SPRING 2025 MATH 6100
CLASSICAL GAUGE FIELD THEORY
LECTURE NOTES

S.T. HARTWIG

Lecture 1

Geometries

On the Origin of Geometry in Physics

In one word, the origin of geometry in physics is *observer-independence*. Physics is the ongoing attempt to describe "nature", which is to say, the collection of phenomena that seem "external" to us; phenomena which we feel would take place even without the presence of humans. It is therefore clear that physical laws, to the extent possible, should be formulated in a way that is independent of any one person. This includes past, current, and future people, all over the world. More generally, due to the feeling that we have a choice about our future, we include *hypothetical* people. The umbrella term for these different perspectives is that of an *observer*. Since we are not really concerned with individual details related only to the observer itself, what matters is how the observer interacts with the "external" environment. This is sometimes described as an observer carrying a clock and a marked rod. This is referring to the observer's notion of the passage of time and extension through space. More generally, we can imagine an observer carrying other collections of measuring apparatuses, or gauges, through which they can probe the world. In fact, it's the only thing we care about when it comes to an observer. Thus, an **observer** can be abstracted to a choice of coordinate system for spacetime, along with a choice of gauges. Collectively, we refer to coordinate systems and choice of gauge as a **reference frame**. These choices extend only locally, as the observer cannot perform measurements far away.

Now we encounter the strange dilemma of how an observer would actually make such a "choice". Consider the situation of a lone observer floating in empty space, carrying a clock and nothing else. Suppose the clock is equipped with a dial, enabling the clock hands to speed up or slow down. Consider two different settings, a slow and a fast mode. What is the difference? You might say that the observer knows roughly how long a second is, and could therefore tell the two settings apart. But this requires internal details about the observer. You could say, surely their heartbeat can be used to compare the two settings, but again, this requires internal details which we are not allowed to refer to. There is no heartbeat, no person, only a clock. You might say, the faster setting would wear out the gears of the clock faster and would allow us to distinguish between the settings. But that again refers to internal structure of the observer (now just a clock). Therefore the observer must be regarded as a disembodied clock without internal structure. There is therefore no physical difference between the settings of the clock. The same reasoning applies to choosing measuring rods, or calibrating various gauges. The inescapable conclusion is that there is actually **no content to a particular observer itself**, **only the relative comparison between two reference frames has physical meaning**. The allowed transformations between reference frames form the notion of geometry we introduce in this lecture.

Def. 1 Let $n > 0$ be an integer.

i) A (global, n-dimensional) coordinate system on a set M is a mapping from M to $\mathbb{R}^n$ denoted $x \longmapsto x^i = (x^1, x^2, \dots, x^n)$ which is one-to-one (i.e. injective).

ii) If $x \longmapsto x^{i'}$ is another coordinate system the map $x^i \longmapsto x^{i'}$ is called a coordinate transformation. It is a one-to-one and onto map from a subset of $\mathbb{R}^n$ (the image of $x \mapsto x^i$) to another subset of $\mathbb{R}^n$ (the image of $x \mapsto x^{i'}$)

iii) An (n-dimensional) geometry $\mathcal{G}$ is a collection of one-to-one and onto functions between subsets of $\mathbb{R}^n$, closed under taking inverses, and function compositions (when defined).

iv) A (global, n-dimensional) $\mathcal{G}$-space is a set M equipped with a set $\mathcal{C}$ of coordinate systems such that:

1. The coordinate transformations between any coordinate systems from $\mathcal{C}$ belong to $\mathcal{G}$
2. Composition defines a map $\mathcal{G} \times \mathcal{C} \to \mathcal{C}$.

Broadly, there are three geometries that play important roles in physics:

1) Galilean / Newtonian

Here $n = 3+1$ and $\mathcal{G}$ consists of three kinds of transformations and their compositions:

i) $\begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix} = \left[\begin{array}{c|c} 1 & 0\ 0\ 0 \\ \hline 0 & \\ 0 & A_i{}^j \\ 0 & \end{array}\right] \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix}$ where $A_i{}^j$ is an orthogonal 3x3 matrix: $\sum_{j,\ell} A_i{}^j A_k{}^\ell \delta_{j\ell} = \delta_{ik}$ (rotations, reflections)

ii) $\begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix} = \left[\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline v_x & 1 & 0 & 0 \\ v_y & 0 & 1 & 0 \\ v_z & 0 & 0 & 1 \end{array}\right] \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix}$ i.e. $\begin{cases} t' = t \\ x^{i'} = x^i + v^i \cdot t \end{cases}$ (Galilean boost)

iii) $\begin{cases} t' = t + b \\ x^{i'} = x^i + a^i \end{cases}$ (translations)

If $p, q \in M \xleftarrow{} $ a $\mathcal{G}$-space $(M, \mathcal{C})$ with coordinates $p^i = (t, x, y, z)$, $q^i = (u, a, b, c)$ then

$\Delta t := t - u$ and $\Delta S^2 := (x-a)^2 + (y-b)^2 + (z-c)^2$

only depend on p and q, not on the choice of coordinate system from $\mathcal{C}$.

Similarly, $\partial/\partial x^{i'} = \sum_j A^j_{i'} \partial/\partial x^j$ so the Laplace operator $\Delta := \sum_{i,j} (\partial/\partial x^i)(\partial/\partial x^j)\delta_{ij}$ is independent of coordinates (Check!)

2) Lorentzian (Special Relativity)

$n = 3+1$, c = constant conversion factor $x^0 = ct$

Here $\mathcal{G}$ consist of (compositions of)

i) $\begin{bmatrix} x^{0'} \\ x^{1'} \\ x^{2'} \\ x^{3'} \end{bmatrix} = \left[\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & & & \\ 0 & & A_i^{\;j} & \\ 0 & & & \end{array}\right]$ orthogonal (as before)

$A_i^{\;j} A_k^{\;l} \delta_{jl} = \delta_{ik}$

(rotations, reflections)

ii) $\begin{bmatrix} x^{0'} \\ x^{1'} \\ x^{2'} \\ x^{3'} \end{bmatrix} = \left[\begin{array}{c|ccc} \cosh\lambda & \sinh\lambda & 0 & 0 \\ \hline \sinh\lambda & \cosh\lambda & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} = \left[\begin{array}{c|ccc} \gamma & \beta\gamma & 0 & 0 \\ \hline \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix}$

Lorentz boost, λ = rapidity

$$\gamma^2 - \beta^2\gamma^2 = 1 \Rightarrow \gamma = (1-\beta^2)^{-1/2}$$

$$\Leftrightarrow \begin{cases} x^{0'} = \gamma \cdot (x^0 + \beta x^1) \\ x^{1'} = \gamma \cdot (\beta x^0 + x^1) \\ x^{2'} = x^2 \\ x^{2'} = x^3 \end{cases} \overset{x^0 = ct}{\Longleftrightarrow} (*) \begin{cases} t' = \gamma(t + \frac{\beta}{c} x^1) \\ x^{1'} = \gamma(\beta c t + x^1) \\ x^{2'} = x^2 \\ x^{3'} = x^3 \end{cases}$$

Note: $v_x := dx^{1'}/dt' = (dx^{1'}/dt)/(dt'/dt) = \beta c$

Thus, as $c \to \infty$, $(*)$ becomes a Galilean boost.

iii) Translations: $x^{\mu'} = x^\mu + a^\mu$ $(\mu = 0,1,2,3)$

<u>Note</u>: Combining ii) with i) we get Lorentz boosts in other directions.

<u>Thm</u> i) + ii) are equivalent to

i') $x^{\mu'} = \sum_{\nu=0}^{3} \Lambda^{\mu'}_{\ \nu} \, x^{\nu}$

where Λ is a 4×4-matrix satisfying

(a) $\sum_{\nu',\tau'} \Lambda^{\nu'}_{\ \mu} \Lambda^{\tau'}_{\ \sigma} \, \eta_{\nu'\tau'} = \eta_{\mu\sigma}$

$\eta_{\nu'\tau'} = \mathrm{diag}(+1,-1,-1,-1)$
"mostly −" convention

(b) $\Lambda^{0}_{\ 0} > 0$ time-direction-preserving
=: orthochronous

(c) $\det \Lambda = 1$

<u>Proof</u>: Exercise ∎

$$\begin{tikzcd}[column sep=small]
 & O(1,3) = \{\Lambda \mid (a) \text{ holds}\} \arrow[dl, dash] \arrow[dr, dash] & \\
\det\Lambda = 1 \quad SO(1,3) \arrow[dr, dash] & & O^{+}(1,3) \quad \Lambda^{0}_{\ 0} > 0 \arrow[dl, dash] \\
 & SO^{+}(1,3) \quad \Lambda^{0}_{\ 0} > 0, \det\Lambda = 1 &
\end{tikzcd}$$

3) The smooth (or general) geometry consists of all one-to-one and onto functions $f : U \to V$, where U and V are open subsets of $\mathbb{R}^n$, and all partial derivatives of f (and f^{-1}) exist to all orders.

Denoting f by

$$x^i \xmapsto{f} x^{i'} = x^{i'}(x^1, \ldots, x^n) =: x^{i'}(x^j)$$

we let

$$(J_i^{j'}) = \left(\frac{\partial x^{j'}}{\partial x^i} \right)$$

be the Jacobian matrix and

$$J = \det(J_i^{j'})$$

We get more restricted geometries by putting conditions on J:

$J > 0$	oriented geometry
$\lvert J \rvert = 1$	unimodular
$J = 1$	proper

Homework 1

1) Show that a matrix $A = (A_i^j)_{i,j=1}^n$ satisfies $A^T \cdot A = I_n$ iff

$$\sum_{j,l} A_i^j A_k^l \delta_{jl} = \delta_{ik} \quad \forall i,k$$

where $\delta_{jl} = \begin{cases} 1 & j = l \\ 0 & j \neq l \end{cases}$

is the Kronecker delta.

2) If $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a symmetric bilinear form with

$$B(e_i, e_j) = B_{ij} \in \mathbb{R}$$

and $A = (A_i^j)_{i,j=1}^n$ is an $n \times n$-matrix then

$$B(Av, Aw) = B(v, w) \quad \forall v, w \in \mathbb{R}^n$$

iff

$$\sum_{j,l} A_i^j A_k^l B_{jl} = B_{ik} \quad \forall i,k$$

3) Show that if $\mathcal{G}$ is a geometry consisting of smooth (that is, $\mathcal{C}^\infty$) maps between open subsets of $\mathbb{R}^n$, and $\phi : M \to \mathbb{R}$ is a function from a $\mathcal{G}$-space $(M, \mathcal{C})$ to $\mathbb{R}$ and the coordinate representation $\phi(x^i) = \phi(x^1, x^2, \ldots, x^n)$ of ϕ in a coord. system x^i is smooth then $\phi(x^{i'})$ is smooth in all other coord. systems $x^{i'}$.
(Then we say ϕ is smooth)

4) Show that the Laplace operator

$$\Delta = \sum_{i,j} \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} \delta_{ij}$$

is invariant under Galilean coordinate transformations.

5) Show the d'Alembert operator

$$\Box = \sum_{\mu,\nu=0}^{3} \frac{\partial}{\partial x^\mu} \frac{\partial}{\partial x^\nu} \eta_{\mu\nu}$$

is invariant under Lorentzian coord. transformations.

6) Newton-Laplace Equation for Gravity is

$$\Delta \phi = 4\pi G \rho \qquad (*)$$

where

- $\phi = \phi(t,x,y,z)$ gravitational potential

  By definition of ϕ,

  $\nabla \phi = -\vec{a}$, where $\vec{a}(t,x,y,z)$ is the acceleration of a test particle, and

- $\rho = \rho(t,x,y,z)$ is a mass density distribution in spacetime.

Show that for $\rho(t,x,y,z) = \delta(0) M$, we recover Newton's Law of Universal Gravitation.

Integrate both sides of $(*)$.

$$\iiint_{B(r,0)} \Delta \phi \, d^3x = \iint_{\partial B(r,0)} \nabla \phi \cdot d\vec{S} \quad \text{by divergence theorem}$$

$$= -\left|(\nabla \phi)(r)\right| 4\pi r^2 \quad \text{by symmetry}$$

$$\iiint_{B(r,0)} 4\pi G \rho \, d^3x = 4\pi G \cdot M$$

test particle mass

$$F = m \cdot |\vec{a}(r)| = G \frac{m \cdot M}{r^2}$$

7) In Lorentzian geometry, show that if a curve is given by
$$x^i = x^i(t) \quad , \quad i=1,2,3$$
in some coordinate system, and
$$\sum_{i=1}^{3} (dx^i/dt)^2 = c^2$$
then the same is true in any other coordinate system.
(This shows the speed of light is the same in all reference frames.)

8) In Lorentzian geometry, if two events $x \neq y \in M$ have the same time coordinates
$$x^0 = y^0$$
in some coord. system, then
$$x^{0'} \neq y^{0'}$$
in some other coord. system.
(Thus simultaneity is lost.)

9) In Lorentzian geometry, the interval between two events $x, y \in M$ is defined by
$$\Delta\tau = \left(\sum_{\mu,\nu=0}^{3} (x^\mu - y^\mu)(x^\nu - y^\nu)\, \eta_{\mu\nu} \right)^{1/2}$$
(in a coord system). Show $\Delta\tau$ is independent of the choice of coords.

Lecture 2

Tensor Fields

① Vector fields; Contravariant Tensors

Suppose $\mathcal{G} \subset \mathcal{G}_{sm}$ is a smooth geometry, i.e. consisting of smooth maps between open subsets of $\mathbb{R}^n$, and let M be a $\mathcal{G}$-space.

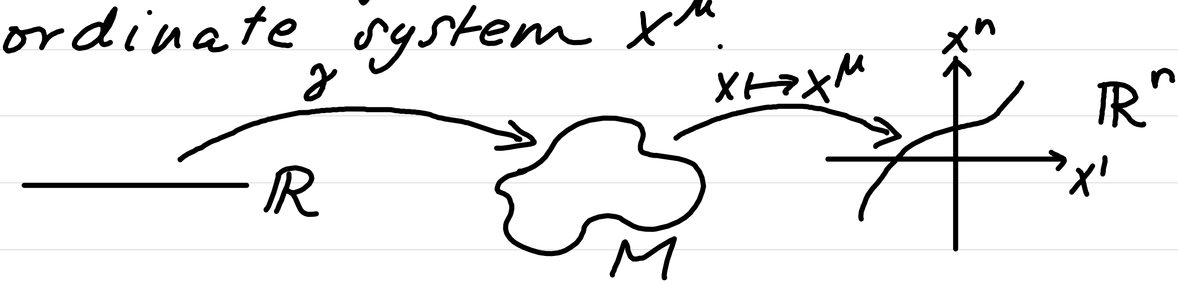
<u>Def</u> A (smooth) curve in M is a function

$$\gamma : \mathbb{R} \longrightarrow M$$

such that the coordinates

$$\gamma^\mu(t) = x^\mu(\gamma(t))$$

are smooth functions of t, in some coordinate system x^μ.



The tangent vector of γ at $\gamma(t)$ is

$$\dot{\gamma}^\mu(t) = \frac{d}{dt}\gamma^\mu(t)$$

<u>Note</u>: If $x^{\mu'}$ is any other coordinate system then $\gamma^{\mu'}(t) = x^{\mu'}(\gamma^\nu(t))$, where $x^{\mu'}(x^\nu)$ denotes the coordinate transformation. Since $x^{\mu'}(x^\nu)$ and $\gamma^\nu(t)$ are smooth, so is $\gamma^{\mu'}(t)$.

Furthermore, by the chain rule,

$$\dot{\gamma}^{\mu'}(t) := \frac{d}{dt}\left(\gamma^{\mu'}(t)\right) = \frac{d}{dt}\left(x^{\mu'}(\gamma^{\nu}(t))\right) =$$
$$= \sum_{\nu} \frac{\partial x^{\mu'}}{\partial x^{\nu}} \cdot \dot{\gamma}^{\nu}(t) = J^{\mu'}_{\nu} \dot{\gamma}^{\nu}(t)$$

where summation over repeated indices (always one upper and one lower) is implied.

Since any vector in $\mathbb{R}^n$ is the tangent vector of some curve, we make the following definition:

<u>Def</u> A (smooth) vector field on M assigns to each coordinate system x^{μ} on M an n-tuple of smooth real-valued functions $A^{\mu} = (A^{\mu})_{\mu=1}^{n}$ (defined on the image of x^{μ}) such that, under any coordinate transformation $x^{\mu} \mapsto x^{\mu'}(x^{\nu})$ we have:

$$A^{\mu'} = J^{\mu'}_{\nu} A^{\nu} \qquad (*)$$

where $A^{\mu'}$ is the n-tuple corresponding to the coordinate system $x^{\mu'}$.

A^{μ} are the components of the vector field.

For $n=2$, (*) written out reads:

$$\begin{cases} A^{1'}(x^{1'},x^{2'}) = \dfrac{\partial x^{1'}}{\partial x^1}(x^1,x^2)\cdot A^1(x^1,x^2) + \dfrac{\partial x^{1'}}{\partial x^2}(x^1,x^2)\cdot A^2(x^1,x^2) \\ A^{2'}(x^{1'},x^{2'}) = \dfrac{\partial x^{2'}}{\partial x^1}(x^1,x^2)\cdot A^1(x^1,x^2) + \dfrac{\partial x^{2'}}{\partial x^2}(x^1,x^2)\cdot A^2(x^1,x^2) \end{cases}$$

Addition and Scaling

If A^μ and B^μ are (the components in an arbitrary coordinate system x^μ of) two vector fields, then

$$A^\mu + B^\mu \quad \text{and} \quad cA^\mu \qquad (c\in\mathbb{R} \text{ const.})$$

define vector fields. Indeed,

$$A^{\mu'} + B^{\mu'} = J^{\mu'}_\nu A^\nu + J^{\mu'}_\nu B^\nu$$
$$= J^{\mu'}_\nu \cdot (A^\nu + B^\nu)$$

$$cA^{\mu'} = c\, J^{\mu'}_\nu A^\nu = J^{\mu'}_\nu \cdot (cA^\nu)$$

Multiplying two vector fields does not yield another vector field:

$$A^{\mu'}B^{\nu'} = J^{\mu'}_\lambda J^{\nu'}_\rho A^\lambda B^\rho$$

From now on we drop "smooth".

Def A (contravariant) r-tensor (field) assigns to each coordinate system x^μ an n^r-tuple of functions $T^{\mu_1 \mu_2 \cdots \mu_r}$ such that under $x^\mu \mapsto x^{\mu'}$,

$$T^{\mu_1' \mu_2' \cdots \mu_r'} = J^{\mu_1'}_{\nu_1} J^{\mu_2'}_{\nu_2} \cdots J^{\mu_r'}_{\nu_r} T^{\nu_1 \nu_2 \cdots \nu_r}.$$

Example: If A^μ, B^ν are vector fields then $A^\mu B^\nu$ is a 2-tensor.

(2) Covector Fields; Covariant Tensors

Def The gradient of a scalar field ϕ is

$$\partial_\mu \phi = \frac{\partial}{\partial x^\mu} \phi(x^\mu)$$

Note: Under $x^\mu \mapsto x^{\mu'}$, the gradient transforms as

$$\begin{aligned}
\partial_{\mu'} \phi &= \frac{\partial}{\partial x^{\mu'}} \phi(x^\lambda) \\
&= \frac{\partial x^\nu}{\partial x^{\mu'}} \frac{\partial}{\partial x^\nu} \phi(x^\lambda) \\
&= J^\nu_{\mu'} \partial_\nu \phi
\end{aligned}$$

Note that $J^\nu_{\mu'} J^{\mu'}_\lambda = \delta^\nu_\lambda$, $J^\nu_{\mu'} J^{\lambda'}_\nu = \delta^{\lambda'}_{\mu'}$.

Def A covector field assigns to each coord. system x^μ an n-tuple B_μ which transform as

$$B_{\mu'} = J^{\nu}_{\mu'} B_\nu$$

under $x^\mu \mapsto x^{\mu'}$.

Def A covariant s-tensor (field) assigns to each coord system an n^s-tuple $B_{\mu_1 \ldots \mu_s}$ which transform as

$$B_{\mu_1' \ldots \mu_s'} = J^{\nu_1}_{\mu_1'} \cdots J^{\nu_s}_{\mu_s'} B_{\nu_1 \ldots \nu_s}$$

3 Mixed Tensors

Def An $\binom{r}{s}$-tensor, or a tensor of rank $\binom{r}{s}$ assigns to each coord system x^μ an n^{r+s}-tuple

$$T^{\mu_1 \cdots}_{\nu_1 \cdots} \quad \begin{matrix} (r \text{ indices}) \\ (s \text{ indices}) \end{matrix}$$

which transform as

$$T^{\mu_1' \cdots}_{\nu_1' \cdots} = J^{\mu_1'}_{\lambda_1} \cdots J^{\rho_1}_{\nu_1'} \cdots T^{\lambda_1 \cdots}_{\rho_1 \cdots}$$

Note The set $\mathcal{T}\binom{r}{s}$ of all $\binom{r}{s}$-tensors forms a vector space. Furthermore if $A^{\mu_1 \cdots}_{\nu_1 \cdots}$ is an $\binom{r}{s}$-tensor and $B^{\mu_1 \cdots}_{\nu_1 \cdots}$ is a $\binom{t}{u}$-tensor, then $A^{\mu_1 \cdots}_{\nu_1 \cdots} B^{\sigma_1 \cdots}_{\tau_1 \cdots}$ is an $\binom{r+t}{s+u}$-tensor.

__Example__ In all coord. systems, define

$$\delta_\mu^\nu := \begin{cases} 1 & \mu = \nu \\ 0 & \text{else} \end{cases}$$

Then δ_μ^ν is a $\binom{1}{1}$-tensor, called the Kronecker delta.

Indeed, we must check

$$\delta_{\mu'}^{\nu'} = J_\sigma^{\nu'} J_{\mu'}^{\tau} \delta_\tau^\sigma$$

The right hand side equals

$$J_\sigma^{\nu'} J_{\mu'}^{\sigma}$$

which is $\delta_{\mu'}^{\nu'}$ as we noted earlier.

Example

The Levi-Civita symbol (in dim n) is defined in all coordinate systems by

$$\varepsilon_{\mu_1\mu_2\cdots\mu_n} = \begin{cases} +1, & \text{if } (\mu_1,\mu_2,\ldots,\mu_n) \text{ is an even permutation of } (1,2,\ldots,n) \\ -1, & \ldots\ldots \text{ odd } \ldots\ldots \\ 0, & \text{otherwise.} \end{cases}$$

Ex. $n=2$

$$\varepsilon_{12} = -\varepsilon_{21} = 1, \quad \varepsilon_{11} = \varepsilon_{22} = 0$$

We have,

$$J^{\nu_1'}_{\mu_1} J^{\nu_2'}_{\mu_2} \cdots J^{\nu_n'}_{\mu_n} \varepsilon_{\nu_1'\nu_2'\cdots\nu_n'} =$$

$$= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) J^{\sigma(1')}_{\mu_1} J^{\sigma(2')}_{\mu_2} \cdots J^{\sigma(n')}_{\mu_n} \varepsilon_{\mu_1\mu_2\cdots\mu_n} =$$

$$= \det(J^{\rho'}_{\lambda})\, \varepsilon_{\mu_1\mu_2\cdots\mu_n} = J' \cdot \varepsilon_{\mu_1\mu_2\cdots\mu_n}$$

Thus, $\varepsilon_{\mu_1\cdots\mu_n}$ is a tensor if and only if we are working in proper geometry ($J'=1$).

In general, it is a so-called relative tensor, but is still useful for expressing determinants of $\binom{1}{1}$-tensors:

$$\det(A^{\rho}_{\lambda})\, \varepsilon_{\mu_1\cdots\mu_n} = A^{\nu_1}_{\mu_1} \cdots A^{\nu_n}_{\mu_n} \varepsilon_{\nu_1\cdots\nu_n}$$

④ Contraction

Setting an upper index equal to a lower index and summing over that index turns an $\binom{r}{s}$-tensor into an $\binom{r-1}{s-1}$-tensor.

Ex. $A^{\lambda}_{\mu\nu}$ contracted along $\lambda=\nu$: $A^{\nu}_{\mu\nu}$

We check it's a $\binom{0}{1}$-tensor:

$$A^{\nu}_{\mu\nu} = J^{\nu}_{\sigma'} J^{\tau'}_{\mu} J^{\lambda'}_{\nu} A^{\sigma'}_{\tau'\lambda'} = \delta^{\lambda'}_{\sigma'} J^{\tau'}_{\mu} A^{\sigma'}_{\tau'\lambda'} = J^{\tau'}_{\mu} A^{\sigma'}_{\tau'\sigma'}$$

Relabeling & permuting indices.

If $A^{\mu\nu\lambda}$ is a tensor then obviously so is $A^{\nu\lambda\mu}$ or $A^{\alpha\beta\gamma}$ etc.

So, $A^{\mu\nu\lambda} + A^{\nu\lambda\mu} + A^{\lambda\mu\nu}$

is a tensor (of rank $\binom{3}{0}$)

Another example, if B^{j}_{i} C^{ℓ}_{k} are tensors then so is

$$B^{j}_{i} C^{\ell}_{k} + B^{\ell}_{i} C^{j}_{k}$$

All operations from linear algebra arise using contraction.

linear alg notion	tensor notion
vector $v = v^i e_i$	vector field v^i
bilinear form b	$\binom{0}{2}$-tensor b_{ij}
linar form ξ	covector ξ_i
linear map T	$\binom{1}{1}$-tensor T^i_j
$T(v)$	$T^i_j v^j$
$b(v,w)$	$b_{ij} v^i w^j$
$u \in V \otimes V$	u^{ij}
Alg mult. $m : V \otimes V \to V$	$\binom{1}{2}$-tensor m^i_{jk}
$m(v \otimes w)$	$m^i_{jk} v^j w^k$
$\mathrm{Tr}(T)$	T^i_i
$\xi(v)$	$\xi_i v^i$
$\det(A)$	$\varepsilon_{i_1 \cdots i_n} A^{i_1}_1 A^{i_2}_2 \cdots A^{i_n}_n$
Kronecker matrix product $A \otimes B$	$A^j_i B^l_k$
$u \in \Lambda^2 V$	$\frac{1}{2}(u^{ij} - u^{ji})$
Anti-symm bil form $\omega : V \times V \to \mathbb{R}$	$\frac{1}{2}(\omega_{ij} - \omega_{ji})$

Choosing the linear geometry $\mathcal{G} = GL_n$ and fixing a point $x \in M$, the above can be made precise. We leave the details as an exercise.

Lecture 3
Covariant Differentiation

① Constant Fields

Fix a smooth geometry $\mathcal{G} \subset \mathcal{G}_{sm}$ and a $\mathcal{G}$-space $\mathcal{M} = (M, \mathcal{C})$.

If $\phi : M \longrightarrow \mathbb{R}$ is a scalar field we say ϕ is (locally) constant if

$$\partial_\mu \phi = 0$$

<u>Remark 1.</u> It makes sense to say that $\phi : M \longrightarrow \mathbb{R}$ is a constant function. This does not require a coordinate system. Every such constant function is locally constant in the above sense. The converse holds if the image of $x \longmapsto x^\mu$ is connected.

<u>Remark 2.</u> The condition $\partial_\mu \phi = 0$ is coordinate independent, because $\partial_\mu \phi$ is a tensor field.

__Remark 3__ If A^μ is a vector field on M, the condition

$$\partial_\mu A^\nu = 0$$

is NOT coordinate independent:

$$\partial_{\mu'} A^{\nu'} = J^{\lambda}_{\mu'} \partial_\lambda \left(J^{\nu'}_{\rho} A^\rho \right) =$$

$$= J^{\lambda}_{\mu'} J^{\nu'}_{\rho} \underbrace{\partial_\lambda A^\rho}_{=0} + \underbrace{J^{\lambda}_{\mu'} J^{\nu'}_{\lambda\rho} A^\rho}_{\substack{\text{no reason} \\ \text{this is zero}}} \qquad (1)$$

where we put

$$J^{\nu'}_{\lambda\rho} = \frac{\partial^2 x^{\nu'}}{\partial x^\lambda \partial x^\rho}$$

This problem is related to the fact that $\partial_\mu A^\nu$ does not transform as a $\binom{1}{1}$-tensor.

Remark 4 The same problem appears for covector fields. However, the expression

$$F_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$$

does transform as a $\binom{0}{2}$-tensor and therefore the equation $F_{\mu\nu} = 0$ is coordinate-independent.

The lesson we draw here is that sometimes, the sum of two non-tensorial terms is tensorial.

This leads to the idea of a counteracting term.

Returning to (1), notice that in the "error" term,

$$\mathcal{J}^{\lambda}_{\mu'} \mathcal{J}^{\nu'}_{\lambda\rho}$$

has the index structure $\binom{1}{2}$ and is contracted against A^ρ.

② Affine Connections

This leads to the following Ansatz:

$$\nabla_\mu A^\nu := \partial_\mu A^\nu + \Gamma^\nu_{\mu\rho} A^\rho \qquad (2)$$

and the question is:

> How should $\Gamma^\nu_{\mu\rho}$ transform under coordinate changes, so as to counteract the error term in (1), making the object $\nabla_\mu A^\nu$ into a tensor of type $\binom{1}{1}$?

What we have is, by (1),

$$\nabla_{\mu'} A^{\nu'} = \partial_{\mu'} A^{\nu'} + \Gamma^{\nu'}_{\mu'\rho'} A^{\rho'} =$$

$$= J^\lambda_{\mu'} J^{\nu'}_\rho \partial_\lambda A^\rho + \left(J^\lambda_{\mu'} J^{\nu'}_{\lambda\tau} + \Gamma^{\nu'}_{\mu'\rho'} J^{\rho'}_\tau \right) A^\tau \qquad (3)$$

We want this to equal

$$J^\lambda_{\mu'} J^{\nu'}_\rho \nabla_\lambda A^\rho =$$

$$= J^\lambda_{\mu'} J^{\nu'}_\rho \partial_\lambda A^\rho + J^\lambda_{\mu'} J^{\nu'}_\rho \Gamma^\rho_{\lambda\tau} A^\tau \qquad (4)$$

Equating the coefficients of A^τ in (3) and (4) gives

$$J^\lambda_{\mu'} J^{\nu'}_{\lambda\tau} + \Gamma^{\nu'}_{\mu'\rho'} J^{\rho'}_{\tau} = J^\lambda_{\mu'} J^{\nu'}_{\rho} \Gamma^{\rho}_{\lambda\tau}$$

Inverting $J^{\rho'}_\tau$ we get (by HW):

$$\boxed{\Gamma^{\nu'}_{\mu'\rho'} = J^\lambda_{\mu'} J^\tau_{\rho'} J^{\nu'}_{\rho} \Gamma^{\rho}_{\lambda\tau} - J^\lambda_{\mu'} J^\tau_{\rho'} J^{\nu'}_{\lambda\tau}}$$

Def An affine connection on M assigns to each coord. system x^μ a n^3-tuple of functions $\Gamma^\nu_{\mu\rho}$ which transform as above under change of coordinates $x^\mu \mapsto x^{\mu'}$.

Remark A better name would be a $\mathcal{G}$-connection.

Def An affine connection space is a $\mathcal{G}$-space equipped with an affine connection:

$$(M, \Gamma) = (M, \mathcal{E}; \Gamma^\nu_{\mu\rho})$$

③ Covariant Derivative

Let (M, Γ) be an affine connection space.

<u>Def</u> The covariant derivative of a vector field A^μ is defined as the $\binom{1}{1}$-tensor

$$\nabla_\mu A^\nu = \partial_\mu A^\nu + \Gamma_\mu{}^\nu{}_\alpha A^\alpha$$

A^μ is (locally) constant if $\nabla_\mu A^\nu = 0$

<u>Notation</u> For a scalar field ϕ we put

$$\nabla_\mu \phi = \partial_\mu \phi$$

as this already is tensorial.

So we have:

ϕ $\binom{0}{0}$-tensor $\Rightarrow$ $\nabla_\mu \phi$ $\binom{0}{1}$-tensor

A^μ $\binom{1}{0}$-tensor $\Rightarrow$ $\nabla_\mu A^\nu$ $\binom{1}{1}$-tensor

What about $\nabla_\mu T^{\alpha\beta}$?

We would want the product rule to hold:

$$\nabla_\mu (A^\lambda B^\rho) = (\nabla_\mu A^\lambda) B^\rho + A^\lambda (\nabla_\mu B^\rho) =$$

$$= (\partial_\mu A^\lambda + \Gamma_\mu{}^\lambda{}_\alpha A^\alpha) B^\rho + A^\lambda (\partial_\mu B^\rho + \Gamma_\mu{}^\rho{}_\alpha B^\alpha)$$

$$= \partial_\mu (A^\lambda B^\rho) + \Gamma_\mu{}^\lambda{}_\alpha A^\alpha B^\rho + \Gamma_\mu{}^\rho{}_\alpha A^\lambda B^\alpha$$

So we guess

$$\nabla_\mu T^{\lambda\rho} = \partial_\mu T^{\lambda\rho} + \Gamma_\mu{}^\lambda{}_\alpha T^{\alpha\rho} + \Gamma_\mu{}^\rho{}_\alpha T^{\lambda\alpha}$$

This does indeed work (check!)

What about covariant tensor fields? Since $A_\mu A^\mu$ is a scalar, we want:

$$\partial_\mu (A_\nu A^\nu) = \nabla_\mu (A_\nu A^\nu) = \nabla_\mu (A_\nu) A^\nu + A_\nu \nabla_\mu A^\nu$$

$$= \nabla_\mu (A_\nu) A^\nu + A_\nu (\partial_\mu A^\nu + \Gamma_\mu{}^\nu{}_\alpha A^\alpha)$$

$$\Leftrightarrow \nabla_\mu (A_\nu) A^\nu = (\partial_\mu A_\nu - \Gamma_\mu{}^\beta{}_\nu A_\beta) A^\nu$$

which suggests

$$\nabla_\mu A_\nu = \partial_\mu A_\nu - \Gamma_\mu{}^\alpha{}_\nu A_\alpha$$

Again, this does give a $\binom{0}{2}$-tensor (check!)

Def The covariant derivative of an $\binom{r}{s}$-tensor field $T^{\mu_1 \cdots \mu_r}_{\nu_1 \cdots \nu_s}$ is

$$\nabla_\lambda T^{\mu_1 \cdots \mu_r}_{\nu_1 \cdots \nu_s} = \partial_\lambda T^{\mu_1 \cdots \mu_r}_{\nu_1 \cdots \nu_s} + \sum_{i=1}^{r} \Gamma^{\mu_i}_{\lambda \alpha} T^{\mu_1 \cdots \overset{\text{i:th pos.}}{\alpha} \cdots \mu_r}_{\nu_1 \cdots \nu_s} - \sum_{j=1}^{s} \Gamma^{\beta}_{\lambda \nu_j} T^{\mu_1 \cdots \mu_r}_{\nu_1 \cdots \underset{\text{j:th pos.}}{\beta} \cdots \nu_s}$$

The result is an $\binom{r}{s+1}$-tensor field.

Def An affine connection is symmetric if

$$\Gamma^{\mu}_{\alpha\beta} = \Gamma^{\mu}_{\beta\alpha}$$

Remark Due to $J^{\lambda}_{\alpha\beta} = J^{\lambda}_{\beta\alpha}$, this condition is coordinate-independent. (Check!)

④ Locally Inertial Coordinates

Now take $g = g_{sm}$.

<u>Thm</u> Let Γ be a symmetric affine connection on M. Let $x \in M$ be any point. Then there exists a coordinate system x^μ in which $\Gamma_\mu{}^\nu{}_\lambda(x) = 0$.

<u>Proof</u> Let x^μ be any coord. system. After a translation if necessary, we may assume $x^\mu = 0$. Consider the coordinate transformation

$$x^{\mu'} = A^{\mu'}_\alpha x^\alpha + \frac{1}{2} B^{\mu'}_{\alpha\beta} x^\alpha x^\beta$$

($A^{\mu'}_\alpha$, $B^{\mu'}_{\alpha\beta}$ <u>constant</u>)

where we assume $B^{\mu'}_{\alpha\beta} = B^{\mu'}_{\beta\alpha}$. We have

$$J^{\mu'}_\alpha = A^{\mu'}_\alpha + B^{\mu'}_{\alpha\beta} x^\beta, \quad J^{\mu'}_{\alpha\beta} = B^{\mu'}_{\alpha\beta}$$

The transformation law for Γ can be written

$$J^{\mu'}_\alpha J^{\nu'}_\beta \Gamma_{\mu'}{}^{\lambda'}{}_{\nu'} = J^{\lambda'}_\gamma \Gamma_\alpha{}^\gamma{}_\beta - J^{\lambda'}_{\alpha\beta}$$

At $x^\mu = 0$ ($\Leftrightarrow x^{\mu'} = 0$) the RHS equals

$$A^{\lambda'}_{\gamma} \Gamma^{\gamma}_{\alpha\beta}(0) - B^{\lambda'}_{\alpha\beta}$$

So $A^{\lambda'}_{\gamma}$ can be any invertible matrix, and we take

$$B^{\lambda'}_{\alpha\beta} = A^{\lambda'}_{\gamma} \Gamma^{\gamma}_{\alpha\beta}(0) \quad \left(\begin{array}{c}\text{requires}\\ \text{symmetric}\\ \text{in } \alpha,\beta\,!\end{array}\right)$$

Then $\Gamma^{\nu'}_{\mu'\lambda'}(0) = 0$, as required. $\blacksquare$

Def A coordinate system x^μ on an affine connection space (M, Γ) is **locally inertial** at $x \in M$ if $\Gamma^{\nu}_{\mu\lambda}(x) = 0$.

Such coordinate systems are also called **normal** and the point x is a **pole**.

Lecture 4

Curvature and Torsion, Metric and Vielbein

① Curvature and Torsion

Let (M, Γ) be an affine connection space (in a smooth geometry $\mathcal{G} \subset \mathcal{G}_{sm}$).

To what extent do the covariant derivative components ∇_μ commute? On a vector field A^μ we have

$$\nabla_\mu \nabla_\nu A^\lambda = \partial_\mu(\nabla_\nu A^\lambda) - \Gamma^\alpha_{\mu\nu} \nabla_\alpha A^\lambda + \Gamma^\lambda_{\mu\alpha} \nabla_\nu A^\alpha =$$

$$= \cancel{\partial_\mu \partial_\nu A^\lambda} + (\partial_\mu \Gamma^\lambda_{\nu\alpha}) A^\alpha + \cancel{\Gamma^\lambda_{\nu\alpha} \partial_\mu A^\alpha}$$

$$- \Gamma^\alpha_{\mu\nu} \nabla_\alpha A^\lambda$$

$$+ \cancel{\Gamma^\lambda_{\mu\alpha} \partial_\nu A^\alpha} + \Gamma^\lambda_{\mu\alpha} \Gamma^\alpha_{\nu\beta} A^\beta$$

Switching μ and ν and subtracting, the 1st, 3rd and 5th terms cancel:

$$(\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu) A^\lambda = -(\underbrace{\Gamma^\alpha_{\mu\nu} - \Gamma^\alpha_{\nu\mu}}_{T^\alpha_{\mu\nu}}) \nabla_\alpha A^\lambda$$

torsion tensor

$$+ (\underbrace{\partial_\mu \Gamma^\lambda_{\nu\beta} - \partial_\nu \Gamma^\lambda_{\mu} + \Gamma^\lambda_{\mu\alpha} \Gamma^\alpha_{\nu\beta} - \Gamma^\lambda_{\nu\alpha} \Gamma^\alpha_{\mu\beta}}_{R^\lambda_{\mu\nu\beta}}) A^\beta$$

Riemann Curvature tensor

Thus

$$(\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu) A^\lambda = R^\lambda_{\ \mu\nu\alpha} A^\alpha - T^\alpha_{\mu\nu} \nabla_\alpha A^\lambda$$

$R^\lambda_{\ \mu\nu\rho}$ is a tensor of type $\binom{1}{3}$

$T^\lambda_{\mu\nu}$ is a tensor of type $\binom{1}{2}$

<u>Note</u> 1) $T^\lambda_{\mu\nu} = 0$ if and only if the connection is symmetric.

2) $R^\lambda_{\ \mu\nu\rho}$ is anti-symmetric in μ, ν

3) The formula for $[\nabla_\mu \nabla_\nu]$ acting on other tensors is similar in form to the covariant derivative.
For ex.:

$$[\nabla_\mu, \nabla_\nu] B_\pi^{\ \sigma\tau} = R^\sigma_{\ \mu\nu\alpha} B_\pi^{\ \alpha\tau} + R^\tau_{\ \mu\nu\alpha} B^{\sigma\alpha} - R^\alpha_{\ \mu\nu\pi} B_\alpha^{\ \sigma\tau} - T^\alpha_{\mu\nu} \nabla_\alpha B_\pi^{\ \sigma\tau}$$

4) The Ricci tensor is obtained by contraction:

$$R_{\mu\nu} = R^\alpha_{\ \alpha\mu\nu}$$

② The Introduction of a Metric

Def i) A $\binom{0}{2}$-tensor $g_{\mu\nu}$ is symmetric if $g_{\mu\nu} = g_{\nu\mu}$.

ii) A symmetric $\binom{0}{2}$-tensor $g_{\mu\nu}$ is non-degenerate if there exists a $\binom{2}{0}$-tensor $\tilde{g}^{\mu\nu}$ such that

called the inverse

$$g_{\mu\alpha}\tilde{g}^{\alpha\nu} = \delta_\mu^\nu$$

iii) A metric $g_{\mu\nu}$ is a symmetric non-degenerate $\binom{0}{2}$-tensor.

Example The flat metric $\eta_{\mu\nu}$ is a metric in Lorentzian geometry.

Note If X^μ and Y^μ are vector fields, then

$$g_{\mu\nu}X^\mu Y^\nu$$

is a scalar field. Thus a metric is a kind of dot product.

Raising and Lowering of Indices.

If Y^μ is a vector field then

$g_{\mu\nu} Y^\lambda$ is a $\binom{1}{2}$-tensor

Contracting gives a $\binom{0}{1}$-tensor

$$Y_\mu := g_{\mu\alpha} Y^\alpha \qquad (= g_{\alpha\mu} Y^\alpha)$$

We say we have lowered the index μ.

The reverse process, using the inverse metric $\tilde{g}^{\mu\nu}$ is called raising an index:

$$A^\mu := \tilde{g}^{\mu\alpha} A_\alpha \quad \text{for a covector } A_\mu.$$

These procedures are mutually inverse:

$$Y^\mu \underset{\text{lower}}{\rightsquigarrow} g_{\mu\alpha} Y^\alpha \underset{\text{raise}}{\rightsquigarrow} \tilde{g}^{\mu\beta} g_{\beta\alpha} Y^\alpha =$$

$$= \delta^\mu_\alpha Y^\alpha = Y^\mu$$

For a tensor $A^{\mu\nu}$, we wish to be able to distinguish

lowering 1^{st} index: $g_{\mu\alpha} A^{\alpha\nu}$ $\quad A_{\mu}^{\nu}$?

lowering 2^{nd} index: $g_{\mu\alpha} A^{\nu\alpha}$ $\quad A_{\mu}^{\nu}$?

which are both $\binom{1}{1}$-tensors, we denote them

$$A_{\mu}{}^{\nu} = g_{\mu\alpha} A^{\alpha\nu}$$

$$A^{\mu}{}_{\nu} = g_{\nu\alpha} A^{\mu\alpha}$$

Unless $A^{\mu\nu}$ is symmetric, these are different $\binom{1}{1}$-tensors.

This makes horizontal placement of indices important.

Similarly, we should avoid writing B^{μ}_{ν} and instead write $B_{\mu}{}^{\nu}$ or $B^{\nu}{}_{\mu}$ when a metric is present, as we might want to lower/raise an index.

<u>Note</u> Raising both indices in the metric we get

$$g^{\mu\nu} = \tilde{g}^{\mu\alpha} \underbrace{\tilde{g}^{\nu\beta} g_{\alpha\beta}}_{=\delta_\alpha^\nu} = \tilde{g}^{\mu\nu} \quad (!)$$

For this reason, we write $g^{\mu\nu}$ for the inverse metric.

<u>Vielbein</u>

If $g_{\mu\nu}$ is a metric, then at each $x \in M$, $g_{\mu\nu}(x)$ can be brought to diagonal form by some invertible matrix $e_a^\mu(x)$ in the sense that

$$g_{\mu\nu}(x)\, e_a^\mu(x)\, e_b^\nu(x) = \eta_{ab}$$

where $\eta_{ab} = \begin{cases} \pm 1 \, , & a=b \\ 0 \, , & a \neq b \end{cases}$ Equivalently,

$$g_{\mu\nu}(x) = \eta_{ab}\, e_\mu^a(x)\, e_\nu^b(x)$$

where e_μ^a is the inverse of e_a^μ. The field e_μ^a is the Vielbein of the metric. It transforms as a covector under $\mathcal{G}$-transformations: $e_{\mu'}^a = J_{\mu'}^\nu e_\nu^a$ and as Lorentz vector: $e_\mu^{a'} = \Lambda_b^{a'} e_\mu^b$.

Horizontal index placement leads to prefix/postfix conventions:

Prefix $T_{\nu_1 \ldots \nu_s}{}^{\mu_1 \ldots \mu_r}$ Lower indices to the left

Consistent with covariant derivative

$$A_{\mu\nu}{}^{\tau} = \nabla_\mu A_\nu{}^{\tau}$$

Connections: $\Gamma_{\mu\nu}{}^{\lambda}$

Curvature tensor: $R_{\mu\nu\alpha}{}^{\beta}$

$\Big\}$ We will follow this

Postfix $T^{\mu_1 \ldots \mu_r}{}_{\nu_1 \ldots \nu_s}$

Derivatives are now written on the right:

$$T^{\mu}{}_{\nu,\lambda} := \partial_\lambda T^{\mu}{}_{\nu}$$

$$T^{\mu}{}_{\nu|\lambda} \text{ or } T^{\mu}{}_{\nu;\lambda} := \nabla_\lambda T^{\mu}{}_{\nu}$$

Connections: $\Gamma_{\mu}{}^{\nu}{}_{\lambda}$ or $\Gamma^{\nu}{}_{\mu\lambda}$

Curvature: $R_{\alpha}{}^{\beta}{}_{\mu\nu}$ or $R^{\beta}{}_{\alpha\mu\nu}$

③ Levi-Civita Connection

Def. i) A $\mathcal{G}$-space M equipped with a metric $g_{\mu\nu}$ is a pseudo-Riemannian $\mathcal{G}$-space. If $g_{\mu\nu}$ is positive definite (i.e. its diagonal form η_{ab} has all $+1$'s on the diagonal) then M is a Riemannian $\mathcal{G}$-space.

ii) A connection $\Gamma^{\lambda}_{\mu\nu}$ on a pseudo-Riemannian $\mathcal{G}$-space $(M, g_{\mu\nu})$ is called metric (or metric-compatible) if

$$\nabla_{\lambda} g_{\mu\nu} = 0$$

where the covariant derivative is computed using $\Gamma^{\lambda}_{\mu\nu}$. Explicitly, $\Gamma^{\lambda}_{\mu\nu}$ is metric if

$$\partial_{\lambda} g_{\mu\nu} - \Gamma^{\alpha}_{\lambda\mu} g_{\alpha\nu} - \Gamma^{\alpha}_{\lambda\nu} g_{\mu\alpha} = 0$$

Theorem (The Fundamental Theorem of Riemannian geometry)

Let $(M, g_{\mu\nu})$ be a pseudo-Riemannian $\mathcal{G}$-space. ($\mathcal{G} \subset \mathcal{G}_{sm}$ any smooth geometry)

Then there exists a unique affine connection $\Gamma^{\lambda}_{\mu\nu}$ on M satisfying the two conditions:

1) $\Gamma^{\lambda}_{\mu\nu}$ is symmetric
   (equivalently, the torsion tensor $T^{\lambda}_{\mu\nu}$ vanishes)

2) $\Gamma^{\lambda}_{\mu\nu}$ is metric.

This affine connection is explicitly given by

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2} g^{\lambda\alpha} \left(\partial_\mu g_{\nu\alpha} + \partial_\nu g_{\mu\alpha} - \partial_\alpha g_{\mu\nu} \right)$$

Proof: Outlined in Homework #4.

<u>Definition</u> This affine connection is called the Levi-Civita connection on $(M, g_{\mu\nu})$. It is sometimes denoted by

$$\overset{g}{\Gamma}{}^{\lambda}_{\mu\nu} \quad \text{or} \quad \gamma^{\lambda}_{\mu\nu} \quad \text{or} \quad \left\{ {\lambda \atop \mu\,\nu} \right\}$$

and the components are called Christoffel symbols.

Lecture 5

Invariant Integration

① Invariant Integration.

Let $K \subset M$ and consider

$$\int_K \phi \, d^n x$$

To make sense of this, we choose a coordinate system $x \mapsto x^\mu$ (from the collection $\mathcal{C}$ that M comes equipped with).

We define

$$\int_K \phi \, d^n x := \int_K \phi(x^\mu) \, dx^1 dx^2 \cdots dx^n$$

where $\phi(x^\mu)$ is the real-valued function expressing ϕ in the coordinate system $x \mapsto x^\mu$, and K on the right really means $\{x^\mu : x \in K\}$.

But, we want this definition to be independent of the choice of coordinates.

If $x \mapsto x^{\mu'}$ is another coordinate system (from $\mathcal{E}$) on M, then they are related by some coordinate transformation

$$x^{\mu} \longmapsto x^{\mu'} = x^{\mu'}(x^{\nu})$$

belonging to the geometry $\mathcal{G}$ under consideration. We assume $\mathcal{J}' > 0$.

By the change-of-variable formula from integral calculus,

$$\int_K \phi(x^{\mu'}) dx^{1'} \cdots dx^{n'} = \int_K \phi(x^{\mu'}) \, \mathcal{J}' \, dx^1 dx^2 \cdots dx^n$$

Therefore, for the integral to be coordinate-independent the two expressions $\phi(x^{\mu})$ and $\phi(x^{\mu'})$ of ϕ should be related by

$$\phi(x^{\mu'}) \, \mathcal{J}' = \phi(x^{\mu})$$

or equivalently:

$$\boxed{\phi(x^{\mu'}) = \mathcal{J}_{,} \cdot \phi(x^{\mu})}$$

(Recall: $\mathcal{J}' = \det(\mathcal{J}^{\mu'}_{\nu})$, $\mathcal{J}_{,} = \det(\mathcal{J}^{\mu}_{\nu'})$).

Remark ϕ is a new object, not quite a scalar field. If ψ is another such object, then ϕ/ψ is a scalar field.
Therefore we call ϕ a relative scalar, or a scalar density.

Since ϕ^2 transforms as

$$\phi^2(x^\mu) \longmapsto \phi^2(x^{\mu'}) = (J_{,})^2 \cdot \phi^2(x^\mu)$$

we say ϕ^2 is a relative scalar of weight 2. The general case:

Definition
A relative tensor field on M of type $\binom{r}{s}$ and weight w is an n^{r+s}-tuple

$$S_{\nu\ldots}{}^{\mu\ldots}$$

of functions, given in each coordinate system x^μ, which are related by

$$S_{\nu'\ldots}{}^{\mu'\ldots} = (J_{,})^w J^{\alpha}_{\nu'} \cdots J^{\mu'}_{\beta} \cdots S_{\alpha\ldots}{}^{\beta\ldots}$$

note

Relative tensors of weight 1 are called tensor densities.

Remark

1) Relative tensors of type $\binom{r}{s}$ and weight w form a vector space.

2) Relative tensors can be multiplied; weights get added (just like the types)

3) The covariant derivative of a relative tensor field of type $\binom{r}{s}$ and weight w is

$$
\begin{aligned}
\nabla_\lambda T_{\mu\ldots}{}^{\nu\ldots} &= \partial_\lambda T_{\mu\ldots}{}^{\nu\ldots} \\
&\left.\begin{aligned} &+ \Gamma_{\lambda\alpha}{}^{\nu} T_{\mu\ldots}{}^{\alpha\ldots} + \ldots \\ &- \Gamma_{\lambda\mu}{}^{\alpha} T_{\alpha\ldots}{}^{\nu\ldots} - \ldots \end{aligned}\right\} \text{same as for usual tensors} \\
&\left.- w\, \Gamma_{\alpha\lambda}{}^{\alpha} T_{\mu\ldots}{}^{\nu\ldots}\right\} \text{a single extra term}
\end{aligned}
$$

The result has type $\binom{r}{s+1}$ and weight w.

4) Product rule works as usual
$$\nabla_\mu (T S) = \cdots$$
even for relative tensors.

<u>Example</u> If j^μ is a vector density its "covariant divergence" is:

$$\nabla_\mu j^\mu = \partial_\mu j^\mu + \cancel{\Gamma_{\mu\alpha}{}^{\mu} j^\alpha} - \cancel{\Gamma_{\beta\mu}{}^{\beta} j^\mu}$$
$$= \partial_\mu j^\mu$$

In particular, $\partial_\mu j^\mu$ is a scalar density!

We will need the following form of the Divergence Theorem:

> For any vector density j^μ vanishing on ∂K we have
>
> $$\int_K \partial_\mu j^\mu \, d^n x = 0$$

Note: We will only need to consider regions $K \subset M$ which in some coordinate system is a closed ball around some point.

② The Levi-Civita Symbols

We define, in all coordinate systems,

$$\varepsilon_{\mu_1\cdots\mu_n} = \varepsilon^{\mu_1\cdots\mu_n} = \begin{cases} +1, & \text{even perm.} \\ -1, & \text{odd perm.} \\ 0, & \text{not a perm.} \end{cases}$$

depending on whether $\mu_1\cdots\mu_n$ is an even, odd, or no permutation of $12\cdots n$.

These are the Levi-Civita symbols.

Example

$n=2$: $\varepsilon_{12}=1$, $\varepsilon_{21}=-1$ $\quad \varepsilon_{11}=\varepsilon_{22}=0$

$n=3$: $\varepsilon_{123}=\varepsilon_{231}=\varepsilon_{312}=1$,

$\varepsilon_{132}=\varepsilon_{321}=\varepsilon_{213}=-1$, $\varepsilon_{112}=0$ etc.

These are convenient for expressing determinants. For example:

$$\begin{aligned}\det(A^\mu_\nu) &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) A^1_{\sigma(1)}\cdots A^n_{\sigma(n)} \\ &= \varepsilon^{\nu_1\cdots\nu_n} A^1_{\nu_1}\cdots A^n_{\nu_n} \\ &= \frac{1}{n!}\varepsilon_{\mu_1\cdots\mu_n}\varepsilon^{\nu_1\cdots\nu_n} A^{\mu_1}_{\nu_1}\cdots A^{\mu_n}_{\nu_n}\end{aligned}$$

Equivalently,

$$\det(A^\mu_\nu)\,\varepsilon^{\mu_1\cdots\mu_n} = \varepsilon^{\nu_1\cdots\nu_n} A^{\mu_1}_{\nu_1}\cdots A^{\mu_n}_{\nu_n} \quad (*)$$

Applying (*) to the Jacobian reveals the tensorial nature of the Levi-Civita symbols:

For any coordinate change $x^\mu \to x^{\mu'}$,

$$J^{\mu_1'}_{\nu_1} \cdots J^{\mu_n'}_{\nu_n} \varepsilon^{\nu_1 \cdots \nu_n} = \det(J^{\mu'}_{\nu}) \varepsilon^{\mu_1' \cdots \mu_n'}$$
$$= J' \varepsilon^{\mu_1' \cdots \mu_n'}$$

Multiplying by J, we get

$$\varepsilon^{\mu_1' \cdots \mu_n'} = J \cdot J^{\mu_1'}_{\nu_1} \cdots J^{\mu_n'}_{\nu_n} \varepsilon^{\nu_1 \cdots \nu_n}$$

Thus, the contravariant Levi-Civita symbol

$$\varepsilon^{\mu_1 \cdots \mu_n}$$

is an $\binom{n}{0}$-tensor density.
(= relative $\binom{n}{0}$-tensor of weight 1)

Likewise, the covariant Levi-Civita symbol

$$\varepsilon_{\mu_1 \cdots \mu_n}$$

is a relative $\binom{0}{n}$-tensor of weight -1.
(Check this!)

When a metric $g_{\mu\nu}$ is present, the raised and lowered Levi-Civita symbols are

$$\varepsilon_g^{\mu_1 \cdots \mu_n} = g^{\mu_1 \nu_1} \cdots g^{\mu_n \nu_n} \varepsilon_{\nu_1 \cdots \nu_n}$$

$$\varepsilon^g_{\mu_1 \cdots \mu_n} = g_{\mu_1 \nu_1} \cdots g_{\mu_n \nu_n} \varepsilon^{\nu_1 \cdots \nu_n}$$

These are in general not the same:

$$\varepsilon^g_{\mu_1 \cdots \mu_n} = g \, \varepsilon_{\mu_1 \cdots \mu_n} , \quad \varepsilon_g^{\mu_1 \cdots \mu_n} = g^{-1} \varepsilon^{\mu_1 \cdots \mu_n}$$

where $g = \det(g_{\mu\nu})$.

For example, for the flat metric $\eta_{ab} = \mathrm{diag}(+,-,-,-)$ in Lorentzian geometry, we have

$$\varepsilon_\eta^{abcd} = \eta \cdot \varepsilon^{abcd} = -\varepsilon^{abcd}$$

[Warning: Conventions differ.]

③ Metrics and Scalar Densities.

We will show that a metric $g_{\mu\nu}$ can be used to construct scalar densities, which is what we can integrate in a coordinate-independent way.

The determinant of $g_{\mu\nu}$ is denoted g. We have

$$g = \frac{1}{n!} \underbrace{\varepsilon^{\mu_1 \dots \mu_n}}_{w=1} \underbrace{\varepsilon^{\nu_1 \dots \nu_n}}_{w=1} g_{\mu_1 \nu_1} \cdots g_{\mu_n \nu_n}.$$

Thus g is a relative scalar of weight 2:

$$g' = (J_,)^2 \, g \qquad \text{under } x^\mu \to x^{\mu'}$$

Assuming $J_, > 0$ for all coord. changes (such $\mathcal{G}$ is oriented),

$$\sqrt{|g|} \text{ is a scalar density}$$

In particular, when $n=4$ and $g_{\mu\nu}$ has signature $(1,3)$ or $(3,1)$: $\sqrt{|g|} = \sqrt{-g}$.

Thus, if ϕ is any scalar field on $(M, g_{\mu\nu})$ then

$$I = \int \phi \sqrt{|g|}\, d^n x$$

has coordinate-independent meaning.

Other examples include

i) $\int (\partial_\mu \phi \underbrace{\partial^\mu \phi}_{(\partial_\nu \phi) g^{\mu\nu}} + \phi^2 + \phi^4) \sqrt{|g|}\, d^n x$

ii) $\int R \sqrt{|g|}\, d^n x$, $\quad R = R_\mu{}^\mu = g^{\mu\nu} R_{\mu\nu} = g^{\mu\nu} R_{\alpha\mu\nu}{}^\alpha$

Scalar curvature

iii) $\int F_{\mu\nu} F^{\mu\nu} \sqrt{|g|}\, d^n x$

where $F_{\mu\nu} = \overset{g}{\nabla}_\mu A_\nu - \overset{g}{\nabla}_\nu A_\mu = \partial_\mu A_\nu - \partial_\nu A_\mu$

and $F^{\mu\nu} = g^{\mu\alpha} g^{\nu\beta} F_{\alpha\beta}$

Recipe: 1) Build a scalar 2) Mult. by $\sqrt{|g|}$
$\Rightarrow$ get a scalar density.

④ Useful formulas

(1) $$\boxed{\partial_\mu g = g \cdot g^{\alpha\beta} \partial_\mu g_{\alpha\beta}}$$

<u>Proof</u> By cofactor expansion,

$$g \delta_\lambda^\rho = g_{\lambda\alpha} C^{\alpha\rho} \qquad (*)$$

where

$$C^{\lambda\rho} = g \cdot g^{\lambda\rho}. \qquad (**)$$

Thus:

$$\partial_\mu g = \frac{\partial g}{\partial g_{\alpha\beta}} \cdot \partial_\mu g_{\alpha\beta} \quad \text{(chain rule)}$$

$$\overset{(*)}{=} C^{\alpha\beta} \partial_\mu g_{\alpha\beta} \overset{(**)}{=} g \cdot g^{\alpha\beta} \cdot \partial_\mu g_{\alpha\beta} \qquad \blacksquare$$

(2) $$\boxed{\partial_\mu \sqrt{-g} = \frac{1}{2} \sqrt{-g} \cdot g^{\alpha\beta} \partial_\mu g_{\alpha\beta}}$$

<u>Proof</u> $$\partial_\mu \sqrt{-g} = \frac{1}{2} (-g)^{-1/2} (-\partial_\mu g) =$$

$$= \frac{1}{2} (-g)^{-1/2} \cdot (-g\, g^{\alpha\beta} \partial_\mu g_{\alpha\beta})$$

$$= \frac{1}{2} \sqrt{-g} \cdot g^{\alpha\beta} \partial_\mu g_{\alpha\beta} \qquad \blacksquare$$

(3) Let $\Gamma_{\mu\nu}{}^{\lambda}$ be the Levi-Civita connection on $(M, g_{\mu\nu})$. Then:

$$\boxed{\Gamma_{\alpha\nu}{}^{\alpha} = \frac{\partial_\nu \sqrt{-g}}{\sqrt{-g}} = \partial_\nu \ln \sqrt{-g}}$$

Proof: $2\Gamma_{\mu\nu\lambda} = \partial_\mu g_{\nu\lambda} + \partial_\nu g_{\mu\lambda} - \partial_\lambda g_{\mu\nu}$

Since $1^{\text{st}} + 3^{\text{rd}} = \partial_\mu g_{\nu\lambda} - \partial_\lambda g_{\nu\mu}$ which is anti-symmetric in μ, λ, those terms vanish upon contraction against the symmetric $g^{\mu\lambda}$:

$$\Gamma_{\alpha\nu}{}^{\alpha} = g^{\mu\lambda}\Gamma_{\mu\nu\lambda} = \frac{1}{2} g^{\mu\lambda} \partial_\nu g_{\mu\lambda} \overset{(2)}{=} \frac{\partial_\nu \sqrt{-g}}{\sqrt{-g}} \quad \blacksquare$$

(4) $\boxed{\nabla_\mu \sqrt{-g} = 0}$

Proof Since $\sqrt{-g}$ is a scalar density,

$$\nabla_\mu \sqrt{-g} = \partial_\mu \sqrt{-g} - \Gamma_{\alpha\mu}{}^{\alpha}\sqrt{-g} \overset{(3)}{=} 0 \quad \blacksquare$$

(5) For any scalar field ϕ:

$$\boxed{\nabla_\mu(\phi\sqrt{-g}) = (\partial_\mu\phi)\sqrt{-g}}$$

Proof: By the product rule,

$$\nabla_\mu(\phi\sqrt{-g}) = (\nabla_\mu\phi)\sqrt{-g} + \phi(\nabla_\mu\sqrt{-g})$$

$\nabla_\mu\phi = \partial_\mu\phi$ since ϕ is a scalar field and $\nabla_\mu\sqrt{-g} = 0$ by (4). $\blacksquare$