what the symmetries of $R_{\alpha\beta\mu\nu}$ are. It is easy to show from the above result that

$$R_{\alpha\beta\mu\nu} = -R_{\beta\alpha\mu\nu} = -R_{\alpha\beta\nu\mu} = R_{\mu\nu\alpha\beta} \tag{6.112}$$

and

$$R_{\alpha\beta\mu\nu} + R_{\alpha\nu\beta\mu} + R_{\alpha\mu\nu\beta} = 0 \ . \tag{6.113}$$

Thus $R_{\alpha\beta\mu\nu}$ is antisymmetric on the final pair and second pair of indices, and symmetric on exchange of the two pairs.

Since these last two equations are valid tensor equations, although they were derived in a local inertial frame, they are valid in all coordinate systems.

We can use these two identities to reduce the number of independent components of $R_{\alpha\beta\mu\nu}$ from 256 to just 20.

A flat manifold is one which has a global definition of parallelism: i.e. a vector can be moved around parallel to itself on an arbitrary curve and will return to its starting point unchanged. This clearly means that

$$R^{\alpha}{}_{\beta\mu\nu} = 0 \; , \tag{6.114}$$

i.e. the manifold is flat [EXERCISE 6.5: try a cylinder!].

An important use of the curvature tensor comes when we examine the consequences of taking two covariant derivatives of a vector field **V**:

$$\begin{aligned} \nabla_\alpha \nabla_\beta V^\mu &= \nabla_\alpha \left(V^\mu{}_{;\beta} \right) \\ &= \left(V^\mu{}_{;\beta} \right)_{,\alpha} + \Gamma^\mu{}_{\sigma\alpha} V^\sigma{}_{;\beta} - \Gamma^\mu{}_{\beta\alpha} V^\mu{}_{;\sigma} \; . \end{aligned} \tag{6.115}$$

As usual we can simplify things by working in a local inertial frame. So in this frame we get

$$\begin{aligned} \nabla_\alpha \nabla_\beta V^\mu &= \left(V^\mu{}_{;\beta} \right)_{,\alpha} \\ &= \left(V^\mu{}_{,\beta} + \Gamma^\mu{}_{\nu\beta} V^\nu \right)_{,\alpha} \\ &= V^\mu{}_{,\beta\alpha} + \Gamma^\mu{}_{\nu\beta,\alpha} V^\nu + \Gamma^\mu{}_{\nu\beta} V^\nu{}_{,\alpha} \; . \end{aligned} \tag{6.116}$$

The second term of this is zero in a local inertial frame, so we obtain

$$\nabla_\alpha \nabla_\beta V^\mu = V^\mu{}_{,\alpha\beta} + \Gamma^\mu{}_{\nu\beta,\alpha} V^\nu \; . \tag{6.117}$$

Consider the same formula with the α and β interchanged:

$$\nabla_\beta \nabla_\alpha V^\mu = V^\mu{}_{,\beta\alpha} + \Gamma^\mu{}_{\nu\alpha,\beta} V^\nu \; . \tag{6.118}$$

If we subtract these we get the commutator of the covariant derivative operators ∇_α and ∇_β:

$$\begin{aligned} \left[\nabla_\alpha, \nabla_\beta \right] V^\mu &= \nabla_\alpha \nabla_\beta V^\mu - \nabla_\beta \nabla_\alpha V^\mu \\ &= \left(\Gamma^\mu{}_{\nu\beta,\alpha} - \Gamma^\mu{}_{\nu\alpha,\beta} \right) V^\nu \; . \end{aligned} \tag{6.119}$$

The terms involving the second derivatives of V^μ drop out because $V^\mu{}_{,\alpha\beta} = V^\nu{}_{,\beta\alpha}$ [partial derivatives commute].

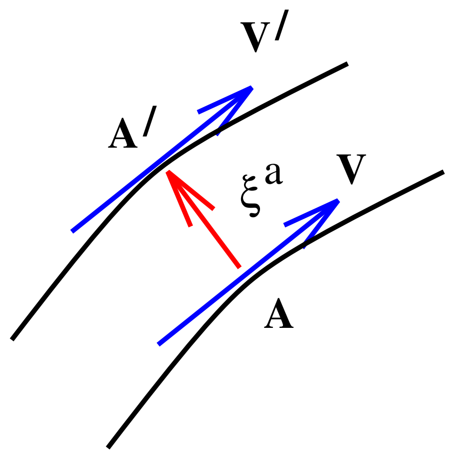

Figure 6.4: Geodesic deviation

Since in a local inertial frame the Riemann tensor takes the form

$$R^{\mu}{}_{\nu\alpha\beta} = \Gamma^{\mu}{}_{\nu\beta,\alpha} - \Gamma^{\mu}{}_{\nu\alpha,\beta} \; , \tag{6.120}$$

we get

$$[\nabla_\alpha, \nabla_\beta]\, V^\mu = R^{\mu}{}_{\nu\alpha\beta} V^\nu \; . \tag{6.121}$$

This is closely related to our original derivation of the Riemann tensor from parallel transport around loops, because the parallel transport problem can be thought of as computing, first the change of **V** in one direction, and then in another, followed by subtracting changes in the reverse order.

6.7 Geodesic deviation

We have shown that in a curved space [for example on the surface of a balloon] parallel lines when extended do not remain parallel. We will now formulate this mathematically in terms of the Riemann tensor.

Consider two geodesics with tangents **V** and **V**′ that begin parallel and near each other at points A and A' [see Figure 6.4]. Let the affine parameter on the geodesics be λ. We define a connecting vector ξ which reaches from one geodesic to another, connecting points at equal intervals in λ.

To simplify things, let's adopt a local inertial frame at A, in which the coordinate x^0 points along the geodesics. Thus at A we have $V^\alpha = \delta^\alpha{}_0$.

The geodesic equation at A is

$$\frac{d^2x^\alpha}{d\lambda^2}|_A = 0 \tag{6.122}$$

since all the connection coefficients vanish at A. The connection does not vanish at A', so the equation of the geodesic at A' is

$$\frac{d^2x^\alpha}{d\lambda^2}|_{A'} + \Gamma^\alpha{}_{00}(A') = 0 \ , \tag{6.123}$$

where again at A' we have arranged the coordinates so that $V^\alpha = \delta^\alpha{}_0$. But, since A and A' are separated by the connecting vector ξ we have

$$\Gamma^\alpha{}_{00}(A') = \Gamma^\alpha{}_{00,\beta}\xi^\beta \ , \tag{6.124}$$

the right hand side being evaluated at A, so

$$\frac{d^2x^\alpha}{d\lambda^2}|_{A'} = -\Gamma^\alpha{}_{00,\beta}\xi^\beta \ . \tag{6.125}$$

Now $x^\alpha(A') - x^\alpha(A) = \xi^\alpha$ so

$$\frac{d^2\xi^\alpha}{d\lambda^2} = \frac{d^2x^\alpha}{d\lambda^2}|_{A'} - \frac{d^2x^\alpha}{d\lambda^2}|_A = -\Gamma^\alpha{}_{00,\beta}\xi^\beta \ . \tag{6.126}$$

This then gives us an expression telling us how the components of ξ change. Consider now the full second covariant derivative $\nabla_{\mathbf{V}}\nabla_{\mathbf{V}}\xi$.

Now

$$\begin{aligned}\nabla_{\mathbf{V}}\nabla_{\mathbf{V}}\xi^\alpha &= \nabla_{\mathbf{V}}\left(\nabla_{\mathbf{V}}\xi^\alpha\right) \\ &= \frac{d}{d\lambda}\left(\nabla_{\mathbf{V}}\xi^\alpha\right) + \Gamma^\alpha{}_{\beta 0}\left(\nabla_{\mathbf{V}}\xi^\beta\right) \ .\end{aligned} \tag{6.127}$$

In a local inertial frame this is

$$\begin{aligned}\nabla_{\mathbf{V}}\nabla_{\mathbf{V}}\xi^\alpha &= \frac{d}{d\lambda}\left(\nabla_{\mathbf{V}}\xi^\alpha\right) \\ &= \frac{d}{d\lambda}\left(\frac{d\xi^\alpha}{d\lambda} + \Gamma^\alpha{}_{\beta 0}\xi^\beta\right) \\ &= \frac{d^2\xi^\alpha}{d\lambda} + \Gamma^\alpha{}_{\beta 0,0}\xi^\beta \ ,\end{aligned} \tag{6.128}$$

where everything is again evaluated at A. Using the result for $\frac{d^2\xi^\alpha}{d\lambda^2}$ we get

$$\begin{aligned} \nabla_{\mathbf{V}}\nabla_{\mathbf{V}}\xi^\alpha &= \left({\Gamma^\alpha}_{\beta 0,0} - {\Gamma^\alpha}_{00,\beta}\right)\xi^\beta \\ &= {R^\alpha}_{00\beta}\xi^\beta \\ &= {R^\alpha}_{\mu\nu\beta}V^\mu V^\nu \xi^\beta \ , \end{aligned} \tag{6.129}$$

since we have chosen $V^\alpha = {\delta^\alpha}_0$.

The final expression is frame invariant, so we have in any basis

$$\nabla_{\mathbf{V}}\nabla_{\mathbf{V}}\xi^\alpha = {R^\alpha}_{\mu\nu\beta}V^\mu V^\nu \xi^\beta \ . \tag{6.130}$$

So geodesics in flat space maintain their separation; those in curved space don't. This is called the equation of geodesic deviation and it shows mathematically that the tidal forces of a gravitational field can be represented by the curvature of spacetime in which particles follow geodesics.

6.8 The Bianchi identities; Ricci and Einstein tensors

In the last section we found that in a local inertial frame the Riemann tensor could be written as

$$R_{\alpha\beta\mu\nu} = \tfrac{1}{2}\left(g_{\alpha\nu,\beta\mu} - g_{\alpha\mu,\beta\nu} + g_{\beta\mu,\alpha\nu} - g_{\beta\nu,\alpha\mu}\right) \ . \tag{6.131}$$

Differentiating with respect to x^λ we get

$$R_{\alpha\beta\mu\nu,\lambda} = \tfrac{1}{2}\left(g_{\alpha\nu,\beta\mu\lambda} - g_{\alpha\mu,\beta\nu\lambda} + g_{\beta\mu,\alpha\nu\lambda} - g_{\beta\nu,\alpha\mu\lambda}\right) \ . \tag{6.132}$$

From this equation, the symmetry $g_{\alpha\beta} = g_{\beta\alpha}$ and the fact that partial derivatives commute, it is easy to show that

$$R_{\alpha\beta\mu\nu,\lambda} + R_{\alpha\beta\lambda\mu,\nu} + R_{\alpha\beta\nu\lambda,\mu} = 0 \ . \tag{6.133}$$

This equation is valid in a local inertial frame, therefore in a general frame we get

$$R_{\alpha\beta\mu\nu;\lambda} + R_{\alpha\beta\lambda\mu;\nu} + R_{\alpha\beta\nu\lambda;\mu} = 0 \ . \tag{6.134}$$

This is a tensor equation, therefore valid in any coordinate system. It is called the Bianchi identities, and will be very important for our work.

6.8.1 The Ricci tensor

Before looking at the consequences of the Bianchi identities, we need to define the Ricci tensor $R_{\alpha\beta}$:

$$R_{\alpha\beta} = R^{\mu}{}_{\alpha\mu\beta} = R_{\beta\alpha} \ . \tag{6.135}$$

It is the contraction of $R^{\mu}{}_{\alpha\nu\beta}$ on the first and third indices. Other contractions would in principle also be possible: on the first and second, the first and fourth, etc. But because $R_{\alpha\beta\mu\nu}$ is antisymmetric on α and β and on μ and ν, all these contractions either vanish or reduce to $\pm R_{\alpha\beta}$. Therefore the Ricci tensor is essentially the only contraction of the Riemann tensor.

Similarly, the Ricci scalar is defined as

$$R = g^{\mu\nu} R_{\mu\nu} = g^{\mu\nu} g^{\alpha\beta} R_{\alpha\beta\mu\nu} \ . \tag{6.136}$$

6.8.2 The Einstein Tensor

Let us apply the Ricci contraction to the Bianchi identities

$$g^{\alpha\mu} \left[R_{\alpha\beta\mu\nu;\lambda} + R_{\alpha\beta\lambda\mu;\nu} + R_{\alpha\beta\nu\lambda;\mu} \right] = 0 \ . \tag{6.137}$$

Since $g_{\alpha\beta;\mu} = 0$ and $g^{\alpha\beta}{}_{;\mu} = 0$, we can take $g^{\alpha\mu}$ in and out of covariant derivatives at will: We get:

$$R^{\mu}{}_{\beta\mu\nu;\lambda} + R^{\mu}{}_{\beta\lambda\mu;\nu} + R^{\mu}{}_{\beta\nu\lambda;\mu} = 0 \ . \tag{6.138}$$

Using the antisymmetry on the indices μ and λ we get

$$R^{\mu}{}_{\beta\mu\nu;\lambda} - R^{\mu}{}_{\beta\mu\lambda;\nu} + R^{\mu}{}_{\beta\nu\lambda;\mu} \ , \tag{6.139}$$

so

$$R_{\beta\nu;\lambda} - R_{\beta\lambda;\nu} + R^{\mu}{}_{\beta\nu\lambda;\mu} = 0 \ . \tag{6.140}$$

These equations are called the contracted Bianchi identities.

Let us now contract a second time on the indices β and ν:

$$g^{\beta\nu} \left[R_{\beta\nu;\lambda} - R_{\beta\lambda;\nu} + R^{\mu}{}_{\beta\nu\lambda;\mu} \right] = 0 \ . \tag{6.141}$$

This gives

$$R^{\nu}{}_{\nu;\lambda} - R^{\nu}{}_{\lambda;\nu} + R^{\mu\nu}{}_{\nu\lambda;\mu} = 0 \ . \tag{6.142}$$

so

$$R_{;\lambda} - 2R^{\mu}{}_{\lambda;\mu} = 0 \ , \tag{6.143}$$

or

$$2R^{\mu}{}_{\lambda;\mu} - R_{;\lambda} = 0 \ . \tag{6.144}$$

Since $R_{;\lambda} = g^{\mu}{}_{\lambda} R_{;\mu}$, we get

$$\left[2R^{\mu}{}_{\lambda} - \tfrac{1}{2} g^{\mu}{}_{\lambda} R\right]_{;\mu} = 0 \ . \tag{6.145}$$

Raising the index λ with $g^{\lambda\nu}$ we get

$$\left[R^{\mu\nu} - \tfrac{1}{2} g^{\mu\nu} R\right]_{;\mu} = 0 \ . \tag{6.146}$$

Defining

$$G^{\mu\nu} = R^{\mu\nu} - \tfrac{1}{2} g^{\mu\nu} R \tag{6.147}$$

we get

$$G^{\mu\nu}{}_{;\mu} = 0 \ . \tag{6.148}$$

The tensor $G^{\mu\nu}$ is constructed only from the Riemann tensor and the metric, and it is automatically divergence free as an identity. It is called the Einstein tensor, since its importance for gravity was first understood by Einstein. We will see in the next chapter that Einstein's field equations for General Relativity are

$$G^{\mu\nu} = \frac{8\pi G}{c^4} T^{\mu\nu} \ , \tag{6.149}$$

where $T^{\mu\nu}$ is the stress-energy tensor.

The Bianchi Identities then imply

$$T^{\alpha\beta}{}_{;\beta} = 0 \ , \tag{6.150}$$

which is the conservation of energy and momentum.