2 Manifolds

2.1 Definition of a manifold

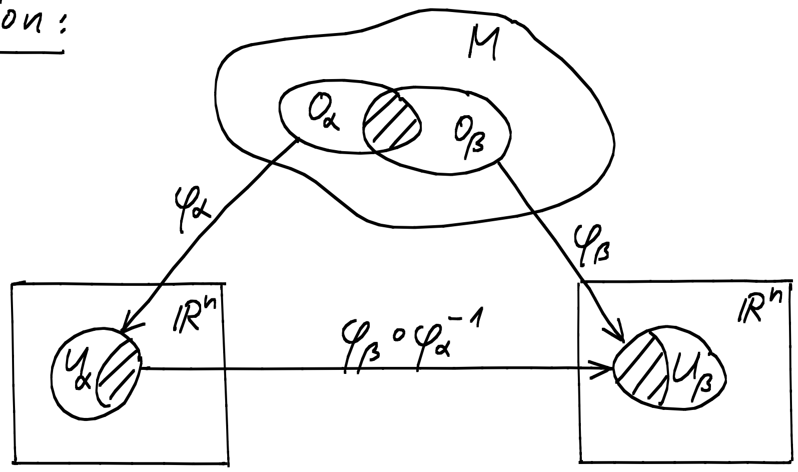
- Roughly speaking, a manifold is a space that locally "looks like" $\mathbb{R}^n$

- <u>Definition:</u> An n-dim., real, C^∞- manifold is a set M together with a collection subsets O_α satisfying

  1) $\bigcup O_\alpha = M$
  2) For each α $\exists$ a one-to-one map $\varphi_\alpha : O_\alpha \to U_\alpha$, where O_α is an open subset of $\mathbb{R}^n$
  3) For O_α, O_β such that $O_\alpha \cap O_\beta \neq \emptyset$, the map $\varphi_\beta \circ \varphi_\alpha^{-1}$ is C^∞. (We also require $\varphi_\alpha(O_\alpha \cap O_\beta)$ and $\varphi_\beta(O_\alpha \cap O_\beta)$ to be open.)

- φ_α is a "chart" or "coordinate system". The collection of all φ_α is an "atlas".

<u>Illustration:</u>

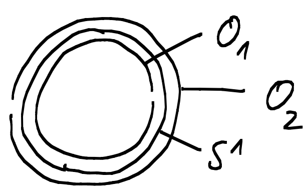


<u>Comment:</u> In the math. literature, one usually requires M to be topological space which is Hausdorff and paracompact. The charts are required to be

homeomorphisms. (This avoids certain pathological cases.) For us, this not important.

- Very simple example: $\underline{\underline{S^1}}$

  Let S^1 be defined by $\{e^{i\alpha}; \alpha \in [0, 2\pi)\} \subset \mathbb{R}^2$.



O_1 : all points except 1 ; O_2 : all points except -1

$\varphi_1 : e^{i\alpha} \longmapsto \alpha$; $\varphi_2 : e^{i(\pi+\alpha)} \longmapsto \alpha$

$(\alpha \in (0, 2\pi))$ $(\alpha \in (0, 2\pi))$

Easy to see: $\varphi_1(O_1 \cap O_2) = (0, \pi) \cap (\pi, 2\pi)$

$$\varphi_2 \circ \varphi_1^{-1} : (0, \pi) \longrightarrow (\pi, 2\pi)$$
$$\alpha \longmapsto \alpha + \pi$$
$$\& \quad (\pi, 2\pi) \longrightarrow (0, \pi)$$
$$\alpha \longmapsto \alpha - \pi$$

This is clearly a "diffeomorphism", as required.

- Less trivial examples include S^2 and T^2

  The existence of an atlas for S^2 is clear from any household world atlas (since the surface of the earth is S^2). For more details see problems.

- For us, the topological features of a manifold are less important since the observed universe is a subset of $\mathbb{R}^4$

$\Rightarrow$ one chart is sufficient.
(although the universe could be, e.g., $\mathbb{R} \times S^3$ with a very large S^3, etc.)

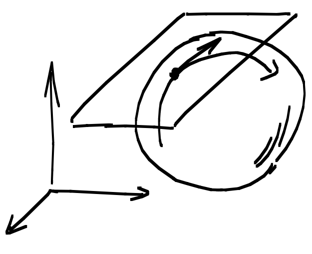
- What is crucial for us is that charts are more or less arbitrary (unphysical) and everything happening on the manifold must be describebable in <u>any chart</u> (i.e. any coordinate system).

- <u>More specifically:</u>

  Let a certain part of our space-time manifold be described by two charts φ_1 & φ_2 (i.e., we are in $O_1 \cap O_2$). Let x^μ be the $\mathbb{R}^4$-coordinates of U_1 and x'^μ be the $\mathbb{R}^4$-coordinates of U_2. Then $\varphi_2 \circ \varphi_1^{-1}$ can be characterized by a set of functions $x'^\mu(x^0, \cdots, x^3)$ or $x'^\mu = x'^\mu(x)$. All the physics has to be indepent of whether we use x or x', i.e., we require <u>diffeomorphism invariance</u>.

2.2 Tangent space

- Motivation: Consider S^2 as a manifold "embedded" in $\mathbb{R}^3$:



  The "velocities" of "trajectories" $x^\mu(\tau)$ going through a point p are clearly all vectors in the tangent-plane at p. (They are simply $dx^\mu/d\tau$.) However, we want a definition of this vector space which is independent of the embedding in $\mathbb{R}^n$.

Any curve through p can be used to define a linear differential operator on functions on M:

Let $f: M \to \mathbb{R}$ be a C^∞-fct. on M. Given $x^\mu(\tau)$, we can define the linear diff. operator at p:

$$f \longmapsto \frac{d f(x(\tau))}{d\tau} .$$

(This is clearly independent of the coordinate choice x^μ since the trajectory maps $\mathbb{R} \to M$ and f maps $M \to \mathbb{R}$; x^μ is never really needed.

<u>Definition:</u> A tangent vector at $p \in M$ is a map

$$v: \text{functions on } M \longrightarrow \mathbb{R}$$

such that 1) $v(af + bg) = a v(f) + b v(g)$

2) $v(fg) = v(f)\, g|_p + f|_p\, v(g)|_p$

(a, b constants; f, g functions)

- In coordinates, any such v can be written as

$$v = v^\mu \frac{\partial}{\partial x^\mu}$$

or $v(f) = v^\mu \dfrac{\partial f}{\partial x^\mu}$

or, more precisely, $v(f) = v^\mu \dfrac{\partial}{\partial x^\mu} (f \circ \varphi^{-1})\Big|_{\varphi(p)}$

relevant chart

- It is clear that a curve $x(\tau)$ does indeed give rise to such a vector:

$$\frac{d}{d\tau} f(x(\tau)) = \frac{\partial f}{\partial x^\mu} \cdot \frac{dx^\mu}{d\tau} = v^\mu \cdot \frac{\partial f}{\partial x^\mu} .$$

<u>Note:</u> curve not necessarily time-like, i.e., not trajectory.

- Thus, we can give an alternative definition of the tangent space at $p \in M$:

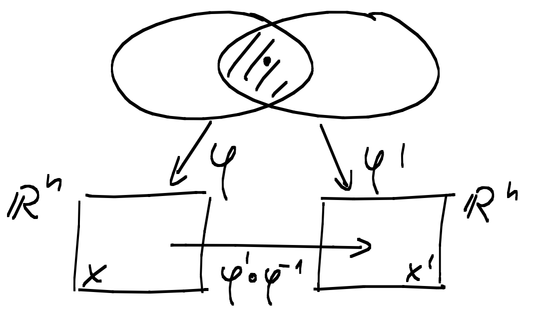
  The tangent space T_p is the space of equivalence classes of curves through p. Two curves ($x(\tau)$ & $y(\tau)$, in some coordinate system) are equivalent if
  $$\left.\frac{dx^\mu}{d\tau}\right|_p = \left.\frac{dy^\mu}{d\tau}\right|_p$$

(This definition is equivalent to the diff.-operator definition given before.)

- In each coord. system, there is a natural basis of the tangent space at p:
$$e_\mu = \left.\frac{\partial}{\partial x^\mu}\right|_{\varphi(p)} \qquad \text{("partial derivatives")}$$

$$\left(\text{recall } v = v^\mu \frac{\partial}{\partial x^\mu} = v^\mu e_\mu\right)$$

- Thus, a coordinate change $x \to x' = x'(x)$ induces a basis trf. on T_p:
$$v(f) = v^\mu \frac{\partial}{\partial x^\mu}(f \circ \varphi^{-1}) = v'^\mu \frac{\partial}{\partial x'^\mu}(f \circ \varphi'^{-1})$$



$$\varphi^{-1} = \varphi'^{-1} \circ \underbrace{(\varphi' \circ \varphi^{-1})}_{}$$
This is precisely the coord. change $x'(x)$

$$\Rightarrow \frac{\partial}{\partial x^\mu}(f \circ \varphi^{-1}) = \frac{\partial}{\partial x^\mu}\left(f \circ \varphi'^{-1} \circ (\varphi' \circ \varphi^{-1})\right) = \frac{\partial}{\partial x^\mu}\left((f \circ \varphi'^{-1})(x'(x))\right)$$
$$= \frac{\partial}{\partial x'^\nu}(f \circ \varphi'^{-1}) \cdot \frac{\partial x'^\nu}{\partial x^\mu}$$

$$\Rightarrow v^{\mu}\left(\frac{\partial x'^{\nu}}{\partial x^{\mu}}\right)\frac{\partial}{\partial x'^{\nu}}(f\circ\varphi'^{-1}) = v'^{\mu}\frac{\partial}{\partial x'^{\mu}}(f\circ\varphi'^{-1})$$

$$\Rightarrow v'^{\mu} = \underbrace{\left(\frac{\partial x'^{\mu}}{\partial x^{\nu}}\right)}\cdot v^{\nu}$$

This is the matrix realizing the basis change, as in our general vector space discussion before.

- An easy way to remember this:

$$\frac{\partial}{\partial x'^{\mu}} = \left(\frac{\partial x^{\nu}}{\partial x'^{\mu}}\right)\frac{\partial}{\partial x^{\nu}} = (M^{-1})^{\nu}{}_{\mu}\frac{\partial}{\partial x^{\nu}}$$

$$\Rightarrow v'^{\mu} = M^{\mu}{}_{\nu}v^{\nu} \quad ; \quad M^{\mu}{}_{\nu} = \frac{\partial x'^{\mu}}{\partial x^{\nu}}$$ is indeed the inverse matrix of $\frac{\partial x^{\mu}}{\partial x'^{\nu}}$.

2.3 Cotangent space & tensors, tensor fields

- The dual vector space to T_p is the <u>cotangent space T_p^*</u>
- The dual basis is denoted by dx^{μ}: $\quad dx^{\mu}\left(\frac{\partial}{\partial x^{\nu}}\right) = \delta^{\mu}{}_{\nu}$
- Elements of T_p^* are called covectors:

$$T_p^* \ni v = v_{\mu}\, dx^{\mu}.$$

- We can also consider tensors at every point $p \in M$:

$$\underbrace{T_p\otimes\cdots\otimes T_p}_{m}\otimes\underbrace{T_p^*\otimes\cdots\otimes T_p^*}_{n} \ni t^{\mu_1\cdots\mu_m}{}_{\nu_1\cdots\nu_n}\frac{\partial}{\partial x^{\mu_1}}\otimes\cdots\otimes\frac{\partial}{\partial x^{\mu_m}}\otimes dx^{\nu_1}\otimes\cdots\otimes dx^{\nu_n}$$

<u>A further comment on tensor spaces:</u>

We had defined tensor products of vector spaces via the basis

$$e_{\mu_1}\otimes\cdots\otimes e_{\mu_m}\otimes e^{\nu_1}\otimes\cdots\otimes e^{\nu_n}.$$

A basis-independent definition can be given as follows:
$\underbrace{V \otimes \cdots \otimes V}_{m} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{n}$ is the space of multilinear

(i.e. linear in every variable) functionals

$$\underbrace{V^* \times \cdots \times V^*}_{m} \times \underbrace{V \times \cdots \times V}_{n} \longrightarrow \mathbb{R}$$

direct product

Problem: Show the equivalence of these definitions.

- trf. law for tensors:

$$t'^{\mu_1 \cdots \mu_m}{}_{\nu_1 \cdots \nu_n} = \left(\frac{\partial x'^{\mu_1}}{\partial x^{\rho_1}}\right) \cdots \left(\frac{\partial x'^{\mu_m}}{\partial x^{\rho_m}}\right) t^{\rho_1 \cdots \rho_m}{}_{\sigma_1 \cdots \sigma_n} \left(\frac{\partial x^{\sigma_1}}{\partial x'^{\nu_1}}\right) \cdots \left(\frac{\partial x^{\sigma_n}}{\partial x'^{\nu_n}}\right)$$

- <u>tensor field</u>: A tensor field (of a certain rank) is

  specified by an element of the appropriate tensor product of T_p & T_p^* for <u>every</u> $p \in M$. It is smooth if the components $t^{\mu_1 \cdots \mu_m}{}_{\nu_1 \cdots \nu_n}$ are smooth functions in every coordinate patch.

  (Vector- & covector fields are special cases. Covector fields are also known as 1-forms, $\omega = \omega_\mu dx^\mu$.)

2.4 Metric

- In analogy to the situation for vector spaces, a metric on a manifold is (non-degenerate) covariant, symmetric, 2nd rank tensor field

$$\underline{\underline{g_{\mu\nu}(x)}}$$ (in some region parametrized by x).

- non-degenerate means that the matrix $g_{\mu\nu}$ is invertible at every point of M. (<u>Problem</u>: find coordinate-free formulation of this statement)

- We can use the metric to measure infinitesimal distances

$$ds^2 = g_{\mu\nu}(x)\, dx^\mu dx^\nu ,$$

(do not confuse with $dx^\nu \in T_p{}^*$)

or distances along a curve:

$$\int d\tau \sqrt{|g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu|} .$$

[Problem: 1) Check explicitly that this expression is diff.-invariant.
2) Formulate it in a coordinate-independent way, so that diff.-invariance becomes trivial]

- Using coordinate changes, we can bring $g_{\mu\nu}$ to the form $\mathrm{diag}(\pm 1, \ldots, \pm 1)$ at any given point $p \in M$ (Problem: prove this). However, this is in general not possible in a finite neighbourhood of p.

- Since we want to recover familiar Minkowski-space physics, we require that, in the above special form, the metric is $\mathrm{diag}(-1, +1, +1, +1)$ for our space-time manifold.

(<u>abstractly</u>: $g_{\mu\nu}$ has one neg. & 3 positive eigenvalues at every $p \in M$.)

We now have a "<u>Lorentzian manifold</u>".

2.5 Connection

- We will need to take partial derivatives, e.g. $\frac{\partial}{\partial x^\mu}(f \circ \varphi^{-1})$ or $\partial_\mu f$ for short. For f a fct. on M, this defines a covector field.

- We want to continue and take derivatives of tensor fields, e.g. $\partial_\mu w_\nu$ or $\partial_\mu w^\nu$ etc. However, these objects will not transform like tensor-components (since ∂_μ acts on the matrix transforming w_ν etc.). This makes our (coordinate-based) definition meaningless.

- What we would like to have is a "covariant" partial derivative D_μ (sometimes also ∇_μ) such that

$$D_\mu t^{\mu_1 \cdots \mu_m}{}_{\nu_1 \cdots \nu_n} \quad \text{is a tensor of increased rank.}$$

- We would like 1) Linearity: $D_\mu(\alpha v^\nu + \beta w^\nu) = \underset{\uparrow}{\alpha} D_\mu v^\nu + \underset{\uparrow}{\beta} D_\mu w^\nu$

constants

2) Leibnitz-rule: $D_\mu(v^\nu w^\rho) = (D_\mu v^\nu) w^\rho + v^\nu (D_\mu w^\rho)$

3) $D_\mu f = \partial_\mu f$ on functions

4) Consistency with contraction

$$\sum_\nu D_\mu t^\nu{}_\nu = D_\mu \left(\sum_\nu t^\nu{}_\nu\right)$$

- Comments concerning "4)"

a) "4)" is not a trivial consequence of linearity since, on the l.h. side, D_μ acts on a 2-index tensor while, on the r.h. side, it acts on a function. These two cases are in no way related to each other by properties 1) – 3).

b) Property 4) can be read as the natural extension of a covariant derivative defined on vectors to covectors:

$D_\mu v_\nu$ is a 2-index covariant tensor field. It is unambiguously defined if we know its action any element of $T_p \times T_p$ at any p. Thus, we need

$$(D_\mu v_\nu) w^\nu u^\mu \underset{\uparrow}{=} \sum_\nu D_\mu (v_\nu w^\nu) u^\mu - v_\nu (D_\mu w^\nu) u^\mu$$

by Leibnitz rule

$$\underset{\uparrow}{=} \partial_\mu (v_\nu w^\nu) u^\mu - v_\nu (D_\mu w^\nu) u^\mu .$$

by property "4)".

Since the last expression is defined, we have effectively defined the action of D_μ on covectors.

- Properties 1) – 4) (as well as the comment on property 4)) extend in an obvious way to higher-rank tensors, e.g.

$$\sum_{\nu_i} D_\mu t^{\nu_1 \cdots \nu_i \cdots \nu_m}{}_{\rho_1 \cdots \rho_{i-1} \nu_i \rho_{i+1} \cdots \rho_n} = D_\mu \sum_{\nu_i} t^{\nu_1 \cdots \nu_i \cdots \nu_m}{}_{\rho_1 \cdots \rho_{i-1} \nu_i \rho_{i+1} \cdots \rho_n} .$$

Important result:

The difference of two connections, $D - D'$, is, at every point $p \in M$, a linear map $T_p{}^* \longrightarrow T_p{}^* \otimes T_p{}^*$. (and correspondingly for vectors and higher tensors).

Demonstration:

We need to show that $D_\mu v_\nu$ depends only on the value of v_μ at p. Thus, we need to show that

$(D_\mu - D'_\mu) v_\nu = 0$ if $v_\nu|_p = 0$.

- Let $b_\nu^{(\rho)}$ be the components of a basis $b^{(\rho)}$ in a neighbourhood of p (i.e., the vector fields $b^{(\rho)}$ give rise to basis of the cotangent space at every point of this neighbourhood.). Then we can write

$$v_\nu = \sum_\rho f^{(\rho)} b_\nu^{(\rho)} \quad \text{with } f^{(\rho)}|_p = 0 \text{ for all } \rho.$$

- $$(D_\mu - D'_\mu)\Big(\sum_\rho f^{(\rho)} b_\nu^{(\rho)}\Big) = \sum_\rho \Big((D_\mu - D'_\mu) f^{(\rho)}\Big) \cdot b_\nu^\rho + \sum_\rho f^{(\rho)}\Big((D_\mu - D'_\mu) b_\nu^{(\rho)}\Big).$$

This vanishes at p since $D_\mu = D'_\mu = \partial_\mu$ on functions and since $f^{(\rho)}|_p = 0$. $\square$

- Thus, $D_\mu - D'_\mu$ applied to a covector field v_ν can be understood as a linear map $T_p^* \to T_p^*$ at every point p. Such a map can be characterized by a matrix $C_\rho{}^\nu$, $v_\rho \to C_\rho{}^\nu$. This map is also characterized by the index μ of the connection:

$$(D_\mu - D'_\mu) v_\nu = C_{\mu\nu}{}^\rho v_\rho$$

- It is easy to check that, given a certain coordinate system, we can define a connection by $\tilde{D}_\mu v_\nu = \partial_\mu v_\nu$. (Check properties 1) – 4)!)

(This connection depends on the coordinate system!)

- An arbitrary connection D_μ can be written as

$$D_\mu v_\nu = [\tilde{D}_\mu - (\tilde{D}_\mu - D_\mu)] v_\nu = \partial_\mu v_\nu - C_{\mu\nu}{}^{\rho} v_\rho .$$

- $C_{\mu\nu}{}^{\rho}$ depends on the connection D_μ and on the coordinate system. The specific $C_{\mu\nu}{}^{\rho}$ defined by taking the difference with $\tilde{D}_\mu = \partial_\mu$ are called Christoffel symbols and are denoted by $\Gamma_{\mu\nu}{}^{\rho}$

- In a given coordinate system, any connection can be specified by $\boxed{D_\mu v_\nu = \partial_\mu v_\nu - \Gamma_{\mu\nu}{}^{\rho} v_\rho}$.

- Note that, in spite of appearances, $\Gamma_{\mu\nu}{}^{\rho}$ does <u>not</u> represent the components of a 3-index tensor field.

<u>Problem</u>: Given a connection D, defined in a certain coordinate system x by its Christoffel symbols $\Gamma_{\mu\nu}{}^{\rho}$, calculate the Christoffel $\Gamma'_{\mu\nu}{}^{\rho}$ of this very same connection in a different coordinate system x'.

- Let us calculate explicitly the action of a connection D_μ on a vector field:

$$D_\mu(v^\nu w_\nu) = \partial_\mu (v^\nu w_\nu) = (\partial_\mu v^\nu) w_\nu + v^\nu (\partial_\mu w_\nu)$$

$$D_\mu(v^\nu w_\nu) = (D_\mu v^\nu) w_\nu + v^\nu (D_\mu w_\nu)$$

$$= (D_\mu v^\nu) w_\nu + v^\nu \partial_\mu w_\nu - v^\nu \Gamma_{\mu\nu}{}^{\rho} w_\rho$$

$$\Rightarrow (D_\mu v^\nu) w_\nu = (\partial_\mu v^\nu) w_\nu + \underbrace{\Gamma_{\mu\nu}{}^{\rho} v^\nu w_\rho}_{\Gamma_{\mu\rho}{}^{\nu} v^\rho w_\nu}$$

- Since w_ν is arbitrary, the equation remains valid dropping w_ν!

$$D_\mu v^\nu = \partial_\mu v^\nu + \Gamma_{\mu\rho}{}^\nu v^\rho$$

- This is easily extended to arbitrary tensor fields:

$$\begin{aligned} D_\mu t^{\mu_1 \cdots \mu_m}{}_{\nu_1 \cdots \nu_n} &= \partial_\mu t^{\mu_1 \cdots \mu_m}{}_{\nu_1 \cdots \nu_n} \\ &+ \Gamma_{\mu\rho}{}^{\mu_1} t^{\rho \cdots \mu_m}{}_{\nu_1 \cdots \nu_n} + \cdots + \Gamma_{\mu\rho}{}^{\mu_m} t^{\mu_1 \cdots \rho}{}_{\nu_1 \cdots \nu_n} \\ &- \Gamma_{\mu\nu_1}{}^\rho t^{\mu_1 \cdots \mu_m}{}_{\rho \cdots \nu_n} - \cdots - \Gamma_{\mu\nu_n}{}^\rho t^{\mu_1 \cdots \mu_m}{}_{\nu_1 \cdots \rho} \end{aligned}$$

2.6 The Riemannian connection

- Given a metric, one can define natural unique connection associated with this metric by demanding that
  1) The connection is torsion-free, i.e.
  $$[D_\mu, D_\nu] f = 0 \quad \text{for any fct. } f \text{ on } M.$$
  2) The metric is covariantly constant, i.e.
  $$D_\mu g_{\nu\rho} = 0 .$$

- $$\begin{aligned} [D_\mu, D_\nu] f &= D_\mu (\partial_\nu f) - D_\nu (\partial_\mu f) \\ &= \partial_\mu \partial_\nu f - \Gamma_{\mu\nu}{}^\rho \partial_\rho f - \partial_\nu \partial_\mu f + \Gamma_{\nu\mu}{}^\rho \partial_\rho f \end{aligned}$$

Thus, torsion-freeness is equivalent to $\Gamma_{\mu\nu}{}^\rho = \Gamma_{\nu\mu}{}^\rho$.

- $$0 = D_\mu g_{\nu\rho} = \partial_\mu g_{\nu\rho} - \Gamma_{\mu\nu}{}^\sigma g_{\sigma\rho} - \Gamma_{\mu\rho}{}^\sigma g_{\nu\sigma}$$

or, defining
$\Gamma_{\mu\nu\rho} = g_{\rho\sigma} \Gamma_{\mu\nu}{}^\sigma$: $\quad \Gamma_{\mu\nu\rho} + \Gamma_{\mu\rho\nu} = \partial_\mu g_{\nu\rho} \qquad (1)$

- By index substitution we also have

$$\Gamma_{\nu\mu\rho} + \Gamma_{\nu\rho\mu} = \partial_\nu g_{\mu\rho} \qquad (2)$$

$$\Gamma_{\rho\mu\nu} + \Gamma_{\rho\nu\mu} = \partial_\rho g_{\mu\nu} \qquad (3)$$

- (1) + (2) − (3) gives, using also torsion freeness:

$$2\,\Gamma_{\mu\nu\rho} = \partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu}$$

or $$\boxed{\Gamma_{\mu\nu}{}^{\rho} = \frac{1}{2} g^{\rho\sigma}\left(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}\right)}$$

This crucial result demonstrates the existence and uniqueness of the Riem. conn.

2.7 Parallel transport

- A vector field satisfying $D_\mu v^\nu = 0$ is called covariantly constant (in the given connection). (This extends to tensor fields in the obvious way.)

- To calculate covariant derivatives, we do not really need a vector field. It is sufficient if v^ν is defined at every point of a curve $x(\tau)$. We can then calculate the covariant derivative along the curve:

$$\frac{dx^\mu}{d\tau} D_\mu v^\nu = \frac{dx^\mu}{d\tau} \cdot \left(\partial_\mu v^\nu + \Gamma_{\mu\rho}{}^{\nu} v^\rho\right)$$

$$= \frac{dv^\nu}{d\tau} + \Gamma_{\mu\rho}{}^{\nu} \frac{dx^\mu}{d\tau} \cdot v^\rho \qquad (*)$$

If this vanishes, we say that v^ν is covariantly constant along the curve $x(\tau)$.

- Given a vector v^ν at some point p of the curve $x(\tau)$, we can view $(*)$ as a linear 1st. order diff. equation with initial value $v^\nu(\tau_p) = v^\nu$, where τ_p corresponds to $p \in M$. Solving this diff. equation (which is always possible), we can define $v^\nu(\tau)$ for all τ. This set of vectors will be covar. constant along $x(\tau)$. The vector $v^\nu(\tau_{p'})$ is called the <u>parallel transport</u> of v^ν from p to p' along $x(\tau)$.
- Thus, a connection allows for the parallel transport of vectors (or tensors) along curves in M.