# Brief Introduction to Differential Geometry and General Relativity for Cosmologists

ABSTRACT: This is the review material of differential geometry and general relativity for students of Lecture: Cosmology

# Contents

## 1. Definition of basic concepts

This material is devoted to the introduction of basic ideas and concepts that are needed for understanding of Cosmology.

We begin with the context of vector in general relativity. We assign to each point $p$ in space-time set of all possible vectors located at that point; this set is known as the **tangent space** at $p$, or $T_p$. We have to emphasize that these vectors are located at a single point. Let us illustrate this idea on simple picture:
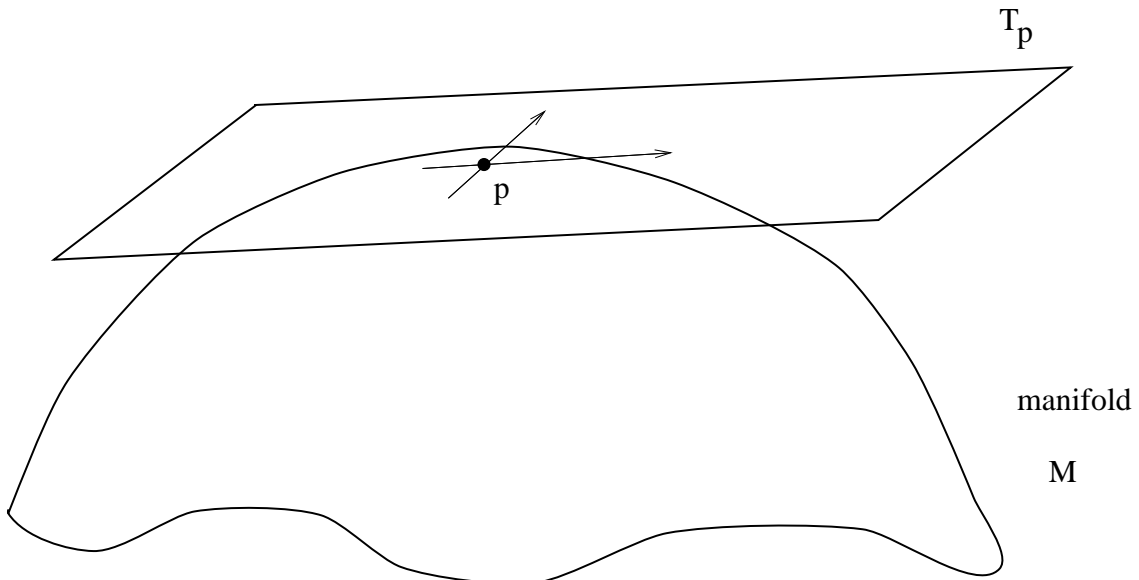


In other words we think of $T_p$ as an abstract vector space for each point in space-time. A **(real) vector space** is a collection of objects ("vectors") which, roughly

speaking, can be added together and multiplied by real numbers in a linear way. Thus, for any two vectors $V$ and $W$ and real numbers $a$ and $b$, we have

$$(a + b)(V + W) = aV + bV + aW + bW . \tag{1.1}$$

Every vector space has an origin, *i.e.* a zero vector which functions as an identity element under vector addition.

It is important to stress that a vector is well-defined geometric object, as is a **vector field**, defined as a set of vectors with exactly one at each point in space-time [1]. However in most of the physical application it is useful to decompose vectors into components with respect to some set of basis vectors. Recall that a **basis** is any set of vectors which both spans the vector space (any vector is a linear combination of basis vectors) and is linearly independent (no vector in the basis is a linear combination of other basis vectors). For any given vector space, there will be an infinite number of legitimate bases, but each basis will consist of the same number of vectors, known as the dimension of the space. Let us presume that at each tangent space we set up a basis of four vectors $e_\mu$, with $\mu \in \{0, 1, 2, 3\}$ as usual. Then any abstract vector $A$ can be written as a linear combination of basis vectors:

$$A = A^\mu e_\mu . \tag{1.2}$$

The coefficients $A^\mu$ are the **components** of the vector $A$. The real vector is an abstract geometrical entity, while the components are just the coefficients of the basis vectors in some convenient basis.

A standard example of a vector in space-time is the tangent vector to a curve. A parameterized curve or path through space-time is specified by the coordinates as a function of the parameter $x^\mu(\lambda)$. The tangent vector $V(\lambda)$ has components

$$V^\mu = \frac{dx^\mu}{d\lambda} . \tag{1.3}$$

Then we denote the entire vector as $V = V^\mu e_\mu$. Once we have set up a vector we can define, known as the **dual vector space**. The dual space is usually denoted by an asterisk, so that the dual space to the tangent space $T_p$ is called the **cotangent space** and denoted $T_p^*$. The dual space is the space of all linear maps from the original vector space to the real numbers. In other words if $\omega \in T_p^*$ is a dual vector, then it acts as a map from $T_p \to R$ such that:

$$\omega(aV + bW) = a\omega(V) + b\omega(W) \in \mathbf{R} , \tag{1.4}$$

where $V$, $W$ are vectors and $a$, $b$ are real numbers. These maps form a vector space themselves; thus, if $\omega$ and $\eta$ are dual vectors, we have

$$(a\omega + b\eta)(V) = a\omega(V) + b\eta(V) . \tag{1.5}$$

---

[1]Mathematically the set of all the tangent spaces of a manifold $M$ is called the **tangent bundle**, $T(M)$.

To proceed we introduce a set of basis dual vectors $\theta^\nu$ by demanding

$$\theta^\nu(e_\mu) = \delta^\nu_\mu .\tag{1.6}$$

Then every dual vector can be written in terms of its components, which we label with lower indices:

$$\omega = \omega_\mu \theta^\nu.\tag{1.7}$$

There is alternative notation when we see elements of $T_p$ (what we have called vectors) referred to as **contravariant vectors**, and elements of $T_p^*$ (what we have called dual vectors) referred to as **covariant vectors**. offended. Another name for dual vectors is **one-forms**, a somewhat mysterious designation which will become clearer soon.

The component notation leads to a simple way of writing the action of a dual vector on a vector:

$$\begin{aligned}\omega(V) &= \omega_\mu V^\nu \theta^\mu(e_\nu) \\ &= \omega_\mu V^\nu \delta^\mu_\nu \\ &= \omega_\mu V^\mu \in \mathbf{R} .\end{aligned}\tag{1.8}$$

The form of (1.35) also suggests that we can think of vectors as linear maps on dual vectors, by defining

$$V(\omega) \equiv \omega(V) = \omega_\mu V^\mu .\tag{1.9}$$

Therefore, the dual space to the dual vector space is the original vector space itself.

(The set of all cotangent spaces over $M$ is the **cotangent bundle**, $T^*(M)$.) In that case the action of a dual vector field on a vector field is not a single number, but a **scalar** (or just "function") on space-time.

In space-time the simplest example of a dual vector is the **gradient** of a scalar function, the set of partial derivatives with respect to the space-time coordinates, which we denote by "d":

$$\mathrm{d}\phi = \frac{d\phi}{dx^\mu}\theta^\mu .\tag{1.10}$$

A straightforward generalization of vectors and dual vectors is the notion of a **tensor**. Just as a dual vector is a linear map from vectors to $\mathbf{R}$, a tensor $T$ of type (or rank) $(k, l)$ is a multilinear map from a collection of dual vectors and vectors to $\mathbf{R}$:

$$\begin{array}{cc}T: & T_p^* \times \cdots \times T_p^* \times T_p \times \cdots \times T_p \to \mathbf{R} \\ & (k \text{ times}) \qquad (l \text{ times})\end{array}\tag{1.11}$$

Here, "$\times$" denotes the Cartesian product, so that for example $T_p \times T_p$ is the space of ordered pairs of vectors. Multilinearity means that the tensor acts linearly in each of its arguments; for instance, for a tensor of type $(1, 1)$, we have

$$T(a\omega + b\eta, cV + dW) = acT(\omega, V) + adT(\omega, W) + bcT(\eta, V) + bdT(\eta, W) .\tag{1.12}$$

3

From this point of view, a scalar is a type $(0,0)$ tensor, a vector is a type $(1,0)$ tensor, and a dual vector is a type $(0,1)$ tensor.

The space of all tensors of a fixed type $(k,l)$ forms a vector space; they can be added together and multiplied by real numbers. In order to construct a basis for this space, we need to define a new operation known as the **tensor product**, denoted by $\otimes$. If $T$ is a $(k,l)$ tensor and $S$ is a $(m,n)$ tensor, we define a $(k+m,l+n)$ tensor $T \otimes S$ by

$$
\begin{aligned}
T \otimes S(\omega^{(1)}, &\ldots, \omega^{(k)}, \ldots, \omega^{(k+m)}, V^{(1)}, \ldots, V^{(l)}, \ldots, V^{(l+n)}) \\
&= T(\omega^{(1)}, \ldots, \omega^{(k)}, V^{(1)}, \ldots, V^{(l)}) S(\omega^{(k+1)}, \ldots, \omega^{(k+m)}, V^{(l+1)}, \ldots, V^{(l+n)}) \;.
\end{aligned}
\tag{1.13}
$$

In other words, first act $T$ on the appropriate set of dual vectors and vectors, and then act $S$ on the remainder, and then multiply the answers. Note that, in general, $T \otimes S \neq S \otimes T$.

Using these rules it is straightforward to construct a basis for the space of all $(k,l)$ tensors. We simply take tensor products of basis vectors and dual vectors. Then this basis will consist of all tensors of the form

$$
e_{\mu_1} \otimes \cdots \otimes e_{\mu_k} \otimes \theta^{\nu_1} \otimes \cdots \otimes \theta^{\nu_l} \;.
\tag{1.14}
$$

In component notation we then write our arbitrary tensor as

$$
T = T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l} e_{\mu_1} \otimes \cdots \otimes e_{\mu_k} \otimes \theta^{\nu_1} \otimes \cdots \otimes \theta^{\nu_l} \;.
\tag{1.15}
$$

The order of the indices is obviously important, since the tensor need not act in the same way on its various arguments.

Now let's turn to some examples of tensors. The most familiar example of a $(0,2)$ tensor in flat Minkowski space-time is the metric, $\eta_{\mu\nu}$. The action of the metric on two vectors is so useful that it gets its own name, the **inner product** (or dot product):

$$
\eta(V,W) = \eta_{\mu\nu} V^\mu W^\nu = V \cdot W \;.
\tag{1.16}
$$

The **norm** of a vector is defined to be inner product of the vector with itself; unlike in Euclidean space, this number is not positive definite: When $\eta_{\mu\nu} V^\mu V^\nu < 0$ then we call $V^\mu$ to be **time-like** for $\eta_{\mu\nu} V^\mu V^\nu = 0$ $V^\mu$ is **null** or **light-like** and for $\eta_{\mu\nu} V^\mu V^\nu > 0$ we call $V^\mu$ as **space-like**.

Another tensor is the Kronecker delta $\delta^\mu_\nu$, of type $(1,1)$, which you already know the components of. Related to this and the metric is the **inverse metric** $\eta^{\mu\nu}$, a type $(2,0)$ tensor defined as the inverse of the metric:

$$
\eta^{\mu\nu} \eta_{\nu\rho} = \eta_{\rho\nu} \eta^{\nu\mu} = \delta^\rho_\mu \;.
\tag{1.17}
$$

There is also the **Levi-Civita tensor**, a $(0,4)$ tensor:

$$
\epsilon_{\mu\nu\rho\sigma} = \begin{cases} +1 & \text{if } \mu\nu\rho\sigma \text{ is an even permutation of } 0123 \\ -1 & \text{if } \mu\nu\rho\sigma \text{ is an odd permutation of } 0123 \\ 0 & \text{otherwise} \;. \end{cases}
\tag{1.18}
$$

Here, a "permutation of 0123" is an ordering of the numbers 0, 1, 2, 3 which can be obtained by starting with 0123 and exchanging two of the digits; an even permutation is obtained by an even number of such exchanges, and an odd permutation is obtained by an odd number. Thus, for example, $\epsilon_{0321} = -1$.

With some examples in hand we can now be a little more systematic about some properties of tensors. First consider the operation of **contraction**, which turns a $(k, l)$ tensor into a $(k - 1, l - 1)$ tensor. Contraction is defined as the sum over one upper and one lower index:

$$S^{\mu\rho}{}_{\sigma} = T^{\mu\nu\rho}{}_{\sigma\nu} \ . \tag{1.19}$$

It is important to stress that we can contract an upper index with a lower index (as opposed to two indices of the same type). It is also important to stress that the order of the indices matters, so that you can get different tensors by contracting in different ways; thus,

$$T^{\mu\nu\rho}{}_{\sigma\nu} \neq T^{\mu\rho\nu}{}_{\sigma\nu} \tag{1.20}$$

in general.

The metric and inverse metric can be used to **raise and lower indices** on tensors. That is, given a tensor $T^{\alpha\beta}{}_{\gamma\delta}$, we can with the help of the metric to define new tensors which we choose to denote by the same letter $T$:

$$
\begin{aligned}
T^{\alpha\beta\mu}{}_{\delta} &= \eta^{\mu\gamma} T^{\alpha\beta}{}_{\gamma\delta} \ , \\
T_{\mu}{}^{\beta}{}_{\gamma\delta} &= \eta_{\mu\alpha} T^{\alpha\beta}{}_{\gamma\delta} \ , \\
T_{\mu\nu}{}^{\rho\sigma} &= \eta_{\mu\alpha} \eta_{\nu\beta} \eta^{\rho\gamma} \eta^{\sigma\delta} T^{\alpha\beta}{}_{\gamma\delta} \ ,
\end{aligned}
\tag{1.21}
$$

Again, it is important that summing does not change the position of an index relative to other indices, and also that "free" indices (which are not summed over) must be the same on both sides of an equation, while "dummy" indices (which are summed over) only appear on one side. As an example, we can turn vectors and dual vectors into each other by raising and lowering indices:

$$
\begin{aligned}
V_{\mu} &= \eta_{\mu\nu} V^{\nu} \\
\omega^{\mu} &= \eta^{\mu\nu} \omega_{\nu} \ .
\end{aligned}
\tag{1.22}
$$

Further we refer to a tensor as **symmetric** in any of its indices if it is unchanged under exchange of those indices. Thus, if

$$S_{\mu\nu\rho} = S_{\nu\mu\rho} \ , \tag{1.23}$$

we say that $S_{\mu\nu\rho}$ is symmetric in its first two indices, while if

$$S_{\mu\nu\rho} = S_{\mu\rho\nu} = S_{\rho\mu\nu} = S_{\nu\mu\rho} = S_{\nu\rho\mu} = S_{\rho\nu\mu} \ , \tag{1.24}$$

we say that $S_{\mu\nu\rho}$ is symmetric in all three of its indices. Similarly, a tensor is **anti-symmetric** (or "skew-symmetric") in any of its indices if it changes sign when those indices are exchanged; thus,

$$A_{\mu\nu\rho} = -A_{\rho\nu\mu} \tag{1.25}$$

means that $A_{\mu\nu\rho}$ is antisymmetric in its first and third indices (or just "antisymmetric in $\mu$ and $\rho$"). If a tensor is (anti-) symmetric in all of its indices, we refer to it as simply (anti-) symmetric (sometimes with the redundant modifier "completely"). As examples, the metric $\eta_{\mu\nu}$ and the inverse metric $\eta^{\mu\nu}$ are symmetric, while the Levi-Civita tensor $\epsilon_{\mu\nu\rho\sigma}$ and the electromagnetic field strength tensor $F_{\mu\nu}$ are antisymmetric.

Given any tensor, we can **symmetrize** (or antisymmetrize) any number of its upper or lower indices. The symmetrization is to defined as the sum of all permutations of the relevant indices and divide by the number of terms:

$$T_{(\mu_1\mu_2\cdots\mu_n)\rho}{}^{\sigma} = \frac{1}{n!} \left(T_{\mu_1\mu_2\cdots\mu_n\rho}{}^{\sigma} + \text{sum over permutations of indices } \mu_1\cdots\mu_n\right) , \tag{1.26}$$

while antisymmetrization comes from the alternating sum:

$$T_{[\mu_1\mu_2\cdots\mu_n]\rho}{}^{\sigma} = \frac{1}{n!} \left(T_{\mu_1\mu_2\cdots\mu_n\rho}{}^{\sigma} + \text{alternating sum over permutations of indices } \mu_1\cdots\mu_n\right) . \tag{1.27}$$

where *alternating sum* we mean that permutations which are the result of an odd number of exchanges are given a minus sign, thus:

$$T_{[\mu\nu\rho]\sigma} = \frac{1}{6} \left(T_{\mu\nu\rho\sigma} - T_{\mu\rho\nu\sigma} + T_{\rho\mu\nu\sigma} - T_{\nu\mu\rho\sigma} + T_{\nu\rho\mu\sigma} - T_{\rho\nu\mu\sigma}\right) . \tag{1.28}$$

Notice that round/square brackets denote symmetrization/antisymmetrization.

There is a special class of tensors that play an important role in physics. These tensors are known as **differential forms** (or just "forms"). *A differential p-form is a $(0,p)$ tensor which is completely antisymmetric.* Thus, scalars are automatically 0-forms, and dual vectors are automatically one-forms. We also have the 2-form $F_{\mu\nu}$ and the 4-form $\epsilon_{\mu\nu\rho\sigma}$. The space of all $p$-forms is denoted $\Lambda^p$, and the space of all $p$-form fields over a manifold $M$ is denoted $\Lambda^p(M)$. The number of linearly independent $p$-forms on an $n$-dimensional vector space is $n!/(p!(n-p)!)$. So at a point on a 4-dimensional space-time there is one linearly independent 0-form, four 1-forms, six 2-forms, four 3-forms, and one 4-form. There are no $p$-forms for $p > n$, since all of the components will automatically be zero by antisymmetry.

Given a $p$-form $A$ and a $q$-form $B$, we can form a $(p+q)$-form known as the **wedge product** $A \wedge B$ by taking the antisymmetrized tensor product:

$$(A \wedge B)_{\mu_1\cdots\mu_{p+q}} = \frac{(p+q)!}{p! \, q!} A_{[\mu_1\cdots\mu_p} B_{\mu_{p+1}\cdots\mu_{p+q}]} . \tag{1.29}$$

For example the wedge product of two 1-forms is

$$(A \wedge B)_{\mu\nu} = 2A_{[\mu}B_{\nu]} = A_\mu B_\nu - A_\nu B_\mu \ . \tag{1.30}$$

Using this definition we obtain an important relation

$$A \wedge B = (-1)^{pq} B \wedge A \ . \tag{1.31}$$

The **exterior derivative** "d" allows us to differentiate $p$-form fields to obtain $(p + 1)$-form fields. It is defined as an appropriately normalized antisymmetric partial derivative:

$$(\mathrm{d}A)_{\mu_1 \cdots \mu_{p+1}} = (p+1)\partial_{[\mu_1} A_{\mu_2 \cdots \mu_{p+1}]} \ . \tag{1.32}$$

The reason why the exterior derivative deserves special attention is that *it is a tensor*, even in curved space-times, unlike its cousin the partial derivative. Another interesting fact about exterior differentiation is that, for any form $A$,

$$d(dA) = 0 \ , \tag{1.33}$$

which is often written $d^2 = 0$. This identity is a consequence of the definition of $d$ and the fact that partial derivatives commute, $\partial_\alpha \partial_\beta = \partial_\beta \partial_\alpha$ (acting on anything).

Let us now introduce another operation on differential forms known as **Hodge duality**. We define the "Hodge star operator" on an $n$-dimensional manifold as a map from $p$-forms to $(n - p)$-forms,

$$(*A)_{\mu_1 \cdots \mu_{n-p}} = \frac{1}{p!} \epsilon^{\nu_1 \cdots \nu_p}{}_{\mu_1 \cdots \mu_{n-p}} A_{\nu_1 \cdots \nu_p} \ , \tag{1.34}$$

mapping $A$ to "$A$ dual". Unlike our other operations on forms, the Hodge dual depends on the metric of the manifold (which should be obvious, since we had to raise some indices on the Levi-Civita tensor in order to define (1.34). Applying the Hodge star twice returns either plus or minus the original form:

$$* * A = (-1)^{s+p(n-p)} A \ , \tag{1.35}$$

where $s$ is the number of minus signs in the eigenvalues of the metric (for Minkowski space, $s = 1$). It is important to stress that the Hodge duality is defined on Manifold where the metric structure is defined.

Now we give brief review of physics in Minkowski space-time that will be useful in what follows. let's review how physics works in Minkowski space-time.

Let us consider the world-line of a single particle. This is specified by a map $\mathbf{R} \to M$, where $M$ is the manifold representing space-time; we usually think of the path as a parameterized curve $x^\mu(\lambda)$. Note that the tangent vector to this path is $dx^\mu/d\lambda$ (note that it depends on the parameterization). Given path is characterized by the norm of the tangent vector. If the tangent vector is timelike/null/spacelike
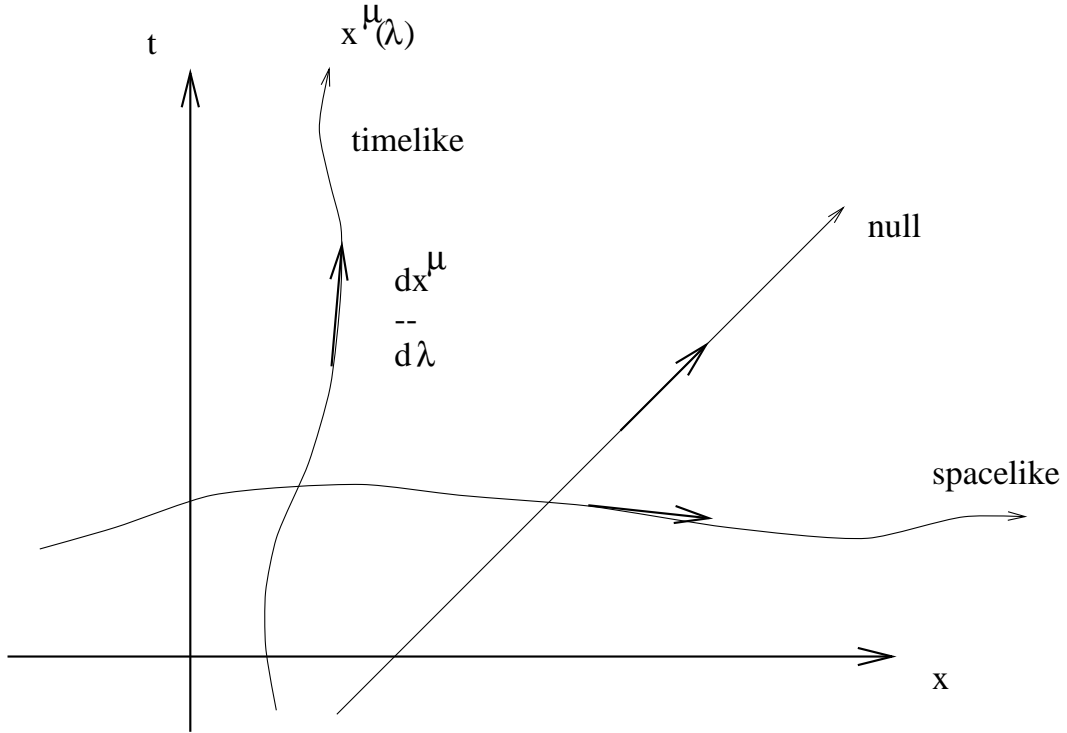
at some parameter value $\lambda$, we say that the path is timelike/null/spacelike at that point. A fundamental object of the theory is the **line element**, or infinitesimal interval:

$$ds^2 = \eta_{\mu\nu}dx^\mu dx^\nu \ . \tag{1.36}$$

From this definition it is tempting to take the square root and integrate along a path to obtain a finite interval. But since $ds^2$ need not be positive, we define different procedures for different cases. For space-like paths we define the **path length**

$$\Delta s = \int \sqrt{\eta_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}} \ d\lambda \ , \tag{1.37}$$

where the integral is taken over the path. For null paths the interval is zero, so no extra formula is required. For time-like paths we define the **proper time**

$$\Delta\tau = \int \sqrt{-\eta_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}} \ d\lambda \ , \tag{1.38}$$

which will be positive.

Let's move from the consideration of paths in general to the paths of massive particles (which will always be time-like). Since the proper time is measured by a clock traveling on a time-like world-line, it is convenient to use $\tau$ as the parameter along the path. That is, we use (1.38) to compute $\tau(\lambda)$, which (if $\lambda$ is a good parameter in the first place) we can invert to obtain $\lambda(\tau)$, after which we can think

of the path as $x^\mu(\tau)$. The tangent vector in this parameterization is known as the **four-velocity**, $U^\mu$:

$$U^\mu = \frac{dx^\mu}{d\tau} \ . \tag{1.39}$$

Since $d\tau^2 = -\eta_{\mu\nu}dx^\mu dx^\nu$ as follows from the invariance of the line element, the four-velocity is automatically normalized:

$$\eta_{\mu\nu}U^\mu U^\nu = -1 \ . \tag{1.40}$$

(It will always be negative, since we are only defining it for time-like trajectories. You could define an analogous vector for space-like paths as well; null paths give some extra problems since the norm is zero.) In the rest frame of a particle, its four-velocity has components $U^\mu = (1, 0, 0, 0)$.

A related vector is the **energy-momentum four-vector**, defined by

$$p^\mu = mU^\mu \ , \tag{1.41}$$

where $m$ is the mass of the particle. The mass is a fixed quantity independent of inertial frame; what you may be used to thinking of as the "rest mass." Although $p^\mu$ provides a complete description of the energy and momentum of a particle, for extended systems it is necessary to go further and define the **energy-momentum tensor** (sometimes called the stress-energy tensor), $T^{\mu\nu}$. This is a symmetric $(2, 0)$ tensor which tells us all we need to know about the energy-like aspects of a system: energy density, pressure, stress, and so forth. A general definition of $T^{\mu\nu}$ is "the flux of four-momentum $p^\mu$ across a surface of constant $x^\nu$". In more details let us consider general form of matter that can be characterized as a **fluid** — a continuum of matter described by macroscopic quantities such as temperature, pressure, entropy, viscosity, etc. In general relativity all interesting types of matter can be thought of as **perfect fluids**, from stars to electromagnetic fields to the entire universe. An alternative definition of a perfect fluid to be one with no heat conduction and no viscosity or it can be defined as a fluid which looks isotropic in its rest frame; these two viewpoints turn out to be equivalent. For our use we can think about a perfect fluid as one which may be completely characterized by its pressure and density.

The most simple example of perfect fluid is **dust**. Dust is defined as a collection of particles at rest with respect to each other, or alternatively as a perfect fluid with zero pressure. Since by definition all particles have an equal velocity in any fixed inertial frame, we can imagine a "four-velocity field" $U^\mu(x)$ defined all over space-time. (Indeed, its components are the same at each point.) Define the **number-flux four-vector** to be

$$N^\mu = nU^\mu \ , \tag{1.42}$$

where $n$ is the number density of the particles as measured in their rest frame. Then $N^0$ is the number density of particles as measured in any other frame, while $N^i$ is the

flux of particles in the $x^i$ direction. Now we can imagine that each of the particles have the same mass $m$. Then in the rest frame the energy density of the dust is given by

$$\rho = nm \ . \tag{1.43}$$

It is important to stress that $\rho$ only measures the energy density in the rest frame. However we want to know the energy density in other frames. To proceed note that both $n$ and $m$ are 0-components of four-vectors in their rest frame. In more details, $N^\mu = (n, 0, 0, 0)$ and $p^\mu = (m, 0, 0, 0)$. Therefore $\rho$ is the $\mu = 0$, $\nu = 0$ component of the tensor $p \otimes N$ as measured in its rest frame. Then it is natural to define the energy-momentum tensor for dust as:

$$T_{\text{dust}}^{\mu\nu} = p^\mu N^\nu = nm U^\mu U^\nu = \rho U^\mu U^\nu \ , \tag{1.44}$$

where $\rho$ is defined as the energy density in the rest frame.

Let us nor return to definition of "perfect" fluid. The natural definition is that it is matter that is "isotropic in its rest frame." This in turn means that $T^{\mu\nu}$ is diagonal — there is no net flux of any component of momentum in an orthogonal direction. Furthermore, due to the isotropy of the perfect fluid in its rest frame the nonzero space-like components must all be equal, $T^{11} = T^{22} = T^{33}$. This fact implies that the only two independent numbers are $T^{00}$ and one of the $T^{ii}$. It is convenient to call the first of these the energy density $\rho$, and the second the pressure $p$. Then the energy-momentum tensor of a perfect fluid therefore takes the following form in its rest frame:

$$T^{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} \ . \tag{1.45}$$

We would like, of course, a formula which is good in any frame. In other words we have to write it in covariant manner. For dust we had $T^{\mu\nu} = \rho U^\mu U^\nu$, so we might begin by guessing $(\rho + p) U^\mu U^\nu$, which gives

$$\begin{pmatrix} \rho + p & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \ . \tag{1.46}$$

using the fact that in the rest frame $U^\mu = (1, 0, 0, 0)$. To get the answer we want we must therefore add

$$\begin{pmatrix} -p & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} \ . \tag{1.47}$$

It turns out however that this matrix has an obvious covariant generalization in the form $p\eta^{\mu\nu}$. Thus, the general form of the energy-momentum tensor for a perfect fluid is

$$T^{\mu\nu} = (\rho + p)U^\mu U^\nu + p\eta^{\mu\nu} \ . \tag{1.48}$$

This is an important formula for applications such as stellar structure and cosmology.

Further examples of the energy-momentum tensors are the energy-momentum tensors of electromagnetism and scalar field theory. We will see their form in the main text.

It is important to stress that $T^{\mu\nu}$ is *conserved* that means vanishing of the "divergence":

$$\partial_\mu T^{\mu\nu} = 0 \ . \tag{1.49}$$

This is a set of four equations, one for each value of $\nu$. The $\nu = 0$ equation corresponds to conservation of energy, while $\partial_\mu T^{\mu k} = 0$ expresses conservation of the $k^{\text{th}}$ component of the momentum.


# 2. Manifolds

In this section we generalize the notion of the flat space to the case of curved space-time. In order to properly understand how it works we have to learn a bit about the mathematics of curved spaces. As the first step we study the notion of manifold. Note that we will work in $n$ dimensions.

A manifold (or sometimes *"differentiable manifold"*) is very important concept in mathematics and physics. One is certainly familiar with the properties of $n$-dimensional Euclidean space, $\mathbf{R}^n$, the set of $n$-tuples $(x^1, \ldots, x^n)$.

We can imagine manifold as a space which may be curved and have a complicated topology, but in local regions looks just like $\mathbf{R}^n$. (Here by "looks like" we do not mean that the metric is the same, but only basic notions of analysis like open sets, functions, and coordinates.) In other words we can imagine that the entire manifold is constructed by smoothly sewing together these local regions. Let us now present more rigorous definition of manifold.

The most elementary notion is that of a **map** between two sets. (We assume you know what a set is.) If we have two sets $M$ and $N$, a map $\phi : M \to N$ is a relationship which assigns, to each element of $M$, exactly one element of $N$. In other words A map is a simple generalization of a function. This fact can be demonstrated on following picture:
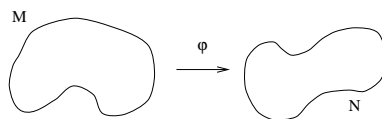
Given two maps $\phi : A \to B$ and $\psi : B \to C$, we define the **composition** $\psi \circ \phi : A \to C$ by the operation $(\psi \circ \phi)(a) = \psi(\phi(a))$. So $a \in A$, $\phi(a) \in B$, and thus $(\psi \circ \phi)(a) \in C$. In pictures:
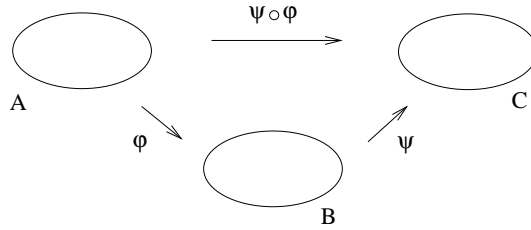


A map $\phi$ is called **one-to-one** (or "injective") if each element of $N$ has at most one element of $M$ mapped into it, and **onto** (or "surjective") if each element of $N$ has at least one element of $M$ mapped into it. (If you think about it, a better name for "one-to-one" would be "two-to-two".)

The set $M$ is known as the **domain** of the map $\phi$, and the set of points in $N$ which $M$ gets mapped into is called the **image** of $\phi$. For some subset $U \subset N$, the set of elements of $M$ which get mapped to $U$ is called the **preimage** of $U$ under $\phi$, or $\phi^{-1}(U)$. A map which is both one-to-one and onto is known as **invertible** (or "bijective"). In this case we can define the **inverse map** $\phi^{-1} : N \to M$ by $(\phi^{-1} \circ \phi)(a) = a$. (Note that the same symbol $\phi^{-1}$ is used for both the preimage and the inverse map, even though the former is always defined and the latter is only defined in some special cases.) Thus:
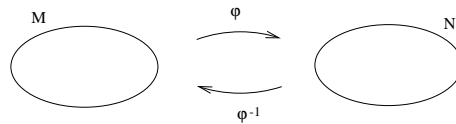


Let us now introduce the notion of **continuity** of a map between topological spaces (and thus manifolds). Luckily the precise mathematical definition is not needed for us so that we can give an intuitive meaning of continuity and differentiability of maps $\phi : \mathbf{R}^m \to \mathbf{R}^n$ between Euclidean spaces are useful. A map from $\mathbf{R}^m$ to $\mathbf{R}^n$ takes an $m$-tuple $(x^1, x^2, \ldots, x^m)$ to an $n$-tuple $(y^1, y^2, \ldots, y^n)$, and can therefore be thought of as a collection of $n$ functions $\phi^i$ of $m$ variables:

$$
\begin{aligned}
y^1 &= \phi^1(x^1, x^2, \ldots, x^m) \\
y^2 &= \phi^2(x^1, x^2, \ldots, x^m) \\
&\quad\cdot \\
&\quad\cdot \\
&\quad\cdot \\
y^n &= \phi^n(x^1, x^2, \ldots, x^m) \ .
\end{aligned}
\tag{2.1}
$$

We will refer to any one of these functions as $C^p$ if it is continuous and $p$-times differentiable, and refer to the entire map $\phi : \mathbf{R}^m \to \mathbf{R}^n$ as $C^p$ if each of its component

functions are at least $C^p$. Thus a $C^0$ map is continuous but not necessarily differentiable, while a $C^\infty$ map is continuous and can be differentiated as many times as you like. $C^\infty$ maps are sometimes called **smooth** and we will consider these functions only. Then we call two sets $M$ and $N$ **diffeomorphic** if there exists a $C^\infty$ map $\phi : M \to N$ with a $C^\infty$ inverse $\phi^{-1} : N \to M$; the map $\phi$ is then called a **diffeomorphism**.

For further purposes we recall the definition of **chain rule**. Let us presume that we have maps $f : \mathbf{R}^m \to \mathbf{R}^n$ and $g : \mathbf{R}^n \to \mathbf{R}^l$, and their the composition $(g \circ f) : \mathbf{R}^m \to \mathbf{R}^l$.
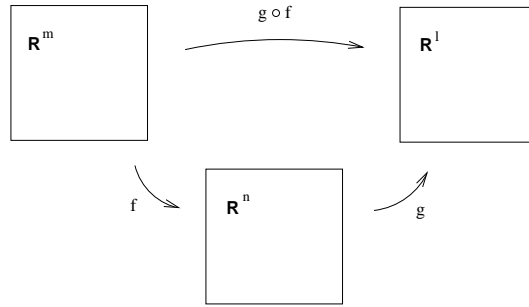


We can represent each space in terms of coordinates: $x^a$ on $\mathbf{R}^m$, $y^b$ on $\mathbf{R}^n$, and $z^c$ on $\mathbf{R}^l$, where the indices range over the appropriate values. The chain rule relates the partial derivatives of the composition to the partial derivatives of the individual maps:

$$\frac{\partial}{\partial x^a}(g \circ f)^c = \sum_b \frac{\partial f^b}{\partial x^a}\frac{\partial g^c}{\partial y^b} \ . \tag{2.2}$$

This relation is usually written as

$$\frac{\partial}{\partial x^a} = \sum_b \frac{\partial y^b}{\partial x^a}\frac{\partial}{\partial y^b} \ . \tag{2.3}$$

It is important that for $m = n$ the determinant of the matrix $\partial y^b / \partial x^a$ is called the **Jacobian** of the map, and the map is called **invertible** whenever the Jacobian is nonzero.

Using these well known definitions we proceed to the definition of manifold. In order to do this we firstly have to define the notion of an open set, on which we can put coordinate systems, and then sew the open sets together in an appropriate way.

We start with the notion of an **open ball**, which is the set of all points $x$ in $\mathbf{R}^n$ such that $|x-y| < r$ for some fixed $y \in \mathbf{R}^n$ and $r \in \mathbf{R}$, where $|x-y| = [\sum_i (x^i - y^i)^2]^{1/2}$. Note that this is a strict inequality — the open ball is the interior of an $n$-sphere of radius $r$ centered at $y$. An **open set** in $\mathbf{R}^n$ is a set constructed from an arbitrary (maybe infinite) union of open balls. In other words, $V \subset \mathbf{R}^n$ is open if, for any $y \in V$, there is an open ball centered at $y$ which is completely inside $V$. Roughly speaking, an open set is the interior of some $(n-1)$-dimensional closed surface (or the
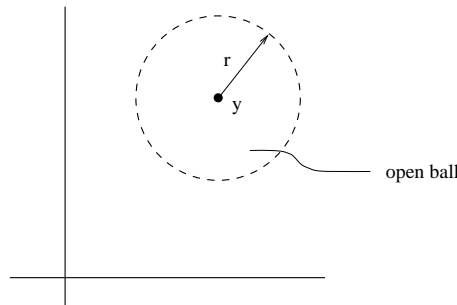
union of several such interiors). By defining a notion of open sets, we have equipped $\mathbf{R}^n$ with a topology — in this case, the "standard metric topology."

A **chart** or **coordinate system** consists of a subset $U$ of a set $M$, along with a one-to-one map $\phi : U \to \mathbf{R}^n$, such that the image $\phi(U)$ is open in $\mathbf{R}$. (Any map is onto its image, so the map $\phi : U \to \phi(U)$ is invertible.) We then can say that $U$ is an open set in $M$. (We have thus induced a topology on $M$, although we will not explore this.)
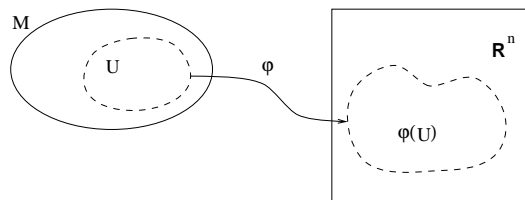


A $C^\infty$ **atlas** is an indexed collection of charts $\{(U_\alpha, \phi_\alpha)\}$ which satisfies two conditions:

1. The union of the $U_\alpha$ is equal to $M$. In other words the $U_\alpha$ cover $M$.

2. The charts are smoothly sewn together. More precisely, if two charts overlap, $U_\alpha \cap U_\beta \neq \emptyset$, then the map $(\phi_\alpha \circ \phi_\beta^{-1})$ takes points in $\phi_\beta(U_\alpha \cap U_\beta) \subset \mathbf{R}^n$ *onto* $\phi_\alpha(U_\alpha \cap U_\beta) \subset \mathbf{R}^n$, and all of these maps must be $C^\infty$ where they are defined. One can see this more clearly from pictures

Now finally, $C^\infty$ *n-dimensional* **manifold** *(or n-manifold for short) is simply a set M along with a "maximal atlas", one that contains every possible compatible chart.* The requirement that the atlas be maximal is so that two equivalent spaces equipped with different atlases don't count as different manifolds. This definition captures in formal terms our notion of a set that looks locally like $\mathbf{R}^n$.

Note that this definition does not rely on an embedding of the manifold in some higher-dimensional Euclidean space. In other words it's important to recognize that the manifold has an individual existence independent of any embedding. In other words there is no reason to believe, for example, that four-dimensional space-time is stuck in some larger space.

these maps are only
defined on the shaded
regions, and must be
smooth there.
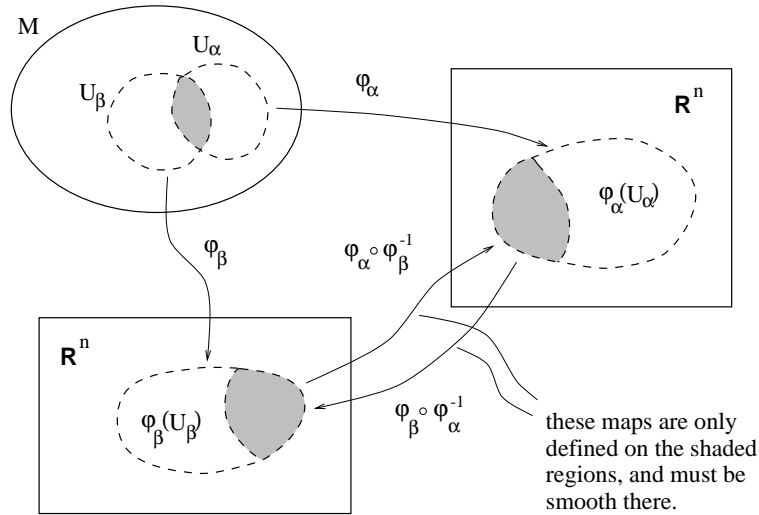
It is important to stress the necessity of charts and atlases: many manifolds cannot be covered with a single coordinate system. The fact that manifolds look locally like $\mathbf{R}^n$, which is most clearly seen from the construction of coordinate charts, introduces the possibility of analysis on manifolds, as for example differentiation and integration.

Let us consider two manifolds $M$ and $N$ of dimensions $m$ and $n$, with coordinate charts $\phi$ on $M$ and $\psi$ on $N$. Imagine we have a function $f : M \to N$,
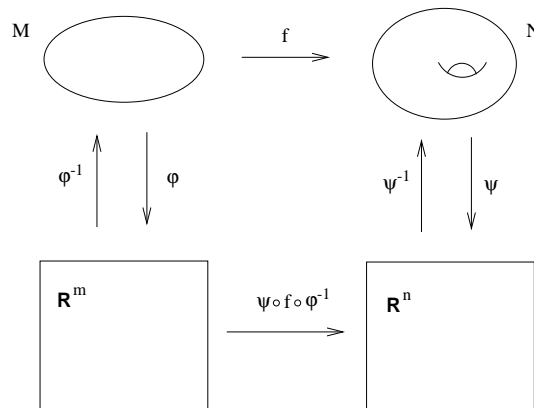


Since $M$ and $N$ are spaces with no appriory definition of differentiation we cannot nonchalantly differentiate the map $f$, since we don't know what such an operation means. But the coordinate charts allow us to construct the map $(\psi \circ f \circ \phi^{-1})$ : $\mathbf{R}^m \to \mathbf{R}^n$. From definition this is just a map between Euclidean spaces, and all of the concepts of advanced calculus apply. For example $f$, thought of as an $N$-valued function on $M$, can be differentiated to obtain $\partial f / \partial x^\mu$, where the $x^\mu$ represent $\mathbf{R}^m$. More precisely, using the definition above we obtain

$$\frac{\partial f}{\partial x^\mu} \equiv \frac{\partial}{\partial x^\mu} (\psi \circ f \circ \phi^{-1})(x^\mu) \ . \tag{2.4}$$

For most practical purposes the short hand notation will be the appropriate one.

Now we can proceed to the construction of various kinds of structure on manifolds. As the first step we begin with the vectors and tangent spaces. Remember the notion of a tangent space— the set of all vectors at a single point in space-time. The natural definition how to introduce the tangent space is to use our intuitive knowledge that there are objects called "tangent vectors to curves" which belong in the tangent space. Then we can consider the set of all parameterized curves through $p$. In other words the space of all (nondegenerative) maps $\gamma : \mathbf{R} \to M$ such that $p$ is in the image of $\gamma$. Then it is natural to define the tangent space as simply the space of all tangent vectors to these curves at the point $p$. To do this let us consider some coordinate system $x^\mu$ any curve through $p$ defines an element of $\mathbf{R}^n$ specified by the $n$ real numbers $\frac{dx^\mu}{d\lambda}$ (where $\lambda$ is the parameter along the curve), but this map is clearly coordinate-dependent, which is not what we want. In order to find coordinate independent formulation we proceed as follows. We define $\mathcal{F}$ to be the space of all smooth functions on $M$ (that is, $C^\infty$ maps $f : M \to \mathbf{R}$). Each curve through $p$ defines an operator on this space, the directional derivative, which maps $f \to \frac{df}{d\lambda}$ (at $p$). Then we claim that *the tangent space $T_p$ can be identified with the space of directional derivative operators along curves through $p$*. It can be shown that the space of directional derivatives is a vector space, and that it is the vector space we want (it has the same dimensionality as $M$, yields a natural idea of a vector pointing along a certain direction, and so on). In fact let us search a basis for the space. Consider again a coordinate chart with coordinates $x^\mu$. Then there is an obvious set of $n$ directional derivatives at $p$, namely the partial derivatives $\partial_\mu$ at $p$.
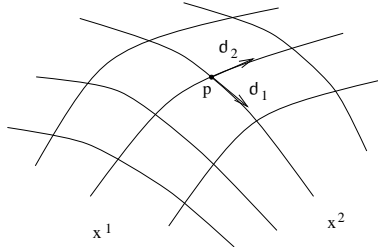


Then the partial derivative operators $\{\partial_\mu\}$ at $p$ form a basis for the tangent space $T_p$. In fact any directional derivative can be decomposed into a sum of real numbers times partial derivatives. To see this let us consider an $n$-manifold $M$, a coordinate chart $\phi : M \to \mathbf{R}^n$, a curve $\gamma : \mathbf{R} \to M$, and a function $f : M \to \mathbf{R}$. Then if $\lambda$ is the parameter along $\gamma$, we want to expand the vector/operator $\frac{d}{d\lambda}$ in terms of the partials $\partial_\mu$. Using the chain rule (2.2), we have

$$\frac{d}{d\lambda}f = \frac{d}{d\lambda}(f \circ \gamma)$$
$$= \frac{d}{d\lambda}[(f \circ \phi^{-1}) \circ (\phi \circ \gamma)]$$

$$= \frac{d(\phi \circ \gamma)^\mu}{d\lambda} \frac{\partial(f \circ \phi^{-1})}{\partial x^\mu}$$

$$= \frac{dx^\mu}{d\lambda} \partial_\mu f \ . \tag{2.5}$$

The first line simply takes the informal expression on the left hand side and rewrites it as an honest derivative of the function $(f \circ \gamma) : \mathbf{R} \to \mathbf{R}$. The second line just comes from the definition of the inverse map $\phi^{-1}$ (and associativity of the operation of composition). The third line is the formal chain rule (2.2), and the last line is a return to the informal notation of the start. Since the function $f$ was arbitrary, we have

$$\frac{d}{d\lambda} = \frac{dx^\mu}{d\lambda} \partial_\mu \ . \tag{2.6}$$

Thus, the partials $\{\partial_\mu\}$ do indeed represent a good basis for the vector space of directional derivatives, which we can therefore safely identify with the tangent space. We can see this in more details on following picture:
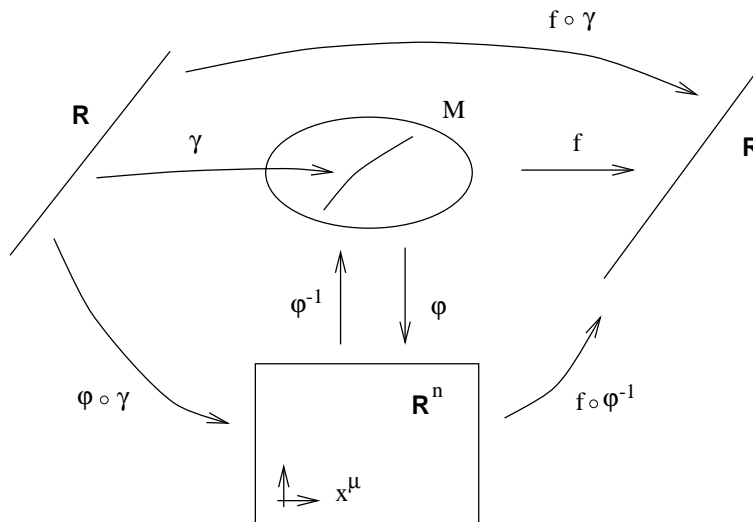


This particular basis $(\hat{e}_{(\mu)} = \partial_\mu)$ is known as a **coordinate basis** for $T_p$. There is no reason why we are limited to coordinate bases when we consider tangent vectors; it is sometimes more convenient, for example, to use orthonormal bases of some sort. However, the coordinate basis is very simple and natural, and we will use it almost exclusively throughout the course.

One of the advantages of this abstract definition of vectors is that the transformation law is immediate. Since the basis vectors are $\hat{e}_{(\mu)} = \partial_\mu$, the basis vectors in some new coordinate system $x^{\mu'}$ are given by the chain rule (2.3) as

$$\partial_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \partial_\mu \ . \tag{2.7}$$

We can get the transformation law for vector components by the same technique used in flat space, demanding the vector $V = V^\mu \partial_\mu$ be unchanged by a change of

basis. We have

$$V^\mu \partial_\mu = V^{\mu'} \partial_{\mu'}$$
$$= V^{\mu'} \frac{\partial x^\mu}{\partial x^{\mu'}} \partial_\mu \ , \tag{2.8}$$

and hence (since the matrix $\partial x^{\mu'}/\partial x^\mu$ is the inverse of the matrix $\partial x^\mu/\partial x^{\mu'}$),

$$V^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu} V^\mu \tag{2.9}$$

Since the basis vectors are usually not written explicitly, the rule (2.13) for transforming components is what we call the "vector transformation law."

As the next step we study the transformation properties of one-forms. Once again the cotangent space $T_p^*$ is the set of linear maps $\omega : T_p \to \mathbf{R}$. The canonical example of a one-form is the gradient of a function $f$, denoted $\mathrm{d}f$. In fact it turns out that the gradients of the coordinate functions $x^\mu$ provide a natural basis for the cotangent space.

$$\mathrm{d}x^\mu(\partial_\nu) = \frac{\partial x^\mu}{\partial x^\nu} = \delta^\mu_\nu \ . \tag{2.10}$$

Therefore the gradients $\{\mathrm{d}x^\mu\}$ are an appropriate set of basis one-forms; an arbitrary one-form is expanded into components as $\omega = \omega_\mu \, \mathrm{d}x^\mu$.

The transformation properties of basis dual vectors and components follow from what is by now the usual procedure. We obtain, for basis one-forms,

$$\mathrm{d}x^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu} \, \mathrm{d}x^\mu \ , \tag{2.11}$$

and for components,

$$\omega_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \omega_\mu \ . \tag{2.12}$$

We will usually write the components $\omega_\mu$ when we speak about a one-form $\omega$.

The transformation law for general tensors follows this same pattern of replacing the Lorentz transformation matrix used in flat space with a matrix representing more general coordinate transformations. A $(k, l)$ tensor $T$ can be expanded

$$T = T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l} \partial_{\mu_1} \otimes \cdots \otimes \partial_{\mu_k} \otimes \mathrm{d}x^{\nu_1} \otimes \cdots \otimes \mathrm{d}x^{\nu_l} \ , \tag{2.13}$$

and under a coordinate transformation the components change according to

$$T^{\mu_1' \cdots \mu_k'}{}_{\nu_1' \cdots \nu_l'} = \frac{\partial x^{\mu_1'}}{\partial x^{\mu_1}} \cdots \frac{\partial x^{\mu_k'}}{\partial x^{\mu_k}} \frac{\partial x^{\nu_1}}{\partial x^{\nu_1'}} \cdots \frac{\partial x^{\nu_l}}{\partial x^{\nu_l'}} T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l} \ . \tag{2.14}$$

Let us demonstrate on following example how to transform forms in practise. Consider a symmetric $(0, 2)$ tensor $S$ on a 2-dimensional manifold, whose components in a coordinate system $(x^1 = x, x^2 = y)$ are given by

$$S_{\mu\nu} = \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} \ . \tag{2.15}$$

or equivalently

$$
\begin{aligned}
S &= S_{\mu\nu}(\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu) \\
&= x(\mathrm{d}x)^2 + (\mathrm{d}y)^2 \ ,
\end{aligned}
\tag{2.16}
$$

where in the last line the tensor product symbols are suppressed for brevity. Now consider new coordinates

$$
\begin{aligned}
x' &= x^{1/3} \\
y' &= e^{x+y} \ .
\end{aligned}
\tag{2.17}
$$

This leads directly to

$$
\begin{aligned}
x &= (x')^3 \\
y &= \ln(y') - (x')^3 \\
\mathrm{d}x &= 3(x')^2 \, \mathrm{d}x' \\
\mathrm{d}y &= \frac{1}{y'} \, \mathrm{d}y' - 3(x')^2 \, \mathrm{d}x' \ .
\end{aligned}
\tag{2.18}
$$

$$
\begin{aligned}
x &= (x')^3 \\
y &= \ln(y') - (x')^3 \\
\mathrm{d}x &= 3(x')^2 \, \mathrm{d}x' \\
\mathrm{d}y &= \frac{1}{y'} \, \mathrm{d}y' - 3(x')^2 \, \mathrm{d}x' \ .
\end{aligned}
\tag{2.19}
$$

We insert these results to the expression above and remember that tensor products don't commute, so $\mathrm{d}x' \, \mathrm{d}y' \neq \mathrm{d}y' \, \mathrm{d}x'$): and we obtain

$$
S = 9(x')^4[1 + (x')^3](\mathrm{d}x')^2 - 3\frac{(x')^2}{y'}(\mathrm{d}x' \, \mathrm{d}y' + \mathrm{d}y' \, \mathrm{d}x') + \frac{1}{(y')^2}(\mathrm{d}y')^2 \ ,
\tag{2.20}
$$

or

$$
S_{\mu'\nu'} = \begin{pmatrix} 9(x')^4[1 + (x')^3] & -3\frac{(x')^2}{y'} \\ -3\frac{(x')^2}{y'} & \frac{1}{(y')^2} \end{pmatrix} \ .
\tag{2.21}
$$

Notice that it is still symmetric. We did not use the transformation law (2.19) directly, but doing so would have yielded the same result, as you can check.

As the next step we study exterior derivative d. The exterior derivative operator d forms an antisymmetric $(0, p + 1)$ tensor when acted on a $p$-form. So the exterior derivative is a legitimate tensor operator; it is not, however, an adequate substitute for the partial derivative, since it is only defined on forms.

The metric tensor in curved space is denoted as $g_{\mu\nu}$ (while $\eta_{\mu\nu}$ is reserved specifically for the Minkowski metric). $g_{\mu\nu}$ is symmetric $(0, 2)$ tensor that is also non-degenerate, meaning that the determinant $g = |g_{\mu\nu}|$ doesn't vanish. This allows us to define the inverse metric $g^{\mu\nu}$ via

$$
g^{\mu\nu} g_{\nu\sigma} = \delta^\mu_\sigma \ .
\tag{2.22}
$$

The symmetry of $g_{\mu\nu}$ implies that $g^{\mu\nu}$ is also symmetric. The metric and its inverse may be used to raise and lower indices on tensors.

The natural object that is directly related to metric tensor is the line element

$$ds^2 = g_{\mu\nu}\, \mathrm{d}x^\mu\, \mathrm{d}x^\nu \ . \tag{2.23}$$

For example, we know that the Euclidean line element in a three-dimensional space with Cartesian coordinates is

$$ds^2 = (\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2 \ . \tag{2.24}$$

We can now change to any coordinate system we choose. For example, in spherical coordinates we have

$$
\begin{aligned}
x &= r\sin\theta\cos\phi \\
y &= r\sin\theta\sin\phi \\
z &= r\cos\theta \ ,
\end{aligned}
\tag{2.25}
$$

which leads directly to

$$ds^2 = \mathrm{d}r^2 + r^2\, \mathrm{d}\theta^2 + r^2\sin^2\theta\, \mathrm{d}\phi^2 \ . \tag{2.26}$$

Obviously the components of the metric look different than those in Cartesian coordinates, but all of the properties of the space remain unaltered.

A good example of a space with curvature is the two-sphere, which can be thought of as the locus of points in $\mathbf{R}^3$ at distance 1 from the origin. The metric in the $(\theta, \phi)$ coordinate system comes from setting $r = 1$ and $\mathrm{d}r = 0$ in (2.32):

$$ds^2 = \mathrm{d}\theta^2 + \sin^2\theta\, \mathrm{d}\phi^2 \ . \tag{2.27}$$

The metric tensor contains all the information we need to describe the curvature of the manifold. In Minkowski space we can choose coordinates in which the components of the metric are constant; but it should be clear that the existence of curvature is more subtle than having the metric depend on the coordinates, since in the example above we showed how the metric in flat Euclidean space in spherical coordinates is a function of $r$ and $\theta$. Later, we shall see that constancy of the metric components is sufficient for a space to be flat, and in fact there always exists a coordinate system on any flat space in which the metric is constant.

A useful characterization of the metric is obtained by putting $g_{\mu\nu}$ into its **canonical form**. In this form the metric components become

$$g_{\mu\nu} = \ \mathrm{diag}\ (-1, -1, \ldots, -1, +1, +1, \ldots, +1, 0, 0, \ldots, 0) \ , \tag{2.28}$$

where "diag" means a diagonal matrix with the given elements. If $n$ is the dimension of the manifold, $s$ is the number of $+1$'s in the canonical form, and $t$ is the number

of $-1$'s, then $s - t$ is the **signature** of the metric (the difference in the number of minus and plus signs), and $s + t$ is the **rank** of the metric (the number of nonzero eigenvalues). If a metric is continuous, the rank and signature of the metric tensor field are the same at every point, and if the metric is nondegenerative the rank is equal to the dimension $n$. We will always deal with continuous, nondegenerative metrics. If all of the signs are positive ($t = 0$) the metric is called **Euclidean** or **Riemannian** (or just "positive definite"), while if there is a single minus ($t = 1$) it is called **Lorentzian** or **pseudo-Riemannian**, and any metric with some $+1$'s and some $-1$'s is called "indefinite." (So the word "Euclidean" sometimes means that the space is flat, and sometimes doesn't, but always means that the canonical form is strictly positive; the terminology is unfortunate but standard.) The space-times of interest in general relativity have Lorentzian metrics.

It can be shown that it is always possible to put the metric into canonical form at some point $p \in M$, but in general it will only be possible at that single point, not in any neighborhood of $p$.

We will now define the **Levi-Civita symbol** to be exactly this $\tilde{\epsilon}_{\mu_1\mu_2\cdots\mu_n}$ — that is, an object with $n$ indices which has the components specified above *in any coordinate system*. This is called a "symbol," of course, because it is not a tensor; it is defined not to change under coordinate transformations. On the other hand we can define the Levi-Civita tensor as

$$\epsilon_{\mu_1\mu_2\cdots\mu_n} = \sqrt{|g|}\, \tilde{\epsilon}_{\mu_1\mu_2\cdots\mu_n} \ . \tag{2.29}$$

It is this tensor which is used in the definition of the Hodge dual, (1.87), which is otherwise unchanged when generalized to arbitrary manifolds. Since this is a real tensor, we can raise indices, etc.

One final appearance of tensor densities is in integration on manifolds. In ordinary calculus on $\mathbf{R}^n$ the volume element $d^n x$ picks up a factor of the Jacobian under change of coordinates:

$$d^n x' = \left| \frac{\partial x^{\mu'}}{\partial x^\mu} \right| d^n x \ . \tag{2.30}$$

There is actually a beautiful explanation of this formula from the point of view of differential forms, which arises from the following fact: *on an n-dimensional manifold, the integrand is properly understood as an n-form.* To see how this works, we should make the identification

$$d^n x \leftrightarrow \mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^{n-1} \ . \tag{2.31}$$

The expression on the right hand side can be misleading, because it looks like a tensor (an $n$-form, actually) but is really a density. Certainly if we have two functions $f$ and $g$ on $M$, then $\mathrm{d}f$ and $\mathrm{d}g$ are one-forms, and $\mathrm{d}f \wedge \mathrm{d}g$ is a two-form. To see this

let us see how (2.31) changes under coordinate transformations. First notice that the definition of the wedge product allows us to write

$$\mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^{n-1} = \frac{1}{n!}\tilde{\epsilon}_{\mu_1\cdots\mu_n}\,\mathrm{d}x^{\mu_1}\wedge\cdots\wedge\mathrm{d}x^{\mu_n}\ , \qquad (2.32)$$

since both the wedge product and the Levi-Civita symbol are completely antisymmetric. Under a coordinate transformation $\tilde{\epsilon}_{\mu_1\cdots\mu_n}$ stays the same while the one-forms change according to their transformation rules leading to

$$\tilde{\epsilon}_{\mu_1\cdots\mu_n}\,\mathrm{d}x^{\mu_1}\wedge\cdots\wedge\mathrm{d}x^{\mu_n} = \tilde{\epsilon}_{\mu_1\cdots\mu_n}\frac{\partial x^{\mu_1}}{\partial x^{\mu_1'}}\cdots\frac{\partial x^{\mu_n}}{\partial x^{\mu_n'}}\,\mathrm{d}x^{\mu_1'}\wedge\cdots\wedge\mathrm{d}x^{\mu_n'}$$

$$= \left|\frac{\partial x^\mu}{\partial x^{\mu'}}\right|\tilde{\epsilon}_{\mu_1'\cdots\mu_n'}\,\mathrm{d}x^{\mu_1'}\wedge\cdots\wedge\mathrm{d}x^{\mu_n'}\ . \qquad (2.33)$$

Multiplying by the Jacobian on both sides recovers (2.30).

It is clear that the naive volume element $d^n x$ transforms as a density, not a tensor, but it is straightforward to construct an invariant volume element by multiplying by $\sqrt{|g|}$:

$$\sqrt{|g'|}\,\mathrm{d}x^{0'}\wedge\cdots\wedge\mathrm{d}x^{(n-1)'} = \sqrt{|g|}\,\mathrm{d}x^0\wedge\cdots\wedge\mathrm{d}x^{n-1}\ , \qquad (2.34)$$

which is of course just $(n!)^{-1}\epsilon_{\mu_1\cdots\mu_n}\,\mathrm{d}x^{\mu_1}\wedge\cdots\wedge\mathrm{d}x^{\mu_n}$. In the interest of simplicity we will usually write the volume element as $\sqrt{|g|}\,d^n x$, rather than as the explicit wedge product $\sqrt{|g|}\,\mathrm{d}x^0\wedge\cdots\wedge\mathrm{d}x^{n-1}$; it will be enough to keep in mind that it's supposed to be an $n$-form.

We finish this section with the introduction Stokes's theorem. Imagine that we have an $n$-manifold $M$ with boundary $\partial M$, and an $(n-1)$-form $\omega$ on $M$. (We haven't discussed manifolds with boundaries, but the idea is obvious; $M$ could for instance be the interior of an $(n-1)$-dimensional closed surface $\partial M$.) Then $\mathrm{d}\omega$ is an $n$-form, which can be integrated over $M$, while $\omega$ itself can be integrated over $\partial M$. Stokes's theorem is then

$$\int_M \mathrm{d}\omega = \int_{\partial M} \omega\ . \qquad (2.35)$$

You can convince yourself that different special cases of this theorem include not only the fundamental theorem of calculus, but also the theorems of Green, Gauss, and Stokes, familiar from vector calculus in three dimensions.

Now we introduce a few extra mathematical techniques. Let us now discuss the problem how different maps between two manifolds $M$ and $N$ carry along tensor fields from one manifold to another. We therefore consider two manifolds $M$ and $N$, possibly of different dimension, with coordinate systems $x^\mu$ and $y^\alpha$, respectively. We imagine that we have a map $\phi : M \to N$ and a function $f : N \to \mathbf{R}$.

It is obvious that we can compose $\phi$ with $f$ to construct a map $(f \circ \phi) : M \to \mathbf{R}$, which is simply a function on $M$. Such a construction is sufficiently useful that it gets its own name; we define the **pullback** of $f$ by $\phi$, denoted $\phi_* f$, by

$$\phi_* f = (f \circ \phi)\ . \qquad (2.36)$$
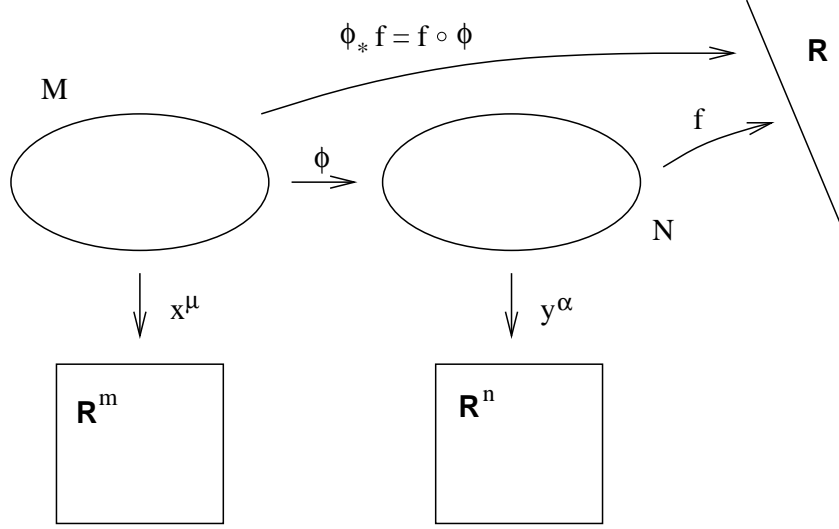
The name makes sense, since we think of $\phi_*$ as "pulling back" the function $f$ from $N$ to $M$.

We can pull functions back, but we cannot push them forward. If we have a function $g : M \to \mathbf{R}$, there is no way we can compose $g$ with $\phi$ to create a function on $N$; the arrows don't fit together correctly. But recall that a vector can be thought of as a derivative operator that maps smooth functions to real numbers. This allows us to define the **pushforward** of a vector; if $V(p)$ is a vector at a point $p$ on $M$, we define the pushforward vector $\phi^* V$ at the point $\phi(p)$ on $N$ by giving its action on functions on $N$:

$$(\phi^* V)(f) = V(\phi_* f) \; . \tag{2.37}$$

So to push forward a vector field we say "the action of $\phi^* V$ on any function is simply the action of $V$ on the pullback of that function."

Let us now give more concrete description. We know that a basis for vectors on $M$ is given by the set of partial derivatives $\partial_\mu = \frac{\partial}{\partial x^\mu}$, and a basis on $N$ is given by the set of partial derivatives $\partial_\alpha = \frac{\partial}{\partial y^\alpha}$. Therefore we would like to relate the components of $V = V^\mu \partial_\mu$ to those of $(\phi^* V) = (\phi^* V)^\alpha \partial_\alpha$. We can find the sought-after relation by applying the pushed-forward vector to a test function and using the chain rule (2.3):

$$(\phi^* V)^\alpha \partial_\alpha f = V^\mu \partial_\mu (\phi_* f) = V^\mu \partial_\mu (f \circ \phi) = V^\mu \frac{\partial y^\alpha}{\partial x^\mu} \partial_\alpha f \; . \tag{2.38}$$

This result shows that the pushforward operation $\phi^*$ can be considered as a matrix operator, $(\phi^* V)^\alpha = (\phi^*)^\alpha{}_\mu V^\mu$, with the matrix being given by

$$(\phi^*)^\alpha{}_\mu = \frac{\partial y^\alpha}{\partial x^\mu} \; . \tag{2.39}$$

23

The behavior of a vector under a pushforward has similar form as the vector transformation law under change of coordinates. However generally, when $\mu$ and $\alpha$ have different allowed values, then there is no reason for the matrix $\partial y^\alpha / \partial x^\mu$ to be invertible.

Let us now discuss transformation properties of one-forms. Since one-forms are dual to vectors then it is natural to expect that one-forms can be pulled back (but not in general pushed forward). In fact, we remember that one-forms are linear maps from vectors to the real numbers. The pullback $\phi_*\omega$ of a one-form $\omega$ on $N$ can therefore be defined by its action on a vector $V$ on $M$, by equating it with the action of $\omega$ itself on the pushforward of $V$:

$$(\phi_*\omega)(V) = \omega(\phi^* V) . \tag{2.40}$$

Once again, there is a matrix description of the pullback operator on forms, $(\phi_*\omega)_\mu = (\phi_*)_\mu{}^\alpha \omega_\alpha$, which we can derive using the chain rule. It is given by

$$(\phi_*)_\mu{}^\alpha = \frac{\partial y^\alpha}{\partial x^\mu} . \tag{2.41}$$

That is, it is the same matrix as the pushforward (2.39).

Let us now consider $(0, l)$ tensor — one with $l$ lower indices and no upper ones. Recall that it is a linear map from the direct product of $l$ vectors to $\mathbf{R}$. Then it is natural to pull back not only one-forms, but tensors with an arbitrary number of lower indices. The definition is simply the action of the original tensor on the pushed-forward vectors:

$$(\phi_* T)(V^{(1)}, V^{(2)}, \ldots, V^{(l)}) = T(\phi^* V^{(1)}, \phi^* V^{(2)}, \ldots, \phi^* V^{(l)}) , \tag{2.42}$$

where $T_{\alpha_1 \cdots \alpha_l}$ is a $(0, l)$ tensor on $N$. We can similarly push forward any $(k, 0)$ tensor $S^{\mu_1 \cdots \mu_k}$ by acting it on pulled-back one-forms:

$$(\phi^* S)(\omega^{(1)}, \omega^{(2)}, \ldots, \omega^{(k)}) = S(\phi_* \omega^{(1)}, \phi_* \omega^{(2)}, \ldots, \phi_* \omega^{(k)}) . \tag{2.43}$$
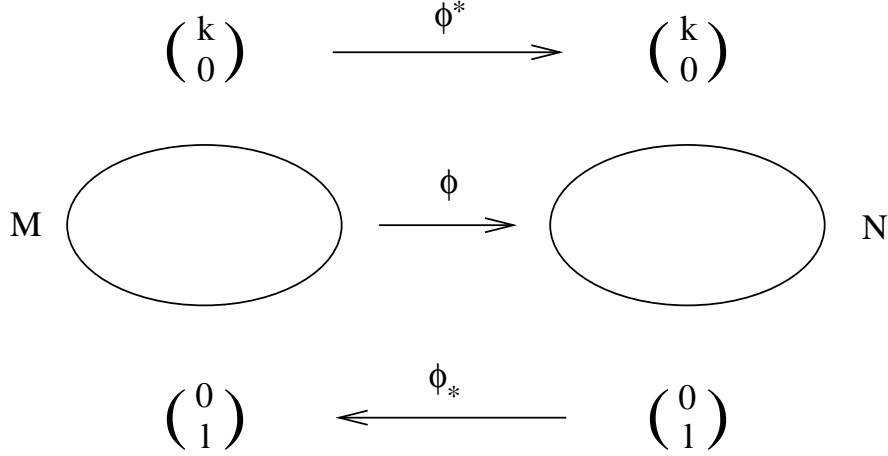
Fortunately, the matrix representations of the pushforward (2.39) and pullback (2.41) extend to the higher-rank tensors simply by assigning one matrix to each index; thus, for the pullback of a $(0, l)$ tensor, we have

$$(\phi_* T)_{\mu_1 \cdots \mu_l} = \frac{\partial y^{\alpha_1}}{\partial x^{\mu_1}} \cdots \frac{\partial y^{\alpha_l}}{\partial x^{\mu_l}} T_{\alpha_1 \cdots \alpha_l} , \tag{2.44}$$

while for the pushforward of a $(k, 0)$ tensor we have

$$(\phi^* S)^{\alpha_1 \cdots \alpha_k} = \frac{\partial y^{\alpha_1}}{\partial x^{\mu_1}} \cdots \frac{\partial y^{\alpha_k}}{\partial x^{\mu_k}} S^{\mu_1 \cdots \mu_k} . \tag{2.45}$$

Our complete picture is therefore:

$$\begin{pmatrix} k \\ 0 \end{pmatrix} \xrightarrow{\ \ \phi^* \ \ } \begin{pmatrix} k \\ 0 \end{pmatrix}$$

$$M \qquad \xrightarrow{\ \ \phi \ \ } \qquad N$$

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \xleftarrow{\ \ \phi_* \ \ } \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

It is important to stress that tensors with both upper and lower indices can generally be neither pushed forward nor pulled back. On the other hand it is important to stress that if $\phi$ is invertible (and both $\phi$ and $\phi^{-1}$ are smooth, which we always implicitly assume), then it defines a diffeomorphism between $M$ and $N$. In this case $M$ and $N$ are the same abstract manifold. The of diffeomorphism is that we can use both $\phi$ and $\phi^{-1}$ to move tensors from $M$ to $N$; this will allow us to define the pushforward and pullback of arbitrary tensors. Specifically, for a $(k,l)$ tensor field $T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \mu_l}$ on $M$, we define the pushforward by

$$(\phi^* T)(\omega^{(1)}, \ldots, \omega^{(k)}, V^{(1)}, \ldots, V^{(l)}) = T(\phi_* \omega^{(1)}, \ldots, \phi_* \omega^{(k)}, [\phi^{-1}]^* V^{(1)}, \ldots, [\phi^{-1}]^* V^{(l)}) , \tag{2.46}$$

where the $\omega^{(i)}$'s are one-forms on $N$ and the $V^{(i)}$'s are vectors on $N$. In components this becomes

$$(\phi^* T)^{\alpha_1 \cdots \alpha_k}{}_{\beta_1 \cdots \beta_l} = \frac{\partial y^{\alpha_1}}{\partial x^{\mu_1}} \cdots \frac{\partial y^{\alpha_k}}{\partial x^{\mu_k}} \frac{\partial x^{\nu_1}}{\partial y^{\beta_1}} \cdots \frac{\partial x^{\nu_l}}{\partial y^{\beta_l}} T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l} . \tag{2.47}$$

Since $\phi$ is invertible the inverse matrix $\frac{\partial x^\nu}{\partial y^\beta}$ is well defined.

Now we are ready to explain the relationship between diffeomorphism and coordinate transformations. We can interpret diffeomorphism as "active coordinate transformations", while traditional coordinate transformations as "passive."

Since a diffeomorphism allows us to pull back and push forward arbitrary tensors, it provides another way of comparing tensors at different points on a manifold. Let us consider a diffeomorphism $\phi : M \to M$ and a tensor field $T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \mu_l}(x)$, we can form the difference between the value of the tensor at some point $p$ and $\phi_* [T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \mu_l}(\phi(p))]$, its value at $\phi(p)$ pulled back to $p$. This fact suggests that we could define another kind of derivative operator on tensor fields that characterized the rate of change of the tensor as it changes under the diffeomorphism. In order to do this we have to introduce a one-parameter family of diffeomorphism, $\phi_t$. This family can be thought of as a smooth map $\mathbf{R} \times M \to M$, such that for each $t \in \mathbf{R}$

$\phi_t$ is a diffeomorphism and $\phi_s \circ \phi_t = \phi_{s+t}$. Note that this last condition implies that $\phi_0$ is the identity map.

It can be show that one-parameter families of diffeomorphisms arises from vector fields (and vice-versa). Let us study what happens to the point $p$ under the entire family $\phi_t$. In fact it is clear that it describes a curve in $M$. Further the same thing will be true of every point on $M$ and these curves fill the manifold (although there can be degeneracies where the diffeomorphisms have fixed points). Then we define a vector field $V^\mu(x)$ as the set of tangent vectors to each of these curves at every point, evaluated at $t = 0$.

We can proceed in opposite direction. We can define a one-parameter family of diffeomorphisms from any vector field. Given a vector field $V^\mu(x)$, we define the **integral curves** of the vector field to be those curves $x^\mu(t)$ which solve

$$\frac{dx^\mu}{dt} = V^\mu . \tag{2.48}$$

This equation should be interpreted in the opposite sense from our usual way — we are given the vectors, from which we define the curves. The diffeomorphisms $\phi_t$ represents "flow down the integral curves," and the associated vector field is defined as the **generator** of the diffeomorphism.

Let us consider vector field $V^\mu(x)$ so that we have a family of diffeomorphisms parameterized by $t$. Then we can ask the question how fast a tensor changes as we travel down the integral curves. For each $t$ we can define this change as

$$\Delta_t T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \mu_l}(p) = \phi_{t*}[T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \mu_l}(\phi_t(p))] - T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \mu_l}(p) . \tag{2.49}$$

Note that both terms on the right hand side are tensors at $p$.
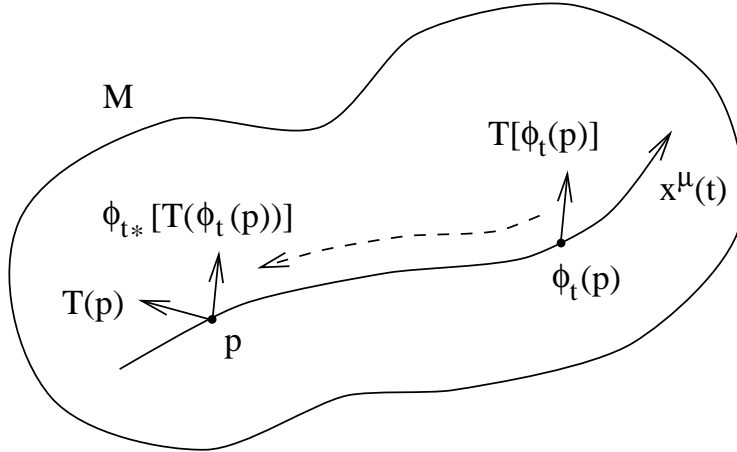


We then define the **Lie derivative** of the tensor along the vector field as

$$\pounds_V T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \mu_l} = \lim_{t \to 0} \left( \frac{\Delta_t T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \mu_l}}{t} \right) . \tag{2.50}$$

The Lie derivative is a map from $(k, l)$ tensor fields to $(k, l)$ tensor fields, which is manifestly independent of coordinates. It is clear this derivative is linear,

$$\pounds_V(aT + bS) = a\pounds_V T + b\pounds_V S \ , \tag{2.51}$$

and obeys the Leibniz rule,

$$\pounds_V(T \otimes S) = (\pounds_V T) \otimes S + T \otimes (\pounds_V S) \ , \tag{2.52}$$

where $S$ and $T$ are tensors and $a$ and $b$ are constants. It is important to stress that definition of Lie derivative does not depend on the metric structure of the manifold. For functions it reduces to the ordinary derivative on functions,

$$\pounds_V f = V(f) = V^\mu \partial_\mu f \ . \tag{2.53}$$

It can be shown that in components the Lie derivative takes the form

$$\pounds_V U^\mu = [V, U]^\mu \ . \tag{2.54}$$

where

$$[V, U]^\mu = V^\nu \partial_\nu U^\mu - U^\nu \partial_\nu V^\mu \tag{2.55}$$

From this definition we immediately see that $\pounds_V S = -\pounds_W V$. It is because of (5.27) that the commutator is sometimes called the "Lie bracket."

To derive the action of $\pounds_V$ on a one-form $\omega_\mu$, begin by considering the action on the scalar $\omega_\mu U^\mu$ for an arbitrary vector field $U^\mu$. First use the fact that the Lie derivative with respect to a vector field reduces to the action of the vector itself when applied to a scalar:

$$\begin{aligned}
\pounds_V(\omega_\mu U^\mu) &= V(\omega_\mu U^\mu) \\
&= V^\nu \partial_\nu(\omega_\mu U^\mu) \\
&= V^\nu(\partial_\nu \omega_\mu)U^\mu + V^\nu \omega_\mu(\partial_\nu U^\mu) \ .
\end{aligned} \tag{2.56}$$

Then use the Leibniz rule on the original scalar:

$$\begin{aligned}
\pounds_V(\omega_\mu U^\mu) &= (\pounds_V \omega)_\mu U^\mu + \omega_\mu(\pounds_V U)^\mu \\
&= (\pounds_V \omega)_\mu U^\mu + \omega_\mu V^\nu \partial_\nu U^\mu - \omega_\mu U^\nu \partial_\nu V^\mu \ .
\end{aligned} \tag{2.57}$$

Setting these expressions equal to each other and requiring that equality hold for arbitrary $U^\mu$, we see that

$$\pounds_V \omega_\mu = V^\nu \partial_\nu \omega_\mu + (\partial_\mu V^\nu)\omega_\nu \ . \tag{2.58}$$

By a similar procedure we can define the Lie derivative of an arbitrary tensor field. The answer can be written

$$\pounds_V T^{\mu_1 \mu_2 \cdots \mu_k}{}_{\nu_1 \nu_2 \cdots \nu_l} = V^\sigma \partial_\sigma T^{\mu_1 \mu_2 \cdots \mu_k}{}_{\nu_1 \nu_2 \cdots \nu_l}$$

$$-(\partial_\lambda V^{\mu_1}) T^{\lambda \mu_2 \cdots \mu_k}{}_{\nu_1 \nu_2 \cdots \nu_l} - (\partial_\lambda V^{\mu_2}) T^{\mu_1 \lambda \cdots \mu_k}{}_{\nu_1 \nu_2 \cdots \nu_l} - \cdots$$
$$+(\partial_{\nu_1} V^\lambda) T^{\mu_1 \mu_2 \cdots \mu_k}{}_{\lambda \nu_2 \cdots \nu_l} + (\partial_{\nu_2} V^\lambda) T^{\mu_1 \mu_2 \cdots \mu_k}{}_{\nu_1 \lambda \cdots \nu_l} + \cdot \quad (2.59)$$

Let us now discuss another aspects of the diffeomorphism.

There is one more use to which we will put the machinery we have set up in this section: symmetries of tensors. We say that a diffeomorphism $\phi$ is a **symmetry** of some tensor $T$ if the tensor is invariant after being pulled back under $\phi$:

$$\phi_* T = T \ . \tag{2.60}$$

It is important to stress that even if discrete symmetry (invariance under reflections) exists it is more common to have a one-parameter family of symmetries $\phi_t$. If the family is generated by a vector field $V^\mu(x)$, then (5.39) leads to

$$\pounds_V T = 0 \ . \tag{2.61}$$

It can be shown that if $T$ is symmetric under some one-parameter family of diffeomorphisms, it is possible to find a coordinate system in which the components of $T$ are all independent of one of the coordinates (the integral curve coordinate of the vector field).

The most important symmetries are those of the metric, for which

$$\phi_* g_{\mu\nu} = g_{\mu\nu} \ . \tag{2.62}$$

A diffeomorphism of this type is called an **isometry**. If a one-parameter family of isometries is generated by a vector field $V^\mu(x)$, then $V^\mu$ is known as a **Killing vector field**. The condition that $V^\mu$ be a Killing vector is thus

$$\pounds_V g_{\mu\nu} = 0 \ , \tag{2.63}$$

or using the relation between Lie derivative and connection

$$\nabla_{(\mu} V_{\nu)} = 0 \ . \tag{2.64}$$

This last version is **Killing's equation**. Then if a space-time has a Killing vector, then we can find a coordinate system in which the metric is independent of one of the coordinates.

It is important to stress that *Killing vectors imply conserved quantities associated with the motion of free particles.* To see this let us consider a motion of a free particle. Then $x^\mu(\lambda)$ is a geodesic with tangent vector $U^\mu = \frac{dx^\mu}{d\lambda}$. Let us denote $V^\mu$ as a Killing vector. Then we have

$$U^\nu \nabla_\nu (V_\mu U^\mu) = U^\nu U^\mu \nabla_\nu V_\mu + V_\mu U^\nu \nabla_\nu U^\mu$$
$$= 0 \ , \tag{2.65}$$

where the first term vanishes from Killing's equation and the second from the fact that $x^\mu(\lambda)$ is a geodesic. Consequently the quantity $V_\mu U^\mu$ is conserved along the particle's world-line. More physically, by definition the metric is unchanging along the direction of the Killing vector. Then we can say that a free particle will not feel any "forces" in this direction, and the component of its momentum in that direction will consequently be conserved.

For next purposes we also give a definition of the concept of a space with maximal symmetry. The **maximally symmetric space** is one which possesses the largest possible number of Killing vectors, which on an $n$-dimensional manifold is $n(n+1)/2$.

It is possible that it may not be simple to actually solve Killing's equation in any given space-time. On the other hand it is frequently possible to write down some Killing vectors by inspection. For example in $\mathbf{R}^2$ with metric $ds^2 = \mathrm{d}x^2 + \mathrm{d}y^2$, independence of the metric components with respect to $x$ and $y$ immediately yields two Killing vectors:

$$
\begin{aligned}
X^\mu &= (1,0) \ , \\
Y^\mu &= (0,1) \ .
\end{aligned}
\tag{2.66}
$$

These clearly represent the two translations. The one rotation would correspond to the vector $R = \partial/\partial\theta$ if we were in polar coordinates; in Cartesian coordinates this becomes

$$
R^\mu = (-y, x) \ .
\tag{2.67}
$$

## 3. Curvature

We know that we can define functions, take their derivatives, consider parameterized paths, set up tensors, and so on any manifolds. On the other hand other concepts, such as the volume of a region or the length of a path, required some additional piece of structure, namely the introduction of a metric. Then we can show how the existence of a metric implies a certain connection, whose curvature may be thought of as that of the metric.

The connection has to be defined when we attempt to address the problem that the partial derivative is not a good tensor operator. Our goal is to introduce the covariant derivative; that is, an operator which reduces to the partial derivative in flat space with Cartesian coordinates, but transforms as a tensor on an arbitrary manifold.

In flat space in Cartesian coordinates, the partial derivative operator $\partial_\mu$ is a map from $(k,l)$ tensor fields to $(k,l+1)$ tensor fields, which acts linearly on its arguments and obeys the Leibniz rule on tensor products. All of this continues to be true in the more general situation we would now like to consider, but the map provided by the partial derivative depends on the coordinate system used. We would therefore

like to define a **covariant derivative** operator $\nabla$ to perform the functions of the partial derivative, but in a way independent of coordinates. We therefore require that $\nabla$ be a map from $(k, l)$ tensor fields to $(k, l+1)$ tensor fields which has these two properties:

1. linearity: $\nabla(T + S) = \nabla T + \nabla S$ ;

2. Leibniz (product) rule: $\nabla(T \otimes S) = (\nabla T) \otimes S + T \otimes (\nabla S)$ .

It can be shown that $\nabla$ takes the form

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda \ , \tag{3.1}$$

where $(\Gamma_\mu)^\rho{}_\sigma$ (an $n \times n$ matrix known as **connection coefficients**, where $n$ is the dimensionality of the manifold, for each $\mu$). Further, we also demand that $\nabla_\mu V^\nu$ transform as $(1, 1)$ tensor

$$\nabla_{\mu'} V^{\nu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \nabla_\mu V^\nu \ . \tag{3.2}$$

This requirement implies following transformation rules for metric coefficients

$$\Gamma^{\nu'}_{\mu'\lambda'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\lambda}{\partial x^{\lambda'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \Gamma^\nu_{\mu\lambda} - \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\lambda}{\partial x^{\lambda'}} \frac{\partial^2 x^{\nu'}}{\partial x^\mu \partial x^\lambda} \ . \tag{3.3}$$

which of course is not the tensor transformation law; the second term on the right spoils it. The covariant derivative of a one-form takes the form

$$\nabla_\mu \omega_\nu = \partial_\mu \omega_\nu - \Gamma^\lambda_{\mu\nu} \omega_\lambda \ . \tag{3.4}$$

Then it is clear that the connection coefficients encode all of the information necessary to take the covariant derivative of a tensor of arbitrary rank. The formula is quite straightforward; for each upper index you introduce a term with a single $+\Gamma$, and for each lower index a term with a single $-\Gamma$:

$$\begin{aligned}
\nabla_\sigma T^{\mu_1 \mu_2 \cdots \mu_k}{}_{\nu_1 \nu_2 \cdots \nu_l} = \ &\partial_\sigma T^{\mu_1 \mu_2 \cdots \mu_k}{}_{\nu_1 \nu_2 \cdots \nu_l} \\
&+ \Gamma^{\mu_1}_{\sigma\lambda} T^{\lambda \mu_2 \cdots \mu_k}{}_{\nu_1 \nu_2 \cdots \nu_l} + \Gamma^{\mu_2}_{\sigma\lambda} T^{\mu_1 \lambda \cdots \mu_k}{}_{\nu_1 \nu_2 \cdots \nu_l} + \cdots \\
&- \Gamma^\lambda_{\sigma\nu_1} T^{\mu_1 \mu_2 \cdots \mu_k}{}_{\lambda \nu_2 \cdots \nu_l} - \Gamma^\lambda_{\sigma\nu_2} T^{\mu_1 \mu_2 \cdots \mu_k}{}_{\nu_1 \lambda \cdots \nu_l} - \cdots \ .
\end{aligned} \tag{3.5}$$

This is the general expression for the covariant derivative.

It turns out that in order to define unique connection on a manifold with a metric $g_{\mu\nu}$ we have to introduce two additional properties:

- torsion-free: $\Gamma^\lambda_{\mu\nu} = \Gamma^\lambda_{(\mu\nu)}$.

- metric compatibility: $\nabla_\rho g_{\mu\nu} = 0$.

This requirement implies that we can express connection coefficients as functions of metric

$$\Gamma^\sigma_{\mu\nu} = \frac{1}{2}g^{\sigma\rho}(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu}) \ . \tag{3.6}$$

This connection we have derived from the metric is the one on which conventional general relativity is based . It is known as **Christoffel** connection or as the **Levi-Civita** connection.

Let us mention once again that the exterior derivative is a well-defined tensor in the absence of any connection. Then if we use a symmetric (torsion-free) connection, the exterior derivative (defined to be the antisymmetrized partial derivative) happens to be equal to the antisymmetrized covariant derivative:

$$\begin{aligned}
\nabla_{[\mu}\omega_{\nu]} &= \partial_{[\mu}\omega_{\nu]} - \Gamma^\lambda_{[\mu\nu]}\omega_\lambda \\
&= \partial_{[\mu}\omega_{\nu]} \ .
\end{aligned} \tag{3.7}$$

Now we define the notion of parallel transport that corresponds to the motion of a vector along a path, keeping constant all the while. The parallel transport is defined whenever we have a connection. In fact the crucial difference between flat and curved spaces is that, in a curved space, *the result of parallel transporting a vector from one point to another will depend on the path taken between the points.* More precisely, let us have a curve $x^\mu(\lambda)$ and define the covariant derivative along the path to be given by an operator

$$\frac{D}{d\lambda} = \frac{dx^\mu}{d\lambda}\nabla_\mu \ . \tag{3.8}$$

We then define **parallel transport** of the tensor $T$ along the path $x^\mu(\lambda)$ to be the requirement that, along the path,

$$\left(\frac{D}{d\lambda}T\right)^{\mu_1\mu_2\cdots\mu_k}{}_{\nu_1\nu_2\cdots\nu_l} \equiv \frac{dx^\sigma}{d\lambda}\nabla_\sigma T^{\mu_1\mu_2\cdots\mu_k}{}_{\nu_1\nu_2\cdots\nu_l} = 0 \ . \tag{3.9}$$

This is a well-defined tensor equation, since both the tangent vector $dx^\mu/d\lambda$ and the covariant derivative $\nabla T$ are tensors. This is known as the **equation of parallel transport**. For a vector it takes the form

$$\frac{d}{d\lambda}V^\mu + \Gamma^\mu_{\sigma\rho}\frac{dx^\sigma}{d\lambda}V^\rho = 0 \ . \tag{3.10}$$

It is clear that the notion of parallel transport depends on the connection, and different connections lead to different answers. On the other hand since we consider the connection that is metric-compatible we obtain that the metric is always parallel transported

$$\frac{D}{d\lambda}g_{\mu\nu} = \frac{dx^\sigma}{d\lambda}\nabla_\sigma g_{\mu\nu} = 0 \ . \tag{3.11}$$

Then we show that the inner product of two parallel-transported vectors is preserved. In fact, if $V^\mu$ and $W^\nu$ are parallel-transported along a curve $x^\sigma(\lambda)$, we have

$$\frac{D}{d\lambda}(g_{\mu\nu}V^\mu W^\nu) = \left(\frac{D}{d\lambda}g_{\mu\nu}\right)V^\mu W^\nu + g_{\mu\nu}\left(\frac{D}{d\lambda}V^\mu\right)W^\nu + g_{\mu\nu}V^\mu\left(\frac{D}{d\lambda}W^\nu\right)$$

$$= 0 \; . \tag{3.12}$$

This means that parallel transport with respect to a metric-compatible connection preserves the norm of vectors, the sense of orthogonality, and so on.

Now we are going to discuss the geodesics. To begin with we recall that the tangent vector to a path $x^\mu(\lambda)$ is $\frac{dx^\mu}{d\lambda}$. The condition that it be parallel transported is thus

$$\frac{D}{d\lambda}\frac{dx^\mu}{d\lambda} = 0 \; , \tag{3.13}$$

or alternatively

$$\frac{d^2 x^\mu}{d\lambda^2} + \Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\lambda}\frac{dx^\sigma}{d\lambda} = 0 \; . \tag{3.14}$$

This is the familiar **geodesic equation**.

It is important to stress that geodesics in general relativity are the paths followed by unaccelerated particles. To see this note that we can think about the geodesic equation as the generalization of Newton's law $\mathbf{f} = m\mathbf{a}$ for the case $\mathbf{f} = 0$. In fact it can be shown that the equation of motion for a particle of mass $m$ and charge $q$ in general relativity takes the form

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau} = \frac{q}{m}F^\mu{}_\nu\frac{dx^\nu}{d\tau} \; . \tag{3.15}$$

An important property of geodesics in a spacetime with Lorentzian metric is that the character (timelike/null/spacelike) of the geodesic (relative to a metric-compatible connection) never changes. This follows from the fact that parallel transport preserves inner products. Consequently the character of the curve does not change since it is determined by the inner product of the tangent vector with itself. There are also null geodesics, which satisfy the same equation, except that the proper time cannot be used as a parameter (some set of allowed parameters will exist, related to each other by linear transformations).

We can also find some interesting application of geodetics. Namely the geodetics can be used for mapping the tangent space at a point $p$ to a local neighborhood of $p$. To begin with notice that any geodesic $x^\mu(\lambda)$ which passes through point $p$ can be specified by its behavior at $p$. We parameterize the geodetics with parameter $\lambda$ and choose to be equal to 0 for $\lambda(p) = 0$. Then the tangent vector at $p$ is

$$\frac{dx^\mu}{d\lambda}(\lambda = 0) = k^\mu \; , \tag{3.16}$$

where $k^\mu$ some vector at $p$ (some element of $T_p$). Then there will be a point on the manifold $M$ which lies on this geodesic where the parameter has the value $\lambda = 1$. We define the **exponential map** at $p$, $\exp_p : T_p \to M$, via

$$\exp_p(k^\mu) = x^\nu(\lambda = 1) \; , \tag{3.17}$$
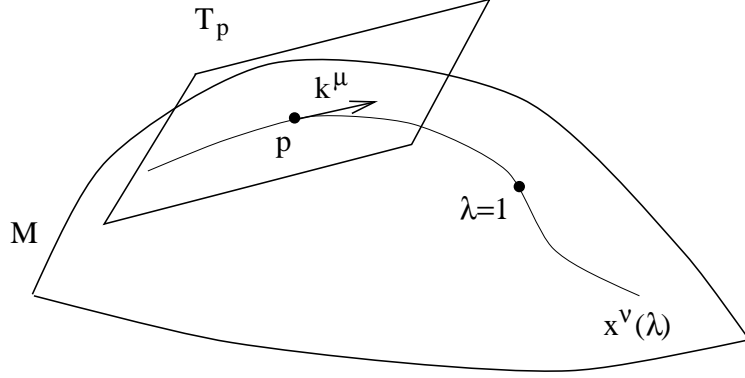
where $x^\nu(\lambda)$ solves the geodesic equation subject to (3.60). For tangent vectors $k^\mu$ near the zero vector, this map will be well-defined, and in fact invertible. Then in the neighborhood of $p$ that is determined by validity of the map on this set of tangent vectors we can say that the tangent vectors themselves define a coordinate system on the manifold. In this coordinate system, any geodesic through $p$ is expressed as

$$x^\mu(\lambda) = \lambda k^\mu \ , \tag{3.18}$$

for some appropriate vector $k^\mu$. Now we will study curvature that is related to the Riemann tensor. that is $(1,3)$ that is antisymmetric in the last two indices:

$$R^\rho{}_{\sigma\mu\nu} = -R^\rho{}_{\sigma\nu\mu} \ . \tag{3.19}$$

We can show that the Riemann tensor is related to covariant derivatives. In fact, let us consider a vector field $V^\rho$, we take

$$
\begin{aligned}
[\nabla_\mu, \nabla_\nu]V^\rho &= \nabla_\mu\nabla_\nu V^\rho - \nabla_\nu\nabla_\mu V^\rho \\
&= \partial_\mu(\nabla_\nu V^\rho) - \Gamma^\lambda_{\mu\nu}\nabla_\lambda V^\rho + \Gamma^\rho_{\mu\sigma}\nabla_\nu V^\sigma - (\mu \leftrightarrow \nu) \\
&= \partial_\mu\partial_\nu V^\rho + (\partial_\mu\Gamma^\rho_{\nu\sigma})V^\sigma + \Gamma^\rho_{\nu\sigma}\partial_\mu V^\sigma - \Gamma^\lambda_{\mu\nu}\partial_\lambda V^\rho - \Gamma^\lambda_{\mu\nu}\Gamma^\rho_{\lambda\sigma}V^\sigma \\
&\quad + \Gamma^\rho_{\mu\sigma}\partial_\nu V^\sigma + \Gamma^\rho_{\mu\sigma}\Gamma^\sigma_{\nu\lambda}V^\lambda - (\mu \leftrightarrow \nu) \\
&= (\partial_\mu\Gamma^\rho_{\nu\sigma} - \partial_\nu\Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma})V^\sigma - 2\Gamma^\lambda_{[\mu\nu]}\nabla_\lambda V^\rho \ . \tag{3.20}
\end{aligned}
$$

The last term is simply the torsion tensor and hence we write

$$[\nabla_\mu, \nabla_\nu]V^\rho = R^\rho{}_{\sigma\mu\nu}V^\sigma - T_{\mu\nu}{}^\lambda\nabla_\lambda V^\rho \ , \tag{3.21}$$

where the Riemann and torsion tensors are identified as

$$R^\rho{}_{\sigma\mu\nu} = \partial_\mu\Gamma^\rho_{\nu\sigma} - \partial_\nu\Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma} \ . \tag{3.22}$$

and

$$T_{\mu\nu}{}^\lambda = \Gamma^\lambda_{\mu\nu} - \Gamma^\lambda_{\nu\mu} \tag{3.23}$$

Since in GR we are mainly interested in Christoffel connection. In this case the connection is derived from the metric and consequently the associated curvature

may be thought of as that of the metric itself. Let us try to show following way. If we are in some coordinate system such that $\partial_\sigma g_{\mu\nu} = 0$ (everywhere, not just at a point), then $\Gamma^\rho_{\mu\nu} = 0$ and $\partial_\sigma \Gamma^\rho_{\mu\nu} = 0$; thus $R^\rho{}_{\sigma\mu\nu} = 0$ by (3.22). However since this is a tensor equation that implies that if it is true in one coordinate system it must be true in any coordinate system. Therefore, the statement that the Riemann tensor vanishes is a necessary condition for it to be possible to find coordinates in which the components of $g_{\mu\nu}$ are constant everywhere. Note that the Riemann tensor obeys **Bianchi identity**:

$$\nabla_{[\lambda} R_{\rho\sigma]\mu\nu} = 0 \ . \tag{3.24}$$

This identity is closely related to the Jacobi identity, since it basically expresses

$$[[\nabla_\lambda, \nabla_\rho], \nabla_\sigma] + [[\nabla_\rho, \nabla_\sigma], \nabla_\lambda] + [[\nabla_\sigma, \nabla_\lambda], \nabla_\rho] = 0 \ . \tag{3.25}$$

It is frequently useful to consider contractions of the Riemann tensor. Firstly we form contraction known as the **Ricci tensor**:

$$R_{\mu\nu} = R^\lambda{}_{\mu\lambda\nu} \ . \tag{3.26}$$

The Ricci tensor associated with the Christoffel connection is symmetric,

$$R_{\mu\nu} = R_{\nu\mu} \ , \tag{3.27}$$

as a consequence of the various symmetries of the Riemann tensor. Using the metric, we can take a further contraction to form the **Ricci scalar**:

$$R = R^\mu{}_\mu = g^{\mu\nu} R_{\mu\nu} \ . \tag{3.28}$$

Another identity related to Riemann tensor is

$$\nabla^\mu R_{\rho\mu} = \frac{1}{2} \nabla_\rho R \ . \tag{3.29}$$

Let us define the **Einstein tensor** as

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \ , \tag{3.30}$$

that obeys

$$\nabla^\mu G_{\mu\nu} = 0 \ . \tag{3.31}$$

The Einstein tensor is very important in general relativity.

Let us illustrate the main ideas and notation on a simple example that is two-sphere, with metric

$$ds^2 = a^2 (d\theta^2 + \sin^2\theta \ d\phi^2) \ , \tag{3.32}$$

where $a$ is the radius of the sphere (thought of as embedded in $\mathbf{R}^3$). It is simple example to calculate the connection coefficients from the metric above and we obtain

$$\Gamma^\theta_{\phi\phi} = -\sin\theta \cos\theta$$

$$\Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = \cot\theta \ . \tag{3.33}$$

and also

$$\begin{aligned}
R^\theta_{\ \phi\theta\phi} &= \partial_\theta \Gamma^\theta_{\phi\phi} - \partial_\phi \Gamma^\theta_{\theta\phi} + \Gamma^\theta_{\theta\lambda}\Gamma^\lambda_{\phi\phi} - \Gamma^\theta_{\phi\lambda}\Gamma^\lambda_{\theta\phi} \\
&= (\sin^2\theta - \cos^2\theta) - (0) + (0) - (-\sin\theta\cos\theta)(\cot\theta) \\
&= \sin^2\theta \ .
\end{aligned} \tag{3.34}$$

It is easy to check that all of the components of the Riemann tensor either vanish or are related to this one by symmetry. We can compute the Ricci tensor and we obtain

$$\begin{aligned}
R_{\theta\theta} &= g^{\phi\phi} R_{\phi\theta\phi\theta} = 1 \\
R_{\theta\phi} &= R_{\phi\theta} = 0 \\
R_{\phi\phi} &= g^{\theta\theta} R_{\theta\phi\theta\phi} = \sin^2\theta \ .
\end{aligned} \tag{3.35}$$

Consequently the Ricci scalar is equal to

$$R = g^{\theta\theta} R_{\theta\theta} + g^{\phi\phi} R_{\phi\phi} = \frac{2}{a^2} \ . \tag{3.36}$$

We see that the Ricci scalar that for a two-dimensional manifold completely characterizes the curvature, is a constant over this two-sphere. This follows from the fact that two-sphere is "maximally symmetric," manifold [2]. In any number of dimensions the curvature of a maximally symmetric space satisfies (for some constant $a$)

$$R_{\rho\sigma\mu\nu} = a^{-2}(g_{\rho\mu}g_{\sigma\nu} - g_{\rho\nu}g_{\sigma\mu}) \ . \tag{3.37}$$

Two-sphere is an example of "positively curved" space-time where the Ricci scalar is positive. We can demonstrate this notation on following example where we can also find meaning of the negative curved space-time.

---

[2] We give precise definition of this notion letter.

positive curvature

negative curvature
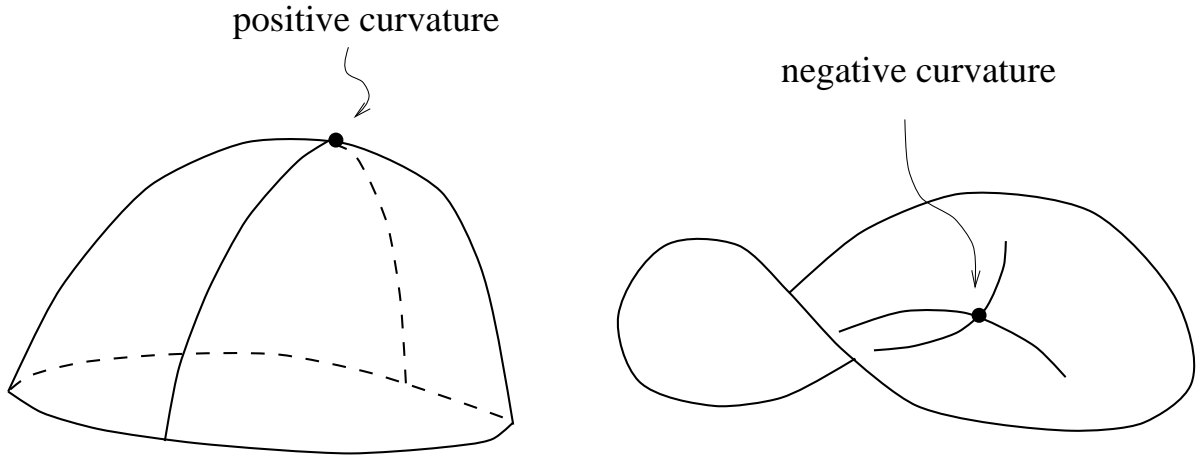
# 4. Gravitation

The main idea of the General theory of relativity is that the spacetime should be described as a curved manifold. In other words the famous Einstein idea is that gravity is a manifestation of spacetime curvature.

Let us again introduce the Einstein tensor

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} \ , \tag{4.1}$$

which always obeys $\nabla^\mu G_{\mu\nu} = 0$. Then the Einstein equation takes the form

$$G_{\mu\nu} = \kappa T_{\mu\nu} \tag{4.2}$$

Note that the right-hand side is a covariant expression of the energy and momentum density in the form of a symmetric and conserved $(0,2)$ tensor, while the left-hand side is a symmetric and conserved $(0,2)$ tensor constructed from the metric and its first and second derivatives.

Let us now contract both sides of (4.2) and we obtain (in four dimensions)

$$R = -\kappa T \ , \tag{4.3}$$

and using this we can rewrite (4.2) as

$$R_{\mu\nu} = \kappa(T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu}) \ . \tag{4.4}$$

This is the same equation, just written slightly differently.

Einstein's equations may be thought of as second-order differential equations for the metric tensor field $g_{\mu\nu}$. There are ten independent equations (since both sides are symmetric two-index tensors), which seems to be exactly right for the ten unknown functions of the metric components. However, the Bianchi identity $\nabla^\mu G_{\mu\nu} = 0$ represents four constraints on the functions $R_{\mu\nu}$, so there are only six

truly independent equations in (4.4). In fact this is appropriate, since if a metric is a solution to Einstein's equation in one coordinate system $x^\mu$ it should also be a solution in any other coordinate system $x^{\mu'}$. This means that there are four unphysical degrees of freedom in $g_{\mu\nu}$ (represented by the four functions $x^{\mu'}(x^\mu)$), and we should expect that Einstein's equations only constrain the six coordinate-independent degrees of freedom.

It is important to stress that as differential equations, these are extremely complicated; the Ricci scalar and tensor are contractions of the Riemann tensor, which involves derivatives and products of the Christoffel symbols, which in turn involve the inverse metric and derivatives of the metric. Furthermore, the energy-momentum tensor $T_{\mu\nu}$ will generally involve the metric as well. The equations are also nonlinear, that implies that two known solutions cannot be superposed to find a third. It is therefore very difficult to solve Einstein's equations in any sort of generality. Then in order to solve them we have to perform some simplifying assumptions. The most popular sort of simplifying assumption is that the metric has a significant degree of symmetry, and we will talk later on about how symmetries of the metric make life easier.

Now we demonstrate how Einstein's equations can be derived from an action principle. The action should be the integral over spacetime of a Lagrange density ("Lagrangian" for short, although strictly speaking the Lagrangian is the integral over space of the Lagrange density):

$$S_H = \int d^n x \mathcal{L}_H \ . \tag{4.5}$$

The Lagrange density is a tensor density, which can be written as $\sqrt{-g}$ times a scalar. What scalars can we make out of the metric? Since we know that the metric can be set equal to its canonical form and its first derivatives set to zero at any one point, any nontrivial scalar must involve at least second derivatives of the metric. The Riemann tensor is of course made from second derivatives of the metric, and we argued earlier that the only independent scalar we could construct from the Riemann tensor was the Ricci scalar $R$. What we did not show, but is nevertheless true, is that any nontrivial tensor made from the metric and its first and second derivatives can be expressed in terms of the metric and the Riemann tensor. Therefore, the only independent scalar constructed from the metric, which is no higher than second order in its derivatives, is the Ricci scalar. Hilbert figured that this was therefore the simplest possible choice for a Lagrangian, and proposed

$$\mathcal{L}_H = \sqrt{-g}R \ . \tag{4.6}$$

The equations of motion should come from varying the action with respect to the metric. In fact let us consider variations with respect to the inverse metric $g^{\mu\nu}$, which are slightly easier but give an equivalent set of equations. Using $R = g^{\mu\nu}R_{\mu\nu}$,

in general we will have

$$\delta S = \int d^n x \left[ \sqrt{-g} g^{\mu\nu} \delta R_{\mu\nu} + \sqrt{-g} R_{\mu\nu} \delta g^{\mu\nu} + R \delta \sqrt{-g} \right]$$
$$= (\delta S)_1 + (\delta S)_2 + (\delta S)_3 . \tag{4.7}$$

The second term $(\delta S)_2$ is already in the form of some expression times $\delta g^{\mu\nu}$; let's examine the others more closely.

Recall that the Ricci tensor is the contraction of the Riemann tensor, which is given by

$$R^\rho_{\ \mu\lambda\nu} = \partial_\lambda \Gamma^\lambda_{\nu\mu} + \Gamma^\rho_{\lambda\sigma} \Gamma^\sigma_{\nu\mu} - (\lambda \leftrightarrow \nu) . \tag{4.8}$$

We perform the variation of the Riemann tensor in such a way that we firstly perform variation of the connection coefficients and then we substitute into this expression. In fact, after some calculations we find the variation of the Riemann tensor in the form

$$\delta R^\rho_{\ \mu\lambda\nu} = \nabla_\lambda (\delta \Gamma^\rho_{\nu\mu}) - \nabla_\nu (\delta \Gamma^\rho_{\lambda\mu}) . \tag{4.9}$$

Therefore, the contribution of the first term in (4.7) to $\delta S$ can be written

$$(\delta S)_1 = \int d^n x \sqrt{-g} \, g^{\mu\nu} \left[ \nabla_\lambda (\delta \Gamma^\lambda_{\nu\mu}) - \nabla_\nu (\delta \Gamma^\lambda_{\lambda\mu}) \right]$$
$$= \int d^n x \sqrt{-g} \, \nabla_\sigma \left[ g^{\mu\sigma} (\delta \Gamma^\lambda_{\lambda\mu}) - g^{\mu\nu} (\delta \Gamma^\sigma_{\mu\nu}) \right] , \tag{4.10}$$

where we have used metric compatibility. However the integral above is an integral with respect to the natural volume element of the covariant divergence of a vector; by Stokes's theorem, this is equal to a boundary contribution at infinity which we can set to zero by making the variation vanish at infinity. Therefore this term does not contribute to the total variation.

In order to calculate the $(\delta S)_3$ term we have to use the variation

$$\delta(g^{-1}) = \frac{1}{g} g_{\mu\nu} \delta g^{\mu\nu} . \tag{4.11}$$

and consequently

$$\delta \sqrt{-g} = -\frac{1}{2} \sqrt{-g} g_{\mu\nu} \delta g^{\mu\nu} . \tag{4.12}$$

If we now return back to (4.7), and remembering that $(\delta S)_1$ does not contribute, we find

$$\delta S = \int d^n x \sqrt{-g} \left[ R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \right] \delta g^{\mu\nu} . \tag{4.13}$$

However this should vanish for arbitrary variations and consequently we derive Einstein's equations in vacuum:

$$\frac{1}{\sqrt{-g}} \frac{\delta S}{\delta g^{\mu\nu}} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 0 . \tag{4.14}$$

However we would like to get the non-vacuum field equations as well. In other words we consider an action of the form

$$S = \frac{1}{8\pi G} S_H + S_M \ , \tag{4.15}$$

where $S_M$ is the action for matter, and we have presciently normalized the gravitational action (although the proper normalization is somewhat convention-dependent). Following through the same procedure as above leads to

$$\frac{1}{\sqrt{-g}} \frac{\delta S}{\delta g^{\mu\nu}} = \frac{1}{8\pi G} \left( R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \right) + \frac{1}{\sqrt{-g}} \frac{\delta S_M}{\delta g^{\mu\nu}} = 0 \ , \tag{4.16}$$

and we recover Einstein's equations if we set

$$T_{\mu\nu} = -\frac{1}{\sqrt{-g}} \frac{\delta S_M}{\delta g^{\mu\nu}} \ . \tag{4.17}$$

In fact (4.17) turns out to be the best way to define a symmetric energy-momentum tensor. We are mainly interested in the existence of solutions to Einstein's equations in the presence of "realistic" sources of energy and momentum. The most common property that is demanded of $T_{\mu\nu}$ is that it represent positive energy densities — no negative masses are allowed. In a locally inertial frame this requirement can be written as $\rho = T_{00} \geq 0$. We write it in the coordinate-independent notation as

$$T_{\mu\nu} V^\mu V^\nu \geq 0 \ , \qquad \text{for all timelike vectors } V^\mu \ . \tag{4.18}$$

This is known as the **Weak Energy Condition**, or WEC. It seems like a reasonable requirement however it is very restrictive. Indeed it is straightforward to show that there are many examples of the classical field theories which violate the WEC, and almost impossible to invent a quantum field theory which obeys it. Nevertheless, it is legitimate to assume that the WEC holds in most cases and it is violated in some extreme conditions. (There are also stronger energy conditions, but they are even less true than the WEC, and we won't dwell on them.)

We continue with the study of the Einstein equations where we now discuss the possibility of the introduction of a cosmological constant. In order to introduce it we add it to the conventional Hilbert action. We therefore consider an action given by

$$S = \int d^n x \sqrt{-g} (R - 2\Lambda) \ , \tag{4.19}$$

where $\Lambda$ is some constant. The resulting field equations are

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu} = 0 \ , \tag{4.20}$$

and of course there would be an energy-momentum tensor on the right hand side if we had included an action for matter. $\Lambda$ is the cosmological constant. In order to

find its meaning it is convenient to move the additional term in (4.20) to the right hand side, and think of it as a kind of energy-momentum tensor, with $T_{\mu\nu} = -\Lambda g_{\mu\nu}$ (it is automatically conserved by metric compatibility). Then $\Lambda$ can be interpreted as the "energy density of the vacuum," a source of energy and momentum that is present even in the absence of matter fields. This interpretation is important because quantum field theory predicts that the vacuum should have some sort of energy and momentum. In ordinary quantum mechanics, an harmonic oscillator with frequency $\omega$ and minimum classical energy $E_0 = 0$ upon quantization has a ground state with energy $E_0 = \frac{1}{2}\hbar\omega$. A quantized field can be thought of as a collection of an infinite number of harmonic oscillators, and each mode contributes to the ground state energy. The result is of course infinite, and must be appropriately regularized, for example by introducing a cutoff at high frequencies. The final vacuum energy, which is the regularized sum of the energies of the ground state oscillations of all the fields of the theory, has no good reason to be zero and in fact would be expected to have a natural scale

$$\Lambda \sim m_P^4 \ , \tag{4.21}$$

where the Planck mass $m_P$ is approximately $10^{19}$ GeV, or $10^{-5}$ grams. Observations of the universe on large scales allow us to constrain the actual value of $\Lambda$, which turns out to be smaller than (4.21) by at least a factor of $10^{120}$. This is the largest known discrepancy between theoretical estimate and observational constraint in physics, and convinces many people that the "cosmological constant problem" is one of the most important unsolved problems today. On the other hand the observations do not tell us that $\Lambda$ is strictly zero, and in fact allow values that can have important consequences for the evolution of the universe.