WEEK 4

Prelude to General Relativity

- Main lesson from 20th century physics:
The Laws of Nature are governed by symmetries

- Special Relativity → invariance under Lorentz transformations

- Electromagnetism (i.e. Maxwell's eqs) is invariant under Lorentz transf. but Newton's law of gravitation is NOT

⇒ Need to develop a theory of gravity that is Lorentz invariant and compatible with experiments

⟶ General Relativity

· Equivalence Principle

1) The motion of a test particle in a gravitational field is independent of its mass and composition

→ equivalence of the gravitational and inertial masses

$$F = -\frac{GMm_g}{r^2} = m_i a \quad \Rightarrow \quad m_g = m_i$$

Galileo's tower experiment

2) All matter and energy is acted on and is a source of the gravitational field

→ it is not possible to shield the gravitational field.

The only other forces that behave in this way are the so-called fictitious forces: centrifugal and Coriolis force.

→ They appear in Newton's equation when

one uses it in a non-inertial (i.e., accelerating) reference frame.

These forces, like gravity, depend on the mass of the particle and made Einstein realise that, locally, gravity and acceleration are equivalent (→ "the happiest thought of my life")

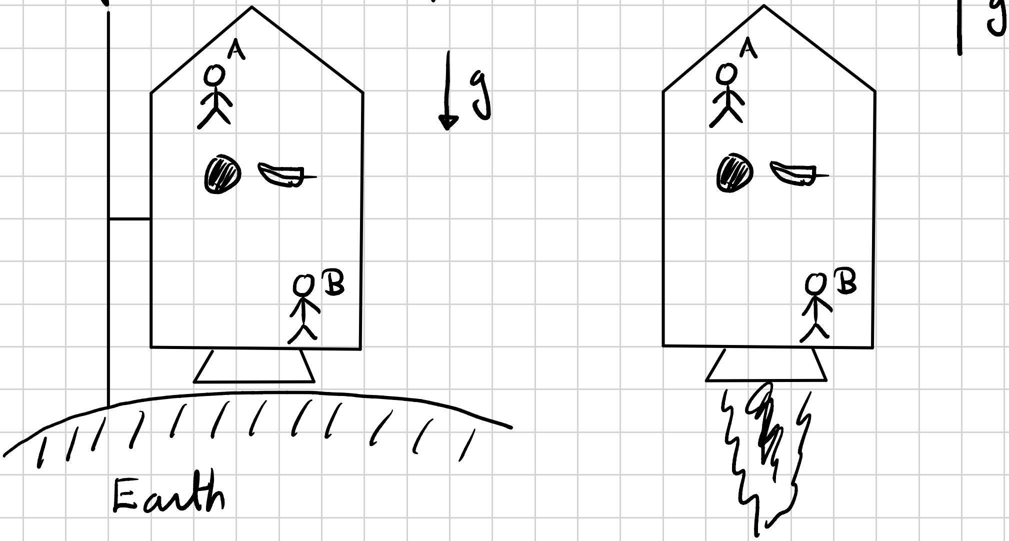
→ An observer that is freely falling does not experience gravity

If the observer drops objects then these will remain relative to him in uniform motion or at rest, independently of their chemical or physical nature.

→ this can only be true if $m_g = m_i$

- Equivalence Principle:

There is no local experiment that can distinguish a uniform acceleration from a uniform gravitational field.



→ It is impossible for A and B to tell whether they are on the surface of the Earth or in outer space accelerating in the rocket

→ Since locally there is no gravity, the spacetime is locally Minkowski (flat) and Special Relativity should apply.

· Summary

- In the presence of gravitational fields, locally (i.e., in small regions) there exist preferred inertial frames (i.e., free falling) in which Special Relativity applies.

- On large scales, no such frames exist and hence the laws of nature should be formulated in a way such that they are valid in any reference frame $\rightarrow$ invariant under arbitrary coordinate transformations.

$\Rightarrow$ The Laws of Nature should be written in tensorial form.

Differential Geometry and Tensor Calculus

- **Manifolds and coordinates**

Manifolds are spaces that locally look like $\mathbb{R}^n$. The entire space is constructed by sewing together these local regions.
E.g., S^n or T^n

- Preliminary definitions.

- Given two sets M and N, a map $\phi : M \to N$ is a relationship that assigns, to each element of M, exactly one element of N.

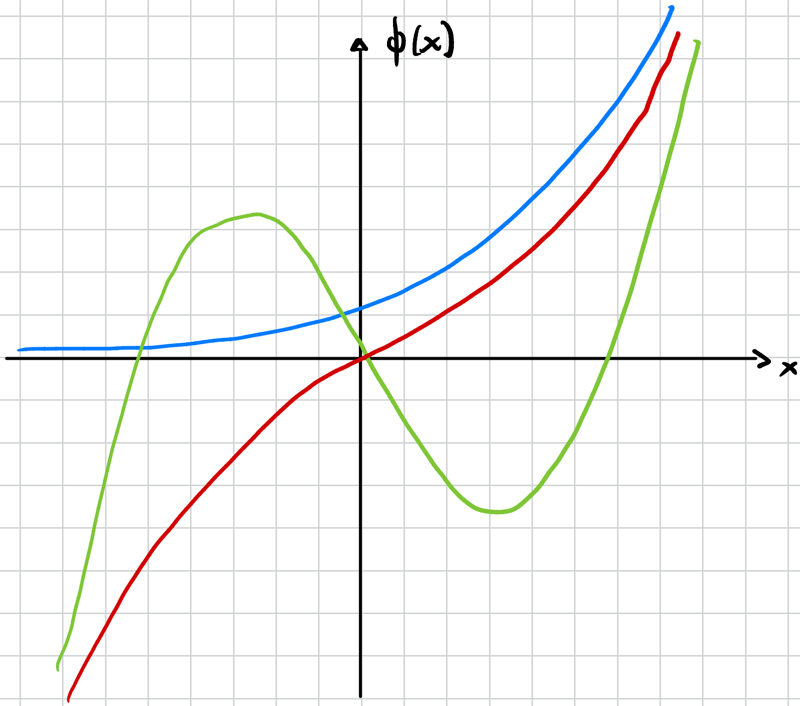
- Given two maps $\phi : A \to B$ and $\psi : B \to C$ the composition $\psi \circ \phi : A \to C$ is defined by $(\psi \circ \phi)(a) = \psi(\phi(a))$ so $a \in A$, $\phi(a) \in B$ and hence $\psi(\phi(a)) \in C$.

- one to one (injective): each element of N has at most one element of M mapped into it.

- onto (surjective): each element of N has at least one element of M mapped into it.

Examples of functions $\phi : \mathbb{R} \to \mathbb{R}$

- $\phi(x) = e^x$: one-to-one but not onto
- $\phi(x) = x^3 - x$: onto but not one to one
- $\phi(x) = x^3$: one to one and onto



- M is the domain of the map ϕ and the elements in N that M gets mapped into is called the image of ϕ.
- For any subset $U \subset N$, the set of elements of M that get mapped to U is called the preimage of U under ϕ or $\phi^{-1}(U)$.
- A map that is both one-to-one and onto is known as invertible (or bijective). In this case we can define the inverse map:

$$\phi^{-1}: N \longrightarrow M \qquad (\phi^{-1} \circ \phi)(a) = \phi^{-1}(\phi(a)) = a$$

- Consider maps between Euclidean spaces $\phi: \mathbb{R}^m \to \mathbb{R}^n$

$$y^1 = \phi^1(x^1, \dots, x^m)$$
$$y^2 = \phi^2(x^1, \dots, x^m)$$
$$\vdots$$
$$y^n = \phi^n(x^1, \dots, x^m)$$

$\rightarrow$ n functions ϕ^i of m variables

We say that the map $\phi: \mathbb{R}^m \to \mathbb{R}^n$ is C^p if the p^{th} derivative of the ϕ^i's exists and is continuous ; C^0 is continuous but not differentiable and C^∞ if it is smooth (i.e., continuous and infinitely differentiable).

- Two sets M and N are said to be diffeomorphic if there exists a C^∞ map $\phi: M \to N$ with a C^∞ inverse $\phi^{-1}: N \to M$; the map ϕ is called a diffeomorphism

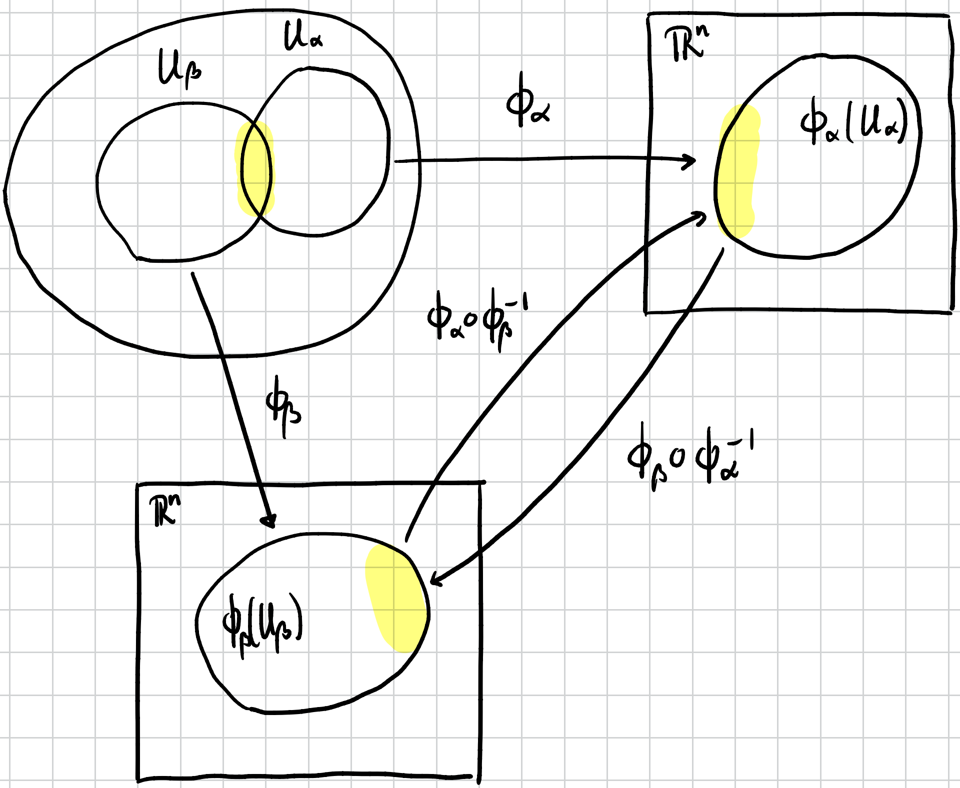
- Def: Manifold

A chart or a coordinate system consists of a subset U of M together with a one-to-one map $\phi: U \to \mathbb{R}^n$ such that the image $\phi(U)$ is open in $\mathbb{R}^n$. Then we say that U is an open subset in M. A C^∞ atlas is an indexed collection of charts $[(U_\alpha, \phi_\alpha)]$ such that :

1) The union of U_α is equal to M : $\bigcup_\alpha U_\alpha = M$
2) The charts are smoothly sewn together.

If $U_\alpha \cap U_\beta \neq \emptyset$, then $(\phi_\alpha \circ \phi_\beta^{-1})$ takes points in $\phi_\beta(U_\alpha \cap U_\beta) \subset \mathbb{R}^n$ onto an open set $\phi_\alpha(U_\alpha \cap U_\beta) \subset \mathbb{R}^n$, and all these maps are C^∞.

Then a C^∞ n-dimensional manifold is the set M together with a maximal atlas, i.e, one that contains every possible chart.



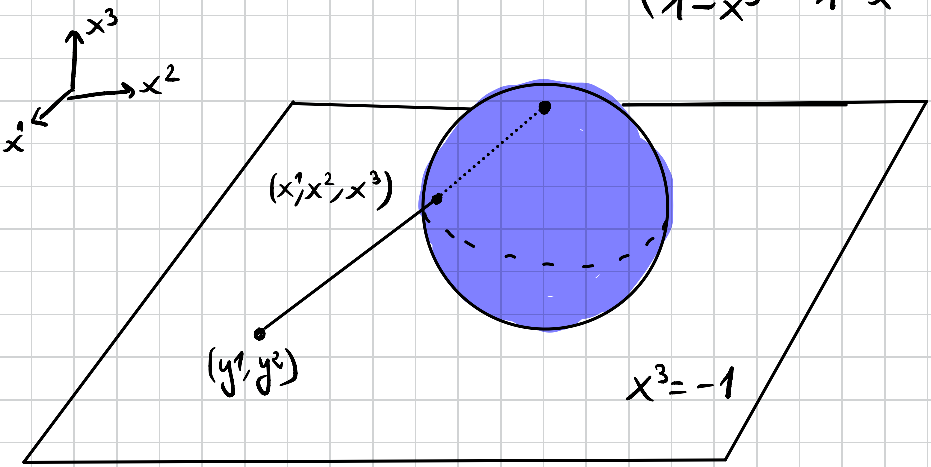
Example of a manifold: S^2

S^2 (of unit radius): $x^i \in \mathbb{R}^3$ s.t. $(x^1)^2 + (x^2)^2 + (x^3)^2 = 1$

U_1: the sphere minus the north pole

(Stereographic projection)

$$\phi_1(x^1, x^2, x^3) \equiv (y^1, y^2) = \left(\frac{2x^1}{1-x^3}, \frac{2x^2}{1-x^3} \right) \in \mathbb{R}^2$$



Another chart (U_2, ϕ_2) can be obtained by projecting from the south pole to the plane $x^3 = +1$. The resulting coordinates cover the sphere minus the south pole:

$$\phi_2(x^1, x^2, x^3) = (z^1, z^2) = \left(\frac{2x^1}{1+x^3}, \frac{2x^2}{1+x^3} \right)$$

Together, ϕ_1 and ϕ_2 cover the entire sphere and they overlap in the region $-1 < x^3 < 1$. The composition $\phi_2 \circ \phi_1^{-1}$ is given by

$$z^i = \frac{4y^i}{\sqrt{(y^1)^2+(y^2)^2}} \quad , \quad i = 1,2$$

which is C^∞ in the overlap region.

Consider the maps $f : \mathbb{R}^m \to \mathbb{R}^n$, $g : \mathbb{R}^n \to \mathbb{R}^\ell$ and the composition map $(g \circ f) : \mathbb{R}^m \to \mathbb{R}^\ell$. We can label points on each space in terms of the usual Cartesian coordinates: $x^a \in \mathbb{R}^m$, $y^b \in \mathbb{R}^n$, $z^c \in \mathbb{R}^\ell$. The chain rule relates the partial derivatives of the composition to the partial derivatives of the individual maps:

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

We write this as $\displaystyle \frac{\partial}{\partial x^a} = \sum_c \frac{\partial y^b}{\partial x^a} \frac{\partial}{\partial y^b}$

If $m=n$ then the determinant of $\dfrac{\partial y^b}{\partial x^a}$ is the Jacobian

· Vectors

We want to construct the tangent space at a point $p \in M$, T_p, with objects that are intrinsic to M. Let $\mathcal{F}$ be the space of all smooth functions on M: C^∞ maps $f: M \to \mathbb{R}$. Each curve through p defines an operator on this space, namely the directional derivative which maps $f \to \frac{df}{d\lambda}$ at p. Then T_p can be identified with the space of directional derivative operators along curves through p.

- Two operators $\frac{d}{d\lambda}$ and $\frac{d}{d\eta}$ representing derivatives along two curves $x^a(\lambda)$, $x^a(\eta)$ through p can be added and scaled by real numbers to give another operator $a\frac{d}{d\lambda} + b\frac{d}{d\eta}$
- The new operator acts linearly on functions and satisfies the Leibniz rule

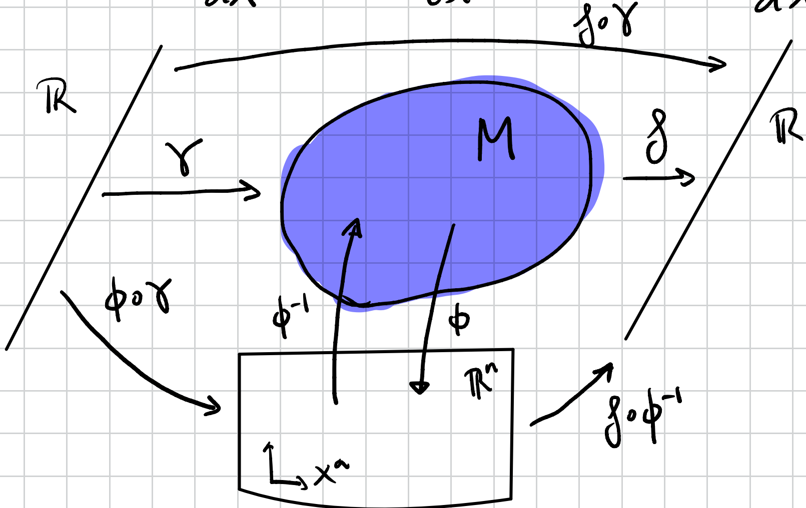
$\Rightarrow$ The set of directional derivatives forms a vector space at p

To identify the vector space of directional derivatives at p with T_p we need to show that they form a basis.

Consider a coordinate chart x^a and a set of n-directional derivatives at p are ∂_a, i.e., the directional derivative along $x^b = \text{const.}$ for $b \neq a$ parametrised by x^a itself. Now we show that any directional derivative can be written as a linear combination of partial derivatives:

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

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



Since the function f is arbitrary, we have

$$\frac{d}{d\lambda} = \frac{dx^a}{d\lambda} \partial_a$$

$\Rightarrow \{\partial_a\}$ form a basis of T_p

- This particular basis $\hat{e}_{(a)} = \partial_a$ is known as a coordinate basis for T_p

- Transformation of vectors under coordinate changes

For the basis vectors we can use the chain rule:

$$x^{a'} = x^{a'}(x^b)$$

$$\partial_{a'} = \frac{\partial x^a}{\partial x^{a'}} \partial_a$$

Remember that a vector is a geometric object and is unchanged under a change of basis:

$$V = V^a \partial_a = V^{a'} \partial_{a'} = V^{a'} \frac{\partial x^a}{\partial x^{a'}} \partial_a$$

$\Rightarrow \quad V^{a'} = \dfrac{\partial x^{a'}}{\partial x^a} V^a \quad$ since $\dfrac{\partial x^{a'}}{\partial x^a}$ is the inverse of $\dfrac{\partial x^a}{\partial x^{a'}}$

This transformation law for vectors is just the generalisation of the transformation of the vector components in Special Relativity under Lorentz transformations:

$$V^{a'} = L^{a'}{}_b V^b \quad : \quad L^{a'}{}_b : \text{Lorentz transf.}$$

Lorentz transf. are a special type of coordinate transf:

$$x^{a'} = L^{a'}{}_b x^b$$

- <u>Commutator</u> of <u>two vectors</u> (Lie derivative)

Given two arbitrary vectors X and Y we can define the commutator $[X, Y]$ by its action on arbitrary function $f(x)$:

$$[X, Y](f) = X(Y(f)) - Y(X(f))$$

- This operator is independent of the choice of coordinates
- The commutator of two vector fields is itself a vector field

$$[X,Y](af+bg) = a[X,Y](f) + b[X,Y](g)$$
$$[X,Y](fg) = f[X,Y](g) + g[X,Y](f)$$
for $a,b \in \mathbb{R}$ and f, g two arbitrary functions

Exercises :

- Show that $[X,Y]^a = X^b \partial_b Y^a - Y^b \partial_b X^a$
- $[X,Y]^a$ transforms as a vector under coordinate transformations :
$$[X,Y]^{a'} = \frac{\partial x^{a'}}{\partial x^a} [X,Y]^a$$