General Relativity

Part III Michaelmas 2019

Lectures by David Tong

Umut C. Özer
uco21@cam.ac.uk

October 21, 2019

1 Introducing Differential Geometry

Structure: will do only differential geometry (maths methods) in the first 12 lectures, first 4 weeks. Afterwards the connection to the physics will be made. We will develop a mathematical language that allows us to write valid equations involving vectors, similar to dimensional analysis for physicists. The physics will be introduced with the action principle.

Extra book: "Geometry, Topology, and Physics" by Nakahara.

Office Hours: Friday $4 \to 5$ in B2.13

Extra stuff not examinable. Lectures are what matters.

1.1 Manifolds

An *n-dimensional manifold*, $\mathcal{M}$, is a space that locally looks like $\mathbb{R}^n$. More precisely, we require

1. For each point $p \in \mathcal{M}$, there is a map

$$\phi\colon \mathcal{O} \to U \tag{1.1.1}$$

   where $\mathcal{O} \subset \mathcal{M}$ is an open set with $p \in \mathcal{O}$ and $U \subset \mathbb{R}^n$. We will think of $\phi(p) = (x^1(p), \ldots, x^n(p))$ as coordinates on $\mathcal{O} \subset \mathcal{M}$.

   This map must be a *homeomorphism*:

   (a) injective (or 1-1) $p \neq q \implies \phi(p) \neq \phi(q)$

   (b) surjective (or onto) $\phi(\mathcal{O}) = U$ These ensure that ϕ^{-1} exists

   (c) both ϕ and ϕ^{-1} are continuous

2. If $\mathcal{O}_\alpha$ and $\mathcal{O}_\beta$ are two open sets with

$$\phi_\alpha : \mathcal{O}_\alpha \to U_\alpha \text{ and } \phi_\beta : \mathcal{O}_\beta \to U_\beta \tag{1.1.2}$$

   then $\phi_\alpha \circ \phi_\beta^{-1} : \phi_\beta(\mathcal{O}_\alpha \cap \mathcal{O}_\beta) \to \phi_\alpha(\mathcal{O}_\alpha \cap \mathcal{O}_\beta)$ is smooth (ie. infinitely differentiable)

   The ϕ_α are called *charts*. The idea is that there may be different ways to assign coordinates to a given point $p \in \mathcal{M}$. The collection of all charts is called an *atlas*. We require that these coordinate systems are mutually compatible. $\phi_\alpha \circ \phi_\beta^{-1}$ are called *transition functions*.

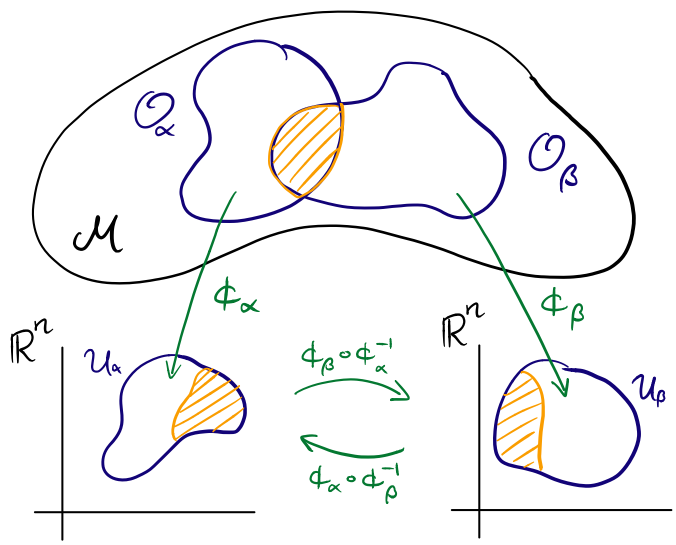

Figure 1.1.1: An illustration of charts on a manifold

Example ($\mathbb{R}^n$)**:** $\mathbb{R}^n$ or any open subset of $\mathbb{R}^n$ is a manifold. You only need a single chart.

Example (S^1)**:** We can view this as $(\cos\theta, \sin\theta) \in \mathbb{R}^2$ with $\theta \in [0, 2\pi)$. The closed set $[0, 2\pi)$ means that we cannot differentiate at 0. This is <u>not</u> a good chart because it is not an open set.

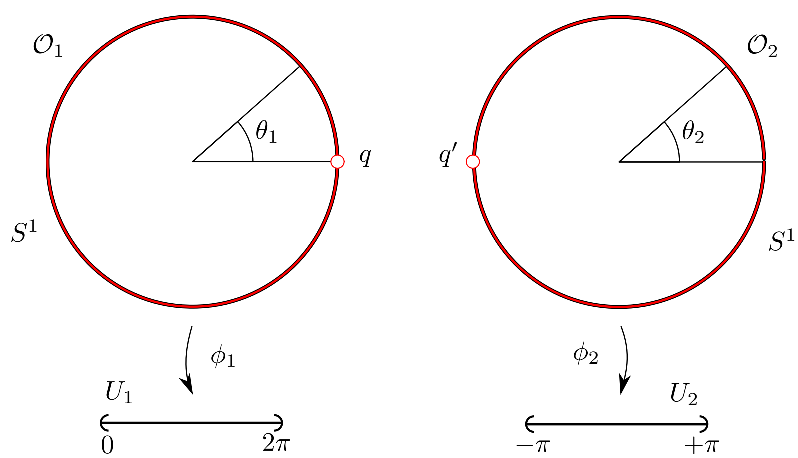

Figure 1.1.2: The two charts ϕ_1, ϕ_2 form an atlas of the manifold S^1.

We need at least two charts, depicted in 1.1.2:

$$\begin{aligned} \phi_1 \colon \mathcal{O}_1 &\to (0, 2\pi) & \qquad \phi_2 \colon \mathcal{O}_2 &\to (-\pi, \pi) \\ p &\mapsto \theta_1 & p &\mapsto \theta_2 \end{aligned} \tag{1.1.3}$$

The transition function is

$$\theta_2 = \phi_2(\phi_1^{-1}(\theta_1)) = \begin{cases} \theta_1, & \theta_1 \in (0, \pi) \\ \theta_1 - 2\pi, & \theta_1 \in (\pi, 2\pi) \end{cases} \tag{1.1.4}$$

Remark: The fact that the coordinates go bad in the case of closed sets, similar to spherical polar coordinates, does not bother us too much for physical applications.

Since we can map $\mathcal{M} \to \mathbb{R}^n$ (at least locally), anything we can do on $\mathbb{R}^n$, we can now also do on $\mathcal{M}$ (e.g. differentiation).

Remark: Note that at the moment, the distance in $\mathbb{R}^n$ cannot be translated back to the manifold. This is because the maps ϕ_α are arbitrary.

Definition 1 (Diffeomorphism)**:** A *diffeomorphism* is a smooth homeomorphism $f : \mathcal{M} \to \mathcal{N}$ between two manifolds. I.e. two manifolds are diffeomorphic if the map $\psi \circ f \circ \phi^{-1} : U \to V$ is smooth for <u>all</u> charts $\phi : \mathcal{M} \to U \subset \mathbb{R}^n$ and $\psi : N \to V \subset \mathbb{R}^n$.

Remark: There are interesting properties of S^7 and $\mathbb{R}^4$. You can find two atlasses of these manifolds such that the two atlasses are not diffeomorphic to each other (maybe I didn't understand this correctly). As far as we know, there is yet no application of this in physics.

1.2 Tangent Spaces

So far, we have the idea of a manifold. It is a space that looks locally like $\mathbb{R}^n$, with a set of charts that allow us to map patches to $\mathbb{R}^n$ and differentiate (and later, integrate).

Let $C^\infty(\mathcal{M})$ denote the set of all smooth functions that assign each point in the manifold $\mathcal{M}$ a real number in $\mathbb{R}$. We know how to differentiate on $\mathbb{R}^n$. Smooth functions allow us to differentiate in $\mathbb{R}^n$ before we map back onto the manifold.

Definition 2 (tangent vector)**:** A *tangent vector* X_p is an object that differentiates functions at some point $p \in \mathcal{M}$ on the manifold. Specifically, it is a map $X_p : C^\infty(\mathcal{M}) \to \mathbb{R}$ with certain properties:

1. linearity: $X_p(f+g) = X_p(f) + X_p(g)$ for all $f, g \in C^\infty(\mathcal{M})$
2. constants vanish: $X_p(f) = 0$ when f is constant
3. Leibniz rule: $X_p(fg) = f(p)X_p(g) + X_p(f)g(p)$ for all $f, g \in C^\infty(\mathcal{M})$

Definition 3 (tangent space)**:** The set of all tangent vectors at p is the *tangent space* $T_p(\mathcal{M})$ at p.

One of the early surprises of differential geometry is thinking of vectors as differential operators. In $\mathbb{R}^3$, we used position vectors as displacement. However, this does not generalise to curved spaces. The second idea of a vector in physics is the velocity of a particle. This change in time is analogous to the definition that we use here in differential geometry. By differentiating, the tangent vector tells us how things change when moving in a particular direction.

Claim: In a chart $\phi = (x^1, \cdots, x^n)$, we can write a tangent vector as

$$X_p = X^\mu \partial_\mu|_p, \tag{1.2.1}$$

with $X^\mu = x_p(x^\mu)$ and $\partial_\mu f = \frac{\partial(f\circ\phi^{-1})}{\partial x^\mu}$ for all functions $f \in C^\infty(\mathcal{M})$. This definition of $\partial_\mu f$ uses the fact that we can differentiate on $\mathbb{R}^n$, which we map to via $f \circ \phi^{-1}$.

Proof. The idea is that given a function f, we use coordinates to push it to $\mathbb{R}^n$, and then we differentiate there. A detailed proof is in the notes. □

> **Remark:** We are going to use the summation convention with indices up/down. The object with coefficients as superscripts are tangent vectors, different from objects with subscript coefficients. Hence, the position of the indices matter even more than in Special Relativity.

Writing $X_p = X^\mu \partial_\mu$ clearly depends on coordinates x^μ. What would happen if we chose different coordinates? If we picked another set of coordinates $\widetilde{x}^\mu$, we could write

$$X_p = X^\mu \frac{\partial}{\partial x^\mu}|_p = \widetilde{X}^\mu \frac{\partial}{\partial \widetilde{x}^\mu}|_p. \tag{1.2.2}$$

Notation We write $\partial_\mu \equiv \frac{\partial}{\partial x^\mu}$.

Acting on a function f, (and dropping the implicit action of going to $\mathbb{R}^n$ via ϕ^{-1})

$$X_p(f) = X^\mu \left.\frac{\partial f}{\partial x^\mu}\right|_p = X^\mu \left.\frac{\partial \widetilde{x}^\nu}{\partial x^\mu}\right|_{\phi(p)} \left.\frac{\partial f}{\partial \widetilde{x}^\nu}\right|_p . \tag{1.2.3}$$

We can view this as a change of basis of $T_p(\mathcal{M})$

$$\left.\frac{\partial}{\partial x^\mu}\right|_p = \left.\frac{\partial \widetilde{x}^\nu}{\partial x^\mu}\right|_p \left.\frac{\partial}{\partial \widetilde{x}^\nu}\right|_p . \tag{1.2.4}$$

Alternatively, we can think of the change of components

$$\widetilde{X}^\nu = X^\mu \left.\frac{\partial \widetilde{x}^\nu}{\partial x^\mu}\right|_{\phi(p)} . \tag{1.2.5}$$

This is called a *contravariant* transformation; the indices tell us which way things transform.

To see why this is called a tangent vector, we are going to consider a path $\sigma(t) : \mathbb{R} \to \mathcal{M}$ in the manifold such that $\sigma(0) = p$. We can think of the parameter as time for example. Given coordinates, this becomes a path $x^\mu(t)$ in $\mathbb{R}^n$. The tangent in $\mathbb{R}^n$ is $X^\mu = \frac{\mathrm{d}x^\mu(t)}{\mathrm{d}t}|_{t=0}$. We can then use this to define $X_p \in T_p(\mathcal{M})$ as

$$X_p = \frac{\mathrm{d}x^\mu(t)}{\mathrm{d}t}|_{t=0} \frac{\partial}{\partial x^\mu}|_p. \tag{1.2.6}$$

In this sense, $T_p(\mathcal{M})$ is the space of all tangent vectors at p. By considering all possible paths, one can describe all possible tangent vectors. Physically, it is like the space of velocities of a particle. The way to think about this is as in fig

> **Remark:** Tangent spaces at two different points are two different vector spaces. It is meaningless to try to add vectors of two different tangent spaces . . . for now.

> **Remark:** These manifolds have intrinsic meaning and do not depend on an embedding. Physically, the $3 + 1$-dimensional spacetime is a manifold that we do not think is embedded in any higher dimensional space.

1.3 Vector Fields

Definition 4 (vector field)**:** A *vector field* X is a smooth assignment of tangent vectors at each point $p \in \mathcal{M}$:

$$X : C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M}) \tag{1.3.1}$$

with $(Xf)(p) = X_p(f)$. In a given coordinate chart we can write

$$X = X^\mu(x)\partial_\mu. \tag{1.3.2}$$

Notation: We will denote the space of all vector fields as $\mathfrak{X}(\mathcal{M})$.

> **Remark:** We will make the notion of *smooth* more precise later. Intuitively, we want neighbouring tangent spaces to have similar directions.

Given two vector fields $X, Y \in \mathfrak{X}(\mathcal{M})$, we can define the *commutator*.

Definition 5 (commutator)**:** The *commutator* $[X, Y] \in \mathfrak{X}(\mathcal{M})$ is a map

$$[X, Y]f = X(Yf) - Y(Xf). \tag{1.3.3}$$

> **Remark:** Both X, Y are first order differential operator. Thus, each term on the RHS is a second order operator. However, by subtracting these terms from each other, the second order terms cancel by commutation of partial derivatives. Thus, the commutator is a first order differential operator.

Exercise 1.3.1: In a coordinate basis, you can check

$$[X, Y] = \left(X^\mu \frac{\partial Y^\nu}{\partial x^\mu} - Y^\mu \frac{\partial X^\nu}{\partial x^\mu} \right) \frac{\partial}{\partial x^\nu}. \tag{1.3.4}$$

> **Remark:** As the course progresses, we will see connections with the *Symmetries, Field and Particles* course and its treatment of Lie algebras though these commutators.

Exercise 1.3.2: Check that the *Jacobi identity* holds:

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0 \tag{1.3.5}$$

1.4 Integral Curves

There is a relationship between vector fields and streamlines, also known as flows, on a manifold $\mathcal{M}$.

Definition 6 (flow)**:** A *flow* on $\mathcal{M}$ is a one-parameter family of diffeomorphisms $\sigma_t : \mathcal{M} \to \mathcal{M}$, labelled by $t \in \mathbb{R}$, such that $\sigma_{t=0} = id$ and $\sigma_s \circ \sigma_t = \sigma_{s+t}$.

This defines *flow lines* on $\mathcal{M}$ as illustrated in For each point on the manifold, there is a unique such streamline that flows through it. Each line has coordinates $x^\mu(t)$.

Definition 7 (vector field)**:** We define a vector field X by taking the tangent vector at each point:

$$X^\mu(x(t)) = \frac{\mathrm{d}x^\mu(t)}{\mathrm{d}t}. \tag{1.4.1}$$

Definition 8 (integral curves)**:** Conversely, given a vector field $X^\mu(x)$, we can integrate (1.4.1) to find $x^\mu(t)$. These flow lines are called *integral curves*.

Example: On S^2, with (θ, ϕ) polar coordinates, consider $X = \frac{\partial}{\partial \phi}$. From (1.4.1), $\frac{\mathrm{d}\theta}{\mathrm{d}t} = 0$ and $\frac{\mathrm{d}\phi}{\mathrm{d}t} = 1$. This means that we have constant $\theta = \theta$.

$$\phi = \phi_0 + t \implies \sigma_t(\theta, \phi) \to (\theta, \phi + t) \tag{1.4.2}$$

The integral curves are lines of constant lattitude.

1.5 Tensors

Definition 9 (Dual Space)**:** Given a vector space V, there exists a dual vector space V^* which is the space of all linear maps $V \to \mathbb{R}$.

The easiest way to define V^* is to come up with basis. Given a basis $\{e_\mu : \mu = 1, \ldots, n\}$ of the first vector space V, we have a dual basis $\{f^\mu\}$ of V^* defined by

$$f^\mu(e_\nu) = \delta^\mu_\nu. \tag{1.5.1}$$

Given a general element $X = X^\mu e_\mu \in V$, the linearity property means that we have $f^\nu(X) = X^\mu f^\nu(e_\mu) = X^\nu$.

> **Remark:** Given a basis of V, this means that we have a natural basis of V^*, giving an isomorphism from V to V^*. But the map defining this isomorphism depends on the basis.

> **Remark:** Most vector spaces that we meet in physics come endowed with an inner product. The vector space V does not (yet) have an inner product structure. If you do have an inner product, defined by a metric, then there is a natural basis-independent isomorphism between V and V^*.

> **Remark:** $V^{**} = V$

Definition 10 (cotangent space)**:** At each point $p \in \mathcal{M}$, the tangent space $T_p(\mathcal{M})$ is a vector space. The dual space $T_p^*(\mathcal{M})$ is called the *cotangent space*.

Definition 11 (1-forms)**:** A smooth assignment of cotangent vectors at each point is called a *cotangent field*. However, most people just call this a *1-form*.

Notation: We denote the space of all 1-forms over $\mathcal{M}$ as $\Lambda^i(\mathcal{M})$.

Definition 12: Given a smooth function $f \in C^\infty(\mathcal{M})$, there is a natural 1-form that we call $df \in \Lambda^1(\mathcal{M})$, defined by

$$\underbrace{df}_{\text{1-form}}\Big(\overbrace{X}^{\text{vector-field}}\Big) = \overbrace{\underbrace{X}_{\text{vector field}}\Big(\underbrace{f}_{\text{function}}\Big)}^{\text{function}} \tag{1.5.2}$$

Example: Suppose we have coordinates x^μ. These define a basis $e_\mu = \partial/\partial x^\mu$ of vector fields. Taking a particular $f = x^\mu$ (this f is a function) gives $dx^\mu(\partial/\partial x^\nu) = \partial/\partial x^\nu = \delta^\mu_\nu$. So $f^\mu = dx^\mu$ (this f^μ is a basis element) provides a basis for $\Lambda^1(\mathcal{M})$. In general, $\omega \in \Lambda^1(\mathcal{M})$ can be written (in a given chart) as

$$\omega = \omega_\mu(x) dx^\mu. \tag{1.5.3}$$

Remark: Note that this only holds locally, not necessarily globally.

In this basis, the 1-form is $df = \frac{\partial f}{\partial x^\mu} dx^\mu$, so that we have

$$df(X) = \frac{\partial f}{\partial x^\mu} dx^\mu (X^\nu \partial_\nu) = X^\mu \frac{\partial f}{\partial x^\mu} = X(f), \qquad \partial_\nu := \frac{\partial}{\partial x^\nu}, \tag{1.5.4}$$

as expected.

If we change coordinates from x^μ to $\tilde{x}^\mu(x)$, we have seen previously that the basis of vector fields changes as

$$\frac{\partial}{\partial \tilde{x}^\mu} = \frac{\partial x^\nu}{\partial \tilde{x}^\mu} \frac{\partial}{\partial x^\nu}. \tag{1.5.5}$$

The basis of 1-forms changes as

$$d\tilde{x}^\mu = \frac{\partial \tilde{x}^\mu}{\partial x^\nu} dx^\nu. \tag{1.5.6}$$

These are called *covariant* transformations.

This ensures that

$$d\tilde{x}^\mu \left(\frac{\partial}{\partial \tilde{x}^\nu} \right) = \frac{\partial \tilde{x}^\mu}{\partial x^\rho} dx^\rho \left(\frac{\partial x^\sigma}{\partial \tilde{x}^\nu} \frac{\partial}{\partial x^\sigma} \right) = \frac{\partial \tilde{x}^\mu}{\partial x^\rho} \frac{\partial x^\sigma}{\partial \tilde{x}^\nu} dx^\rho \left(\frac{\partial}{\partial x^\sigma} \right) = \frac{\partial \tilde{x}^\mu}{\partial x^\rho} \frac{\partial x^\rho}{\partial x^\nu} = \delta^\mu_\nu \tag{1.5.7}$$

as required.

Remark: This transformation looks very much like the Jacobian incurred by a change of variables in a multi-variate integral. We will see the relation between 1-forms and integration in the future.

The question is now, how do we differentiate 1-forms? Given a vector field $X \in \mathfrak{X}(\mathcal{M})$ and $\omega \in \Lambda^1(\mu)$, we can take the Lie derivative $\mathcal{L}_X \omega$.

First, note that 1-forms are pulled-back: If $\varphi : \mathcal{M} \to \mathcal{N}$ and $\omega \in \Lambda^1(\mathcal{N})$, then $\varphi^* \omega \in \Lambda^1(\mathcal{M})$ with

$$(\varphi^* \omega) X = \omega(\varphi_* X), \qquad \varphi_* X \in \mathfrak{X}(\mathcal{N}), \tag{1.5.8}$$

with $(\varphi^* \omega)_\mu = \omega_\alpha \frac{\partial y^\alpha}{\partial x^\mu}$, where y^α are coordinates on $\mathcal{N}$ and x^μ are coordinates on $\mathcal{M}$.

Now we have the Lie derivative

$$\mathcal{L}_X \omega = \lim_{t \to 0} \left[\frac{(\sigma_t^* \omega)_p - \omega_p}{t} \right]. \tag{1.5.9}$$

Now $x_\mu(t) = x^\mu(0) + X^\mu(x) t + \ldots$, which means that the push-forward of a 1-form

$$\sigma_t^* dx^\mu = \left(\delta^\mu_\nu + t \frac{\partial X^\mu}{\partial x^\nu} + \ldots \right) dx^\nu \tag{1.5.10}$$

and the Lie derivative takes the form

$$\mathcal{L}_X \omega = (X^\nu \partial_\nu \omega_\mu + \omega_\nu \partial_\mu X^\nu) dx^\mu. \tag{1.5.11}$$

This is the *Lie derivative* of a 1-form.

1.5.1 Tensor Fields

Definition 13 (tensor)**:** A *tensor* of rank (r, s), $r, s \in \mathbb{N}_0^+$, at a point $p \in \mathcal{M}$ is a multi-linear map

$$T : \underbrace{T_p^*(\mathcal{M}) \times \cdots \times T_p^*(\mathcal{M})}_{r} \underbrace{T_p(\mathcal{M}) \times \cdots \times T_p(\mathcal{M})}_{s} \to \mathbb{R}. \tag{1.5.12}$$

Definition 14 (rank)**:** The *total rank* of the tensor is $r + s$.

Example: A cotangent vector has rank $(0, 1)$.

A tangent vector has rank $(1, 0)$ (since $T_p^{**}(\mathcal{M}) = T_p(\mathcal{M})$)

Definition 15 (tensor field)**:** A *tensor field* is a smooth assignment of an (r, s) tensor to each point $p \in \mathcal{M}$.

Claim: Given a basis $\{e_\mu\}$ for tangent vectors and a dual basis $\{f^\mu\}$, a tensor has components

$$T^{\mu_1 \cdots \mu_r}{}_{\mu_1 \cdots \mu_s} = T(f^{\mu_1}, \ldots, f^{\mu_r}, e_{\nu_1}, \ldots, e_{\nu_s}). \tag{1.5.13}$$