GENERAL RELATIVITY II

ALEXANDER ZAHRER

CONTENTS

ABSTRACT

> Braided cords represent the flow of time itself. They converge and take shape. They twist, tangle, sometimes unravel, break, and then connect again. Musubi - knotting. That's time.
>
> (From "Your name")

The following is an attempt to provide a more mathematically oriented introduction to the lecture course GR 2 of the physics master at the university of Vienna. We aim to provide the same content as the original lecture, however, we will try to go about these concepts in a more purely mathematical manner first. In the beginning we will fully avoid the use of the usual physics notation and only slowly introduce it, as the said notational conventions are what makes the life of a mathematician truly hard whenever enrolled in a physics course. We also provide more details on some of the mathematical contents as in the original course, that is, we introduce more abstract nonsense. Those passages are highlighted by $(\star)$ and are not at all relevant for the material covered in the lecture course. However, for the mathematically inclined those sections might be interesting. As references we used both [1] and [2] extensively, and we sometimes even shamelessly copied some parts.

1 INTRODUCTION TO TENSOR CALCULUS AND MANIFOLDS

1.1 *Manifolds*

We will start by first introducing the objects which will lay the foundation for our theory of relativity.

Definition 1. A d-dimensional topological manifold $\mathscr{M}$ is a second countable, Hausdorff topological space, which is locally Euclidean in the sense that each point $p \in \mathscr{M}$ has an open neighborhood which is homeomorphic to an open subset of $\mathbb{R}^d$. Moreover, we introduce the following concepts:

1. A chart on a topological manifold $\mathscr{M}$ is a pair $(\mathscr{U}, \phi)$, where $\mathscr{U} \subset \mathscr{M}$ is an open subset and ϕ is a homeomorphism from $\mathscr{U}$ onto an open subset $\phi(\mathscr{U}) \subset \mathbb{R}^d$.
2. For $k \in \mathbb{N} \cup \{\infty\}$, two charts $(\mathscr{U}_\lambda, \phi_\lambda), (\mathscr{U}_\mu, \phi_\mu)$ are called $\mathscr{C}^k$-compatible if $\mathscr{U}_{\lambda\mu} := \mathscr{U}_\lambda \cap \mathscr{U}_\mu$ is either empty or $\phi_{\lambda\mu} := \phi_\lambda \circ \phi_\mu^{-1}$ is a $\mathscr{C}^k$-diffeomorphism between the open subsets $\phi_\mu(\mathscr{U}_{\lambda\mu})$ and $\phi_\lambda(\mathscr{U}_{\lambda\mu})$.
3. A collection of such pairs $\{(\mathscr{U}_\lambda, \phi_\lambda) \mid \lambda \in \Lambda\}$ of mutually $\mathscr{C}^k$-compatible charts on $\mathscr{M}$ such that $\mathscr{M} = \bigcup \mathscr{U}_\lambda$ is called a $\mathscr{C}^k$-atlas on $\mathscr{M}$.

It is quite non-trivial to prove that the dimension of a manifold $\mathscr{M}$ is well defined. In fact, the common proof for this fact relies on homology theory from algebraic topology. For our purposes it is very much sufficient to always assume $\mathscr{C}^\infty$-compatibility.

The main point of defining all this *madness* is that it allows us to talk about differentiable functions and maps, and therefore we may do analysis:

Definition 2. Let $\mathscr{M}$ and $\mathscr{N}$ be smooth manifolds and let $f\colon \mathscr{M} \to \mathscr{N}$ be a map.

- f is called smooth if and only if for any point $p \in \mathscr{M}$, there are charts $(\mathscr{U}, \phi)$ for $\mathscr{M}$ and $(\mathscr{V}, \zeta)$ for $\mathscr{N}$ such that $p \in \mathscr{U}, f(\mathscr{U}) \subset \mathscr{V}$ and such that
$$\zeta \circ f \circ \phi^{-1}\colon \phi(\mathscr{U}) \to \mathbb{R}^d \tag{1}$$
is smooth in the usual sense.
- f is called a diffeomorphism if and only if f is smooth and bijective and the inverse $f^{-1}\colon \mathscr{N} \to \mathscr{M}$ is smooth, too.

1.2 *Scalar Functions*

A scalar f on a manifold $\mathscr{M}$ is simply defined to be a (real-valued) function on $\mathscr{M}$, that is $f\colon \mathscr{M} \to \mathbb{R}$, which is assumed to be as differentiable as the differentiable structure of the manifold allows. The set of scalars, or put differently, the set of smooth real-valued functions on $\mathscr{M}$ can be denoted by $\mathscr{C}^\infty(\mathscr{M}, \mathbb{R})$. From the definition of smoothness it follows readily that $f\colon \mathscr{M} \to \mathbb{R}$ is smooth if and only if for each $p \in \mathscr{M}$ there is some chart $(\mathscr{U}, \phi)$ with $p \in \mathscr{U}$ such that $f \circ \phi^{-1}\colon \phi(\mathscr{U}) \to \mathbb{R}$ is smooth.

1.3 *Tangent vectors and Tangent spaces in a point for abstract manifolds* $(\star)$

For a fixed point $p \in \mathscr{M}$ we want to define the tangent space $T_p\mathscr{M}$ at the point p. In order to do so in the setting of abstract manifolds we need to go via the concept of germs at first.

Definition 3. Let $\mathscr{M}$ be a smooth manifold, and let $p \in \mathscr{M}$ be a point and let $\mathscr{C}_p^\infty(\mathscr{M},\mathbb{R})$ be the algebra of germs of smooth functions at p. Then we define the tangent space $T_p\mathscr{M}$ to $\mathscr{M}$ at the point p by

$$T_p\mathscr{M} := \{X_p \in \mathscr{L}(\mathscr{C}_p^\infty(\mathscr{M},\mathbb{R}),\mathbb{R}) | X_p \text{ is a derivation at } p\}. \tag{2}$$

To assume X_p be a derivation at p means that we have the Leibniz rule

$$X_p(fg) = X_p(f)g(p) + f(p)X_p(g) \tag{3}$$

More details can be found in [1].

Of course it is evident from the definition that $T_p\mathscr{M}$ is a vector space.

Theorem 1. *Let $\mathscr{M}$ and $\mathscr{N}$ be smooth manifolds and let $f\colon \mathscr{M} \to \mathscr{N}$ be a smooth map. For a point $p \in \mathscr{M}$ let $f_p^*\colon \mathscr{C}_{f(p)}^\infty(\mathbb{N},\mathbb{R}) \to \mathscr{C}_p^\infty(\mathscr{M},\mathbb{R})$ be given by $[g] \mapsto [g \circ f]$ (so precomposition with f and then taking the class of germs at p of which $g \circ f$ is a member of).*

1. *The map*

$$T_pf\colon T_p\mathscr{M} \to T_{f(p)}\mathscr{N} \qquad X_p \mapsto X_p \circ f_p^* \tag{4}$$

is a well defined linear map.

2. *If f is a diffeomorphism or the embedding of an open subset, then T_pf is a linear isomorphism for each $p \in \mathscr{M}$. In particular, if $\mathscr{M}$ has dimension d, then for each $p \in \mathscr{M}$, the vector space $T_p\mathscr{M}$ has dimension d.*

3. *If $g\colon \mathscr{N} \to \mathscr{P}$ is another smooth map between manifolds, then we get the chain rule*

$$T_p(g \circ f) = T_{f(p)}g \circ T_pf \tag{5}$$

Proving this is rather easy, so we will skip it.

1.4 *Tangent vectors in local charts* $(\star)$

We can now try to interpret tangent vectors in local charts. For a local chart $(\mathscr{U},\phi)$ for $\mathscr{M}$ and $p \in \mathscr{U}$ we know that $T_p\phi\colon T_p\mathscr{M} \to T_{\phi(p)}\phi(\mathscr{U}) \cong \mathbb{R}^d$ is a linear isomorphism. We can now define the tangent vectors $\frac{\partial}{\partial\phi^i}|_p$ by means of

$$\frac{\partial}{\partial\phi^i}|_p = T_{\phi(p)}\phi^{-1}(e_i) \tag{6}$$

where $e_i \in T_{\phi(p)}\phi(\mathscr{U})$ is the derivation $e_i(f) := \partial_i f(\phi(p))$ for $f \in \mathscr{C}^\infty_{\phi(p)}(\phi(\mathscr{U}),\mathbb{R})$. Note that the set $\{e_i\}_{i=1}^d$ forms a basis for $T_{\phi(p)}\phi(\mathscr{U})$. We shall now relate this picture for different charts $(\mathscr{U}_\lambda, \phi_\lambda), (\mathscr{U}_\mu, \phi_\mu)$. We then know that we can write

$$\phi_\lambda = \phi_{\lambda\mu} \circ \phi_\mu \tag{7}$$

on $\mathscr{U}_{\lambda\mu}$ and by taking the tangent map at p we obtain

$$T_p\phi_\lambda = T_p(\phi_{\lambda\mu} \circ \phi_\mu) = T_{\phi_\mu(p)}\phi_{\lambda\mu} \circ T_p\phi_\mu \tag{8}$$

and we can easily convince ourselves that $T_{\phi_\mu(p)}\phi_{\lambda\mu}$ is simply the ordinary derivative $D\phi_{\lambda\mu}(\phi_\mu(p))$ (if we develop the linear map $T_{\phi_\mu(p)}\phi_{\lambda\mu}$ in the basis $\{e_i\}$). Alternatively, we can rewrite the relation (8) as

$$(T_p\phi_\mu)^{-1} = (T_p\phi_\lambda)^{-1} \circ D\phi_{\lambda\mu}(\phi_\mu(p)) \tag{9}$$

and therefore

$$\frac{\partial}{\partial \phi^i_\mu}|_p = (T_p\phi_\mu)^{-1}(e_i) = (T_p\phi_\lambda)^{-1} \circ D\phi_{\lambda\mu}(\phi_\mu(p))(e_i) \tag{10}$$

$$= \partial_i\phi^j_{\lambda\mu}(\phi_\mu(p))\frac{\partial}{\partial \phi^j_\lambda}|_p \tag{11}$$

In particular, if we are given a tangent vector $X_p \in T_p\mathscr{M}$, then we can write

$$X_p = c^i_\lambda \frac{\partial}{\partial \phi^i_\lambda}|_p \tag{12}$$

for $c^i_\lambda \in \mathbb{R}$ (note that the set $\{\frac{\partial}{\partial \phi^i_\lambda}|_p\}$ forms a basis for $T_p\mathscr{M}$) and therefore

$$X_p = c^i_\lambda \partial_i \phi^j_{\mu\lambda}(\phi_\lambda(p))\frac{\partial}{\partial \phi^j_\mu}|_p \tag{13}$$

Thus

$$c^j_\mu = c^i_\lambda \partial_i \phi^j_{\mu\lambda}(\phi_\lambda(p)) \tag{14}$$

1.5 *Tangent bundle and tangent maps* ($\star$)

If we are given a manifold $\mathscr{M}$ we could consider the disjoint union

$$T\mathscr{M} := \bigcup T_p\mathscr{M} \tag{15}$$

and ask ourselves if we could turn this into a smooth manifold too. This is indeed the case and we refer to $T\mathscr{M}$ as the tangent bundle of $\mathscr{M}$.

Theorem 2. *For any smooth manifold $\mathscr{M}$ the space $T\mathscr{M}$ can be naturally made into a smooth manifold such that the projection $p\colon T\mathscr{M} \to \mathscr{M}$ is smooth. In particular, charts for $T\mathscr{M}$ are of the form $(p^{-1}(\mathscr{U}), T\phi)$ where $T\phi\colon p^{-1}(\mathscr{U}) \to \phi(\mathscr{U}) \times \mathbb{R}^d$ is given by*

$$T\phi(X_p) := (\phi(p), T_p\phi(X_p)) \tag{16}$$

Proposition 1. *For a smooth map $f\colon \mathscr{M} \to \mathscr{N}$ the tangent map $Tf\colon T\mathscr{M} \to T\mathscr{N}$ given by $Tf(X_p) := T_pf(x_p)$ is smooth. In particular, if $g\colon \mathscr{N} \to \mathscr{P}$ is another smooth map between manifolds, we have the chain rule*

$$T(g \circ f) = Tg \circ Tf \tag{17}$$

1.6 *Vector fields*

1.6.1 *A mathy perspective* ($\star$)

Usually if one goes about vector fields in a very mathematical manner, we could define vector fields as smooth maps $\xi\colon \mathscr{M} \to T\mathscr{M}$ such that $p \circ \xi = id$, where $p\colon T\mathscr{M} \to \mathscr{M}$ is the canonical projection. Note that $p \circ \xi = id$ simply means $\xi(p) \in T_p\mathscr{M}$ for all $p \in \mathscr{M}$. The set of vector fields on $\mathscr{M}$ will be denoted by $\mathscr{X}(\mathscr{M})$. With this definition in mind we can prove:

Proposition 2. *Let $\mathscr{M}$ be a smooth manifold of dimension d with tangent bundle $p\colon T\mathscr{M} \to \mathscr{M}$.*

- *For a chart $(\mathscr{U}, \phi)$ for $\mathscr{M}$ the maps $\frac{\partial}{\partial \phi^i}\colon \mathscr{U} \to T\mathscr{M}$ define local smooth vector fields on $\mathscr{U}$.*
- *For a fixed vector field $\xi \in \mathscr{X}(\mathscr{M})$ and a chart $(\mathscr{U}, \phi)$ there are smooth functions $\xi^i\colon \mathscr{U} \to \mathbb{R}$ such that $\xi|_{\mathscr{U}} = \xi^i \frac{\partial}{\partial \phi^i}$.*
- *Fix two charts $(\mathscr{U}_\lambda, \phi_\lambda), (\mathscr{U}_\mu, \phi_\mu)$, then the coordinate change for vector fields reads as*

$$\xi^i_\lambda(p) = \partial_j \phi^i_{\lambda\mu}(\phi_\mu(p))\xi^j_\mu(p) \tag{18}$$

A most important result is then the following:

Theorem 3. *Let $\mathscr{M}$ be a smooth manifold and let $\xi \in \mathscr{X}(\mathscr{M})$.*

1. *Let $(\mathscr{U}, \phi)$ be a chart for $\mathscr{M}$ and expand $\xi|_{\mathscr{U}} = \xi^i \frac{\partial}{\partial \phi^i}$ as seen before. For $f \in \mathscr{C}^\infty(\mathscr{M}, \mathbb{R})$, the local coordinate representation $\xi(f) \circ \phi^{-1}$ (where $\xi(f)(p) := \xi(p)(f)$) is given by*

$$\xi(f) \circ \phi^{-1} = (\xi^i \circ \phi^{-1})\partial_i(f \circ \phi^{-1}) \tag{19}$$

In particular, for any $f \in \mathscr{C}^\infty(\mathscr{M}, \mathbb{R})$ the function $\xi(f)$ is smooth.

2. *Moreover, ξ induces a linear map $\mathscr{C}^\infty(\mathscr{M},\mathbb{R}) \to \mathscr{C}^\infty(\mathscr{M},\mathbb{R})$, which is a derivation in the sense that $\xi(fg) = \xi(f)g + f\xi(g)$ for all $f,g \in \mathscr{C}^\infty(\mathscr{M},\mathbb{R})$.*

3. *Conversely, if $\mathscr{D}\colon \mathscr{C}^\infty(\mathscr{M},\mathbb{R}) \to \mathscr{C}^\infty(\mathscr{M},\mathbb{R})$ is a derivation, then there is a unique vector field $\xi \in \mathscr{X}(\mathscr{M})$ such that $\xi(f) = \mathscr{D}(f)$ for all $f \in \mathrm{C}^\infty(\mathscr{M},\mathbb{R})$. Put differently, we have a linear isomorphism*

$$\left\{\textit{Derivations}\ \mathscr{C}^\infty(\mathscr{M},\mathbb{R}) \to \mathscr{C}^\infty(\mathscr{M},\mathbb{R})\right\} \cong \mathscr{X}(\mathscr{M}) \tag{20}$$

1.6.2 *From the physics point of view*

The notion of a vector field finds its roots in the notion of tangents to a curve, say $\mathbb{R} \supset I \ni s \mapsto \gamma(s) \in \mathscr{M}$. If we use local coordinates, say ϕ^i, or if we would like to use more physicsy notation local coordinates x^i, then we can write $\gamma(s)$ as $(\gamma^1(s),...,\gamma^d(s)) := x \circ \gamma(s)$, the tangent to that curve at the point $\gamma(s)$ is defined as the set of numbers

$$(\dot\gamma^1(s),...,\dot\gamma^d(s)) \tag{21}$$

This physics notation is a bit irritating in my opinion, as it is easily misunderstood. However, we will stick to this notation more and more, as this is after all a physics lecture. But before using that notation too heavily, we will do it in a most precise fashion first. So consider a curve $\gamma\colon I \to \mathscr{M}$ and fix charts $(\mathscr{U}_\lambda,\phi_\lambda),(\mathscr{U}_\mu,\phi_\mu)$ such that $\gamma(I) \subset \mathscr{U}_\lambda \cap \mathscr{U}_\mu$. Now we can look at the equation

$$\phi_\lambda \circ \gamma(s) = \phi_{\lambda\mu} \circ (\phi_\mu \circ \gamma)(s) \tag{22}$$

and differentiate both sides with respect to s to obtain

$$\frac{d}{ds}(\phi_\lambda \circ \gamma)(s) = D\phi_{\lambda\mu}(\phi_\mu(\gamma(s)))\frac{d}{ds}(\phi_\mu \circ \gamma)(s) = \partial_j\phi_{\lambda\mu}(\phi_\mu(\gamma(s)))\dot\gamma^j_\mu(s) \in \mathbb{R}^d \tag{23}$$

with $\gamma^j_\mu = (\phi_\mu \circ \gamma)^j$ and we have used the summation convention for the index j. What we did in (22) is referred to as a change of coordinates and in physics notation one usually writes $x^i \to y^j(x^i)$, where $x = (x^i) = \phi_\mu$ and $y = (y^j) = \phi_\lambda$. So what we have calculated is that in these new coordinates y^j the curve γ is represented by the functions $y^j(\gamma_\mu(s)) := \phi^j_{\lambda\mu} \circ (\phi_\mu \circ \gamma)(s) = \phi^j_\lambda \circ \gamma(s)$. Thus again in the notation of physics equation (23) reads as

$$\frac{dy^j}{ds} = \frac{\partial y^j}{\partial x^i}\dot\gamma^i \tag{24}$$

where $\frac{\partial y^i}{\partial x^j} := \partial_j\phi^i_{\lambda\mu}$ and $\gamma^i = (\phi_\mu \circ \gamma)^i$. Equation (24) defines what is called the transformation law of vectors: Given a point $x = (x^i)$ and a set of numbers $X = (X^i)$, the set (X^i) is called a vector at x if, under a change of coordinates $x^i \to y^j(x^i)$ the set (X^i) transforms as

$$X^i(x) \to \overline{X}^j(y(x)) = \frac{\partial y^j}{\partial x^i}(x)X^i(x) \tag{25}$$

or equivalently

$$\overline{X}^j(y) = \frac{\partial y^j}{\partial x^i}(x(y))X^i(x(y)) \tag{26}$$

A convenient way of representing vectors is using first order homogenous differential operators. So consider the linear differential operator

$$X := X^i \frac{\partial}{\partial_i} \tag{27}$$

where i runs from 1 to d and X_i are smooth scalars which transform using the transformation rule. Now if $f \in \mathscr{C}^\infty(\mathscr{M}, \mathbb{R})$ we have

$$X(f)(x) = X^i(x)\frac{\partial f(x)}{\partial x^i} \tag{28}$$

and thus

$$X(f)(x) = X^i(x)\frac{\partial f(x)}{\partial x^i} = X^i(x)\frac{\partial \overline{f}(y(x))}{\partial x^i} = X^i(x)\frac{\partial \overline{f}(y(x))}{\partial y^k}\frac{\partial y^k}{\partial x^i}(x) \tag{29}$$

$$= \overline{X}^k(y(x))\frac{\partial \overline{f}(y(x))}{\partial y^k} = (\overline{X}^k\frac{\partial \overline{f}}{\partial y^k})(y(x)) \tag{30}$$

As we have seen in the maths section on vector fields, there is a linear isomorphism between derivations on $\mathscr{C}^\infty(\mathscr{M}, \mathbb{R})$ and the space of vector fields $\mathscr{X}(\mathscr{M})$. Thus it is perfectly fine to define a vector field by means of derivation. The *Lie bracket* of two vector fields $\xi, \eta \in \mathscr{X}(\mathscr{M})$ is defined by its action

$$[\xi, \eta](f) := \xi(\eta(f)) - \eta(\xi(f)) \tag{31}$$

It is easily checked that this yields a derivation and thus defines a vector field $[\xi, \eta] \in \mathscr{X}(\mathscr{M})$.

1.7 *Why the heck would we define vector fields as derivations?* ($\star$)

If one usually stumbles upon the term of a vector field in some physics literature, the given definition will usually be that a vector field $V: \mathbb{R}^d \to \mathbb{R}^d$ is nothing more than a smooth function, that is, $V \in \mathscr{C}^\infty(\mathbb{R}^d, \mathbb{R}^d)$. However, in our more general definition, a vector field $X \in \mathscr{X}(\mathscr{M})$ (for some manifold $\mathscr{M}$) can be interpreted as a derivation on the space $\mathscr{C}^\infty(\mathscr{M}, \mathbb{R})$, that is, a X is a linear map $\mathscr{C}^\infty(\mathscr{M}, \mathbb{R}) \to \mathscr{C}^\infty(\mathscr{M}, \mathbb{R})$ which satisfies the Leibniz rule:

$$\forall f, g \in \mathscr{C}^\infty(\mathscr{M}, \mathbb{R}) : X(fg) = X(f)g + fX(g) \tag{32}$$

So how do these two pictures of vector fields relate? In order to answer this, let $V \in \mathscr{C}^\infty(\mathbb{R}^d, \mathbb{R}^d)$ be a vector field and define the map $\widetilde{V}: \mathscr{C}^\infty(\mathbb{R}^d, \mathbb{R}) \to \mathscr{C}^\infty(\mathbb{R}^d, \mathbb{R})$ by

$$\widetilde{V}(f)(x) := Df(x)(V(x)) \tag{33}$$

for all $f \in \mathscr{C}^\infty(\mathbb{R}^d,\mathbb{R})$ and for all $x \in \mathbb{R}^d$. Certainly $\widetilde{V}$ is a linear map and of course by the usual product rule we have

$$\widetilde{V}(fg) = \widetilde{V}(f)g + f\widetilde{V}(g) \tag{34}$$

which means that $\widetilde{V}$ is a derivation. This means that every vector field V naturally induces a derivation $\widetilde{V}$. Even more is true:

Theorem 4. *The space of vector fields $\mathscr{C}^\infty(\mathbb{R}^d,\mathbb{R}^d)$ is linearly isomorphic to the space of derivations $\mathscr{D}$ on $\mathscr{C}^\infty(\mathbb{R}^d,\mathbb{R})$. In particular, the isomorphism is given by*

$$\zeta : \mathscr{C}^\infty(\mathbb{R}^d,\mathbb{R}^d) \to \mathscr{D} \qquad V \mapsto \widetilde{V} \tag{35}$$

Proof. We have already established that ζ is a well defined linear map $\mathscr{C}^\infty(\mathbb{R}^d,\mathbb{R}) \to \mathscr{D}$, so all that is left to show is injectivity and surjectivity. So let us assume that $\zeta(V) = \zeta(W)$. Then

$$\zeta(V) = \zeta(W) \iff \widetilde{V} = \widetilde{W} \iff \forall f\colon \forall x\colon Df(x)(V(x)) = Df(x)(W(x)) \tag{36}$$

$$\overset{\text{take } f=x^i}{\Longrightarrow} \forall x\colon \forall i\colon V^i(x) = W^i(x) \iff V = W \tag{37}$$

So ζ is injective. So now let $\phi \in \mathscr{D}$ be a derivation. We have to find a vector field V such that $\widetilde{V} = \phi$. First note that if $1 \in \mathscr{C}^\infty(\mathbb{R}^d,\mathbb{R})$ represents the constant 1-function, then by the Leibniz rule for derivations we have

$$\phi(1) = \phi(1\cdot 1) = \phi(1)\cdot 1 + 1\cdot\phi(1) = 2\phi(1) \tag{38}$$

So $\phi(1) = 0$. In particular, by linearity $\phi(\lambda) = 0$ for all $\lambda \in \mathbb{R}^d$. Now let $f \in \mathscr{C}^\infty(\mathbb{R}^d,\mathbb{R})$ and $x \in \mathbb{R}^d$ be arbitrary. Then by the fundamental theorem of calculus we have

$$f(y) = f(x) + \int_0^1 \frac{d}{dt}[f(x+t(y-x))]dt \tag{39}$$

$$= f(x) + \sum_i (y^i - x^i) \underbrace{\int_0^1 \partial_i f(x+t(y-x))dt}_{h_i(y)} \tag{40}$$

But then

$$\phi(f) = \underbrace{\phi(\underbrace{f(x)}_{\in\mathbb{R}})}_{=0} + \sum_i \{\underbrace{\phi(y^i - x^i)}_{=\phi(y^i)} h_i + (y^i - x^i)\phi(h_i)\} \tag{41}$$

and therefore

$$\phi(f)(x) = \sum_i \{\phi(y^i)(x) \underbrace{h_i(x)}_{=\partial_i f(x)} + (x^i - x^i)\phi(h_i)(x)\} = \sum_i \partial_i f(x)\phi(y^i)(x) \tag{42}$$

So if we then set $V^i := \phi(y^i) \in \mathscr{C}^\infty(\mathbb{R}^d,\mathbb{R})$, then the vector field V with component functions V^i satisfies by equation (42)

$$\phi = \widetilde{V} = \zeta(V) \tag{43}$$

that is, ζ is surjective and thus an isomorphism. □

1.8 *Covectors*

At any given point $p \in \mathscr{M}$ the set of covectors is defined to be the dual space of $T_p\mathscr{M}$, that is, $T_p^*(\mathscr{M})$. The collection of all cotangent spaces $T_p^*\mathscr{M}$

$$T^*\mathscr{M} := \bigcup T_p^*(\mathscr{M}) \tag{44}$$

is called the cotangent bundle of $\mathscr{M}$ and it can be shown that this yields a smooth manifold in the following way: We again have a canonical projection $p\colon T^*\mathscr{M} \to \mathscr{M}$ and charts for $T^*\mathscr{M}$ are of the form $(p^{-1}(\mathscr{U}), T^*\phi)$ with

$$T^*\phi\colon p^{-1}(\mathscr{U}) \to \phi(\mathscr{U}) \times \mathbb{R}^{d*} \qquad T^*\phi(\omega_p) := (\phi(p), \underbrace{((T_p\phi)^{-1})^*(\omega_p))}_{\text{dual map of } (T_p\phi)^{-1}} \tag{45}$$

In mathematics one usually defines covectors as smooth maps $\omega\colon \mathscr{M} \to T^*\mathscr{M}$ which satisfy $p \circ \omega = id$. However, we will not go into this further and look at covectors from the physics point of view. Here we are a little bit lazy and we will immediately define the basic objects which are the coordinate differentials dx^i, defined by its action on vector fields:

$$dx^i(X^j \frac{\partial}{\partial x^j}) := X^i \tag{46}$$

Equivalently, we could write this as

$$dx^i(\frac{\partial}{\partial x^j}) = \delta^i_j \tag{47}$$

The set $\{dx^i\}$ forms a basis for the space of (local) covectors $\Omega(\mathscr{U})$: Indeed, if $\omega \in \Omega(\mathscr{U})$, then for $X \in \mathscr{X}(\mathscr{U})$

$$\omega(X) = \omega(X^i \frac{\partial}{\partial x^i}) = X^i \underbrace{\omega(\frac{\partial}{\partial x^i})}_{:=\omega_i} = \omega_i dx^i(X) \tag{48}$$

that is

$$\omega = \omega_i dx^i \tag{49}$$

In particular, for $\omega \in \mathscr{X}(\mathscr{M})$ we have $\omega|_{\mathscr{U}} = \omega_i dx^i$. In particular, if we go for the same spiel of coordinates as before we easily see that

$$\omega = \omega_i^\lambda \frac{\partial}{\partial \phi^i_\lambda} = \omega_i^\lambda (\partial_i \phi^j_{\lambda\mu} \circ \phi_\mu) \frac{\partial}{\partial \phi^j_\mu} \tag{50}$$

that is

$$\omega_j^\mu = \omega_i^\lambda (\partial_i \phi^j_{\lambda\mu} \circ \phi_\mu) \tag{51}$$

So in the physics notation this reads as

$$\overline{\omega}_j = \omega_i \frac{\partial y^i}{\partial x^j} \tag{52}$$

Now for $f \in \mathscr{C}^\infty(\mathscr{M}, \mathbb{R}), p \in \mathscr{M}$ and $X_p \in T_p\mathscr{M}$ we may define

$$df(p)(X_p) := T_p f(X_p) \in T_{f(p)}\mathbb{R} \cong \mathbb{R} \tag{53}$$

We observe that $df(\xi) = \xi(f)$ for all $\xi \in \mathscr{X}(\mathscr{M})$, and from that one can deduce that $df \in \Omega(\mathscr{M})$. From a physicist's perspective one simply sets

$$df = \frac{\partial f}{\partial x^i} dx^i \tag{54}$$

Last but not least, since $\frac{\partial}{\partial \phi^i_\mu} = \partial_i \phi^j_{\mu\lambda} \frac{\partial}{\partial \phi^j_\lambda}$ we have $d\phi^r_\lambda(\frac{\partial}{\partial \phi^i_\mu}) = \partial_i \phi^r_{\mu\lambda}$ and thus

$$d\phi^r_\mu = \partial_i \phi^r_{\mu\lambda} d\phi^i_\lambda \tag{55}$$

or in the notation of physics

$$dy^j = \frac{\partial y^j}{\partial x^i} dx^i \tag{56}$$

1.9 *Bilinear maps and 2-covariant tensors*

Let $g \colon \mathscr{X}(\mathscr{M}) \times \mathscr{X}(\mathscr{M}) \to \mathbb{R}$ be bilinear. Then

$$g(X,Y) = g(X^i\partial_i, Y^j\partial_j) = X^iY^j g(\partial_i, \partial_j) = g_{ij} dx^i(X) dx^j(Y) \tag{57}$$

where $\partial_i := \frac{\partial}{\partial x^i}$ for a chart $(\mathscr{U}, x)$. We then say that g is a covariant tensor of valence two. A symmetric bilinear tensor field is said to be nondegenerate if $\det(g_{ij})$ has no zeros. Equivalently,

$$\forall Y \colon g(X,Y) = 0 \implies X = 0 \tag{58}$$

By Sylvester's inertia theorem, if g is a symmetric bilinear form, there exists a basis of covectors $\{\varphi^j\}$ such that

$$\forall X, Y \colon g(X,Y) = -\sum_{j=1}^{s} \varphi^j(X)\varphi^j(Y) + \sum_{j=s+1}^{r+s} \varphi^j(X)\varphi^j(Y) \tag{59}$$

The pair (s, r) is called the signature of g. If $r = d$, in dimension d, then g is said to be a Riemannian metric tensor. Thus, a Riemannian metric on a manifold $\mathscr{M}$ is a field of symmetric nondegenerate bilinear forms with signature $(0, \dim \mathscr{M})$.

1.10 *Tensor products*

1.10.1 *A mathy perspective* ($\star$)

Definition 4. An $\binom{l}{k}$-tensor field on a smooth manifold $\mathscr{M}$ is a map T that associates to each $p \in \mathscr{M}$ and element $T_p \in \bigotimes^k T_p^*\mathscr{M} \otimes \bigotimes^l T_p\mathscr{M}$ such that for all $\xi_1, ..., \xi_k \in \mathscr{X}(\mathscr{M})$ and for all $\omega^1, ..., \omega^l \in \Omega(\mathscr{M})$ the function $T(\xi_1, ..., \xi_k, \omega^1, ..., \omega_l) \colon \mathscr{M} \to \mathbb{R}$ is smooth. Note that we can interpret the space $\bigotimes^k T_p^*\mathscr{M} \otimes \bigotimes^l T_p\mathscr{M}$ as the space of $(k+l)$-linear maps $\underbrace{T_p\mathscr{M} \times ... \times T_p\mathscr{M}}_{k \text{ times}} \times \underbrace{T_p^*\mathscr{M} \times ... \times T_p^*\mathscr{M}}_{l \text{ times}} \to \mathbb{R}$. The space of all such tensor fields will be denoted by $\mathscr{T}_k^l(\mathscr{M})$.

Given $T \in \mathscr{T}_k^l(\mathscr{M})$ and a chart $(\mathscr{U}, \phi)$ fix some $(k+l)$-tuple $(i_1, ..., i_k, j_1, ..., j_l)$ of integers in $\{1, ..., n\}$, then we get a smooth function

$$T_{i_1 \dots i_k}^{j_1 \cdots j_l} := T\left(\frac{\partial}{\partial \phi^{i_1}}, ..., \frac{\partial}{\partial \phi^{i_k}}, d\phi^{j_1}, ..., d\phi^{j_l}\right) \in \mathscr{C}^\infty(\mathscr{U}, \mathbb{R}) \tag{60}$$

On the other hand, we obtain a tensor field

$$d\phi^{i_1} \otimes ... \otimes d\phi^{i_k} \otimes \frac{\partial}{\partial \phi^{j_1}} \otimes ... \otimes \frac{\partial}{\partial \phi^{j_l}} \in \mathscr{T}_k^l(\mathscr{U}) \tag{61}$$

given by

$$d\phi^{i_1} \otimes ... \otimes d\phi^{i_k} \otimes \frac{\partial}{\partial \phi^{j_1}} \otimes ... \otimes \frac{\partial}{\partial \phi^{j_l}} \left(\frac{\partial}{\partial \phi^{r_1}}, ..., \frac{\partial}{\partial \phi^{r_k}}, d\phi^{s_1}, ..., d\phi^{s_l}\right) := \delta_{r_1 \dots r_k}^{i_1 \dots i_k} \delta_{j_1 \dots j_l}^{s_1 \dots s_l} \tag{62}$$

By construction we then have

$$T|_{\mathscr{U}} = T_{i_1 \dots i_k}^{j_1 \cdots j_l} d\phi^{i_1} \otimes ... \otimes d\phi^{i_k} \otimes \frac{\partial}{\partial \phi^{j_1}} \otimes ... \otimes \frac{\partial}{\partial \phi^{j_l}} \tag{63}$$

We could then go on about changing coordinates, or that we could prove that $\bigotimes^k T\mathscr{M} \otimes \bigotimes^l T^*\mathscr{M}$ can also be turned into a smooth manifold. For details on this we refer to [1]. Something which is also quite interesting is the following:

Theorem 5. *Let* $\Phi \colon \mathscr{X}(\mathscr{M})^k \times \Omega(\mathscr{M})^l \to \mathscr{C}^\infty(\mathscr{M}, \mathbb{R})$ *be* $(k+l)$*-linear. Then* Φ *is induced by a tensor field* $T \in \mathscr{T}_k^l(\mathscr{M})$ *if and only if* Φ *is linear over* $\mathscr{C}^\infty(\mathscr{M}, \mathbb{R})$ *in each variable.*

1.10.2 *From the perspective of physics*

If we are given covector fields $\varphi, \theta \in \Omega(\mathscr{M})$, then we can clearly define a bilinear map using the formula

$$(\varphi \otimes \theta)(X, Y) := \varphi(X)\theta(Y) \tag{64}$$

Note that this definition is somehow fueled or motivated by theorem 5. For example

$$(dx^i \otimes dx^j)(X,Y) = X^i Y^j \tag{65}$$

Using this notation we have that if g is a bilinear form on $\mathscr{X}(\mathscr{M})$, then

$$g = g_{ij} dx^i \otimes dx^j \tag{66}$$

with $g_{ij} = g(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j})$. Next we will define the symmetric product

$$dx^i dx^j := \frac{1}{2}(dx^i \otimes dx^j + dx^j \otimes dx^i) \tag{67}$$

and we immediately note that if the bilinear form g is symmetric, then

$$g = g_{ij} dx^i dx^j \tag{68}$$

This formula allows one to read-off, without even having to think, the transformation law of a metric tensor under coordinate changes:

$$g_{ij}(x) \underbrace{dx^i}_{\frac{\partial x^i}{\partial y^k} dy^k} \underbrace{dx^j}_{\frac{\partial x^j}{\partial y^l} dy^l} = \underbrace{g_{ij}(x) \frac{\partial x^i}{\partial y^k} \frac{\partial x^j}{\partial y^l}}_{\bar{g}_{ij}} dy^k dy^l \tag{69}$$

A most important notion of duality is that we may apply vector fields to covector fields. Indeed, if $\varphi \in \Omega(\mathscr{M})$ is a covector field and $X \in \mathscr{X}(\mathscr{M})$ is some fixed vector field, then we may define

$$X(\varphi) := \varphi(X) \in \mathscr{C}^\infty(\mathscr{M}, \mathbb{R}) \tag{70}$$

In particular, if we write $X = X^j \frac{\partial}{\partial x^j}$ and $\varphi = \varphi_i dx^i$ for some chart $(\mathscr{U}, x)$, then (locally) we have

$$X(\varphi) = \varphi(X) = \varphi_i dx^i (X^j \frac{\partial}{\partial x^j}) = \varphi_i X^i \tag{71}$$

We will now stick to the convention $\partial_i := \frac{\partial}{\partial x^i}$ for a local chart $(\mathscr{U}, x)$. It then also makes sense to define $\partial_i \otimes \partial_j$ as the clearly bilinear map

$$\partial_i \otimes \partial_j (\varphi, \theta) := \varphi_i \theta_j \tag{72}$$

Putting everything together, we may define general tensor fields of valence (k,l) by first fixing a $(k+l)$-tuple $(i_1, ..., i_k, j_1, ..., j_l)$ of natural numbers in $\{1, ..., d\}$ (where d is the dimension of our manifold $\mathscr{M}$) and then define local basis tensor fields by

$$(dx^{i_1} \otimes ... \otimes dx^{i_k} \otimes \partial_{j_1} \otimes ... \otimes \partial_{j_l})(X_1, ..., X_k, \varphi^1, ..., \varphi^l) := \prod_r X_r^{i_r} \prod_s \varphi^s_{j_s} \tag{73}$$

where $X_1, ..., X_k \in \mathscr{X}(\mathscr{M}), \varphi^1, ..., \varphi^l \in \Omega(\mathscr{M})$ are written as $X_r = X_r^l \partial_l, \varphi^s = \varphi^s_t dx^t$ in local coordinates. A general tensor field T of valence (k,l) is then given by (locally)

$$T = \underbrace{T^{j_1, ..., j_l}_{i_1, ..., i_k}}_{\in \mathscr{C}^\infty(\mathscr{U}, \mathbb{R})} (dx^{i_1} \otimes ... \otimes dx^{i_k} \otimes \partial_{j_1} \otimes ... \otimes \partial_{j_l}) \tag{74}$$

A coordinate change for such tensor would then read as

$$T = T^{j_1...j_l}_{i_1...i_k}\big(\underbrace{dx^{i_1}}_{\frac{\partial x^{i_1}}{\partial y^{s_{i_1}}}dy^{s_{i_1}}} \otimes ... \otimes dx^{i_k} \otimes \underbrace{\partial_{x_{j_1}}}_{\frac{\partial y^{m_{j_1}}}{\partial x^{j_1}}\partial_{y_{m_{j_1}}}} \otimes ... \otimes \partial_{x_{j_l}}\big) \tag{75}$$

$$= \underbrace{T^{j_1...j_l}_{i_1...i_k}\prod_v \frac{\partial x^{i_v}}{\partial y^{s_{i_v}}}\prod_u \frac{\partial y^{m_{j_u}}}{\partial x^{j_u}}}_{T^{m_{j_1}...m_{j_l}}_{s_{i_1}...s_{i_k}}} dy^{s_{i_1}} \otimes ... \otimes dy^{s_{i_k}} \otimes \partial_{y_{m_{j_1}}} \otimes ... \otimes \partial_{y_{m_{j_l}}} \tag{76}$$

If we are given two tensors T and S then we are certainly also able to define a tensor field $T \otimes S$ in the obvious way, that is,

$$(T \otimes S)(\psi,...,\varphi,...) := T(\psi,...)S(\varphi,...) \tag{77}$$

1.11 *Contractions*

The simplest example of a contraction applies to tensor fields , say $S^j_i dx^i \otimes \partial_j$, with one index up and one index down. We then simply perform the sum S^i_i which is of course a scalar since $\bar{S}^l_k dy^k \otimes \partial_{y^l} = \underbrace{\bar{S}^l_k \frac{\partial y^k}{\partial x^s}\frac{\partial x^r}{\partial y^l}}_{S^s_r} dx^s \otimes \partial_{x^r}$ and therefore $S^r_r = \bar{S}^l_k \underbrace{\frac{\partial y^k}{\partial x^r}\frac{\partial x^r}{\partial y^l}}_{\delta^k_l} = \bar{S}^r_r$.

More generally, we may define the map C^s_r as a map which takes tensor fields $T = T^{j_1...j_l}_{i_1...i_k}\left(dx^{i_1} \otimes ... \otimes dx^{i_k} \otimes \partial_{j_1} \otimes ... \otimes \partial_{j_l}\right)$ and gives a new resulting tensor field given by

$$C^s_r(T) = T^{j_1...r...j_l}_{i_1...s...i_k} dx^{i_1} \otimes ... \otimes \widehat{dx^r} \otimes ... \otimes dx^{i_k} \otimes \partial_{j_1} \otimes ... \otimes \widehat{\partial_s} \otimes ... \otimes \partial_{j_l} \tag{78}$$

where putting a hat over a symbol means omission. Thus C^s_r takes the sth factor in the vector field part, inserts it into the rth factor of the covector field part and multiplies the resulting function with the tensor product of the remaining elements (in the original order).

1.12 *Raising and Lowering Indices*

Let g be a symmetric two-covariant tensor field, that is, for each $p \in \mathscr{M}$ $g_p : T_p\mathscr{M} \times T_p\mathscr{M} \to \mathbb{R}$ is bilinear and symmetric. The symbol g will be reserved to nondegenerate symmetric two-covariant tensor fields. We will sometimes write g_p for $g(p)$ when referencing p. Sylvester's inertia theorem tells us that at each $p \in \mathscr{M}$ the map g will have a well-defined signature; clearly this signature will be point-independent on a connected manifold when g is nondegenerate. A pair $(\mathscr{M}, g)$ is said to be a Riemannian manifold when the signature of g is $(0, \dim \mathscr{M})$; equivalently, when g is a positive definite bilinear form on every product $T_p\mathscr{M} \times T_p\mathscr{M}$. A pair $(\mathscr{M}, g)$ is said to be a Lorentzian manifold when the signature of g is $(1, \dim \mathscr{M} - 1)$. One talks about pseudo-Riemannian metrics and manifolds whatever the signature of g, as long as g is nondegenerate. Any pseudo-Riemannian metric g defines an isomorphism

$$\mathfrak{g} : T_p\mathscr{M} \to T^*_p\mathscr{M} \qquad X \mapsto g(X,.) \tag{79}$$

which in local coordinates reads as

$$g(X,.) = \underbrace{g_{ij}X^i}_{X_j} dx^j \tag{80}$$

This last equality defines X_j, the vector X^j with the index j lowered. The inverse of this operation will be denoted by # and is called raising of indices. In order to get to # we first take some $\varphi \in T_p^*\mathscr{M}$, then since $\mathfrak{g}$ is an isomorphism there exists $X \in T_p\mathscr{M}$ such that $\mathfrak{g}(X) = \varphi$. Thus

$$\varphi^\# = (\mathfrak{g}(X))^\# = (g_{ij}X^i dx^j)^\# = g_{ij}X^i dx^{j\#} \tag{81}$$

Hence, in order to recover X from φ we may simply define $dx^{j\#} := g^{jl}\partial_j$, which defines # locally (where g^{jl} is the inverse to g_{jl}). We may then also define a scalar product $g^\#$ on $T_p^*\mathscr{M}$ by

$$g^\#(\varphi, v) := g(\varphi^\#, v^\#) \tag{82}$$

In other words, $g^\#(dx^i, dx^j) = g^{ij}$. Now the gradient ∇f of a function f is a vector field obtained by raising the indices on the differential df, that is, ∇f locally looks like

$$\nabla f := df^\# = \left(\frac{\partial f}{\partial x^i} dx^i\right)^\# = \frac{\partial f}{\partial x^i} g^{il}\partial_l \tag{83}$$

1.12.1 *The musical isomorphism* $(\star)$

In any Hilbert space V with scalar product g we have an identification of vectors in V with covectors in V^* via

$$V \ni v \mapsto \langle . \mid v \rangle \in V^* \tag{84}$$

This construction extends to Semi-Riemannian manifolds providing an identification of vector fields and one forms.

Theorem 6 (Musical Isomorphism). *Let $(\mathscr{M}, g)$ be a semi-Riemannian manifold. For any $X \in \mathscr{X}(\mathscr{M})$ define $X^\wp \in \Omega^1(\mathscr{M})$ by*

$$X^\wp := g(X,.) \tag{85}$$

Then the mapping $\Psi : X \mapsto X^\wp$ is a $\mathscr{C}^\infty(\mathscr{M}, \mathbb{R})$-linear isomorphism $\mathscr{X}(\mathscr{M}) \to \Omega^1(\mathscr{M})$.

Proof. See [3] □

Hence in semi-Riemannian geometry we can always identify vectors and vector fields with covectors and one forms, respectively: X and $X^\wp$ contain the same information. One also writes $\varphi^\# = \Psi^{-1}(\varphi)$ and this notation is the source of the name 'musical isomorphism'.

1.13 *Covariant derivatives*

When dealing with $\mathbb{R}^d$ we have at our disposal the canonical trivialization $\{\partial_i\}$ of $T\mathbb{R}^d$ (a globally defined set of vector fields which, at every point, form a basis of the tangent space), together with its dual trivialization $\{dx^i\}$ of $T^*\mathbb{R}^d$. We may expand a tensor field T of valence (k,l) in terms of those bases

$$T = T^{i_1 \dots i_k}_{j_1 \dots j_l} \partial_{i_1} \otimes \dots \otimes \partial_{i_k} \otimes dx^{j_1} \otimes dx^{j_l} \tag{86}$$

If we then differentiate each component $T^{i_1 \dots i_k}_{j_1 \dots j_l}$ separately we get

$$X(T) := X^i \partial_{x^i}(T^{i_1 \dots i_k}_{j_1 \dots j_l}) \partial_{i_1} \otimes \dots \otimes \partial_{i_k} \otimes dx^{j_1} \otimes dx^{j_l} \tag{87}$$

The resulting object does, however, not behave as a tensor under coordinate transformations. Therefore we must think of a differentiate notion of differentiation which must be a map which to a tensor field T and a vector field X assigns a tensor field of the same type as T, denoted by $\nabla_X T$, with the following properties:

1. $\nabla_X T$ is linear concerning addition with respect to both X and T:

$$\nabla_{X+Y} T = \nabla_X T + \nabla_Y T \qquad \nabla_X (T+S) = \nabla_X T + \nabla_X S \tag{88}$$

2. $\nabla_X T$ is linear with respect to multiplication of X by functions $f \in \mathscr{C}^\infty(\mathscr{M}, \mathbb{R})$:

$$\nabla_{fX} T = f \nabla_X T \tag{89}$$

3. $\nabla_X T$ satisfies the Leibniz rule:

$$\nabla_X (fT) = f \nabla_X T + X(f) T \tag{90}$$

   for smooth $f \in \mathscr{C}^\infty(\mathscr{M}, \mathbb{R})$.

1.13.1 *Covariant derivative on functions*

The canonical covariant derivative on functions $f \in \mathscr{C}^\infty(\mathscr{M}, \mathbb{R})$ is defined by

$$\nabla_X (f) = X(f) \tag{91}$$

for a vector field $X \in \mathscr{X}(\mathscr{M})$. It is clear that this map satisfies all the properties of a covariant derivative.

1.13.2 *Covariant derivative on vector fields*

Let us, at first, not worry about the existence of covariant derivatives for vector fields and simply work from there. Next we assume that we are working on a subset $\mathscr{U} \subset \mathscr{M}$ over which we have a trivialization of the tangent bundle $T\mathscr{U} \subset T\mathscr{M}$. In other words, for $1 \leq a \leq d$ there exist vector fields $e_a \in \mathscr{X}(\mathscr{M})$ such that at every point $p \in \mathscr{M}$ the fields $e_a|_p \in T_p\mathscr{U} = T_p\mathscr{M}$ form a basis of $T_p\mathscr{M}$. Of course such trivializations do exist in general, say on charts $(\mathscr{U},\phi)$ with $e_l = \frac{\partial}{\partial \phi^l}$. However, trivializations do not exist globally in general. Let θ^a denote the dual trivializations of $T^*\mathscr{M}$, that is,

$$\theta^a(e_b) = \delta^a_b \tag{92}$$

Given a covariant derivative ∇ on vector fields we set

$$\Gamma^c_{ab} := \theta^c(\nabla_{e_b} e_a) \iff \nabla_{e_b} e_a = \Gamma^c_{ab} e_c \tag{93}$$

The locally defined functions Γ^c_{bc} are called connection coefficients. If $\{e_a\}$ is the coordinate basis $\{\partial_\mu\}$ we shall write

$$\Gamma^\mu_{\varphi\nu} = dx^\mu(\nabla_{\partial_\nu}\partial_\varphi) \tag{94}$$

In this case the connection coefficients will be called the Christoffel symbols. Given vector fields X,Y and using the Leibniz rule we find

$$\nabla_X Y = \nabla_X(Y^a e_a) = \nabla_X(Y^a)e_a + Y^a\nabla_X e_a = X(Y^a)e_a + Y^a X^b \nabla_{e_b} e_a \tag{95}$$

$$= X(Y^c)e_c + Y^a X^b \Gamma^c_{ab} e_c = \left[X(Y^c) + Y^a X^b \Gamma^c_{ab}\right] e_c \tag{96}$$

We will often make use of the notation $\nabla_a := \nabla_{e_a}$. The one-covariant, one-contravariant tensor field ∇Y is defined as

$$\nabla Y := \underbrace{(\nabla_a Y^b)}_{\theta^b(\nabla_{e_a} Y)} \theta^a \otimes e_b \tag{97}$$

Hence

$$\nabla_a Y^b = \theta^b(\nabla_a Y) = \theta^b(\nabla_a(Y^c e_c)) = \theta^b(e_a(Y^c)e_c + Y^c \nabla_a e_c) \tag{98}$$

$$= e_a(Y^b) + Y^c \theta^b(\Gamma^d_{ca} e_d) = e_a(Y^b) + Y^c \Gamma^b_{ca} \tag{99}$$

In particular, we note that

$$(\nabla Y)(X,.) = \nabla_X Y \tag{100}$$

1.13.3 *Transformation Law*

We shall now enquire about the transformation law of the connection coefficients Γ^i_{jk} with respect to a coordinate basis ∂_{x^i}. Let us denote by $\widehat{\Gamma}^i_{jk}$ the connection coefficients with respect to some other coordinates y.

$$\Gamma^i_{jk} = dx^i(\nabla_{\partial_{x^k}}\partial_{x^j}) = dx^i(\nabla_{\partial_{x^k}} \frac{\partial y^l}{\partial x^j}\frac{\partial}{\partial y^l}) = \frac{\partial x^i}{\partial y^s} dy^s(\nabla_{\partial_{x^k}} \frac{\partial y^l}{\partial x^j}\frac{\partial}{\partial y^l})$$

$$= \frac{\partial x^i}{\partial y^s} dy^s\left(\frac{\partial^2 y^l}{\partial x^j \partial x^k}\frac{\partial}{\partial y^l} + \frac{\partial y^l}{\partial x^j}\nabla_{\frac{\partial}{\partial x^k}}\frac{\partial}{\partial y^l}\right) = \frac{\partial x^i}{\partial y^s} dy^s\left(\frac{\partial^2 y^l}{\partial x^j \partial x^k}\frac{\partial}{\partial y^l} + \frac{\partial y^l}{\partial x^j}\frac{\partial y^r}{\partial x^k}\nabla_{\frac{\partial}{\partial y^r}}\frac{\partial}{\partial y^l}\right)$$

$$= \frac{\partial x^i}{\partial y^s} dy^s\left(\frac{\partial^2 y^l}{\partial x^j \partial x^k}\frac{\partial}{\partial y^l} + \frac{\partial y^l}{\partial x^j}\frac{\partial y^r}{\partial x^k}\widehat{\Gamma}^c_{lr}\frac{\partial}{\partial y^c}\right)$$

$$= \underbrace{\frac{\partial x^i}{\partial y^s}\frac{\partial^2 y^s}{\partial x^j \partial x^k}}_{\text{nonhomogenous term}} + \frac{\partial x^i}{\partial y^s}\frac{\partial y^l}{\partial x^j}\frac{\partial y^r}{\partial x^k}\widehat{\Gamma}^s_{lr}$$

Thus, the Γ^i_{jk}'s do not form a tensor; instead the transform as a tensor plus a nonhomogenous term containing second derivatives.

1.13.4 *Torsion*

Because the inhomogenous term in the previous transformation formula for the connection coefficients Γ^i_{jk} is symmetric in k and j, it immediately follows that

$$T^i_{jk} := \Gamma^i_{kj} - \Gamma^i_{jk} \tag{101}$$

does transform as a tensor, called the torsion tensor of ∇. There is also an index-free definition of torsion, which proceeds as follows: Let ∇ be some covariant vector field defined for vector fields, the torsion tensor T is defined by the formula

$$T(X,Y) := \nabla_X Y - \nabla_Y X - [X,Y] \tag{102}$$

We obviously have $T(X,Y) = -T(Y,X)$. Let us check that this defines a tensor field. Multilinearity with respect to addition is obvious. To check what happens under multiplication by functions, in view of the antisymmetry of T, it is sufficient to do the calculation for the first slot of T. We then have

$$T(fX,Y) = \nabla_{fX} Y - \nabla_Y fX - [fX,Y] = f(\nabla_X Y - \nabla_Y X) - Y(f)\nabla_Y X - [fX,Y] \tag{103}$$

In order to work out the last term we compute, for any smooth $\varphi \in \mathscr{C}^\infty(\mathscr{M},\mathbb{R})$

$$[fX,Y](\varphi) = fX(Y(\varphi)) - \underbrace{Y(fX(\varphi))}_{Y(f)X(\varphi)+fY(X(\varphi))} = \Big(f[X,Y] - Y(f)X\Big)(\varphi) \tag{104}$$

which shows $T(fX,Y) = fT(X,Y)$. In a coordinate basis $\{\partial_\mu\}$ we have $[\partial_\mu,\partial_\nu] = 0$ for all μ,ν and thus one finds

$$T(\partial_\mu,\partial_\nu) = \nabla_{\partial_\mu}\partial_\nu - \nabla_{\partial_\nu}\partial_\mu = (\Gamma^\sigma_{\nu\mu} - \Gamma^\sigma_{\mu\nu})\partial_\sigma \tag{105}$$

which as we have seen transforms as a tensor and agrees with our previous definition.

1.13.5 *Covariant derivative on Covector fields*

Suppose we are given a covariant derivative ∇ on vector fields, then there is a natural way of inducing a covariant derivative on one-forms by imposing the condition that the duality operation be compatible with the Leibniz rule: Given two vector fields X and Y together with a field of one-forms φ, one sets

$$(\nabla_X \varphi)(Y) := X(\varphi(Y)) - \varphi(\nabla_X Y) \tag{106}$$

Let us now check that this indeed defines a field of one-forms. The linearity in the variable Y with respect to addition is clear. Next for any smooth function $f \in \mathscr{C}^\infty(\mathscr{M},\mathbb{R})$ we have

$$\nabla_X \varphi(fY) = X(\varphi(fY)) - \varphi(\nabla_X(fY)) \tag{107}$$

$$= X(f)\varphi(Y) + fX(\varphi(Y)) - \varphi\big(X(f)Y + f\nabla_X Y\big) = f(\nabla_X\varphi)(Y) \tag{108}$$

Next we need to check that this definition actually yields a covariant derivative on covectors. Again multilinearity with respect to addition is obvious, as well as linearity with respect to multiplication of X by a function. Finally,

$$\nabla_X(f\varphi)(Y) = X(f\varphi(Y)) - f\varphi(\nabla_X Y) = X(f)\varphi(Y) + f\nabla_X\varphi(Y) \tag{109}$$

So the Leibniz rule is satisfied. The duality pairing

$$T_p^*\mathcal{M} \times T_p\mathcal{M} \to \mathbb{R} \qquad (\varphi, X) \mapsto \varphi(X) \tag{110}$$

is a special case of the contraction operation. The operation ∇ on one-forms has been defined so as to satisfy the Leibniz rule under duality pairing:

$$X(\varphi(Y)) = \nabla_X\varphi(Y) + \varphi(\nabla_X Y) \tag{111}$$

Next we observe that

$$\nabla_X\varphi = X^a(\nabla_a\varphi_b)\theta^b \tag{112}$$

with

$$\nabla_a\varphi_b := (\nabla_{e_a}\varphi)(e_b) = e_a(\varphi(e_b)) - \varphi(\nabla_{e_a}e_b) = e_a(\varphi_b) - \Gamma^c_{ba}\varphi_c \tag{113}$$

1.13.6 *Higher order Tensors*

We can now extend ∇ to tensors of arbitrary valence: If T is r covariant and s contravariant one sets

$$(\nabla_X T)(X_1,...,X_r,\varphi^1,...,\varphi^s) := X(T(X_1,...,X_r,\varphi^1,...,\varphi^s)) \tag{114}$$

$$-T(\nabla_X X_1,X_2,...,X_r,\varphi^1,...,\varphi^s) - ... - T(X_1,X_2,...,\nabla_X X_r,\varphi^1,...,\varphi^s) \tag{115}$$

$$-T(X_1,X_2,...,X_r,\nabla_X\varphi^1,...,\varphi^s) - ... - T(X_1,X_2,...,X_r,\varphi^1,...,\nabla_X\varphi^s) \tag{116}$$

In order to verify that this indeed yields a covariant derivative on tensors of valence (r,s) we solely need to verify that $\nabla_X T$ is a tensor of valence (r,s) and that ∇_X satisfies the Leibniz rule, as the other properties are obvious. However, that $\nabla_X T$ is a tensor of the same valence as T is clear, for $\nabla_X T$ is linear over $\mathscr{C}^\infty(\mathcal{M},\mathbb{R})$ in every coordinate. An easy calculation also shows that

$$\nabla_X(fT) = X(f)\nabla_X T + f\nabla_X T \tag{117}$$

Locally, we may write

$$\nabla_X T = X^a\nabla_a T^{b_1...b_s}_{a_1...a_r}\theta^{a_1}\otimes...\otimes\theta^{a_r}\otimes e_{b_1}\otimes...\otimes e_{b_s} \tag{118}$$

with

$$\nabla_a T^{b_1...b_s}_{a_1...a_r} := (\nabla_a T)(e_{a_1},...e_{a_r},\theta^{b_1},...,\theta^{b_s}) \tag{119}$$

$$= e_a(T^{b_1...b_s}_{a_1...a_r}) - \Gamma^c_{a_1a}T^{b_1...b_s}_{ca_2...a_r} \tag{120}$$

$$-... - \Gamma^c_{a_ra}T^{b_1...b_s}_{a_1...a_{r-1}c} - \Gamma^{b_1}_{ca}T^{cb_2...b_s}_{a_1...a_r} - ... - \Gamma^{b_s}_{ca}T^{b_1...b_{s-1}c}_{a_1...a_r} \tag{121}$$

We may also define

$$\nabla T := (\nabla_a T^{b_1...b_s}_{a_1...a_r})\theta^a\otimes\theta^{a_1}\otimes...\otimes\theta^{a_r}\otimes e_{b_1}\otimes...\otimes e_{b_s} \tag{122}$$

and clearly then $(\nabla T)(X) = \nabla_X T$.

2 EXAM QUESTIONS:

2.1 *Lie bracket, Jacobi identity, Levi-Civita connection, Riemann curvature tensor and its properties*

2.1.1 *Lie Bracket and Jacobi identity*

We recall that if we are given a manifold $\mathscr{M}$ the set of vector fields $\mathscr{X}(\mathscr{M})$ is linearly isomorphic to the space of derivations on $\mathscr{C}^\infty(\mathscr{M},\mathbb{R})$. Thus for fixed vector fields $\xi,\zeta \in \mathscr{X}(\mathscr{M})$ we may consider the map $[\xi,\zeta]: \mathscr{C}^\infty(\mathscr{M},\mathbb{R}) \to \mathscr{C}^\infty(\mathscr{M},\mathbb{R})$ given by

$$f \mapsto \xi(\zeta(f)) - \zeta(\xi(f)) \tag{123}$$

An easy calculation immediately shows that this linear map indeed is a derivation on $\mathscr{C}^\infty(\mathscr{M},\mathbb{R})$, that is,

$$[\xi,\zeta](fg) = [\xi,\zeta](f)g + f[\xi,\zeta](g) \tag{124}$$

Using the before mentioned isomorphism we infer that $[\xi,\zeta] \in \mathscr{X}(\mathscr{M})$ can be interpreted as a vector field on $\mathscr{M}$. This vector field is called the Lie bracket of ξ,ζ. The Lie bracket $[,]$ itself can be understood as an antisymmetric, bilinear map

$$[,]: \mathscr{X}(\mathscr{M}) \times \mathscr{X}(\mathscr{M}) \to \mathscr{X}(\mathscr{M}) \tag{125}$$

The Lie bracket satisfies the Jacobi identity: For all $\xi,\zeta,\nu \in \mathscr{X}(\mathscr{M})$ we have

$$[[\xi,\zeta],\nu] = [\xi,[\zeta,\nu]] - [\zeta,[\xi,\nu]] \tag{126}$$

Proving this identity is also rather easy, since it poses only computational obstacles.

2.1.2 *Levi-Civita Connection*

We have seen that we can build a covariant derivative on tensors of any valence from only having a covariant derivative on vector fields. One of the fundamental results in pseudo Riemannian geometry is the existence of a torsion-free connection which preserves the metric. This is called the Levi Civita connection and it is of utmost importance.

Theorem 7 (Levi Civita Connection). *Let g be a two-covariant symmetric nondegenerate tensor field on a manifold $\mathscr{M}$. Then there exists a unique connection ∇ such that*

1. $\nabla g = 0$
2. *The torsion tensor $T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]$ of ∇ vanishes.*

Proof. We first show that if such a Levi Civita connection exists, it must be unique. To see this, we first observe that if $\nabla g = 0$, then $\nabla_X g = 0$ for all $X \in \mathscr{X}(\mathscr{M})$. Thus

$$X(g(Y,Z)) = (\nabla_X g)(Y,Z) + g(\nabla_X Y,Z) + g(Y,\nabla_X Z) \tag{127}$$

turns into

$$X(g(Y,Z)) = g(\nabla_X Y,Z) + g(Y,\nabla_X Z) \tag{128}$$

We may rewrite this equation by applying cyclic permutations to X,Y and Z with a minus sign for the last equation. By adding all these terms we obtain

$$X(g(Y,Z)) + Y(g(X,Z)) - Z(g(X,Y)) \tag{129}$$

$$= g(\nabla_X Y + \nabla_Y X,Z) + g(\nabla_Y Z - \nabla_Z Y,X) + g(\nabla_X Z - \nabla_Z X,Y) \tag{130}$$

Because the torsion tensor vanishes the right hand side of the previous equation turns into

$$2g(\nabla_X Y,Z) - g([X,Y],Z) + g([Y,Z],X) + g([X,Z],Y) \tag{131}$$

which leads us to Koszul's formula

$$g(\nabla_X Y,Z) = \tfrac{1}{2}\Big\{X\big(g(Y,Z)\big) + Y\big(g(Z,X)\big) \tag{132}$$

$$-Z\big(g(X,Y)\big) + g([X,Y],Z) - g([Y,Z],X) - g([X,Z],Y)\Big\} \tag{133}$$

Hence, if a Levi-Civita connection exists, it must be unique, because Z is arbitrary, g is non degenerate, and the right hand side does not depend on ∇. To prove existence of a Levi-Civita connection we note that for given vector fields X,Y the right hand side of Koszul's formula is linear over $\mathscr{C}^\infty(\mathscr{M},\mathbb{R})$ in the variable Z. Thus it must define a covector field $\zeta(X,Y) \in \Omega(\mathscr{M})$. By the musical isomorphism 6 there exists a vector field, which we suggestively denote by $\nabla_X Y$, such that $\zeta(X,Y)(Z) = g(\nabla_X Y,Z)$ for all vector fields $Z \in \mathscr{X}(\mathscr{M})$. By using Koszul's formula one then checks for all vector fields X,Y,Z and all smooth functions f that

$$g(\nabla_X(Y_1+Y_2),Z) = g(\nabla_X Y_1,Z) + g(\nabla_X Y_2,Z) \tag{134}$$

$$g(\nabla_X(fY),Z) = X(f)g(Y,Z) + fg(\nabla_X Y,Z) \tag{135}$$

$$g(\nabla_X Y,Z) + g(\nabla_X Z,Y) = X\Big((g(Y,Z)\Big) \tag{136}$$

$$g(\nabla_X Y,Z) - g(\nabla_Y X,Z) = g([X,Y],Z) \tag{137}$$

Nondegeneracy of g thus implies that ∇ indeed yields a covariant derivative on vector fields, which is torsion free and preserves the metric g. □

Remark 1. *Note that with minor variations the same proof shows that there is a unique connection that is compatible with the metric and has prescribed torsion.*

Remark 2. *In the previous proof we have made use of the musical isomorphism. To make the idea here more concrete, since $\zeta(X,Y) \in \Omega^1(\mathscr{M})$ is a covector field, we may write $\zeta(X,Y) = \zeta_i dx^i$ locally for a chart $(\mathscr{U},x)$. We may raise the indices $\zeta^j := \zeta_i g^{ij}$ and define (locally) $\xi(X,Y)$ by $\zeta^j \partial_{x_j}$. But then*

$$g(\xi(X,Y),Z) = g^{ij}\zeta_i Z^k \underbrace{g(\partial_j,\partial_k)}_{g_{jk}} = g^{ij}g_{jk}\zeta_i Z^k = \delta^i_k \zeta_i Z^k = \zeta_i Z^i = \zeta(X,Y)(Z) \tag{138}$$

Hence defining $\nabla_X Y := \xi(X,Y)$ does the trick.

At last we note that there is a nice way to represent the Christoffel symbols in terms of the metric coefficients of g. Indeed, by Koszul's formula and the fact that $[\partial_\mu, \partial_\lambda] = 0$ for all μ, λ we have

$$g(\nabla_\gamma \partial_\beta, \partial_\sigma) = g_{\zeta\sigma}\Gamma^\zeta_{\beta\gamma} = \frac{1}{2}(\partial_\gamma g_{\beta\sigma} + \partial_\beta g_{\gamma\sigma} - \partial_\sigma g_{\beta\gamma}) \tag{139}$$

Thus multiplying this equation by $g^{\alpha\sigma}$ yields

$$\underbrace{g^{\alpha\sigma} g_{\zeta\sigma}}_{\delta^\alpha_\zeta}\Gamma^\zeta_{\beta\gamma} = \Gamma^\alpha_{\beta\gamma} = \frac{g^{\alpha\sigma}}{2}(\partial_\gamma g_{\beta\sigma} + \partial_\beta g_{\gamma\sigma} - \partial_\sigma g_{\beta\gamma}) \tag{140}$$

2.1.3 *Riemann curvature tensor and its properties*

We shall begin by introducing the Riemann curvature tensor in an index-free fashion.

Definition 5. Let ∇ be a torsionless covariant derivative defined for vector fields. Then the mapping

$$R\colon \mathscr{X}(\mathscr{M})^3 \to \mathscr{X}(\mathscr{M}) \qquad R(X,Y)(Z) := [\nabla_X, \nabla_Y]Z - \nabla_{[X,Y]}Z \tag{141}$$

is called the Riemann curvature tensor of $\mathscr{M}$.

Lemma 1. *The Riemann curvature tensor R is a tensor of valence* $(1,3)$.

Proof. We solely have to prove that R is $\mathscr{C}^\infty$-multilinear. So let $f \in \mathscr{C}^\infty(\mathscr{M}, \mathbb{R})$. We then have $[X, fY] = X(f)Y + f[X,Y]$ and thus

$$R(X,fY)(Z) = \nabla_X(f\nabla_Y Z) - f\nabla_Y\nabla_X Z - X(f)\nabla_Y Z - f\nabla_{[X,Y]}Z \tag{142}$$

$$= X(f)\nabla_Y Z - X(f)\nabla_Y Z + fR(X,Y)(Z) = fR(X,Y)(Z) \tag{143}$$

Since by definition we have $R(X,Y)(Z) = -R(Y,X)(Z)$ we also have $R(fX,Y)(Z) = fR(X,Y)(Z)$. Analogously, one easily verifies that $R(X,Y)(fZ) = fR(X,Y)(Z)$. □

Lemma 2 (Coordinate Representation for R)**.** *Let $(\mathscr{U}, x)$ be a chart. Then we have*

$$R(\partial_k, \partial_l)(\partial_j) = R^i_{jkl}\partial_i \tag{144}$$

where

$$R^i_{jkl} = \frac{\partial}{\partial x^l}\Gamma^i_{kj} - \frac{\partial}{\partial x^k}\Gamma^i_{lj} + \Gamma^i_{lm}\Gamma^m_{kj} - \Gamma^i_{km}\Gamma^m_{lj} \tag{145}$$

Proof. Since $[\partial_i, \partial_j] = 0$ for all i,j we have

$$R(\partial_k, \partial_l)(\partial_j) = \nabla_{\partial_k}\nabla_{\partial_l}\partial_j - \nabla_{\partial_l}\nabla_{\partial_k}\partial_j \tag{146}$$

But we also know that

$$\nabla_{\partial_k}\nabla_{\partial_l}\partial_j = \nabla_{\partial_k}(\Gamma^m_{jl}\partial_m) = \frac{\partial}{\partial x^k}\Gamma^m_{jl}\partial_m + \Gamma^m_{jl}\Gamma^r_{mj}\partial_r \tag{147}$$

Now exchanging k and l and subtracting the respective terms gives the assertion. □

On the other hand, if we consider the induced covariant derivative on covector fields we observe that

$$\nabla_{\partial_k} dx^j = (\nabla_{\partial_k} dx^j)(\partial_a) dx^a = \underbrace{\Big(\partial_k(\underbrace{dx^j(\partial_a)}_{=\delta^j_a})}_{=0} - dx^j(\nabla_{\partial_k}\partial_a)\Big) dx^a = -\Gamma^j_{ak} dx^a \tag{148}$$

and therefore

$$\nabla_{\partial_l}\nabla_{\partial_k} dx^j = -\nabla_{\partial_l}(\Gamma^j_{ak} dx^a) = -\Big[\partial_l \Gamma^j_{ak} dx^a + \Gamma^j_{ak} \underbrace{\nabla_{\partial_l} dx^a}_{=-\Gamma^a_{bl} dx^b}\Big] \tag{149}$$

$$= \Big[\Gamma^j_{ak}\Gamma^a_{bl} - \partial_l \Gamma^j_{bk}\Big] dx^b \tag{150}$$

Therefore, we obtain

$$\nabla_{\partial_l}\nabla_{\partial_k} dx^j - \nabla_{\partial_k}\nabla_{\partial_l} dx^j = \Big[\Gamma^j_{ak}\Gamma^a_{bl} - \partial_l \Gamma^j_{bk} - \Gamma^j_{al}\Gamma^a_{bk} + \partial_k \Gamma^j_{bl}\Big] dx^b = -R^j_{blk} dx^b \tag{151}$$

For a general tensor T and torsion free connection, each tensor index comes with a corresponding Riemann tensor term:

$$\nabla_{\partial_\mu}\nabla_{\partial_\nu} T^{b_1 \dots b_s}_{a_1 \dots a_r} - \nabla_{\partial_\nu}\nabla_{\partial_\mu} T^{b_1 \dots b_s}_{a_1 \dots a_r} \tag{152}$$

$$= -R^{\sigma}_{a_1 \mu\nu} T^{b_1 \dots b_s}_{\sigma \dots a_r} - \dots - R^{\sigma}_{a_r \mu\nu} T^{b_1 \dots b_s}_{a_1 \dots \sigma} + R^{b_1}_{\sigma\mu\nu} T^{\sigma \dots b_s}_{a_1 \dots a_r} + \dots + R^{b_s}_{\sigma\mu\nu} T^{b_1 \dots \sigma}_{a_1 \dots a_r} \tag{153}$$

From now on we assume a Levi Civita connection.

Theorem 8. *There exists a coordinate system in which the metric tensor field has vanishing second derivatives at p if and only if its Riemann tensor vanishes at p. Furthermore, there exists a coordinate system in which the metric tensor field has constant entries near p if and only if the Riemann tensor vanishes near p.*

We now list some symmetries of the curvature tensor of the Levi-Civita connection:

1. Obviously we have

$$R^{\delta}_{\gamma\alpha\beta} = -R^{\delta}_{\gamma\beta\alpha} \tag{154}$$

2. The first Bianchi identity:

$$R^{\delta}_{\gamma\alpha\beta} + R^{\delta}_{\alpha\beta\gamma} + R^{\delta}_{\beta\gamma\alpha} = 0 \tag{155}$$

or put differently

$$R(X,Y)(Z) + R(Y,Z)(X) + R(Z,X)(Y) = 0 \tag{156}$$

for all $X, Y, Z \in \mathscr{X}(\mathscr{M})$.

Proof. For any $\mathbb{R}$-linear map $F\colon \mathscr{X}^3 \to \mathscr{X}$ define the mapping $\mathfrak{S}(F)\colon \mathscr{X}^3 \to \mathscr{X}$ by

$$\mathfrak{S}(F)(X,Y,Z) := F(X,Y,Z) + F(Y,Z,X) + F(Z,X,Y) \tag{157}$$

Then a cyclic permutation of X,Y,Z obviously leaves $\mathfrak{S}(F)(X,Y,Z)$ unchanged. Thus

$$\begin{aligned}\mathfrak{S}(R)(X,Y,Z) = \mathfrak{S}\nabla_Y\nabla_X Z - \mathfrak{S}\nabla_X\nabla_Y Z &= \mathfrak{S}\nabla_X\nabla_Z Y - \mathfrak{S}\nabla_X\nabla_Y Z \\ &= \mathfrak{S}\nabla_X \underbrace{(\nabla_Z Y - \nabla_Y Z)}_{-[Y,Z]}\end{aligned} \tag{158}$$

However, since R is a tensor and thus is $\mathscr{C}^\infty$-multilinear it suffices to prove $\mathfrak{S}R(\partial_j,\partial_k)(\partial_l) = 0$. But from what we have seen in (158) this follows immediately from $[\partial_k,\partial_l] = 0$. □

3. The pair interchange symmetry:

$$R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta} \tag{159}$$

where $R_{\gamma\delta\alpha\beta} = g_{\gamma\sigma}R^{\sigma}_{\delta\alpha\beta}$.

Proof. In order to prove this, we shall suppose that our given metric is twice differentiable. We then know that for any point $p \in \mathscr{M}$ there exists a coordinate system in which the connection coefficients $\Gamma^{\alpha}_{\beta\gamma}$ vanish at p. We thus have

$$R^{\alpha}_{\beta\gamma\delta} = \partial_\gamma\Gamma^{\alpha}_{\beta\delta} - \partial_\delta\Gamma^{\alpha}_{\beta\gamma} = \frac{1}{2}\Big\{g^{\alpha\sigma}\partial_\gamma(\partial_\sigma g_{\sigma\beta} + \partial_\beta g_{\sigma\delta} - \partial_\sigma g_{\beta\delta}) \tag{160}$$

$$-g^{\alpha\sigma}\partial_\delta(\partial_\gamma g_{\sigma\beta} + \partial_\beta g_{\sigma\gamma} - \partial_\sigma g_{\beta\gamma})\Big\} \tag{161}$$

$$= \frac{g^{\alpha\sigma}}{2}\Big\{\partial_\gamma\partial_\beta g_{\sigma\delta} - \partial_\gamma\partial_\sigma g_{\beta\delta} - \partial_\delta\partial_\beta g_{\sigma\gamma} + \partial_\delta\partial_\sigma g_{\beta\gamma}\Big\} \tag{162}$$

and from the last expression, after having lowered the index α, that is, $R^{\alpha}_{\beta\gamma\delta} = g_{\sigma\alpha}R^{\sigma}_{\beta\gamma\delta}$, the above equation yields the claim. □

4. We have $R_{\alpha\beta\gamma\delta} = -R_{\beta\alpha\gamma\delta}$.

Proof. Indeed,

$$R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta} = -R_{\gamma\delta\beta\alpha} = -R_{\beta\alpha\gamma\delta} \tag{163}$$

□

5. The second Bianchi identity:

$$\nabla_\alpha R_{\sigma\beta\gamma\delta} + \nabla_\sigma R_{\beta\alpha\gamma\delta} + \nabla_\beta R_{\alpha\sigma\gamma\delta} = 0 \tag{164}$$

Proof. We again work in coordinates in which the derivatives of the metric vanish at $p \in \mathcal{M}$. A calculation similar to the one for the pair interchange symmetry yields

$$\nabla_\alpha R_{\sigma\beta\gamma\delta} = \partial_\alpha R_{\sigma\beta\gamma\delta} \tag{165}$$

$$= \frac{1}{2}\left\{\partial_\alpha\partial_\gamma\partial_\beta g_{\sigma\delta} - \partial_\alpha\partial_\gamma\partial_\sigma g_{\beta\sigma} - \partial_\alpha\partial_\delta\partial_\beta g_{\sigma\gamma} + \partial_\alpha\partial_\delta\partial_\sigma g_{\beta\gamma}\right\} \tag{166}$$

From this the result will follow after inspecting the terms of the sum on the left hand side of (164). □

6. The Ricci tensor is defined as

$$R_{\alpha\beta} := R^\sigma_{\alpha\sigma\beta} \tag{167}$$

The pair interchange property implies that the Ricci tensor is symmetric, since

$$R_{\alpha\beta} = g^{\delta\sigma} R_{\delta\alpha\sigma\beta} = g^{\delta\sigma} R_{\sigma\beta\delta\alpha} = R_{\beta\alpha} \tag{168}$$

2.2 *Local inertial coordinates, Geodesic deviation(Jacobi equation), tidal forces*

2.2.1 *Local inertial coordinates*

Proposition 3. *Let g be a Lorentzian metric on $\mathcal{M}$.*

1. *For every $p \in \mathcal{M}$ there exists a neighborhood of p with a coordinate system such that $g_{\mu\nu} = \eta_{\mu\nu} = diag(-1, 1, ..., 1)$ at p.*

2. *If g is differentiable, then the coordinates can be further chosen so that*

$$\partial_\sigma g_{\alpha\beta} = 0 \iff \Gamma^\alpha_{\beta\gamma} = 0 \tag{169}$$

at p.

Proof. For the first point, let y be any local coordinate system around p. We may assume without loss of generality, by shifting by a constant vector, that p corresponds to $y(p) = 0$. Let $e_a = e^\mu_a \frac{\partial}{\partial y^\mu}$ be any frame at p such that $g(e_a, e_b) = \eta_{ab}$. Existence of such a frame follows from Gram-Schmidt. Calculating the determinant of both sides of the equation

$$g_{\mu\nu} e^\mu_a e^\nu_b = \eta_{ab} \tag{170}$$

yields, at p,

$$\det(g_{\mu\nu}) \det(e^\mu_a)^2 = -1 \tag{171}$$

So the determinant of $e = (e^\mu_a)_{a,\mu}$ is nonvanishing. Thus e is a local diffeomorphism, so we may define a new chart x implicitly by the equation

$$y = e \circ x \iff y^\mu = e^\mu_a x^a \qquad \forall \mu \tag{172}$$

But then we note that

$$\frac{\partial}{\partial x^a}\Big|_p = T_{x(p)}x^{-1}(e_a) \overset{y^{-1}\circ e = x^{-1}}{=} T_{e(x(p))}y^{-1}\circ \underbrace{T_{x(p)}e(e_a)}_{(e_a^\mu)_\mu} = e_a^\mu \frac{\partial}{\partial y^\mu}\Big|_{y(p)} \tag{173}$$

and therefore

$$g(\partial_{x^a}, \partial_{x^b}) = e_a^\mu e_b^\nu g(\partial_{y^\mu}, \partial_{y^\nu}) = \eta_{ab} \tag{174}$$

For the second claim we will use the formula

$$\Gamma^i_{jk} = \frac{\partial x^i}{\partial y^s}\frac{\partial^2 y^s}{\partial x^j \partial x^k} + \frac{\partial x^i}{\partial y^s}\frac{\partial y^l}{\partial x^j}\frac{\partial y^r}{\partial x^k}\widehat{\Gamma}^s_{lr} \tag{175}$$

Let x be the coordinate chart constructed in the first part of the proof and recall that p lies at the origin of those coordinates. The new coordinates $\widehat{x}^j$ will be implicitly defined by the equations

$$x^i = \widehat{x}^i + \frac{1}{2}A^i_{jk}\widehat{x}^j\widehat{x}^k \tag{176}$$

where $\{A^i_{jk}\}$ is a set of constants, symmetric with respect to the interchange of j and k. In these coordinates we have

$$\widehat{\Gamma}^i_{jk} = \frac{\partial \widehat{x}^i}{\partial x^s}\frac{\partial^2 x^s}{\partial \widehat{x}^j \partial \widehat{x}^k} + \frac{\partial \widehat{x}^i}{\partial x^s}\frac{\partial x^l}{\partial \widehat{x}^j}\frac{\partial x^r}{\partial \widehat{x}^k}\Gamma^s_{lr} \tag{177}$$

Proof to be finished.... □

2.2.2 *Calculus of Variations and the Euler-Lagrange equation* (⋆)

The calculus of variations deals with the problem ofminimizing (in fact, extremizing) nonlinear functionals of the form

$$\mathscr{L}(f) = \int L(f, \partial_1 f, \partial_2 f, ..., \partial_n f)(x) d\lambda(x) \tag{178}$$

where the Lagrangian L is a function $\mathbb{R}\times\mathbb{R}^n \to \mathbb{R}$. If we assume now that f extremizes the functional $\mathscr{L}$, we will certainly have

$$\frac{d}{dt}\Big|_{t=0}\mathscr{L}(f+tg) = 0 \tag{179}$$

for any other (reasonable) function g (this is because the map $t \mapsto \mathscr{L}(f+tg)$ has a local extremum at $t=0$). By the chain rule we have

$$\frac{d}{dt}\mathscr{L}(f+tg) = \int \partial_1 L(...)g + \int \underbrace{\partial^{1+j}L(...)\partial_{x^j}g}_{\text{summation convention}} \tag{180}$$

and integration by parts yields

$$\frac{d}{dt}\mathscr{L}(f+tg) = \int \left\{\partial_1 L(...) - \partial_{x_j}\partial^{1+j}L(...)\right\} g \tag{181}$$

Since this is supposed to hold for all g we get the Euler Lagrange equation

$$\partial_1 L(f,\partial_1 f,...,\partial_n f) - \partial_{x_j}\partial^{1+j} L(f,\partial_1 f,...,\partial_n f) = 0 \tag{182}$$

In physics notation this reads as

$$\frac{\partial L}{\partial f} - \partial^j \frac{\partial L}{\partial(\partial_j f)} = 0 \tag{183}$$

2.2.3 *Geodesic equation*

In a Riemannian manifold $\mathscr{M}$ with metric tensor g, the length L of a continuously differentiable curve γ: $[a,b] \to \mathscr{M}$ is defined by

$$L(\gamma) = \int_a^b \sqrt{g_{\gamma(t)}(\dot\gamma(t),\dot\gamma(t))}\,dt \tag{184}$$

In Riemannian geometry, all geodesics are locally distance-minimizing paths, but the converse is not true. In fact, only paths that are both locally distance minimizing and parameterized proportionately to arc-length are geodesics. Another equivalent way of defining geodesics on a Riemannian manifold, is to define them as the minima of the following action or energy functional

$$E(\gamma) = \frac{1}{2}\int_a^b g_{\gamma(t)}(\dot\gamma(t),\dot\gamma(t))\,dt \tag{185}$$

An application of the Cauchy Schwarz inequality immediately yields

$$L(\gamma)^2 \le 2(b-a)E(\gamma) \tag{186}$$

All minima of E are also minima of L, but the minima of L form a bigger set since paths that are minima of L can be arbitrarily re-parameterized (without changing their length), while minima of E cannot. The geodesic equation is the associated Euler-Lagrange equation with respect to the energy functional E. In this case the Euler Lagrange equation turns into

$$\frac{\partial L}{\partial \gamma} = \frac{d}{ds}\frac{\partial L}{\partial \dot\gamma} \tag{187}$$

Writing this in coordinates $x = (x^\mu)$ we get with $\gamma^\mu = x^\mu \circ \gamma$

$$\frac{\partial L}{\partial \gamma^\lambda} = \frac{1}{2}\partial_\lambda g_{\mu\nu}\dot\gamma^\mu\dot\gamma^\nu \tag{188}$$

$$\frac{\partial L}{\partial \dot\gamma^\lambda} = g_{\mu\lambda}\dot\gamma^\mu \tag{189}$$

and

$$\frac{d}{ds}\left\{g_{\mu\lambda}(\gamma)\dot\gamma^\mu\right\} = \partial_\nu g_{\mu\nu}\dot\gamma^\mu\dot\gamma^\nu + g_{\mu\nu}\ddot\gamma^\mu \tag{190}$$

$$= \frac{1}{2}\partial_\nu g_{\mu\lambda}\dot\gamma^\mu\dot\gamma^\nu + \frac{1}{2}\partial_\mu g_{\nu\lambda}\dot\gamma^\mu\dot\gamma^\nu + g_{\mu\lambda}\ddot\gamma^\mu \tag{191}$$

Plugging in all the expressions and putting it all together the geodesic equation reads as

$$g_{\mu\lambda}\ddot{\gamma}^{\mu} = \frac{1}{2}\Big\{\partial_{\lambda}g_{\mu\nu} - \partial_{\nu}g_{\mu\lambda} - \partial_{\mu}g_{\nu\lambda}\Big\}\dot{\gamma}^{\mu}\dot{\gamma}^{\nu} = -\Gamma_{\lambda\mu\nu}\dot{\gamma}^{\mu}\dot{\gamma}^{\nu} \tag{192}$$

and by raising the index λ, that is, by multiplying the previous equation with $g^{\sigma\lambda}$, we obtain

$$\ddot{\gamma}^{\lambda} + \Gamma^{\lambda}_{\mu\nu}\dot{\gamma}^{\mu}\dot{\gamma}^{\nu} = 0 \tag{193}$$

The physics way of writing this is

$$\frac{d^2x^{\mu}}{ds^2} + \Gamma^{\mu}_{\alpha\beta}\frac{dx^{\alpha}}{ds}\frac{dx^{\beta}}{ds} = 0 \tag{194}$$

which, in my opinion, seems hellishly confusing.

This gives a very convenient way of calculating the Christoffel symbols: given a metric g, write down L, work out the Euler-Lagrange equations, and identify the Christoffels as the coefficients of the first derivative terms in those equations.

2.2.4 *Derivatives of curves as vector fields* ($\star$)

Let $\gamma\colon I \to \mathscr{M}$ be a smooth curve. If $f \in \mathscr{C}^{\infty}(\mathscr{M},\mathbb{R})$ is a smooth function on $\mathscr{M}$, then the function $f \circ \gamma$ is a smooth map $I \to \mathbb{R}$. Thus we can take the derivative

$$\frac{d}{ds}(f \circ \gamma)(s) \tag{195}$$

Now define $\dot{\gamma}(s)$ by

$$\dot{\gamma}(s)(f) := (f \circ \gamma)(s) \qquad (f \in \mathscr{C}^{\infty}(\mathscr{M},\mathbb{R})) \tag{196}$$

Then of course $\dot{\gamma}(s) \in T_{\gamma(s)}\mathscr{M}$ for all $s \in I$.

2.2.5 *Geodesic Deviation - Jacobi equation*

How are extended bodies, as opposed to point objects, affected by the gravitational field? In order to understand this, we shall consider a one-parameter family of geodesics

$$\Big(\gamma(.,\lambda)\Big)_{\lambda\in\mathbb{R}} \tag{197}$$

Throughout, we will assume every bit of necessary smoothness. Now set

$$Z(s,\lambda) := \frac{\partial\gamma(s,\lambda)}{\partial\lambda} \equiv \frac{\partial\gamma^{\alpha}(s,\lambda)}{\partial\lambda}\partial_{\alpha} \tag{198}$$

For each λ this defines a vector field Z along $\gamma(s,\lambda)$, which measures how nearby geodesics deviate from each other, since

$$\gamma^{\alpha}(s,\lambda) - \gamma^{\alpha}(s,\lambda_0) = Z^{\alpha}(\lambda - \lambda_0) + O((\lambda - \lambda_0)^2) \tag{199}$$

To measure how a vector field $W = W^\mu \partial_\mu$ changes along $s \mapsto \gamma(s,\lambda)$, one introduces the differential operator $\frac{D}{ds}$ defined by

$$\frac{DW}{ds} := \dot{\gamma}^\beta \nabla_\beta W = \dot{\gamma}^\beta \partial_\beta W^\mu \partial_\mu + \dot{\gamma}^\beta W^\mu \nabla_\beta \partial_\mu \tag{200}$$

$$= \underbrace{\dot{\gamma}^\beta (\partial_\beta W^\mu)}_{\frac{\partial (W^\mu \circ \gamma)}{\partial s}} \partial_\mu + \dot{\gamma}^\beta W^\mu \Gamma^\zeta_{\mu\beta} \partial_\zeta \tag{201}$$

Analogously, we can define

$$\frac{DW}{d\lambda} := Z^\beta \nabla_\beta W = \frac{\partial \gamma^\beta}{\partial \lambda} (\partial_\beta W^\mu) \partial_\mu + \frac{\partial \gamma^\beta}{\partial \lambda} W^\mu \Gamma^\zeta_{\mu\beta} \partial_\zeta \tag{202}$$

Now assume that each member of the family of geodesics $(\gamma(.,\lambda))_\lambda$ is an integral curve of a vector field ζ_λ. By definition this means

$$\forall s\colon \dot{\gamma}(s,\lambda) = \zeta_\lambda(\gamma(s,\lambda)) \tag{203}$$

If we want to express this very sloppily, we would write

$$\dot{\gamma} = \zeta \tag{204}$$

We then have

$$\frac{d^2\gamma^\mu}{ds}(s,\lambda) = \frac{d\zeta^\mu_\lambda(\gamma(s,\lambda))}{ds} = \frac{\partial \zeta^\mu_\lambda(s,\lambda)}{\partial x^\nu} \dot{\gamma}^\nu(s,\lambda) \tag{205}$$

so that we can define

$$\frac{D^2\gamma}{ds} := \frac{d^2\gamma^\mu}{ds^2} \partial_\mu + \Gamma^\mu_{\alpha\beta} \dot{\gamma}^\alpha \dot{\gamma}^\beta \partial_\mu = \frac{\partial \zeta^\mu}{\partial x^\alpha} \zeta^\alpha \partial_\mu + \Gamma^\mu_{\alpha\beta} \zeta^\alpha \zeta^\beta \partial_\mu = \zeta^\alpha \nabla_\alpha \zeta \tag{206}$$

and if we are again sloppy, this reads as

$$\frac{D^2\gamma}{ds} = \dot{\gamma}^\alpha \nabla_\alpha \dot{\gamma} \tag{207}$$

Since $s \mapsto \gamma(s,\lambda)$ is a geodesic we have by (206)

$$\frac{D^2\gamma}{ds} = 0 \tag{208}$$

We then have (note that we drastically abuse notation here)

$$\frac{DZ}{ds} = \frac{\partial^2 \gamma^\mu}{\partial \lambda} \partial_\mu + \Gamma^\mu_{\alpha\beta} \dot{\gamma}^\alpha \frac{\partial \gamma^\beta}{\partial \lambda} \partial_\mu = \frac{D\dot{\gamma}}{d\lambda} \tag{209}$$

Abusing notation yet again,

$$\nabla_{\dot{\gamma}} Z = \dot{\gamma}^\nu \nabla_\nu Z = \dot{\gamma}^\nu \nabla_\nu \left(\frac{\partial \gamma^\mu}{\partial \lambda} \partial_\mu \right) = \frac{\partial^2 \gamma^\mu}{\partial s \partial \lambda} \partial_\mu + \Gamma^\mu_{\alpha\beta} \dot{\gamma}^\alpha \frac{\partial \gamma^\beta}{\partial \lambda} = Z^\beta \nabla_\beta \dot{\gamma} = \nabla_Z \dot{\gamma} \tag{210}$$

Thus, $\nabla_{\dot{\gamma}} Z = \nabla_Z \dot{\gamma}$. One then also calculates quite easily (by accepting the abusing conventions and faulty definitions) that

$$\left(\frac{D}{ds}\frac{D}{d\lambda} - \frac{D}{d\lambda}\frac{D}{ds}\right)W = \left[R^{\mu}_{\delta\alpha\beta}\dot{\gamma}^{\alpha}Z^{\beta}W^{\delta}\right]\partial_{\mu} \tag{211}$$

which is referred to as the Jacobi equation, or as the geodesic deviation equation. If $W^{\mu} = \dot{\gamma}^{\mu}$, then by equation (207) we have $\frac{D}{ds}\dot{\gamma} = 0$ and thus the second term of at the left-hand side of (211) is zero, and from $\frac{D}{d\lambda}\dot{\gamma} = \frac{D}{ds}Z$ we obtain

$$\frac{D^2 Z}{ds^s} = \left[R^{\mu}_{\delta\alpha\beta}\dot{\gamma}^{\alpha}Z^{\beta}\dot{\gamma}^{\delta}\right]\partial_{\mu} \tag{212}$$

This is the so called geodesic deviation equation. In the index-free notation this reads as

$$\frac{D^2 Z}{ds^s} = R(\dot{\gamma}, Z)\dot{\gamma} \tag{213}$$

Solutions to the geodesic deviation equation are called Jacobi fields along γ. The previous equation shows that curvature causes relative acceleration between neighboring geodesics. Keeping in mind that gravitational force and acceleration are indistinguishable, we say that curvature produces a "gravitational tidal force" between freely falling nearby observers.

2.3 *Einstein equations and matter: examples of energy-momentum tensors, dust in general relativity, the continuity equation*

It is known from lectures on special relativity that Einstein's equation in vaccuum read

$$R_{\mu\nu} = 0 \tag{214}$$

where $R_{\mu\nu}$ is the Ricci tensor $R_{\mu\nu} = R^{\sigma}_{\mu\sigma\nu}$. It is natural to anticipate, in the presence of matter, that the right-hand side of the above equation might not be zero, but rather will be an object describing the density of energy of matter fields. Our idea is: **energy produces curvature.**

2.3.1 *Dust in general Relativity*

By definition, dust is a cloud of noninteracting particles, whose velocities vary smoothly from point to point in spacetime. At each point we have a smooth function $\rho \in \mathscr{C}^{\infty}(\mathscr{M})$ which represents the density of the dust: this is the mass per unit volume measured in a frame in which the particles are at rest. For example, if there are n particles per unit volume and each has rest mass m, then $\rho = nm$. A rest frame is a frame in which the particles do not move, so that their space velocity is zero. If we assumed we were in special relativity, particles in such a rest frame would have velocity four-vector

$$u = u^{\mu}\partial_{\mu} = \partial_t \iff (u^{\mu}) = (1, \mathbf{0}) \tag{215}$$

Let an observer move with space velocity $v \in \mathbb{R}^3$ with respect to the dust, so she has a four velocity vector

$$(v^{\mu}) = \gamma(1, v) \tag{216}$$

where $\gamma := \frac{1}{\sqrt{1-|v|^2}}$ is the Lorentz contraction factor. Recall that

$$L = \frac{1}{\gamma(v)} L_0 \tag{217}$$

where

- L is the length observed by an observer in motion relative to the object
- L_0 is the proper length (the length of the object in its rest frame)
- $\gamma(v)$ is the Lorentz factor, defined as

$$\frac{1}{\sqrt{1-|v|^2/c^2}} \tag{218}$$

- v is the relative velocity between the observer and the moving object
- c is the speed of light

As is standard, we shall assume $c = 1$. Continuing with (216): We choose a coordinate system so that the velocity is aligned along the x-axis and is pointing in the positive direction. Then the observer has velocity v along the x-axis. Let there be n particles of rest mass m in a box with sides dx, dy and dz in the reference frame of the dust. The observer sees n particles of rest mass m in a box with sides $\gamma^{-1}dx$ (Lorentz contraction factor!), dy and dz, with space velocity $-v$, and therefore energy

$$mn\gamma \tag{219}$$

in a volume $\gamma^{-1}dxdydz$ and hence a density

$$mn\gamma^2 = \rho(\underbrace{u_\mu v^\mu}_{=\gamma})^2 = \underbrace{\rho u_\mu u_\nu}_{T_{\mu\nu}} v^\mu v^\nu \tag{220}$$

This gives rise to the so-called energy momentum tensor of dust with energy density ρ and four velocity vector u given by

$$T = T_{\mu\nu}dx^\mu \otimes dx^\nu = (\rho u_\mu u_\nu)dx^\mu \otimes dx^\nu \tag{221}$$

where, of course, $u_\mu = \eta_{\zeta\mu}u^\zeta$ and analogously for u_ν. This tensor is used to measure the energy density of dust in general frames. The above, of course carries over immediately to general relativity, via the correspondence principle. Recall that the correspondence principle is a philosophical guideline for the selection of new theories in physical science, requiring that they explain all the phenomena for which a preceding theory was valid. Thus, what needs to be done is to replace the special relativistic normalization $\eta_{\mu\nu}u^\mu u^\nu = -1$ of the four velocity vector with $g_{\mu\nu}u^\mu u^\nu = -1$. Moreover, in any relevant equations indices are raised and lowered by means of the metric g rather than with the Minkowski metric η, while partial derivatives are replaced with covariant ones. To put the preceding discussion very mathematically, the so-called energy momentum tensor, or more generally the stress energy tensor of a relativistic pressureless fluid, can be written in the simple form

$$T = \rho u^\mu u^\nu \partial_\mu \otimes \partial_\nu \tag{222}$$

where the world lines of the dust particles are the integral curves of the four-velocity $u = (u^\mu)$ and the matter density is given by the scalar function $\rho \in \mathscr{C}^\infty$. Provide some Examples!

2.3.2 *The continuity equation*

Energy momentum tensors in special relativity satisfy a conservation identity:

$$\partial_\nu T^{\mu\nu} = 0 \tag{223}$$

In order to verify this equation for dust, we need to know what the equations are first. Since we assume that the particles are non-interacting, in special relativity each of them moves along a straight line. Now, straight lines are geodesics in Minkowski spacetimes, so if u is the vector field tangent to geodesics followed by the particles, normalized so that $g(u,u)=-1$, we have seen in equation (206) that

$$u^\mu \nabla_\mu u = 0 \tag{224}$$

Since the number of particles is conserved, we also have the conservation equation

$$\nabla_\mu (\rho u)^\mu = 0 \tag{225}$$

Equation (225) is called the continuity equation. Whether in curved spacetime or not, let us calculate the divergence of the energy momentum tensor T:

$$\nabla_\mu T^{\mu\nu} = (\nabla_\mu T)(dx^\mu, dx^\nu) = \partial_\mu T^{\mu\nu} - T(\nabla_\mu dx^\mu, dx^\nu) - T(dx^\mu, \nabla_\mu dx^\nu) \tag{226}$$

Now recall that

$$\nabla_\mu dx^\omega = -\Gamma^\omega_{\sigma\mu} dx^\sigma \tag{227}$$
$$\nabla_\mu \partial_\omega = \Gamma^\sigma_{\omega\mu} \partial_\sigma \tag{228}$$

and thus we see that equation (226) transforms to be

$$\nabla_\mu T^{\mu\nu} = \partial_\mu T^{\mu\nu} + \Gamma^\mu_{\mu\omega} T^{\omega\nu} + \Gamma^\nu_{\mu\omega} T^{\mu\omega} \tag{229}$$

On the other hand, we have

$$\nabla_\mu (\rho u)^\mu = \left[\nabla_\mu (\rho u)\right](dx^\mu) = \partial_\mu (\rho u^\mu) - \rho u(\nabla_\mu dx^\mu) \tag{230}$$
$$= \partial_\mu (\rho u^\mu) + \rho u^\omega \Gamma^\mu_{\mu\omega} \tag{231}$$

and similarly

$$u^\mu \nabla_\mu u^\nu = u^\mu (\nabla_\mu u)(dx^\nu) = u^\mu \partial_\mu u^\nu + u^\mu u^\omega \Gamma^\nu_{\mu\omega} \tag{232}$$

Therefore,

$$\nabla_\mu (\rho u)^\mu u^\nu + \rho u^\mu \nabla_\mu u^\nu \tag{233}$$
$$= \underbrace{u^\nu \partial_\mu (\rho u^\mu) + \rho u^\mu (\partial_\mu u^\nu)}_{\partial_\mu (u^\nu u^\mu \rho)} + \underbrace{u^\nu \rho u^\omega \Gamma^\mu_{\mu\omega} + \rho u^\mu u^\omega \Gamma^\nu_{\mu\omega}}_{\Gamma^\mu_{\mu\omega} T^\nu \omega + \Gamma^\nu_{\mu\omega} T^{\mu\omega}} \tag{234}$$

Hence in total we obtain

$$\nabla_\mu T^{\mu\nu} = \nabla_\mu (\rho u)^\mu u^\nu + \rho u^\mu \nabla_\mu u^\nu \tag{235}$$

where the first term on the right hand side vanishes due to continuity equation and the left hand side vanishes because of the geodesic equation (224). It is noteworthy that if $T^\mu \nu$ is interpreted as a function to which we apply the covariant derivative, then

$$\nabla_\mu (T^\mu \nu) = \partial_\mu (\rho u^\mu u^\nu) = \partial_\mu (\rho u^\mu) u^\nu + \rho u^\mu \partial_\mu u^\nu \tag{236}$$

which is a good memorization tool for the previous equations. The vanishing of the divergence of T in (235) is actually equivalent to equation (224) + (225) in regions where the density ρ does not vanish. Indeed, if we remember the condition

$$g(u,u) = u_\nu u^\nu = -1 \tag{237}$$

then we have

$$0 = \partial_\mu(u_\nu u^\nu) = \partial_\mu(g_{\alpha\nu}u^\alpha u^\nu) = \partial_\mu(g(u,u)) = \underbrace{(\nabla_\mu g)(u,u)}_{=0} + 2g(\nabla_\mu u, u) \tag{238}$$

$$= 2u^\alpha g_{\alpha\beta}\nabla_\mu u^\beta = 2u_\beta \underbrace{\nabla_\mu u^\beta}_{dx^\beta(\nabla_\mu u)} \tag{239}$$

Therefore, multiplying equation (235) by u_ν (and using the assumption $\nabla_\mu T^\mu = 0$) yields

$$0 = u_\nu \nabla_\mu T^{\mu\nu} = \nabla_\mu(\rho u)^\mu \underbrace{u^\nu u_\nu}_{=-1} + \rho u^\mu \underbrace{u_\nu \nabla_\mu u^\nu}_{=0} \tag{240}$$

from which we immediately infer that $\nabla_\mu(\rho u)^\mu = 0$. Thus it also immediately follows that the geodesic equation is satisfied.

2.3.3 *Some bullshitery of index manipulation* ($\star$)

We want to extend the musical isomorphism (6) to arbitrary tensors.

Definition 6. Let $T \in \mathscr{T}^l_k$ be a tensor. We define $\downarrow^j_i T \in \mathscr{T}^{l-1}_{k+1}$ via

$$(\downarrow^j_i T)(X_1,...,X_{k+1},\varphi^1,...,\varphi^{l-1}) :=$$
$$T(X_1,\ldots,X_{i-1},X_{i+1},\ldots X_{k+1},\varphi^1,\ldots,\underbrace{\Psi(X_i)}_{j\text{-th slot}},\ldots\varphi^{l-1}) \qquad \left(\varphi^s \in \Omega, X_r \in \mathscr{X}\right)$$

where $\Psi\colon \mathscr{X} \to \Omega$ was the musical isomorphism defined in (6) (this is just lowering of indices). Recall that in the physics notation $\Psi(X)$ for a vector field X is simply given by

$$\Psi(X^\mu \partial_\mu) = X^\mu g_{\mu\sigma} dx^\sigma \tag{241}$$

locally (which is just lowering of indices). Locally the coefficients of the tensor $\downarrow^j_i T \in \mathscr{T}^{l-1}_{k+1}$ are given by

$$(\downarrow^j_i T)^{j_1\ldots j_{l-1}}_{i_1\ldots i_{k+1}} = (\downarrow^j_i T)(\partial_{i_1},...,\partial_{i_{k+1}},dx^{j_1},\ldots,dx^{j_{l-1}}) \tag{242}$$

$$= T(\partial_{i_1},\ldots,\partial_{i_{i-1}},\partial_{i_{i+1}},\ldots\partial_{i_{k+1}},dx^{j_1},\ldots,\underbrace{\Psi(\partial_{i_i})}_{j\text{th-slot}},\ldots dx^{j_{l-1}}) \tag{243}$$

$$= T^{j_1\ldots \overbrace{\zeta}^{j\text{-th slot}} \ldots j_{l-1}}_{i_1\ldots i_{i-1},i_{i+1},\ldots i_{k+1}} g_{i_i\zeta} \tag{244}$$

On the other hand, we may also define the inverse to the operation $\downarrow_i^j$: Let $T \in \mathcal{T}_k^l$ be given, we define $\uparrow_i^j T \in \mathcal{T}_{k-1}^{l+1}$ by

$$\uparrow_i^j T(X_1,\ldots,X_{k-1},\varphi^1,\ldots,\varphi^{l+1}) := T(X_1,\ldots,\underbrace{\Psi^{-1}(\varphi^j)}_{i\text{-th slot}},\ldots X_{k-1},\varphi^1,\ldots\varphi^{j-1},\varphi^{j+1},\ldots,\varphi^{l+1}) \qquad \left(\varphi^s \in \Omega, X_r \in \mathcal{X}\right)$$

The coefficients of $\uparrow_i^j T$ are given by

$$(\uparrow_i^j T)^{j_1\ldots j_{l+1}}_{i_1\ldots i_{k-1}} = T^{j_1,\ldots,j_{j-1},j_{j+1},\ldots j_{l+1}}_{i_1\ldots\underbrace{\zeta}_{i\text{th-slot}}\ldots i_{k-1}} g^{j_j\zeta} \tag{245}$$

Theorem 9. *Suppose we are working on a smooth Riemannian manifold $(\mathcal{M},g)$ and let an arbitrary tensor $T \in \mathcal{T}_k^l$ be given. Let $X \in \mathcal{X}$, then*

$$\nabla_X(\downarrow_i^j T) = \downarrow_i^j (\nabla_X T) \qquad \nabla_X(\uparrow_i^j T) = \uparrow_i^j (\nabla_X T) \tag{246}$$

Proof. See [3]. □

2.3.4 *Einstein equations with Sources*

The energy momentum tensor T provides a good candidate for the source term in Einstein's theory of gravitation. The energy momentum tensor of matter fields will be described by a symmetric tensor satisfying

$$\nabla_\mu T^{\mu\nu} = 0 \tag{247}$$

or equivalently (by Theorem 9)

$$\nabla^\mu T_{\mu\nu} = 0 \tag{248}$$

Theorem 10. *We have the following identity:*

$$\nabla^\mu (R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}) = 0 \tag{249}$$

where $R := R^\alpha_\alpha = R^{\alpha\beta}_{\alpha\beta}$ and $R_{\mu\nu} = R^\sigma_{\mu\sigma\nu}$ is the Ricci tensor.

Proof. We recall the second Bianchi identity

$$\nabla_\mu R_{\nu\rho\alpha\beta} + \nabla_\nu R_{\rho\mu\alpha\beta} + \nabla_\rho R_{\mu\nu\alpha\beta} = 0 \tag{250}$$

Raising indices by multiplying with $g^{\mu\alpha} g^{\nu\beta}$ (this is yet again justified by Theorem 9) we obtain

$$\nabla^\alpha R^\beta{}_{\rho\alpha\beta} + \nabla^\beta R_\rho{}^\alpha{}_{\alpha\beta} + \nabla_\rho R^{\alpha\beta}{}_{\alpha\beta} = 0 \tag{251}$$

The previous equation is actually equivalent to the one we wanted to prove, so we are done. □

Furthermore, we note that rescaling by any constant Λ we have

$$\nabla^{\mu}(\Lambda g_{\mu\nu}) = 0 \tag{252}$$

and therefore we are led to an equation compatible with (248):

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = \kappa T_{\mu\nu} \tag{253}$$

We will se that κ will be made out to be $\frac{8\pi G}{c^4}$. The constant Λ is called the cosmological constant and current state-of-the art observations indicate strongly that Λ is nonzero:

$$\Lambda \simeq 10^{-121} \text{ Planck units} \tag{254}$$

Hence we will mostly assume $\Lambda = 0$ and use units $G = c = 1$, so that equation (253) boils down to

$$G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi T_{\mu\nu} \tag{255}$$

When writing up these equations Einstein was not aware of the following theorems by Lovelock which show that (253) is the only reasonable tensor equation in which the energy momentum tensor appears as as a source:

Theorem 11. *In four spacetime dimensions, all symmetric tensors $E_{\mu\nu}$ built-out of the metric, its first and its second derivatives, and satisfying identically $\nabla^{\mu}E_{\mu\nu} = 0$ are of the form*

$$E_{\mu\nu} = \alpha G_{\mu\nu} + \beta g_{\mu\nu} \tag{256}$$

with constants α, β.

One might be interested in deriving the Einstein equations from a variational principle. In this context another enlightening result of Lovelock reads:

Theorem 12. *In spacetime dimension four, all coordinate invariant Lagrange functions $\mathscr{L}$ which lead to second order field equations for a metric are of the form*

$$\mathscr{L} = (\alpha R + \beta)\sqrt{|\det g_{\mu\nu}|} \tag{257}$$

where α, β are constants.

2.4 *The Schwarzschild metric: Eddington-Finkelstein extension; Time functions; the black hole; what happens at r = 0;*

2.4.1 *The Schwarzschild metric and Birkhoff's Theorem:*

In Einstein's theory of general relativity, the Schwarzschild metric (also known as the Schwarzschild solution) is an exact solution to the Einstein field equations that describes the gravitational field outside a spherical mass, on the assumption that the electric charge of the mass, angular momentum of the mass, and universal cosmological constant are all zero. The solution is a useful approximation for describing slowly rotating astronomical objects such as many stars and planets, including Earth and the Sun. It was found by Karl Schwarzschild in 1916, and around the same time independently by Johannes Droste, who published his much more complete and modern-looking discussion only four months after Schwarzschild.

According to Birkhoff's theorem, the Schwarzschild metric is the most general spherically symmetric vacuum solution of the Einstein field equations. A Schwarzschild black hole or static black hole is a black hole that has neither electric charge nor angular momentum. A Schwarzschild black hole is described by the Schwarzschild metric, and cannot be distinguished from any other Schwarzschild black hole except by its mass.

The Schwarzschild black hole is characterized by a surrounding spherical boundary, called the event horizon, which is situated at the Schwarzschild radius, often called the radius of a black hole. The boundary is not a physical surface, and a person who fell through the event horizon (before being torn apart by tidal forces), would not notice any physical surface at that position; it is a mathematical surface which is significant in determining the black hole's properties. Any non-rotating and non-charged mass that is smaller than its Schwarzschild radius forms a black hole. The solution of the Einstein field equations is valid for any mass M, so in principle (according to general relativity theory) a Schwarzschild black hole of any mass could exist if conditions became sufficiently favorable to allow for its formation. The Schwarzschild metric is given by

$$g = -\left(1-\frac{2m}{r}\right)dt^2 + \frac{dr^2}{1-\frac{2m}{r}} + r^2 d\Omega^2 \tag{258}$$

where $t \in \mathbb{R}$, $r \neq 2m, 0$ and $d\Omega^2$ denotes the metric of the round unit 2-sphere

$$d\Omega^2 = d\theta^2 + \sin^2\theta d\varphi^2 \tag{259}$$

Birkhoff's theorem now states the following:

Theorem 13 (Birkhoff). *In a vacuum region, away from the set $\{r = 2m\}$, any spherically symmetric metric can locally be written in the Schwarzschild form, for some mass parameter m.*

In view of Birkhoff's theorem, we conclude that the hypothesis of spherical symmetry implies in vacuum, at least locally, the existence of two further symmetries: translation in t and reflections in t (that is $t \mapsto -t$). More precisely, we obtain time translations and time reflections. In the case where $m = 0$ the Schwarzschild metric reduces to the Minkowski metric in spherical coordinates:

$$g|_{m=0} = \eta := -dt^2 + dr^2 + r^2 d\Omega^2 \equiv -dt^2 + dx^2 + dy^2 + dz^2 \tag{260}$$

2.4.2 *What happens at $r = 0$?*

Our Schwarzschild metric has problems when the sets $\{r = 0\}$ and $\{r = 2m\}$ are approached. We start our analysis with the former, which is called a singularity. One can calculate that

$$\underbrace{R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta}}_{K(r)} = \frac{48m^2}{r^6} \tag{261}$$

which shows that the Kretschmann scalar $R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta}$ satisfies

$$\lim_{r\to 0} K(r) = \infty \tag{262}$$

This is true regardless of the sign of m. However, the sign of m still makes a difference: If $m > 0$, then any continuous curve starting in the region $\{r < 2m\}$ has to cross $\{r = 2m\}$, before reaching the exterior world, where r is allowed to grow without bound. But the value $r = 2m$ is not allowed at this stage of our analysis. But when $m < 0$ nothing prevents a continuous curve starting near $\{r = 0\}$ to reach any value of r. What does this even mean? Since $K(r)$ is a scalar, we see that the curvature of the metric grows without bounds when $\{r = 0\}$ is approached independently of the coordinate system used. In other words, there is no coordinate system in which the metric remains twice differentiable (as needed to define the curvature tensor) and in which all components of the Riemann tensor would remain bounded when approaching the set $\{r = 0\}$. Our assumption will, from now on, always be that $m > 0$ unless explicitly stated otherwise.

2.4.3 *Time functions, Time orientation, etc.*

We start with the concept of time orientation. In special relativity this is taken for granted: in coordinates where the Minkowski metric η takes the form

$$\eta = -dt^2 + dx^2 + dy^2 + dz^2 \tag{263}$$

a timelike vector $X^\mu \partial_\mu$ is said to be past-pointing if $X^0 < 0$. Note that this is solely a convention. We will, however, shortly stumble across a situation where such a decision will have to be made, namely when trying to distinguish black holes from white holes. The above special relativistic notion of time orientation carries over to a general manifold as follows:

Definition 7. At every point $p \in \mathscr{M}$ the set of time-like vectors, defined as

$$\left\{X \in T_p\mathscr{M} \mid g(X,X) < 0\right\} \tag{264}$$

splits into two components. One can see this upon noting that we can find a coordinate system near p such that the metric at p coincides with the Minkowski metric. In this coordinate system timelike vectors $X = (X^0, \mathbf{X})$ at p satisfy either $X^0 > |\mathbf{X}|$, or $X^0 < -|\mathbf{X}|$. The time orientation is then defined as the choice of which timelike vectors will be called future or past pointing. If this can be done continuously throughout the manifold, we say that the manifold is time orientable. There are Lorentzian manifolds for which this cannot be done, but as we are in physics we will ignore all mathematical pathologies and assume that we are always in luck to work with time orientable stuff. Such Lorentzian manifolds are called spacetimes. One defines null vectors as nonzero vectors X such that $g(X,X) = 0$. A vector is causal if it is timelike or null. One can likewise talk about past or future directed null or causal vectors. All remaining vectors are called spacelike.

Definition 8 (time function). A function $f \in \mathscr{C}^\infty$ will be called time function if ∇f is everywhere timelike past pointing. A cooordinate, say x^0, is said to be a time coordinate if x^0 is a time function.

For example, $f = t$ on Minkowski spacetime is a time function: indeed,

$$\nabla t = \eta^{\mu\nu}\partial_\mu t \partial_\nu = \eta^{0\nu}\partial_\nu = -\partial_t \tag{265}$$

and

$$\eta(\nabla t, \nabla t) = -1 \tag{266}$$

On the other hand, consider $f = t$ in the Schwarzschild metric: the inverse of the Schwarzschild metric reads

$$g^{\mu\nu}\partial_{\mu\nu} = -\frac{1}{1-\frac{2m}{r}}\partial_t^2 + \left(1-\frac{2m}{r}\right)\partial_r^2 + r^{-2}\left(\partial_\theta^2 + \sin^{-2}(\theta)\partial_\varphi^2\right) \tag{267}$$

and thus

$$\nabla t = g^{\mu\nu}\partial_\mu t \partial_\nu = g^{0\nu}\partial_\nu = -\frac{1}{1-\frac{2m}{r}}\partial_t \tag{268}$$

But then we have

$$g(\nabla t, \nabla t) = \frac{g(\partial_t, \partial_t)}{(1-\frac{2m}{r})^2} = -\frac{1}{1-\frac{2m}{r}} \tag{269}$$

and therefore the length of ∇t is < 0 whenever $r > 2m$ and > 0 whenever $r < 2m$. So we see that t is a time function in the region $\{r > 2m\}$, while it is not a time function on $\{r < 2m\}$. A similar calculation yields $\nabla r = \left(1-\frac{2m}{r}\right)\partial_r$ and therefore

$$g(\nabla r, \nabla r) = \left(1-\frac{2m}{r}\right)^2 g(\partial_r, \partial_r) = 1-\frac{2m}{r} \tag{270}$$

So r is a time function in the region $\{r < 2m\}$. Recall next that a differentiable curve is called timelike if its tangent vector is timelike everywhere. There are obvious corresponding definitions of null, causal, or spacelike curves. Causal curves can further be future directed, or past directed, according to the time orientation of their tangents. A basic axiom of general relativity is that massive physical objects move along timelike future directed curves.

Lemma 3. *Let X be timelike and Y be causal. Then $g(X,Y) < 0$ if both X and Y are consistently time-oriented, while $g(X,Y) > 0$ if they have opposite time orientations.*

Proof. Given a point $p \in \mathscr{M}$ in spacetime, we can find a coordinate system so that $g_{\mu\nu}$ is diagonal at p with entries $(-1,1,...,1)$, and (apparently) in which X is proportional to ∂_0, that is $X = X^0\partial_0$. Then of course

$$g(X,Y) = -X^0Y^0 \tag{271}$$

Since Y is causal we have $Y^0 \neq 0$ and therefore $g(X,Y) < 0$ if X and Y are consistently time-oriented, and $g(X,Y) > 0$ if the opposite is the case. □

With this neat little fact in our toolbox we can prove the following:

Theorem 14. *Time functions are strictly increasing along future directed causal curves.*

Proof. Let γ be a future directed timelike curve and let f be a time function. Then

$$\frac{d(f\circ\gamma)}{ds} = \dot{\gamma}^\mu\partial_\mu f = \dot{\gamma}^\mu g_{\mu\nu}g^{\sigma\nu}\partial_\sigma f = g_{\mu\nu}\dot{\gamma}^\mu\nabla^\nu f = g(\nabla f,\dot{\gamma}) \tag{272}$$

and since ∇f is timelike by assumption and $\dot{\gamma}$ is causal but oppositely time-directed, their scalar product is positive. In other words,

$$\frac{d(f\circ\gamma)}{ds} > 0 \tag{273}$$

and so f is strictly increasing along γ. □

2.5 *The Schwarzschild metric: Stationary observers, the interpretation of m, the Flamm paraboloid*

2.6 *The Kruskal-Szekeres extension of the Schwarzschild metric*

2.7 *The Schwarzschild metric: Conformal Carter-Penrose diagram*

2.8 *The Schwarzschild metric: Geodesics, the interpretation of E, circular timelike geodesics*

2.9 *The Schwarzschild metric: Circular null geodesics, gravitational redshift, weak field light bending*

2.10 *The Schwarzschild metric: Perihelion/periastron precession*

2.11 *The Lie derivative, an axiomatic approach, relation to isometries*

2.11.1 *The Lie derivative - an axiomatic approach*

In differential geometry, the Lie derivative, named after Sophus Lie by Władysław Ślebodziński, evaluates the change of a tensor field (including scalar functions, vector fields and one-forms), along the flow defined by another vector field. This change is coordinate invariant and therefore the Lie derivative is defined on any differentiable manifold. In particular, it doesn't require a predetermined connection or something of the like. We will give a purely algebraic definition of the Lie derivative, completely omitting the geometric interpretation thereof. So fix a vector field $X \in \mathscr{X}$. The algebraic definition for the Lie derivative $\mathscr{L}_X$ of a tensor field follows from the following four axioms:

1. **Axiom 1.** The Lie derivative of a function is equal to the directional derivative of the function:
$$\mathscr{L}_X(f) = X(f) \tag{274}$$
2. **Axiom 2.** The Lie derivative obeys the following version of Leibniz's rule: For any tensor fields S and T, we have
$$\mathscr{L}_X(S \otimes T) = \mathscr{L}_X S \otimes T + S \otimes \mathscr{L}_X T \tag{275}$$
3. **Axiom 3:** The Lie derivative obeys the Leibniz rule with respect to contractions:
$$\mathscr{L}_X(T(Y_1, \ldots, Y_n)) \tag{276}$$
$$= (\mathscr{L}_X T)(Y_1, \ldots, Y_n) + T((\mathscr{L}_X Y_1), \ldots, Y_n) + \cdots + T(Y_1, \ldots, (\mathscr{L}_X Y_n)) \tag{277}$$
4. **Axiom 4:** The Lie derivative commutes with exterior derivative on functions:
$$[\mathscr{L}_X, d] = 0 \tag{278}$$

5. **Axiom 5.** The Lie derivative of a vector field $Y \in \mathscr{X}$ is

$$\mathscr{L}_X Y = [X,Y] \tag{279}$$

6. **Axiom 6.** The Lie derivative of a covector field $\varphi \in \Omega$ is

$$L_X \varphi(Y) = \mathscr{L}_X(\varphi(Y)) - \varphi(\mathscr{L}_X Y) \tag{280}$$

It can be shown that we actually only need Axioms 1-4 in order to define the Lie derivative. Let us check that $\mathscr{L}_X \varphi$, defined in Axiom 6, indeed defines a one-form. In order to show this, it suffices to prove that $\mathscr{L}_X \varphi$ is $\mathscr{C}^\infty$-linear over $\mathscr{X}$. Let $f \in \mathscr{C}^\infty$ and let $Y \in \mathscr{X}$. From

$$\mathscr{L}_X(fY) = [X, fY] = X(f)Y + f\mathscr{L}_X Y \tag{281}$$

we deduce

$$\mathscr{L}_X \varphi(fY) = \mathscr{L}_X(\varphi(fY)) - \varphi(\mathscr{L}_X(fY)) \tag{282}$$

$$= X(f)\varphi(Y) + fX(\varphi(Y)) - X(f)\varphi(Y) - f\varphi(\mathscr{L}_X Y) = f\mathscr{L}_X \varphi(Y) \tag{283}$$

In coordinates we have

$$(\mathscr{L}_X \varphi)_a = \partial_a(\mathscr{L}_X \varphi) = X(\varphi(\partial_a)) - \varphi(\mathscr{L}_X \partial_a) = X(\varphi_a) - \varphi([X^b \partial_b, \partial_a]) \tag{284}$$

$$= X^b \partial_b \varphi_a + \varphi(\partial_a X^b \partial_b) = X^b \partial_b \varphi_a + \varphi_b \partial_a X^b \tag{285}$$

For general tensor fields T we note that

$$\mathscr{L}_X T(\varphi_1, \varphi_2, \ldots, X_1, X_2, \ldots) = X(T(\varphi_1, \varphi_2, \ldots, X_1, X_2, \ldots)) \tag{286}$$

$$-T(\mathscr{L}_X \varphi_1, \varphi_2, \ldots, X_1, X_2, \ldots) - T(\varphi_1, \mathscr{L}_X \varphi_2, \ldots X_1, X_2, \ldots) \tag{287}$$

$$- \ldots - T(\varphi_1, \varphi_2, \ldots, \mathscr{L}_X X_1, X_2, \ldots) - T(\varphi_1, \varphi_2, \ldots, X_1, \mathscr{L}_X X_2, \ldots) - \ldots \tag{288}$$

In particular, we get

$$\mathscr{L}_X T^{b_1 \ldots b_s}_{a_1 \ldots a_p} = X^a \partial_a (T^{b_1 \ldots b_s}_{a_1 \ldots a_p}) + \partial_{a_1} X^a T^{b_1 \ldots b_s}_{a a_2 \ldots a_p} + \ldots \tag{289}$$

$$+\partial_{a_p} X^a T^{b_1 \ldots b_s}_{a_1 \ldots a} - \partial_b X^{b_1} T^{b b_2 \ldots b_s}_{a_1 \ldots a_p} - \ldots - \partial_b X^{b_s} T^{b_1 \ldots b}_{a_1 \ldots a_p} \tag{290}$$

since

$$\mathscr{L}_X \partial_s = -\partial_s X^a \partial_a \qquad \mathscr{L}_X \theta^c = \partial_l X^c \theta^l \tag{291}$$

with $\theta^l := dx^l$.

Theorem 15. *The Lie derivative satisfies*

$$\mathscr{L}_{[X,Y]} = [\mathscr{L}_X, \mathscr{L}_Y] \tag{292}$$

for all $X, Y \in \mathscr{X}$.

Proof. The statement is clear if (292) is evaluated for functions $f \in \mathscr{C}^\infty$. Moreover, for vector fields Z equation (292) is simply the jacobi identity. For $\varphi \in \Omega$ we calculate

$$\mathscr{L}_{[X,Y]}\varphi(Z) = \mathscr{L}_{[X,Y]}(\varphi(Z)) - \varphi(\mathscr{L}_{[X,Y]}Z) \tag{293}$$

$$= [\mathscr{L}_X, \mathscr{L}_Y](\varphi(Z)) - \varphi([\mathscr{L}_X, \mathscr{L}_Y]Z) = [\mathscr{L}_X, \mathscr{L}_Y]\varphi(Z) \tag{294}$$

In particular, by axiom 2, we infer that

$$\mathscr{L}_{[X,Y]}(S \otimes T) = \mathscr{L}_{[X,Y]}S \otimes T + S \otimes \mathscr{L}_{[X,Y]}T \tag{295}$$

$$= [\mathscr{L}_X, \mathscr{L}_Y]S \otimes T + S \otimes [\mathscr{L}_X, \mathscr{L}_Y]T = [\mathscr{L}_X, \mathscr{L}_Y](S \otimes T) \tag{296}$$

for tensor fields S, T. This concludes the proof since a general tensor field T is of the form

$$T = T^{b_1 \dots b_s}_{a_1 \dots a_p} \theta^{a_1} \otimes \dots \otimes \theta^{a_p} \otimes \partial_{b_1} \otimes \dots \otimes \partial_{b_s} \tag{297}$$

□

2.11.2 *Relation to Isometries*

Let $\varphi : \mathscr{M} \to \mathscr{N}$ be a smooth map between smooth manifolds. For a vector $X \in T_p\mathscr{M}$ the pushforward of X under φ is defined by

$$(\varphi_* X)_p(f) := T_p\varphi(X)(f) = X(f \circ \varphi) \tag{298}$$

for an arbitrary $f \in \mathscr{C}^\infty(\mathscr{N})$. $(\varphi_* X)_p$ is of course a derivation on $\mathscr{N}$ at $\varphi(p)$, that is, $(\varphi_* X)_p \in T_{\varphi(p)}\mathscr{N}$. For a vector field X on $\mathscr{M}$ we might think that if we define $\varphi_* X$ in the same spirit as for usual vectors, this would yield a vector field on $\mathscr{N}$. However, this is not the case in general. If a point $y \in \mathscr{N}$ has more than one pre-image, say $\varphi(p_1) = \varphi(p_2)$, then we cannot ensure in general that

$$T_{p_1}\varphi(X(p_1))(f) = T_{p_2}\varphi(X(p_2))(f) \tag{299}$$

which is equivalent to

$$X(p_1)(f \circ \varphi) = X(p_2)(f \circ \varphi) \tag{300}$$

So we might run into the trouble of ambiguity. Nonetheless, if φ is a local diffeomorphism all these problems vanish. Indeed, for $y \in \mathscr{N}$ we may define the vector field $\varphi_* X \in \mathscr{X}(\mathscr{N})$ by

$$\varphi_* X(y) := T_p\varphi(X(p)) \tag{301}$$

with $p := \varphi^{-1}(y)$. Directly from the definition it follows that $(\psi \circ \varphi)_* = \psi_* \circ \varphi_*$ and thus also $(\varphi^{-1})_* = \varphi_*^{-1}$.

2.12 *Isometries, Killing vectors, maximally symmetric space-times*

2.13 *FRWL metrics: Hubble law, cosmological red-shift formula, the red shift-factor z and distance, the deceleration parameter*

2.14 *Einstein equations for a FRWL metric*

REFERENCES

[1] Andreas Cap. *Analysis on manifolds*. Lecture notes, 2021.

[2] Piotr T Chruściel. *Elements of general relativity*. Springer Nature, 2020.

[3] Michael Kunzinger and Roland Steinbauer. *Riemannian Geometry*. Lecture notes, 2020.