General Relativity - FI034

Personal Study Notes by Rodrigo Ramos

About these notes

The present notes comprise a one-semester graduate course on General Relativity taught by professor Ricardo Mosna at State University of Campina's Institute of Mathematics, Statistics and Scientific Computing in 2023. As I (Rodrigo) am writing these notes while attending the course, I won't have any proper time to review possible typos, changes of notations and conventions, or whatsoever problem one may find in this document's content. Whenever it is feasible, I will try to elaborate a bit on mathematical aspects of the theory, since (from personal experience) these were not so well emphasised in the undergraduate-level course I attended. The references I might use to do so are: (Lee, 2018, 2012; Penrose, 1987). Finally, the books I might take a peek at along the course are: Wald (2010); Straumann (2013); Carroll (2019).

Campinas
August 5, 2023

Contents

LECTURE **1**

Special Relativity Redux

Besides the lecture's content, I'll try to also include material from the first 13 lectures present in Schuller (2015) and how to derive the Lie algebra of the proper orthochronous Lorentz group. If I have time, there will be also a section on Galilean spacetime, but can also read about it here.

Chapter Contents

1.1 Why do we need manifolds and metric tensors?

Before we even dare to step into the mathematical zoo of differentiable manifolds, tensors, Lie groups and so on, bear in mind that **the main goal of this course is to learn *general relativity***. In other words, we will not have enough time to spend on the effort of making every little statement, definition or proof as mathematically sound as possible. I highly recommend that you read appendices A and B if you feel too unfamiliar with the terminology of this section. Rather than presenting all definitions and theorems relegated to these appendices, the aim here is to argue why we should model spacetime as a smooth manifold and why insert additional data – such as a metric tensor – in it. Without further ado, let us start our discussions.

Physics, as I perceive it, is all about trying to make sense of very fundamental aspects of what we conceive as Nature[1] at distinct scales. For that purpose, physicists make use of mathematics to create models which try to capture the (essential) features of natural phenomena, and in order to find out which models are less wrong than others we usually resort to experiments. Whether these models truly lead us closer to understanding what is real or they are just convenient tools that actually do not cast light on hidden aspects of Nature, it is a matter of debate between realists and instrumentalists. I sincerely do not know yet to which of these views I commit.

If we want to construct a mathematical framework that (classically) models spacetime, we must first and foremost consider that it is a set, which we denote as M and whose elements we name as **events**. Now, with a good set of rulers and clocks, we can measure space distances and time intervals with accuracy sufficient enough that we may treat these entities (space and time) as with no discontinuities. This means that, if we prescribe time and space coordinates $(t, x(t))$ to describe trajectories of objects in spacetime, then such trajectories should also have no discontinuities. Now, it would be totally fine if another group of people had chosen a different set of time and space coordinates $(t', x'(t'))$ to describe these same trajectories. This means that trajectories of objects in spacetime

[1]Not the journal :P.

must be independent of the choice of coordinates made. Due to this simple conclusion, we must ensure that M has enough structure such that it makes sense to define continuous maps $\gamma : \mathbb{R} \to M$ and such that we can map subsets of M to $\mathbb{R}^4$.

Since a trajectory in spacetime is coordinate-independent, the velocity at each point p of this trajectory must also be coordinate-independent. Moreover, in order to compute these velocities, we must have a way to define well-behaved limits and derivations in M. The former can be done by requiring that M is a Hausdorff topological space. These considerations mean that M should at least be a Hausdorff topological space that is locally euclidean, which is not that short behind from requiring that M should also be a topological manifold. Hence, we require M to be such kind of manifold.

What about the smooth structure that makes M a smooth manifold? My best explanation for adding a smooth structure to M is because we want to be able to describe M with as many distinct coordinate systems as we wish, but in a way that we can move from one coordinate system to another in a smooth way[2].

Metric Tensors

If $\gamma : [0, 1] \to M$ is a smooth curve that describes the trajectory of a physical object – namely, γ is an worldline – we are able to compute its length $L[\gamma]$ provided that we have a way to assign length to its velocity vectors at each point $\gamma(\lambda) \in M$. For our purposes, this requirement is equivalent to considering M as a pseudo-Riemannian smooth manifold. In other words, we should equip each tangent space with a map $g_{\gamma(\lambda)} : T_{\gamma(\lambda)}M \times T_{\gamma(\lambda)}M \to \mathbb{R}$ that is bilinear, symmetric and non-degenerate, where the last condition means that $g_{\gamma(\lambda)}$ has signature (p, q) with $p + q = 4$. That is equivalent to saying that $g_{\gamma(\lambda)}$ is a symmetric (0,2)-tensor on the tangent space at $\gamma(\lambda)$ and that g is a symmetric, non-degenerate (0,2)-tensor field with signature (p,q) on M – to put it shortly, g is a **metric tensor on** M. How do we choose the numbers p and q? This is the point where physics comes in and that will be done in the next section

1.2 Minkowski Spacetime

Why is Spacetime not Galilean?

I believe there are different approaches to answer the question above. The way I would answer it would go more or less along the following stream of thought:

> Maxwell's equations imply that light propagates with a velocity c. Initially, it was thought that this propagation was relative to a medium name *aether*. Hence, by assuming that we live in a Galilean spacetime, the Galilean velocity addition formula should hold when changing between two reference frames. It happens, though, that Maxwell's wave equations are not invariant under Galilean boosts. From this observation we have two possibilites: detect this aether, which means we should reformulate EM to be Galilean invariant, or do not detect the aether and conclude that light moves at the same speed in all inertial frames. This second path leads to the correct transformations between inertial frames that preserve Maxwell's equations.

It is fundamental do emphasise that it was only by means of experimental verification that we ruled out the hypothesis that *aether* exists – for more details, please read about the Michelson-Morley experiment. The non-existence of *aether* forces us to replace our description of spacetime by some other than the Galilean one. One of such replacements gives rise to *special relativity*, which is based on the following principles:

P1) The speed of light is the same to all inertial observers, irrespective of the relative motion between them and the light source.

P2) Any two distinct inertial observers will describe the laws of Physics in the same manner.

[2]Well, we could have required it to be just a k-differentiable manifold for sufficient large k, but smooth does the job too.

The Minkowski Metric

In this subsection, we consider that spacetime is a smooth 4-dimensional manifold globally diffeomorphic to $\mathbb{R}^4$. Under this assumption, consider that an inertial observer, Alice, describes $\mathbb{R}^4$ with the coordinate system (R^4, id), where $\mathrm{id} : \mathbb{R}^4 \to \mathbb{R}^4$ is the identity function. Suppose that Alice has a light-source and it emits a light ray at event $A = (0,0,0,0)$. If she describes the trajectory of this light ray by a curve $\gamma(t) = (ct, x(t), y(t), z(t))$, considering that **P1** holds for any value of t, then the velocity vector of the light ray must satisfy:

$$\dot{x}^2 + \dot{y}^2 + \dot{z}^2 = c^2,$$

where the dot over the letters in the LHS is a short notation for the derivative with respect to t. By reordering the equality above, we obtain

$$-c^2 + \dot{x}^2 + \dot{y}^2 + \dot{z}^2 = 0. \tag{1.1}$$

Equation (1.1) is equivalent to the condition that

$$\eta_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t)) = 0, \tag{1.2}$$

where $\eta_{\gamma(t)}$ is the (0,2)-tensor that is written as

$$\eta_{\gamma(t)} = -c^2 dt \otimes dt + dx \otimes dx + dy \otimes dy + dz \otimes dz. \tag{1.3}$$

We can present equation (1.3) in a more familiar way by mapping it to the following matrix:

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{1.4}$$

and in the above equation I have shamelessly use a widespread physicist notation for tensors, i.e., I identified a tensor for its components in the LHS. In terms of equation (1.4), we rewrite the RHS of equation (1.2) as

$$\eta_{\mu\nu}\dot{\gamma}^{\mu}\dot{\gamma}^{\nu} = \dot{\gamma}^{T}\eta\dot{\gamma}.$$

I purposely omitted the dependence on t as a way to keep notation a bit cleaner.

One may check that $\eta_{\gamma(t)}$ is a symmetric, non-degenerate bilinear map defined on $T_{\gamma(t)}\mathbb{R}^4$. Moreover, it has signature $(3,1)$, where 3 and 1 tell, respectively, how many plus and minus ones appear in equation (1.4). Due to Alice's choice of coordinate system, it is easy to see that at each tangent space $T_p\mathbb{R}^4$, one can assign the same tensor as we did in equation (1.3), because its components do not depend on which point p we are evaluating it. Hence, we have defined a metric tensor field with signature $(3,1)$, η, on $\mathbb{R}^4$. This specific signature – where the number of plus ones equals to $\dim M - 1$ – characterises $(\mathbb{R}^4, \eta)$ as a **Lorentzian manifold** and we shall name this metric as the **Minkowski metric**[3].

The Proper, Orthocronous Lorentz Group

In the previous subsection, we only discussed how Alice would define η at each point of spacetime in one specific coordinate system. It so happens that we have stated two guiding principles: first, that the speed of light is the same to all inertial observers; second, that different inertial observers describe the same laws of Physics. So, how would another inertial observer, Bob, describe the same light ray that Alice emitted? Well, first of all, Bob is free to choose whichever coordinate system he wants, and, since we are not dealing with Galilean spacetimes, this means that the time he assigns to events under the coordinate system $(\mathbb{R}^4, ct', x', y', z')$ has no *a priori* right to be same as Alice's.

[3]It may look like that η is just a convenient mathematical apparatus to compute lengths of tangent vectors, but, as we will show later, it is a central object to the theory of general relativity because it encodes the geometry of spacetime itself and it will dictate how matter shall rearrange by virtue of Einstein's field equations.

Now, if $\alpha(t') = (ct', x'(t'), y'(t'), z'(t')$ is Bob's description of the light ray's trajectory, as consequence of P1:

$$\eta_{\alpha(t')}(\dot{\alpha}(t'), \dot{\alpha}(t') = 0, \tag{1.5}$$

where dot now represents derivatives with respect to t'. In view of the change of coordinates transformation for vector components, we can write:

$$\dot{\alpha}^{\mu}(t') = \left.\frac{\partial x'^{\mu}}{\partial x^{\nu}}\right|_{\gamma(t)} \dot{\gamma}^{\nu}(t),$$

at the point $\alpha(t') = \gamma(t)$. For cleaner notation, let us denote $\left.\frac{\partial x'^{\mu}}{\partial x^{\nu}}\right|_{\gamma(t)} = \Lambda^{\mu}{}_{\nu}(\gamma(t))$.If we use these transformations in equation (1.5) and equate the result to equation (1.2), we obtain

$$\dot{\gamma}^{T}\eta\,\dot{\gamma} = \dot{\gamma}^{T}\Lambda^{T}(\gamma(t))\eta\,\Lambda(\gamma(t))\,\dot{\gamma}.$$

This means that the transformations between the coordinates systems established by inertial observers Alice and Bob must satisfy the following condition for any tangent vectors $u, v \in T_p\mathbb{R}^4$:

$$u^{T}[\Lambda^{T}(p)\,\eta_p\,\Lambda(p)]\,v = u^{T}\eta_p\,v.$$

Observe that, in general, the transformation Λ might depend on the specific point of our manifold. Since Bob has established a global chart, this means that the transformation $\Lambda(p)$ is an isomorphism by property c) of proposition B.1 and that implies

$$\Lambda^{T}\,\eta\,\Lambda = \eta, \tag{1.6}$$

where the equality above must be interpret point-wisely. Equation (1.6) is a very especial equation, it defines the elements Λ of the **Lorentz group**, defined precisely as the subset of invertible 4x4 matrices:

$$O(3,1) \doteq \{\Lambda \in GL(4,\mathbb{R}) : \Lambda^{T}\,\eta\,\Lambda = \eta\}.$$

Endowed with usual matrix multiplication, $O(3,1)$ is a group, and, more than that, it is a Lie group. Given that $\det(AB) = \det(A)\det(B)$ for any square matrices A and B, by taking the determinant on both sides of equation (1.6), we deduce that

$$\det\Lambda = \pm 1. \tag{1.7}$$

Well, equation (1.7) gives us two choices for the determinant. Which one should we pick? If we want to study the transformations that are small and continuous perturbations from the identity, that is, $\Lambda \approx \mathbb{1} + \varepsilon M$, then we should choose $\det\Lambda = +1$, simply because $\det\mathbb{1} = 1$. Consequently, this choice yield a Lie subgroup of the Lorentz group, namely the **proper Lorentz group**, denoted as

$$SO(3,1) \doteq \{\Lambda \in O(3,1) : \det\Lambda = 1\}.$$

These are the Lorentz transformations that preserve spacetime orientations. If we had chosen the other signal for the determinant, then we would have ended up with the Lorentz transformations that reverse space orientations, such as $x \mapsto -x$.

We have one final restriction to make. If we go back to equation (1.6), then component-wise:

$$-(\Lambda^{0}{}_{0})^2 + \sum_{i=1}^{3}(\Lambda^{i}{}_{0})^2 = -1,$$

which means that

$$\Lambda^{0}{}_{0} = \pm\sqrt{1 + \sum_{i=1}^{3}(\Lambda^{i}{}_{0})^2}. \tag{1.8}$$

Since the term inside the square-root is always greater than 1, equation (1.8) essentially boils down to the choice $\Lambda^0_0 > 0$ or $\Lambda^0_0 < 0$. 'What is the point of even bothering about such choice?', one could justifiably ask. Recall that Λ is a transformation that relates the coordinates systems of two distinct inertial observers Alice and Bob. This means that if Alice describes an event by assigning to it the labels (ct, x, y, z), then this same event in terms of Bob's labels is given by:

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \Lambda \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix},$$

where

$$\Lambda = \begin{pmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 & \Lambda^0{}_2 & \Lambda^0{}_3 \\ \Lambda^1{}_0 & \Lambda^1{}_1 & \Lambda^1{}_2 & \Lambda^1{}_3 \\ \Lambda^2{}_0 & \Lambda^2{}_1 & \Lambda^2{}_2 & \Lambda^2{}_3 \\ \Lambda^3{}_0 & \Lambda^3{}_1 & \Lambda^3{}_2 & \Lambda^3{}_3 \end{pmatrix}$$

Hence, for any event $(ct, 0, 0, 0)$[4]:

$$ct' = \Lambda^0{}_0(ct).$$

From the previous discussion, we can choose either $\Lambda^0{}_0 > 0$ or $\Lambda^0{}_0 < 0$. The former will preserve time orientation, while the latter will reverse it. Hence, we stick to the former.

Therefore, we define the **proper, orthochronous Lorentz group** as the set of all Lorentz transformations that preserve space and time orientations. We denote such group[5] as

$$SO(3,1)^{\uparrow} \doteq \{\Lambda \in O(3,1) : \det \Lambda = +1 \text{ and } \Lambda^0{}_0 > 0\}. \tag{1.9}$$

Of course, we could have picked the proper, non-orthochronous transformations, which belong to

$$SO(3,1)^{\downarrow} \doteq \{\Lambda \in O(3,1) : \det \Lambda = +1 \text{ and } \Lambda^0{}_0 < 0\}.$$

But why stop here? We could even have picked the improper transformations – the ones which $\det \Lambda = -1$ – both the orthochronous and non-orthochronous ones, which belong respectively to the sets

$$O(3,1)^{\uparrow}_{-} \doteq \{\Lambda \in O(3,1) : \det \Lambda = -1 \text{ and } \Lambda^0{}_0 > 0\},$$

$$O(3,1)^{\downarrow}_{-} \doteq \{\Lambda \in O(3,1) : \det \Lambda = -1 \text{ and } \Lambda^0{}_0 < 0\}.$$

From this, we see that the full Lorentz group has four disconnected components:

$$O(3,1) = SO(3,1)^{\uparrow} \sqcup SO(3,1)^{\downarrow} \sqcup O(3,1)^{\uparrow}_{-} \sqcup O(3,1)^{\downarrow}_{-}.$$

From these components, the one that is always presented in a first introduction to special relativity is $SO(3,1)^{\uparrow}$, simply because it is the one that contains the usual Lorentz transformations that we are so familiar with. We shall derive these transformations in the next subsection.

Before we proceed to the derivation of the Lorentz transformations, here is a nice result about the other three components of the Lorentz group:

Proposition 1.1. *Let $(\mathbb{R}^4, \eta)$ be the Minkowski spacetime. At each tangent space $T_p\mathbb{R}^4$, define the matrices*

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

[4]Take Alice's position, for example.

[5]By this point, you should convince yourself that all these restrictions of elements $\Lambda \in O(3,1)$ still results in a (Lie) group.

and

$$T_{rev} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then

$$\begin{aligned} SO(3,1)^{\downarrow} &= \{T_{rev} P\Lambda : \ \Lambda \in SO(3,1)^{\uparrow}\}; \\ O(3,1)^{\uparrow}_{-} &= \{P\Lambda : \Lambda \in SO(3,1)^{\uparrow}\}; \\ O(3,1)^{\downarrow}_{-} &= \{T_{rev}\Lambda : \ \Lambda \in SO(3,1)^{\uparrow}\}. \end{aligned}$$

Proof. This is a very long proof. If you really want to read it, check Barata (2023, chap. 21). ∎

The Lie algebra $\mathfrak{so}(3,1)^{\uparrow}$

In the last subsection, we showed that the possible transformations between the coordinates established by two inertial observers are those which belong to the Lorentz group $O(3,1)$. Furthermore, by imposing physical constraints of preservation of space and time orientations, we obtained the proper Lorentz group $SO(3,1)^{\uparrow}$. This subgroup of $O(3,1)$ is a Lie group by its own merit and due to some Lie group wizardry, the exponential map $\exp : \mathfrak{so}(3,1)^{\uparrow} \to SO(3,1)^{\uparrow}$ is surjective. This means that any Lorentz transformation Λ in this component can be expressed[6] as

$$\Lambda = e^{-iA},$$

for some $A \in \mathfrak{so}(3,1)^{\uparrow}$. The task of this subsection is to identify the **infinitesimal generators** of the Lie algebra $\mathfrak{so}(3,1)^{\uparrow}$ and use them to obtain the expressions for the usual Lorentz transformations that we encounter on a first course on special relativity.

The Infinitesimal Generators

First of all, any element $A \in \mathfrak{so}(3,1)^{\uparrow}$ is traceless, that is,

$$\operatorname{Tr} A = 0.$$

To check that, we just have to use the condition that $e^{-iA} \in SO(3,1)^{\uparrow}$, which particularly means that

$$\det\left(e^{-iA}\right) = e^{-i \operatorname{Tr} A} = 1.$$

Given that any Λ will satisfy equation (1.6), then the following equivalence holds

$$e^{-iA^{T}}\,\eta\, e^{-iA} = \eta \iff e^{-iA^{T}} = \eta\, e^{iA}\,\eta^{-1}. \tag{1.10}$$

By means of elementary matrix exponential properties, the RHS side in equation (1.10) implies that

$$-A^{T} = \eta\, A\, \eta^{-1},$$

or equivalently:

$$A^{T}\,\eta = -\eta\, A. \tag{1.11}$$

Component-wise, equation (1.11) means that

$$A^{\sigma}{}_{\mu}\eta_{\sigma\nu} = -\eta_{\mu\sigma}A^{\sigma}{}_{\nu}.$$

Now, the contractions above imply that $\eta\, A$ is a $(0,2)$ antisymmetric tensor with components

$$A_{\mu\nu} = -A_{\nu\mu}.$$

[6] I am a physicist, so my exponential maps **always** have an $-i$.

Due to Proposition A.6, the $(0, 2)$ antisymmetric tensors belong to a vector space that has dimension $\binom{4}{2} = 6$. Let us choose a basis for this vector space and label its elements as

$$\{M^{12}, M^{13}, M^{23}, M^{01}, M^{02}, M^{03}\}.$$

Moreover, for $i, j \in \{1, 2, 3\}$ and $\mu \in \{0, 1, 2, 3\}$, define

$$\begin{aligned}
&M^{ji} \doteq -M^{ij}, \text{ whenever } i < j, \\
&M^{i0} \doteq -M^{0i}, \\
&M^{\mu\mu} \doteq 0.
\end{aligned}$$

Then, there are real coefficients $\omega_{\mu\nu}$ such that

$$\eta\, A = \frac{1}{2}\omega_{\mu\nu}M^{\mu\nu}.$$

Component-wise, the decomposition above translates to

$$A_{\alpha\beta} = \frac{1}{2}\omega_{\mu\nu}(M^{\mu\nu})_{\alpha\beta}.$$

If we raise the index α, then an expression for the components of an element $A \in \mathfrak{so}(3, 1)^{\uparrow}$ is

$$A^{\alpha}{}_{\beta} = \frac{1}{2}\omega_{\mu\nu}(M^{\mu\nu})^{\alpha}{}_{\beta}. \tag{1.12}$$

If we know the components $(M^{\mu\nu})^{\alpha}{}_{\beta}$ of each $M^{\mu\nu}$, then our job is essentially over, because we can write any Lorentz transformation as

$$\Lambda = e^{-\frac{i}{2}\omega_{\mu\nu}M^{\mu\nu}}.$$

Luckily, we can construct the $M^{\mu\nu}$ by rewriting $A^{\alpha}{}_{\beta}$ in a clever manner. Bear with me:

$$\begin{aligned}
A^{\alpha}{}_{\beta} &= \eta^{\alpha\mu}\,\delta^{\nu}{}_{\beta}\,A_{\mu\nu} \\
&= \frac{1}{2}(\eta^{\alpha\mu}\delta^{\nu}{}_{\beta}A_{\mu\nu} + \eta^{\alpha\mu}\delta^{\nu}{}_{\beta}A_{\mu\nu}) \\
&= \frac{1}{2}(A_{\mu\nu}\eta^{\alpha\mu}\delta^{\nu}{}_{\beta} + A_{\mu\nu}\eta^{\alpha\mu}\delta^{\nu}{}_{\beta}) \\
&= \frac{1}{2}A_{\mu\nu}\eta^{\alpha\mu}\delta^{\nu}{}_{\beta} - \frac{1}{2}A_{\nu\mu}\eta^{\alpha\mu}\delta^{\nu}{}_{\beta} \\
&= \frac{1}{2}A_{\mu\nu}\eta^{\alpha\mu}\delta^{\nu}{}_{\beta} - \frac{1}{2}A_{\mu\nu}\eta^{\alpha\nu}\delta^{\mu}{}_{\beta} \\
&= \frac{1}{2}A_{\mu\nu}(\eta^{\alpha\mu}\delta^{\nu}{}_{\beta} - \eta^{\alpha\nu}\delta^{\mu}{}_{\beta}) \\
&= -\frac{i}{2}A_{\mu\nu}i(\eta^{\alpha\mu}\delta^{\nu}{}_{\beta} - \eta^{\alpha\nu}\delta^{\mu}{}_{\beta}) \\
&\doteq -\frac{i}{2}A_{\mu\nu}(M^{\mu\nu})^{\alpha}{}_{\beta}.
\end{aligned}$$

In order to keep everything real, the $A_{\mu\nu}$ must be purely imaginary numbers. If we set $\omega_{\mu\nu} \doteq -iA_{\mu\nu}$, then we've found a decomposition for A accordingly to equation (1.12). Just to be explicit, we have constructed the $M^{\mu\nu}$ such that

$$(M^{\mu\nu})^{\alpha}{}_{\beta} \doteq i(\eta^{\alpha\mu}\delta^{\nu}{}_{\beta} - \eta^{\alpha\nu}\delta^{\mu}{}_{\beta}). \tag{1.13}$$

The objects in equation (1.13) define the **infinitesimal generators** of the Lie algebra of $SO(3,1)^{\uparrow}$. They satisfy the following commutation relation:

$$[M^{\mu\nu}, M^{\rho\sigma}] = i(\eta^{\nu\rho}M^{\mu\sigma} - \eta^{\nu\sigma}M^{\mu\rho} + \eta^{\mu\sigma}M^{\nu\rho} - \eta^{\mu\rho}M^{\nu\sigma}). \tag{1.14}$$

I have purposely omitted the derivation of such commutator, since it is just a boring multiplication involving tons of indices.

So far, everything has been very, very abstract for my liking. Let us use equation (1.14) to compute $\left[M^{32}, M^{13}\right]$, $\left[M^{21}, M^{32}\right]$ and $\left[M^{13}, M^{21}\right]$. Since the η is diagonal, we obtain

$$\left[M^{32}, M^{13}\right] = iM^{21},$$

$$\left[M^{21}, M^{32}\right] = iM^{13},$$

$$\left[M^{13}, M^{21}\right] = iM^{32}.$$

Therefore, if we set $M^{32} \doteq J_1$, $M^{13} \doteq J_2$ and $M^{21} \doteq J_3$, the equalities above translate to

$$[J_m, J_n] = i\varepsilon^{mnp}J_p. \tag{1.15}$$

Lo and behold! The equalities comprised in Equation (1.15) are precisely the angular momentum commutation relations! This means that rotations around any given axis $\hat{n}$ are proper orthocronous Lorentz transformations, because we can always write such rotation as

$$R_{\hat{n}}(\alpha) = e^{-i\alpha\hat{n}\cdot\vec{J}}.$$

What about the remaining commutators? Well, if we denote $Ki \doteq M^{0i}$, then the following are also true:

$$[K_m, K_n] = -i\varepsilon^{mnp}J_p, \quad [J_m, K_n] = i\varepsilon^{mnp}K_p. \tag{1.16}$$

The commutation relations established in equations (1.15) and (1.16) fully define the Lie algebra $\mathfrak{so}(3,1)^{\uparrow}$. On top of that, we have seen that the J_i's are the generators of rotations around fixed axes. Now, we must look at what are the transformations generated by the K_i's.

The Lorentz Boosts

Let us assume that we have two inertial observers, Alice and Bob, such that Alice is at rest with respect to (w.r.t.) the Earth and Bob is inside a train that moves with constant velocity v w.r.t. Alice. For operational simplicity, they choose coordinate systems such that their (spatial) axes have the same orientation as shown in figure 1.1. Moreover, we assume that Bob's move along Alice's x-axis and that they synchronised their clocks such that when each clock reads 0, the origins of the spatial axes coincide.

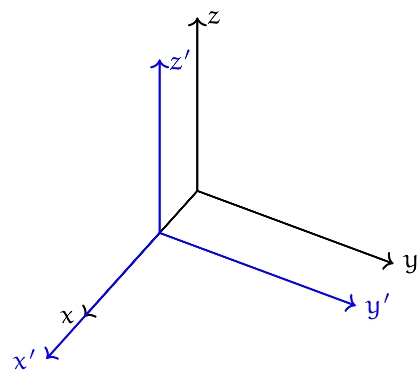

Figure 1.1: Alice and Bob spatial coordinate-axes. The black axes represent Alice's choice, whereas the blue ones represent Bob's choice.

Under this setting, let us compute e^{-iwK_1}. First, recall that $K_i \doteq M^{0i}$. Thus, by our definition of the $M^{\mu\nu}$ (equation (1.13)), we have:

$$K_1 = \begin{pmatrix} 0 & -i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad K_2 = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad K_3 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}. \tag{1.17}$$

By definition of the matrix exponential:

$$\begin{aligned} e^{-iwK_i} &\doteq \sum_{m=0}^{\infty} \frac{(-iwK_1)^m}{m!} \\ &= \mathbb{1} - iwK_1 + \frac{1}{2!}w^2(-iK_1)^2 + \frac{1}{3!}w^3(-iK_1)^3 + \frac{1}{4!}w^4(-iK_1)^4 + \frac{1}{5!}w^5(-iK_1)^5 \dots . \end{aligned}$$

We can regroup the RHS above into the identity plus the sums over the even and over the odd terms:

$$e^{-iwK_1} = \mathbb{1} + \sum_{j=1}^{\infty} \frac{w^{2j}(-iK_1)^{2j}}{(2j)!} + \sum_{j=0}^{\infty} \frac{w^{2j+1}(-iK_1)^{2j+1}}{(2j+1)!}. \tag{1.18}$$

Computing the infinite sums above could be a nightmare. Lucky us, the following is true:

Lemma 1.1. *Let K_1 be the generator in equation (1.17). For any natural numbers $m \geqslant 2$ and $n \geqslant 1$:*

$$(-iK_1)^{2m} = (-iK_1)^2, \text{ and } \quad (-iK_1)^{2n+1} = -iK_1.$$

Proof. It is a simple proof by induction. ∎

By virtue of the previous lemma, equation (1.18) reduces to:

$$e^{-iwK_1} = \mathbb{1} + \left[\sum_{j=1}^{\infty} \frac{w^{2j}}{(2j)!}\right](-iK_1)^2 + \left[\sum_{j=0}^{\infty} \frac{w^{2j+1}}{(2j+1)!}\right](-iK_1).$$

The series above result, respectively, in $\cosh(w) - 1$ and $\sinh(w)$. After computing the matrices $(-iK_1)^2$ and $-iK_1$, we finally obtain

$$e^{-iwK_1} = \begin{pmatrix} \cosh(w) & -\sinh(w) & 0 & 0 \\ -\sinh(w) & \cosh(w) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.19}$$

Equation (1.19) tells us that if Alice labels a spacetime event by ct, x, y, z, then Bob's labels will be:

$$\begin{cases} ct' = \cosh(w)ct - \sinh(w)x, \\ x' = \cosh(w)x - \sinh(w)ct, \\ y' = y, \\ z' = z. \end{cases}$$

Now, since Alice's and Bob's frames coincided when their clock readings were 0, Bob's velocity in Alice's coordinate system can always be expressed

$$v = \frac{x_B}{t_B},$$

where $(t_B, x_B, 0, 0)$ is an event that characterises Bob's position with respect to her reference frame. Naturally, with respect to his own frame, x'_B is always zero, thus:

$$0 = \cosh(w)x_B - \sinh(w)ct_B \implies \tanh(w) = \frac{x_B}{ct_B} \equiv \frac{v}{c}.$$

Hence, the parameter that generates the transformation given by equation (1.19) is the quantity

$$w = \arctan\frac{v}{c}$$

named *rapidity*.

Due to the identity $1 - \tanh^2(w) = \operatorname{sech}^2(w)$ and the positivity of $\cosh(w)$, it is easy to show that

$$\cosh(w) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{and} \quad \sinh(w) = \frac{v}{c}\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

By making the obvious identification $\cosh(w) = \gamma$ and $\frac{v}{c} \equiv \beta$, equation (1.19) assumes a more familiar manner:

$$e^{-iwK_1} \equiv B_x(v) = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{1.20}$$

and the coordinates transformations from Alice's to Bob's systems are

$$\begin{cases} ct' = \gamma ct - \beta\gamma x, \\ x' = \gamma x - \beta\gamma ct, \\ y' = y, \\ z' = z. \end{cases} \tag{1.21}$$

Ta-dah! After some hard work, we recovered our good and old expression for the Lorentz transformations between moving frames. The notation $B_x(v)$ denotes that we are performing what is called a **Lorentz boost** in the x-axis to a reference frame that moves with constant velocity v. If we had chosen to compute the transformations generated by K_2 or K_3, we would have ended up with the boosts in the y and z directions, respectively. The expressions for those are

$$B_y(v) = \begin{pmatrix} \gamma & 0 & -\beta\gamma & 0 \\ 0 & 1 & 0 & 0 \\ -\beta\gamma & 0 & \gamma & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad B_z(v) = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix}. \tag{1.22}$$

Consequences of Lorentz Boosts

Now that we have derived the Lorentz boosts , we can look for the their physical consequences. In order to do that with illustrations, I will restrict our analysis only to the cases in dimension 1+1. Therefore, events have two labels (x^0, x^1), where x^0 denotes the temporal label and x^1 denotes the spatial label.

Minkowski Diagrams, Wordlines and Lengths

Dynamics

1.3 Bibliography

Frederic Schuller. A thorough introduction to the theory of general relativity, 2015. URL https://www.youtube.com/watch?v=mpbWQbkl8_g#t=20m15s.

J. C. A. Barata. Notas para um curso de física-matemática, 2023. URL http://denebola.if.usp.br/~jbarata/Notas_de_aula/capitulos.html.

APPENDIX **A**

Tensors: Universal Properties and Multilinear Maps

A.1 Tensor Product of Vector Spaces

A question that popped up once in a while during my bachelor years was "what is a tensor?". Some colleagues would jokingly answer that "a tensor is something that transforms as a tensor". Ok, this kind of answer did not clarify a single thing to me back then. It was not until I enrolled in a course on group theory for physicists that a proper answer was given to me (though I did not understand it the first few times). If you too do not know what a tensor is, here a brief and rough explanation:

Given a bilinear map $\varphi : V_1 \times V_2 \to W$ between real vector spaces, the tensor product $V_1 \otimes V_2$ is the unique real vector space such that there exists a linear map $\tilde{\varphi} : V_1 \otimes V_2 \to W$ that is compatible with φ. A tensor is nothing more than an element of this new vector space.

There are a few things lacking in the explanation above – sure – but it already clarifies something: tensors are elements of a vector space that satisfies a very important property, *i.e.*, it is the vector space which "linearises" bilinear maps from $V_1 \times V_2$ to any other vector space W. The remainder of this appendix aims to make these ideas more precise. I will try to proceed as carefully as possible during the construction of the tensor product, but there can be some faults on my reasoning, since I am not a mathematician.

Definition A.1. *Let X be a set. The **freely generated vector space by** X is defined as the set*

$$\mathcal{F}(X) \doteq \{f : X \to \mathbb{R} | \operatorname{supp} f \textit{ is finite}\},$$

where $\operatorname{supp} f$ denotes the set of values $x \in X$ such that $f(x) \neq 0$.

One could (rightfully) argue that $\mathcal{F}(X)$ is not yet a real vector space. Though that is true, one does not need to be that imaginative to see that $\mathcal{F}(X)$ becomes a real vector space when we define addition and scalar multiplication just as we do with functions.

The set X is a basis for its freely generated vector space because any element of $\mathcal{F}(X)$ has the following decomposition:

$$f = \sum_{p=1}^{r} f(a_p) a_p,$$

for some $r \in \mathbb{N}$, where the first $a_p \in \operatorname{supp} f$ and the second a_p, seen as an element of $\mathcal{F}(X)$, denotes the function that has as support the set $\{a_p\}$.

Proposition A.1. *Let X be a set, $\mathcal{F}(X)$ be its freely generated vector space, and W be a real vector space. If $T : X \to W$ is a function, then there exists a unique linear map $\tilde{T} : \mathcal{F}(X) \to W$ for which the diagram below is commutative,*

$$\begin{tikzcd}
X \arrow[r, "T"] \arrow[d, hook, "\iota"'] & W \\
\mathcal{F}(X) \arrow[ur, dashed, "\tilde{T}"'] &
\end{tikzcd}$$

where $\iota : X \to \mathcal{F}(X)$ *denotes the inclusion map defined as* $\iota(x) = x$ *for all* $x \in X$.

Proof. One just needs to define the map $\tilde{T} : \mathcal{F}(X) \to W$ by

$$\tilde{T}(f) \doteq \sum_{p=1}^{r} f(a_p)T(a_p),\ \forall f \in \mathcal{F}(X),$$

and be a bit careful when proving that $\tilde{T}(f+g) = \tilde{T}(f)+\tilde{T}(g)$. To prove uniqueness, let $M : \mathcal{F}(X) \to W$ be another linear map that satisfies the diagram above. Then, for an arbitrary $f \in \mathcal{F}(X)$:

$$M(f) = \sum_{a\in \operatorname{supp} f} f(a)M(a) \text{ and } T(a) = (M \circ \iota)(a) = M(a)$$

$$\implies M(f) = \sum_{a\in \operatorname{supp} f} f(a)T(a)$$

$$\implies M(f) = \tilde{T}(f).$$

Since f was arbitrary, one deduces that $M = \tilde{T}$. ∎

The previous proposition was just a restatement of a result usually seen in linear algebra courses: a linear transformation between vector spaces is fully determined by the way it acts on a basis.

Now, let us recall our previous case: we had a set $X = V_1 \times V_2$ and a function $\varphi : V_1 \times V_2 \to W$ which was assumed bilinear if we had treated the set as a vector space. Then, we could be mislead by proposition A.1 and treat $\mathcal{F}(V_1 \times V_2)$ as the tensor product $V_1 \otimes V_2$, after all, for any function (in particular bilinear ones), the proposition gives us a unique linear map from a new vector space to our target space W. However, if $\tilde{\varphi}$ is linear map, it still is not compatible with the bilinearity of φ, for example:

$$\varphi(v_1 + u_1, v2) = \varphi(v_1, v_2) + \varphi(u_1, v_2)$$

for any $v_1, u_1 \in V_1$ and $v_2 \in V_2$, but in general

$$\tilde{\varphi}((v_1 + u_1, v_2)) \neq \tilde{\varphi}((v_1, v_2)) + \tilde{\varphi}((u_1, v_2)).$$

If this looks a bit confusing, remember that in the LHS above $(v_1 + u_1, v_2)$ is the function that is zero everywhere but in $(v_1 + u_1, v_2)$, while in the RHS (v_1, v_2) and (u_1, v_2) are functions defined analogously as the function $(v_1 + u_1, v_2)$. Unless $v_2 = 0$, these functions have little to do with each other, and so $\tilde{\varphi}$ has little to no right to mimic the bilinearity of φ. In order to fix this problem (and similar ones), first we have to consider the vector subspace $\mathcal{R} \subseteq \mathcal{F}(V_1 \times V_2)$ that consists of linear combinations of the following elements:

1. $(v_1 + u_1, v_2) - (v_1, v_2) - (u_1, v_2)$,
2. $(v_1, v_2 + u_2) - (v_1, v_2) - (v_1, u_2)$,
3. $\alpha(v_1, v_2) - (\alpha v_1, v_2)$,
4. $\alpha(v_1, v_2) - (v_1, \alpha v_2)$,

for all $v_1, u_1 \in V_1$, $v_2, u_2 \in V_2$ and $\alpha \in \mathbb{R}$.

Second, recall that given a vector space V and a subspace $S \subseteq V$, one can always construct a new vector space V/S, called the quotient space, by introducing the equivalence relation $v \sim u \iff v - u \in S$ for $u, v \in V$. In our specific scenario, we introduce the following notation:

$$\mathcal{F}(V_1 \times V_2)/\mathcal{R} \doteq V_1 \otimes V_2.$$

Now, the elements of this vector space are equivalence classes of elements of $\mathcal{F}(V_1 \times V_2)$ under $\sim$. By construction of the quotient space as a vector space, it follows, for example, that:

$$[(v_1 + u_1, v_2)] = [(v_1, v_2)] + [(u_1, v_2)],$$

because $(v_1 + u_1, v_2) - (v_1, v_2) - (u_1, v_2)$ clearly belongs to $\mathcal{R}$. If we introduce the notation $[(v, v')] \doteq v \otimes v'$, then the equality above translates to

$$(v_1 + u_1) \otimes v_2 = v_1 \otimes v_2 + u_1 \otimes v_2,$$

which is a bit more familiar to someone who has been previously exposed to tensors. Similarly, one can also show that:

1. $v_1 \otimes (v_2 + u_2) = v_1 \otimes v_2 + v_1 \otimes u_2$,
2. $\alpha(v_1 \otimes v_2) = (\alpha v_1) \otimes v_2$,
3. $\alpha(v_1 \otimes v_2) = v_1 \otimes (\alpha v_2)$.

Third, define $i : V_1 \times V_2 \to V_1 \otimes V_2$ as $i(v_1, v_2) \doteq v_1 \otimes v_2$. It is not that hard to show that it is a bilinear map. After all these steps, the story so far can be described by the following diagram:

$$\begin{tikzcd}
 & \mathcal{F}(V_1 \times V_2) \arrow[d, dashed, "\tilde{\varphi}"] \\
V_1 \times V_2 \arrow[ur, hook, "\iota"] \arrow[r, "\varphi"] \arrow[dr, "i"] & W \\
 & V_1 \otimes V_2.
\end{tikzcd}$$

We are missing one arrow! In order to obtain this arrow, we make use of the following lemma, whose proof we momentarily skip for sake of text continuity:

Lemma A.1. *Let* $T : U \to W$ *be a linear map between vector spaces and* $\mathcal{R} \subseteq U$ *a subspace satisfying* $\mathcal{R} \subseteq \ker T$*. Denote the canonical projection map* $u \mapsto [u]$ *by* π*. There exists a unique linear transformation* $\tilde{T} : U/R \to W$ *such that the following diagram*

$$\begin{tikzcd}
U \arrow[r, "T"] \arrow[d, "\pi"] & W \\
U/\mathcal{R} \arrow[ur, dashed, "\tilde{T}"] &
\end{tikzcd}$$

commutes.

By virtue of the this lemma, if we set $U = \mathcal{F}(V_1 \times V_2)$ and $T = \tilde{\varphi}$, then there exists a unique linear map $\tilde{T} \equiv \varphi_{\otimes} : V_1 \otimes V_2 \to W$ such that

$$\begin{tikzcd}
 & \mathcal{F}(V_1 \times V_2) \arrow[d, dashed, "\exists! \tilde{\varphi}"] \arrow[dd, bend left=60, "\pi"] \\
V_1 \times V_2 \arrow[ur, hook, "\iota"] \arrow[r, "\varphi"] \arrow[dr, "i"] & W \\
 & V_1 \otimes V_2. \arrow[u, dashed, "\exists! \varphi_{\otimes}"]
\end{tikzcd}$$

In other words, we have essentially proven the existence part of the following theorem:

Theorem A.1. *Let V_1, V_2 be real vector spaces. There exists a unique vector space $V_1 \otimes V_2$ and a unique bilinear map $i : V_1 \times V_2 \to V_1 \otimes V_2$ with the following universal property:*

$$\begin{tikzcd} V_1 \times V_2 \arrow[r, "\forall\ \mathtt{bilinear}\ \varphi"] \arrow[d, hook, "i"] & \forall\, W \\ V_1 \otimes V_2 \arrow[ur, dashed, "\exists!\ \mathtt{linear}\ \varphi_\otimes"'] & \end{tikzcd}$$

Proof of theorem A.1. All that is left is to prove uniqueness. Assume there's another vector space M and other bilinear map $m : V_1 \times V_2 \to M$ that also satisfies the diagram above. This means that for $\varphi = i$ and $W = V_1 \otimes V_2$, there exists a unique linear map $i' : M \to V_1 \otimes V_2$ such that $i' \circ m = i$. Similarly, since $(i, V_1 \otimes V_2)$ has the same universal property, there exists a unique linear map $m' : V_1 \otimes V_2 \to M$ such that $m' \circ i = m$. Pictorially, we have the following commutative diagram:

$$\begin{tikzcd} & V_1 \otimes V_2 \arrow[d, dashed, "m'"] \\ V_1 \times V_2 \arrow[ur, hook, "i"] \arrow[r, "m"] \arrow[dr, hook, "i"'] & M \arrow[d, dashed, "i'"] \\ & V_1 \otimes V_2. \end{tikzcd}$$

This means that $i' \circ m' : V_1 \otimes V_2 \to V_1 \otimes V_2$ is a linear map such that $(i' \circ m') \circ i = i$. By making use of the universal property of the pair $(i, V_1 \otimes V_2)$ to itself, it follows that $i' \circ m' = 1_{V_1 \otimes V_2}$, so m' has a left inverse, which means that it is an injective map. By similar reasoning, $m' \circ i' = 1_M$, so m' has a right inverse, which means it is a surjective map. Since the left and right inverses are the same and m' is linear, we have that m' is a vector space isomorphism. ■

Now, all that is left is to prove the lemma aforementioned.

Proof of lemma A.1. Define the map $\tilde{T} : U/\mathcal{R} \to W$ as

$$\tilde{T}([u]) \doteq T(u).$$

If $\tilde{T}$ is well-defined, then, first of all, $\tilde{T}$ is a linear map that obeys the commutative diagram. To show that it is indeed well-defined, let $u' \in [u]$:

$$\begin{aligned} u' \in [u] &\iff u - u' \in \mathcal{R} \subseteq \ker T \\ &\implies T(u - u') = 0_W \\ &\implies T(u) = T(u') \\ &\implies \tilde{T}([u]) = \tilde{T}([u']). \end{aligned}$$

Hence, $\tilde{T}$ does not depend on the representative of the equivalence class. To prove uniqueness, let $M : U/\mathcal{R} \to W$ be another linear map that obeys the diagram. Then, for any $[u]$, $M([u]) = T(u) \doteq \tilde{T}([u])$. Thus, $M = \tilde{T}$. ■

All work done so far can easily be generalised to the case where we have $k \in \mathbb{N}$ real vector spaces $V_1, \ldots, V_k$ and a multilinear map $\varphi : V_1 \times \ldots \times V_k \to W$. I will just state the theorem, but the idea of the proof is entirely analogous to what we did to the case $k = 2$.

Theorem A.2. *Let $V_1, \ldots V_k$, and W be real vector spaces, $k \in \mathbb{N}$. There exists a unique vector space $V_1 \otimes \ldots \otimes V_k$ and a unique multilinear map $\varphi_{\otimes} : V_1 \otimes \ldots \otimes \ldots V_k \to W$ with the following universal property:*

$$\begin{tikzcd}
V_1 \times \ldots \times V_k \arrow[r, "\forall\ \text{multilinear}\ \varphi"] \arrow[d, hook, "i"] & \forall W \\
V_1 \otimes \ldots \otimes V_k . \arrow[ur, dashed, "\exists!\ \text{linear}\ \varphi_{\otimes}"'] &
\end{tikzcd}$$

So far so good. Now, it is time to verify how these objects look like when we introduce ordered basis in each of the vector spaces $V_1, \ldots, V_k$. For simplicity, from now on all vector spaces under consideration will be **finite dimensional**.

Proposition A.2 (A basis for $V_1 \otimes \ldots \otimes V_k$)**.** *Let $V_1, \ldots, V_k$ be real vector spaces and for each $1 \leqslant j \leqslant k$, let $(e_1^{(j)}, \ldots, e_{n_j}^{(j)})$ be an ordered basis for the n_j-dimensional vector space V_j. The set*

$$\mathcal{B} = \{e_{i_1}^{(1)} \otimes \ldots \otimes e_{i_k}^{(k)} : i_1 \in [n_1], \ldots, i_k \in [n_k]\}$$

defines a basis in $V_1 \otimes \ldots \otimes V_k$.

Proof. It is evident that $\mathrm{span}(\mathcal{B}) \subseteq V_1 \otimes \ldots \otimes V_k$. To show the opposite inclusion, consider a tensor of the form $v_1 \otimes \ldots \otimes v_k$. By expanding each v_j in terms of the ordered basis chosen in V_j, it is evident that $v_1 \otimes \ldots \otimes v_k \in \mathrm{span}(\mathcal{B})$. Since any tensor in $V_1 \otimes \ldots \otimes V_k$ is a linear combination of elements like $v_1 \otimes \ldots \otimes v_k$, it is clear that the converse inclusion must also hold. Hence, $V_1 \otimes \ldots \otimes V_k = \mathrm{span}(\mathcal{B})$.

Now, we must prove linear independence. To avoid typing sums, we adopt the Einstein summation convention (repeated upper and lower indices are being summed).

Let $\alpha^{i_1 \ldots i_k} \in \mathbb{R}$ be coefficients such that

$$\alpha^{i_1 \ldots i_k} e_{i_1}^{(1)} \otimes \ldots \otimes e_{i_k}^{(k)} = 0.$$

Now, for each $k - \mathrm{uple}$ $(m_1, \ldots, m_k$ we construct a multilinear map $\psi^{m_1, \ldots, m_k} : V_1 \times \ldots \times V_k \to \mathbb{R}$ as

$$\psi^{m_1 \ldots m_k}(v_1, \ldots, v_k) \doteq e_{(1)}^{m_1}(v_1) \ldots e_{(k)}^{m_k}(v_k),$$

where $e_{(j)}^{i_j}$ represents an element of the dual basis to $e_{i_j}^{(j)}$. By virtue of theorem A.2, there is a linear map $\tilde{\psi}^{m_1 \ldots m_k} : V_1 \otimes \ldots \otimes V_k \to \mathbb{R}$ that $\tilde{\psi}^{m_1, \ldots, m_k} \circ i = \psi^{m_1 \ldots m_k}$. Hence, by virtue of each such map:

$$\begin{aligned}
\tilde{\psi}^{m_1 \ldots m_k}(\alpha^{i_1 \ldots i_k} e_{i_1}^{(1)} \otimes \ldots \otimes e_{i_k}^{(k)}) &= \alpha^{i_1 \ldots i_k} \psi^{m_1 \ldots m_k}(e_{i_1}^{(1)}, \ldots, e_{i_k}^{(k)}) \\
&\doteq \alpha^{i_1 \ldots i_k} e_{(1)}^{m_1}(e_{i_1}^{(1)}) \ldots e_{(k)}^{m_k}(e^{(k)} i_k) \\
&= \alpha^{i_1 \ldots i_k} \delta_{i_1}^{m_1} \ldots \delta_{i_k}^{m_k} \\
&= \alpha^{m_1 \ldots m_k} \\
&= 0.
\end{aligned}$$

Linear independence easily follows from that. ∎

Corollary A.1. $\dim(V_1 \otimes \ldots \otimes V_k) = \prod_{j=1}^{k} n_j$.

A.2 Multilinear Maps and (p, q)-tensors

Hitherto, tensors are extremely abstract objects obtained by taking quotients of vector spaces. In the case where all these spaces are finite-dimensional, the tensor product is actually isomorphic to "more concrete" space, name the vector space of real-valued multilinear maps $L(V_1, \dots, V_k; \mathbb{R})$.

Proposition A.3. *Let $V_1, \dots, V_k$ be real vector spaces and for each $1 \leqslant j \leqslant k$, let $(e_1^{(j)}, \dots, e_{n_j}^{(j)})$ be an ordered basis for the n_j-dimensional vector space V_j. Denote the ordered dual basis for V_j^* as $(e^1_{(j)}, \dots, e^{n_j}_{(j)})$. The set*

$$\mathcal{B}^* \doteq \{e^{i_1}_{(1)} \otimes \dots \otimes e^{i_k}_{(k)} \, : \, i_1 \in [n_1], \dots, i_k \in [n_k]\}$$

is a basis for $L(V_1, \dots, V_k; \mathbb{R})$.

Corollary A.2. $\dim L(V_1, \dots, V_k; \mathbb{R}) = \prod_{j=1}^{k} n_j$.

Proposition A.4 (Tensors as Multilinear Maps)**.** *Let $V_1, \dots, V_k$ be finite-dimensional real vector spaces. It is true that*

$$V_1^* \otimes \dots \otimes V_k^* \simeq L(V_1, \dots, V_K; \mathbb{R}). \tag{A.1}$$

Proof. Let $V_1, \dots, V_k$ be vector spaces as stated above. By corollaries A.1 and A.2, $V_1 \otimes \dots \otimes V_k$ and $L(V_1, \dots, V_k; \mathbb{R})$ are both finite-dimensional real vector spaces with the same dimension. Consequently, they are isomorphic. ∎

By virtue of the canonical isomorphism $V_j^{**} \simeq V_j$, proposition A.4 also implies that

$$V_1 \otimes \dots \otimes V_k \simeq L(V_1^*, \dots, V_k^*; \mathbb{R}). \tag{A.2}$$

We are now ready to talk what covariant, contravariant and mixed tensors are. These definitions are extremely important in general relativity contexts, for they make it clear what kind of object we are working with[1].

Definition A.2. *Let p and q be non-negative integers and V be a real vector space. The (p, q)-tensor product of V is defined as the tensor product*

$$T^{(p,q)}(V) \doteq \underbrace{V \otimes \dots \otimes V}_{p \text{ times}} \otimes \underbrace{V^* \otimes \dots \otimes V^*}_{q \text{ times}}.$$

An element of such tensor product is called a (p,q)-tensor or, alternatively, a rank (p,q) tensor.

Covariant tensors are the $(0, q)$-tensors and **contravariant tensors** are the $(p, 0)$-tensors. We denote these tensor products respectively as

$$T^{(0,q)}(V) = T^q(V^*),$$

$$T^{(p,0)}(V) = T^p(V).$$

Since these tensors products can be reinterpreted as multilinear maps, we also have that

$$T^{(0,q)}(V) \simeq L(\underbrace{V, \dots, V}_{q \text{ times}}; \mathbb{R}),$$

$$T^{(p,0)}(V) \simeq L(\underbrace{V^*, \dots, V^*}_{p \text{ times}}; \mathbb{R}).$$

This means that **covariant tensors eat vectors**, while **contravariant tensors eat covectors**.

[1]Besides, some objects, such the electromagnetic field tensor, may change how they look like if we work in terms of covariant or contravariant components.

Definition A.3. *Let* V *be finite-dimensional real vector space. A rank k covariant tensor* $A \in T^k(V^*)$ *is* ***symmetric*** *if and only if for any* $v_1, \dots, v_k \in V$:

$$A(v_1, \dots, v_i, \dots, v_j, \dots, v_k) = \alpha(v_1, \dots, v_j, \dots, v_i, \dots, v_k)$$

whenever $1 \leqslant i < j \leqslant k$.

The set of all symmetric, rank k, covariant tensors is a vector subspace of $T^k(V^*)$.

Let S_k be the permutation group of $\{1, 2, \dots, k\}$. We can always turn a tensor $A \in T^k(V^*)$ into a symmetric one by virtue of the Symmetrization operator $\mathrm{Sim} : T^k(V^*) \to \Sigma^k(V^*)$, which is defined via:

$$\mathrm{Sym}\, A \doteq \frac{1}{k!} \sum_{\sigma \in S_k} \sigma A.$$

Proposition A.5. *Let* V *be a real vector space and* $A \in T^k(V^*)$. *The following statements are equivalent:*

a) A *is symmetric.*

b) For any $v_1, \dots, v_k \in V$, $A(v_1, \dots, v_k)$ *remains unaltered by any rearrangement of* $v_1, \dots, v_k$.

c) In any basis for V, *the components* $A_{i_1 \dots i_k}$ *remain unaltered by any rearrangement of their indices.*

Exercise A.1. *Let* V *be a real vector space and* $A \in T^k(V^*)$. *Then*

a) $\mathrm{Sym}\, A$ *is symmetric.*

b) $\mathrm{Sym}\, A = A \iff A \in \Sigma^k(V^*)$.

Let $A \in \Sigma^k(V^*)$ and $B \in \Sigma^l(V^*)$, there is no guarantee that $A \otimes B \in T^{(k+l)}(V^*)$ will be a symmetric tensor too. However, by virtue of the operator Sym, we can define a so-called **symmetric product**, which acts on A and B as:

$$AB \doteq \mathrm{Sym}(A \otimes B).$$

This means that

$$\alpha\beta(v_1, \dots, v_{k+l}) = \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \alpha(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \beta(v_{\sigma(k+1)}, \dots, v_{\sigma(k+l)}).$$

For example, if A and B are rank 1 covariant tensors (which are always symmetric), then

$$AB = \frac{1}{2}(A \otimes B + B \otimes A),$$

is a symmetric (0,2)-tensor.

A.3 Antisymmetric Tensors

Just like we can define symmetric covariant tensors, we can define antisymmetric tensors. Such tensors are somewhat special[2] and deserve their own section in this appendix.

Definition A.4. *Let* V *be a finite-dimensional real vector space. A rank* k *covariant tensor* $A \in T^k(V^*)$ *is* ***antisymmetric*** *if and only if for any* $v_1, \dots, v_k \in V$:

$$A(v_1, \dots, v_i, \dots, v_j, \dots, v_k) = -A(v_1, \dots, v_j, \dots, v_i, \dots, v_k),$$

whenever $1 \leqslant i < j \leqslant k$.

These objects go under many other names: k-covectors, covectors or exterior forms, for example. The set of all such tensors is denoted by $\Lambda^k(V^*)$ and it is also a vector subspace of $T^k(V^*)$.

[2]They will help us define differential forms, which are the objects that one integrates on a manifold.

Claim 1. *Let V be a real vector space and $A \in T^k(V^*)$. The following statements are equivalent:*

a) A *is antisymmetric.*

b) *For any $v_1, \dots, v_k \in V$ and any element $\sigma \in S_k$:*

$$A(v_{\sigma(1)}, \dots, v_{\sigma(k)}) = (\mathrm{sgn}\ \sigma) A(v_1, \dots, v_k),$$

with $\mathrm{sgn}\ \sigma \doteq (-1)^m$, where m is the number of transpositions obtained by decomposing σ.

c) *In any basis for V, the components $A_{i_1 \dots i_k}$ change sign for any transposition of indices.*

Furthermore, these are also equivalent:

a') A *is antisymmetric.*

b') $A(v_1, \dots, v_k) = 0$ *if the k-uple $(v_1, \dots, v_k)$ is linearly dependent.*

c') $A(v_1, \dots, v_k) = 0$ *whenever two entries of $(v_1, \dots, v_k)$ are equal.*

The version of the Sym operator for antisymmetric tensors is the operator $\mathrm{Antisym} : T^k(V^*) \to \Lambda^k(V^*)$, named **antisymmetrization**, defined by:

$$\mathrm{Antisym}\, A \doteq \frac{1}{k!} \sum_{\sigma \in S_k} (\mathrm{sgn}\ \sigma)(\sigma A).$$

Explicitly, this means that:

$$(\mathrm{Antisym}\, A)(v_1, \dots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} (\mathrm{sgn}\ \sigma) A(v_{\sigma(1)}, \dots, v_{\sigma(k)}).$$

Clearly, for $A \in T^{(0,2)}(V^*)$:

$$(\mathrm{Antisym}\, A)(v, w) = \frac{1}{2}(A(v, w) - A(w, v)).$$

Exercise A.2. *Let V be a real vector space and $A \in T^k(V^*)$. Then*

a) $\mathrm{Antisym}\, A$ *is antisymmetric.*

b) $\mathrm{Antisym}\, A = A \iff A \in \Lambda^k(V^*)$.

Some Antisymmetric Tensors

A **k multi-index** is defined as a ordered k-uple $(i_1, \dots, i_k)$ of positive integers. If $\sigma \in S_k$, define I_σ as

$$I_\sigma \doteq (i_{\sigma(1)}, \dots, i_{\sigma(k)}).$$

Fix an ordered basis $(e^1, \dots, e^n)$ on V^*. For any k multi-index I, define the (0,k)-tensor ε^I as follows:

$$\varepsilon^I(v_1, \dots, v_k) \doteq \det \begin{pmatrix} e^{i_1}(v_1) & \dots & e^{i_1}(v_k) \\ \vdots & \ddots & \vdots \\ e^{i_k}(v_1) & \dots & e^{i_k}(v_k) \end{pmatrix} = \det \begin{pmatrix} v_1^{i_1} & \dots & v_k^{i_1} \\ \vdots & \ddots & \vdots \\ v_1^{i_k} & \dots & v_k^{i_k} \end{pmatrix}.$$

Clearly that is an antisymmetric tensor, let us call it an elementary alternating tensor.

For k multi-indices I e J a **generalized Kronecker delta** by setting

$$\delta^I_J \doteq \det \begin{pmatrix} \delta^{i_1}_{j_1} & \dots & \delta^{i_1}_{j_k} \\ \vdots & \ddots & \vdots \\ \delta^{i_k}_{j_1} & \dots & \delta^{i_k}_{j_k} \end{pmatrix}.$$

One can show that it satisfies

$$\delta^I_J = \begin{cases} \operatorname{sgn} \sigma, & \text{if neither I nor J have repeated indices and } J = I_\sigma,\ \sigma \in S_k. \\ 0, & \text{otherwhise.} \end{cases}$$

A way do prove the property above is by rewriting the RHS of the definition of δ^I_J as:

$$\delta^I_J = \sum_{\eta \in S_k} (\operatorname{sgn} \eta) \delta^{i_1}_{j_{\eta(1)}} \dots \delta^{i_k}_{j_{\eta(k)}}.$$

Claim 2. *Fix a basis $\{e_i\}_{i=1}^n$ on V and let $\{e^i\}_{i=1}^n$ be its dual basis. For any k multi-indices I and J:*

a) If I has repeated indices, then $\varepsilon^I = 0$.

b) If $J = I_\sigma$ for some $\sigma \in S_k$, then $\varepsilon^I = (\operatorname{sgn} \sigma)\varepsilon^J$.

c) For any k-uple of basis elements

$$\varepsilon^I(e_{j_1}, \dots, e_{j_k}) = \delta^I_J.$$

Proposition A.6. *Let V be a real vector space. If $\{e^i\}_{i=1}^n$ is a basis for V^*, then for every positive integer $k \leqslant \dim V = n$, the set*

$$\mathcal{E} = \{\varepsilon^I : i_1 < \dots < i_k\}$$

is a basis for $\Lambda^k(V^*)$ and

$$\dim \Lambda^k(V^*) = \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Proof.

Let $k > n$, then by statement b′ from claim 1, we have $\mathcal{E} = \{0\}$, for any k-uple of vectors is linearly dependent in this situation.

Now consider the case $k \leqslant n$. If $A \in \Lambda^k(V^*)$ and $I = (i_1, \dots, i_k)$, define

$$a_I \doteq A(e_{i_1}, \dots, e_{i_k}).$$

Thus, for any other k multi-index J,

$$\sum_I{}' a_I \varepsilon^I(e_{j_1}, \dots, e_{j_k}) \doteq \sum_{\{I:\ i_1 < \dots < i_k\}} a_I \varepsilon^I(e_{j_1}, \dots, e_{j_k}) = a_J.$$

This means that $\sum_I' \alpha_I \varepsilon^I = A$, which implies $\Lambda^k(V^*) = \operatorname{span} \mathcal{E}$.

To verify linear independence, let β_I be real numbers such that

$$\sum_I{}' \beta_I \varepsilon^I = 0.$$

Then for any increasing k multi-index J:

$$\sum_I{}' \beta_I \varepsilon^I(e_{j_1}, \dots, e_{j_k}) = \beta_J = 0.$$

■

Proposition A.7. *Let V be an n-dimensional real vector space and $A \in \Lambda^n(V^*)$. For any linear operator $T:\ V \to V$ and any n-uple $(v_1, \dots, v_n)$,*

$$A(T(v_1), \dots, T(v_n)) = (\det T) A(v_1, \dots, v_k).$$

This result is very similar to the change of coordinates formula for integrals. In fact, I believe it generalises that formula, since the Jacobian matrix is a particular linear operator.

Proof.
Let $\{e_i\}_{i=1}^n$ be a basis for V and $\{e^i\}_{i=1}^n$ be its dual basis. Moreover, let (T^i_j) be the matrix representation of the linear operator T in the basis $\{e_i\}$. Due to multilinearity and antisymmetry of A, we just have to check the equality holds in the case $(v_1, \dots, v_n) = (e_1, \dots, e_n)$.

By proposition A.6, it follows that $\dim \Lambda^n(V^*) = 1$. Therefore, $A = a\varepsilon^{1\dots n}$ for some $a \in \mathbb{R}$. Hence, we can write the RHS above as:

$$(\det T) a\varepsilon^{1\dots n}(e_1, \dots, e_n) = a(\det T).$$

On the other hand, the LHS can be written as:

$$a\varepsilon^{1\dots n}(T(E_1), \dots, T(E_n)) = a \det\big(\varepsilon^j T(E_i)\big) = a \det T.$$

Since they both agree, the proof is essentially finished. ■

Exterior Product

If one wants to construct antisymmetric tensors of higher ranks starting from previous ones, one could do something similar to the symmetric product of symmetric tensors: let $\omega \in \Lambda^k(V^*)$ and $\eta \in \Lambda^l(V^*)$, the **exterior product** of these antisymmetric tensors is the $(k+l)$-covector defined as:

$$\omega \wedge \eta \doteq \frac{(k+l)!}{k!l!} \mathcal{Antisym}\,(\omega \otimes \eta). \tag{A.3}$$

The coefficients in front of $\mathcal{Alt}\,(\omega \wedge \eta)$ are justified by the following result:

Lemma A.2. *Let V be an n-dimensional real vector space and let $\{e^i\}_{i=1}^n$ be a basis for V^*. For any multi-indices I e J:*

$$\varepsilon^I \wedge \varepsilon^J = \varepsilon^{IJ}, \tag{A.4}$$

for $IJ \doteq (i_1, \dots, i_k, j_1, \dots, j_l)$.

Proof.
Let $\{E_i\}_{i=1}^n$ be the basis for V dual to $\{e^i\}_{i=1}^n$ and $P = (p_1, \dots, p_{k+l})$. Due to multilinearity, it is sufficient to check equation (A.4) in each case below.

Case I) If P has repeated indices, then Claim 1 implies that $\varepsilon^I \wedge \varepsilon^J(E_{p_1}, \dots, E_{p_{k+l}}) = 0$ and $\varepsilon^{IJ}(E_{p_1}, \dots, E_{p_{k+l}}) = 0$.

Case II) If P has a index that does not appear in I or J, then

$$\varepsilon^I \wedge \varepsilon^J(E_{p_1}, \dots, E_{p_{k+l}}) \doteq \frac{(k+l)!}{k!l!} \mathcal{Antisym}\,(\varepsilon^I \otimes \varepsilon^J)(E_{p_1}, \dots, E_{p_{k+l}}) = 0,$$

because P is neither a permutation of I nor J.

On the other hand, recall that $\mathcal{Antisym}(\varepsilon^{IJ})$ is a linear combination of $\varepsilon^{\sigma(IJ)}$, where σ is a permutation of IJ. Each term of this sum evaluated at $(E_{p_1}, \dots, E_{p_{k+l}})$ yields a $\delta^{\sigma(IJ)}_P$, which is zero because $IJ \neq P$. Hence:

$$\varepsilon^{IJ}(E_{p_1}, \dots, E_{p_{k+l}}) = 0$$

too.

Case III) $P = IJ$ and there are no repeated indices. Then $\varepsilon^{IJ}(E_{p_1}, \dots, E_{p_{k+l}}) = 1$. On the other hand:

$$\varepsilon^I \wedge \varepsilon^J(E_{p_1}, \dots, E_{p_{k+l}}) =$$

$$\frac{(k+l)!}{k!l!} \mathcal{Antisym}\,(\varepsilon^I \otimes \varepsilon^J)(E_{p_1}, \dots, E_{p_{k+l}}) =$$

$$\frac{1}{k!l!} \sum_{\sigma \in S_{k+l}} (\mathrm{sgn}\,\sigma) \varepsilon^I(E_{p_{\sigma(1)}}, \dots, E_{p_{\sigma(k)}}) \varepsilon^J(E_{p_{\sigma(k+1)}}, \dots, E_{p_{\sigma(k+l)}}) = 1.$$

Case IV) $P = \sigma(IJ)$ for some $\sigma \in S_{k+l}$ and there are no repeated indices. Then, we are in a case equal to the previous one, up to a multiplication by $\mathrm{sgn}\sigma$.

■

The exterior product has the following properties:

Claim 3 (Properties of exterior product). *Suppose that ω, ω', η, η' and ξ are multicovectors on a vector space V. Then*

a) *For any $a, a' \in \mathbb{R}$:*

$$(a\omega + a'\omega') \wedge \eta = a(\omega \wedge \eta) + a'(\omega' \wedge \eta)$$

$$\eta \wedge (a\omega + a'\omega') = a(\eta \wedge \omega) + a'(\eta \wedge \omega).$$

b) $\omega \wedge (\eta \wedge \xi) = (\omega \wedge \eta) \wedge \xi$.

c) *If $\omega \in \Lambda^k(V^*)$ and $\eta \in \Lambda^l(V^*)$ then*

$$\omega \wedge \eta = (-1)^{kl} \eta \wedge \omega.$$

d) *Se $\{e^i\}_{i=1}^n$ is a basis on V^* and I is a k multi-index, then*

$$\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_k} = \varepsilon^I.$$

e) *If $\omega^1, \ldots, \omega^k \in V^*$ e $v_1, \ldots, v_k \in V$, then*

$$\omega^1 \wedge \ldots \wedge \omega^k(v_1, \ldots, v_k) = \det\big(\omega^i(v_j)\big).$$

It is clear from the claim above that the vector space

$$\Lambda(V^*) = \bigoplus_{k=1}^{n} \Lambda^k(V^*).$$

is an associative, anticommutative algebra which we will name as the **exterior algebra**[3].

[3] Also named a **Grassmann algebra**. The case when $V = \mathcal{C}$ is of utter importance for the construction of a classical theory of fermionic fields.

Manifolds (Speedrun any %)

B.1 Smooth Manifolds

Since this text is not aimed for mathematicians, I will purposely omit tons of topological technicalities when talking about smooth manifolds. Nonetheless, we proceed as follows: an n-dimensional smooth manifold is a pair $(M, \mathcal{A})$, where M denotes an n-dimensional topological manifold and $\mathcal{A}$ is a smooth structure (alternatively, maximal atlas) on M. Alright, what do all these words mean? Roughly speaking, by n-dimensional topological manifold, I mean the following:

a) M has a collection of subsets τ, whose elements are called *open sets*, that translates the notion of "local".In other words, whenever we say "locally valid near a point", it just means that we've chosen an open subset that cointains said point.

b) Every point $p \in M$ belongs to some $U \in \tau$, that is, every point has a so-called open neighbourhood U.

c) For every $p \in M$ there is an open neighbourhood U and a continuous bijection $\varphi : U \to \mathbb{R}^n$ whose inverse is also continuous[1]. This feature means that M is *locally euclidean* and we call the pair (U, φ) a *coordinate system around* p.

To put it shortly, a **topological manifold is a space that locally resembles euclidean space**[2]. Now what do we mean by *smooth structure*? Well, it is just a maximal collection of coordinate systems $\mathcal{A} = \{(U_\alpha, \varphi_\alpha)\}$ that satisfies:

a) For each $p \in M$ there is an open neighbourhood U_α for some α.

b) If (U, φ) e (V, ψ) are overlapping coordinate systems, then the maps $\psi \circ \varphi^{-1}$ and $\varphi \circ \psi^{-1}$ are C^∞ in the appropriate domains. This conditions is called *compatibility between transition maps*.

In essence, **a smooth manifold is a space that locally resembles euclidean space and whose transition maps are smooth**. This is illustrated in Figure B.1.

A map $F : M \to \mathcal{N}$ betwenn smooth manifolds is smooth if its representations via coordinate systems are always smooth maps between subsets of $\mathbb{R}^n$. More precisely for any coordinates systems (U, φ) and (V, ψ) in M and $\mathcal{N}$, respectively. the map $\psi \circ F \circ \varphi^{-1} : \varphi(U \cap F^{-1}(V)) \to \psi(F(U) \cap V) \subset \mathbb{R}^n$ is smooth. Take notice that this only makes sense when we look at $p \in U$ such that $F(p) \in V$.

[1]In this context, continuous means that pre-image of open sets in one space are open sets in the other space.

[2]Hausdorff and second countable conditions won't be very important in our context, so I have avoided to talk about them

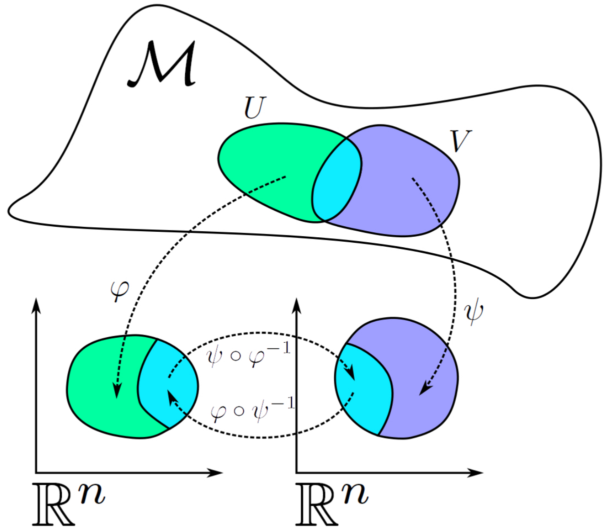

Figure B.1: Depiction of transition maps. Adapted from Stomatapoll (2012).

Example B.1. *$\mathbb{R}^n$ is, perhaps, the most trivial example of what a smooth manifold is. We can always establish the (global) coordinate system $(\mathbb{R}^n, \mathtt{id})$, where $\mathtt{id} : \mathbb{R}^n \to \mathbb{R}^n$ denotes the identity map. As an maximal atlas always exists (and is unique), then all we have to do is to consider the maximal atlas which contains this coordinate system.*

Example B.2. *Another elucidative example of what a smooth manifold is:*

$$S^1 \doteq \{(x, y) \in \mathbb{R}^2 : \ x^2 + y^2 = 1\}.$$

Which coordinate systems describe S^1? Well, for $i \in \{1, 2\}$, we define $(U_i^{\pm}, \varphi_i^{\pm})$ as following::

$$U_i^+ \doteq \{(x_1, x_2) \in S^1 : \ x_i > 0\}$$
$$U_i^- \doteq \{(x_1, x_2) \in S^1 : \ x_i < 0\},$$

where $\varphi_1^{\pm}(x_1, x_2) \doteq y$ and $\varphi_2^{\pm}(x_1, x_2) \doteq x$. One may generalize these ideas to the set $S^n \subset \mathbb{R}^{n+1}$, $n \in \mathbb{N}$, and infer that S^n is also a smooth manifold.

Example B.3 (Somewhat thecnical?)**.** *The group $GL(n, \mathbb{R}) = \{A \in M_{n\times n}(\mathbb{R}) : \ \det(A) \neq 0\}$. The subset $\mathbb{R} \setminus \{0\}$ is open and the determinant is a continuous map, so $GL(n, \mathbb{R}) = \det^{-1}(\mathbb{R} \setminus \{0\})$ is an open set $\mathbb{R}^{n^2}$. Since any open set U in a manifold M is also a manifold, we conclude that $GL(n, \mathbb{R})$ is a smooth manifold.*

It is worth mentioning that one can construct smooth manifolds by considering finite Cartesian products of smooth manifolds and considering atlases like this one:

$$\mathcal{A} = \{(U_\alpha \times V_\beta, \varphi_\alpha \times \psi_\beta) : \ (U_\alpha, \varphi_\alpha) \in \mathcal{A}_1, \ (V_\beta, \psi_\beta) \in \mathcal{A}_2\}.$$

and extending them to the maximal atlas.

From this remark, it follows that a cylinder $S^1 \times \mathbb{R}$ and the n-torus $\mathbb{T}^n = S^1 \times \ldots \times S^1$ are also smooth manifolds.

B.2 Tangent Spaces and Tangent Bundles

To any point $p \in M$ in a smooth manifold, one may associate a n-dimensional real vector space T_pM called the **tangent space at p**. Geometrically, this space contains all velocity vectors at p of curves which pass through p. Since velocity vectors are derivatives at some point, we may also name them as **derivations at p**. More precisely:

Definition B.1. *Let* M *be a n-dimensional smooth manifold. Given* $p \in M$*, a derivation at* p *is a linear map* $v : C^\infty(M) \to \mathbb{R}$ *such that:*

$$v(fg) = f(p)vg + g(p)vf, \; \forall f, g \in C^\infty(M),$$

where $C^\infty(M) \doteq \{f : M \to \mathbb{R} | \; f \textit{ is smooth}\}$.

Definition B.2. *Let* M *be a n-dimensional smooth manifold and* $p \in M$*. The* ***tangent space to*** M ***at p*** *is the real vector space*

$$T_pM \doteq \{v : C^\infty(M) \to \mathbb{R} \mid v \textit{ is a derivation at } p\}.$$

Claim 4. *For every point* p *in an* n*-dimensional manifold* M*,* $\dim T_pM = n$*.*

One may verify the claim above by first establishing a coordinate system around p[3].

Let (U, φ) be a coordinate system around $p \in M$ and $f \in C^\infty(M)$. It follows that $f \circ \varphi^{-1} : \varphi(U) \to \mathbb{R}$ is smooth in$\varphi(U) \subset \mathbb{R}^n$. Thus we define the **partial derivative of** f **with respect to** μ**-th coordinate at p** as:

$$\left.\frac{\partial f}{\partial x^\mu}\right|_p \doteq \left.\frac{\partial (f \circ \varphi^{-1})}{\partial x^\mu}\right|_{\varphi(p)}.$$

Two important remarks: first, the partial derivative is chart-dependent; second, in the RHS the partial derivative of $f \circ \varphi^{-1}$ is the usual one from Calculus, whereas the object in the LHS is just a *symbol* to represent the number obtained from the RHS. Nonetheless, we may treat the partial derivative of $f \in C^\infty(M)$ as the action of a vector in T_pM on f: define $\partial/\partial x^\mu$ as

$$\left(\left.\frac{\partial}{\partial x^\mu}\right|_p\right) f \doteq \left.\frac{\partial f}{\partial x^\mu}\right|_p.$$

This means that a coordinate system induces a basis in T_pM named the **coordinate basis**. In such basis, any tangent vector is written as:

$$v = v^\mu \left.\frac{\partial}{\partial x^\mu}\right|_p.$$

Example B.4 (Change of coordinates). *Let*(U, φ) *and* (V, ψ) *be overlapping coordinate systems. Each of these yield a basis in* T_pM*,* $p \in U \cap V$*. Denote these bases as* $\left\{\partial/\partial x^\mu|_p\right\}$ *and* $\left\{\partial/\partial \tilde{x}^\nu|_p\right\}$*, respectively. If* $v \in T_pM$*, then relative to the* (U, φ) *basis:*

$$\begin{aligned} v(f) &= v^\mu \left.\frac{\partial f}{\partial x^\mu}\right|_p \\ &= \left.\frac{\partial[((f \circ \psi^{-1}) \circ (\psi \circ \varphi^{-1})]}{\partial x^\mu}\right|_{\hat{p}}, \end{aligned}$$

where $f \in C^\infty(M)$ *and* $\hat{p} \equiv \varphi(p)$*. Now, since* $(f \circ \psi^{-1}) \in C^\infty(\mathbb{R}^n; \mathbb{R})$ *and* $(\psi \circ \varphi^{-1}) \in C^\infty(\mathbb{R}^n; \mathbb{R}^n)$*, by the usual chain rule:*

$$\left.\frac{\partial f}{\partial x^\mu}\right|_p = \left.\frac{\partial (f \circ \psi^{-1})}{\partial (\psi \circ \varphi^{-1})^\nu}\right|_{\psi^{-1}(p)} \left.\frac{\partial (\psi \circ \varphi^{-1})^\nu}{\partial x^\mu}\right|_{\hat{p}}.$$

Since $(\psi \circ \varphi^{-1})^\nu = \tilde{x}^\nu$ *and* f *is arbitrary:*

$$\left.\frac{\partial f}{\partial x^\mu}\right|_p = \left.\frac{\partial \tilde{x}^\nu}{\partial x^\mu}\right|_p \left.\frac{\partial f}{\partial \tilde{x}^\nu}\right|_p \iff \left.\frac{\partial}{\partial x^\mu}\right|_p = \left.\frac{\partial \tilde{x}^\nu}{\partial x^\mu}\right|_p \left.\frac{\partial}{\partial \tilde{x}^\nu}\right|_p.$$

Therefore, the components of v *in the* (V, ψ) *chart relate to the components of* v *in* (U, φ) *chart as*

$$\tilde{v}^\nu = \left.\frac{\partial \tilde{x}^\nu}{\partial x^\mu}\right|_p v^\mu. \tag{B.1}$$

[3]The proof is a bit lengthy, but you may check it inWald (2010), Lee (2012) or Tu (2011).

Example B.5. *Let* M *be a 2-dimensional manifold. Consider the charts* $(U, \varphi) \equiv (U, r, \theta)$ *and* $(V, \psi) \equiv (V, x, y)$ *around* $p \in U \cap V$. *Then:*

$$\left.\frac{\partial}{\partial r}\right|_p = \frac{\partial x}{\partial r}(p) \left.\frac{\partial}{\partial x}\right|_p + \frac{\partial y}{\partial r}(p) \left.\frac{\partial}{\partial y}\right|_p$$

$$\left.\frac{\partial}{\partial \theta}\right|_p = \frac{\partial x}{\partial r}(p) \left.\frac{\partial}{\partial x}\right|_p + \frac{\partial y}{\partial \theta}(p) \left.\frac{\partial}{\partial y}\right|_p$$

In terms of φ *and* ψ*:*

$$\frac{\partial x}{\partial r}(p) = \left.\frac{\partial(\psi \circ \varphi^{-1})^1}{\partial r}\right|_{\varphi(p)} \quad e \quad \frac{\partial y}{\partial r}(p) = \left.\frac{\partial(\psi \circ \varphi^{-1})^2}{\partial r}\right|_{\varphi(p)}.$$

If we define $A = \{(x, 0) \in \mathbb{R}^2 : x \in \mathbb{R}\} \cup \{(0, y) \in \mathbb{R}^2 : y \in \mathbb{R}\}$, *the* $\psi \circ \varphi^{-1}$ *is a map from* $(0, \infty) \times (0, 2\pi)$ *to* $\mathbb{R}^2 \setminus A$, *and* $(x(r, \theta), y(r, \theta)) = (r \cos\theta, r \sin\theta)$. *Set* $\varphi(p) = (r_0, \theta_0)$, *then*

$$\left.\frac{\partial}{\partial r}\right|_p = \cos\theta_0 \left.\frac{\partial}{\partial x}\right|_p + \sin\theta_0 \left.\frac{\partial}{\partial y}\right|_p$$

$$\left.\frac{\partial}{\partial \theta}\right|_p = -r_0 \sin\theta_0 \left.\frac{\partial}{\partial x}\right|_p + r_0 \cos\theta_0 \left.\frac{\partial}{\partial y}\right|_p.$$

Definition B.3. *Let* $F : M \to \mathcal{N}$ *be a smooth map. For each* $p \in M$, *define* $dF_p : T_pM \to T_{F(p)}\mathcal{N}$ *as:*

$$dF_p(v)(f) \doteq v(f \circ F), \ \forall v \in T_pM, \ \forall f \in C^\infty(\mathcal{N}).$$

This linear map is called **differential of F at p**.

Claim 5. *In coordinate bases for* T_pM *and* $T_{F(p)}\mathcal{N}$, *the matrix representation of* dF_p *is equal to the Jacobian matrix.*

Proposition B.1. *Let* M, $\mathcal{N}$ *e* $\mathcal{P}$ *be smooth manifolds, and* $F : M \to \mathcal{N}$, $G : \mathcal{N} \to \mathcal{P}$ *be smooth maps.For* $p \in M$ *fixed:*

a) $d(G \circ F)_p = dG_{F(p)} \circ dF_p : T_pM \to T_{G(F(p))}\mathcal{P}$ *(**chain rule**).*

b) $d(\mathbb{1}_M) = \mathbb{1}_{T_pM} : T_pM \to T_pM$.

c) If F *is a diffeomorphism, then* $dF_p : T_pM \to T_{F(p)}\mathcal{N}$ *is an isomorphism and* $(dF_p)^{-1} = dF^{-1}_{F(p)}$.

Proof.

a) Let $v \in T_pM$ and $f \in C^\infty(\mathcal{P})$.

$$\begin{aligned} d(G \circ F)_p (v)f \doteq v\left(f \circ (G \circ F)\right) &\iff d(G \circ F)_p (v)f = dF_p(v)(f \circ G) \\ &\iff d(G \circ F)_p (v)f = \left(dG_{F(p)} \circ dF_p\right)(v)f. \end{aligned}$$

Thus $d(G \circ F)_p = dG_{F(p)} \circ dF_p$.

b) $d(\mathbb{1}_M)_p(v)f \doteq v(f \circ \mathbb{1}_M) = vf = \mathbb{1}_M(v)f$.

■

Exercise B.1. *Prove the last item.*

We now present a different depiction of tangent vectors.

Definition B.4. *Let* $J \subset \mathbb{R}$ *be an interval. A curve on a smooth manifold* M *is a continuous map* $\gamma : J \to M$.

For $t_0 \in J$, the **velocity vector** of a curve γ at t_0 is

$$\gamma'(t_0) \equiv \dot{\gamma}(t_0) \doteq d\gamma\left(\left.\frac{d}{dt}\right|_{t_0}\right) \in T_{\gamma(t_0)}M.$$

Despite the odd appearance, its components in a coordinate basis are quite simple:

$$\gamma'(t) = \frac{dx^\mu}{dt}(t)\left.\frac{\partial}{\partial x^\mu}\right|_p.$$

Claim 6. *Let M be an n-dimensional smooth manifold. For each $p \in M$, any element of T_pM is the velocity vector of some curve γ.*

Proposition B.2. *Let $F : M \to \mathcal{N}$ be a smooth map and $\gamma : J \to M$ a smooth curve. For any $t_0 \in J$, the velocity vector of $F \circ \gamma : J \to \mathcal{N}$ is*

$$(F \circ \gamma)'(t_0) = dF(\gamma'(t_0)).$$

Corollary B.1.

$$dF_p(v) = (F \circ \gamma)'(0)$$

for any smooth curve $\gamma : J \to M$ such that $\gamma(0) = p$ and $\gamma'(0) = v$.

Exercise B.2. *Prove proposition B.2 and its corollary.*

If one considers the collection of all tangent vectors to a smooth manifold M, it is worth to define the so-called **tangent bundle to** M:

$$TM \doteq \{(p, v) :\ p \in M \text{ e } v \in T_pM\} \equiv \bigsqcup_{p \in M} T_pM.$$

The symbol $\bigsqcup$ just emphasizes that we are considering a *disjoint* union. One of the simplest tangent bundles is the TS^1, which is depicted in Figure B.2.

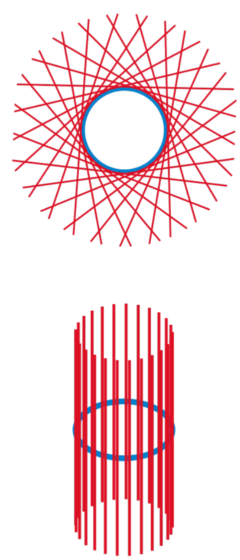

Figure B.2: TS^1. Each lines represents a tangent space to the unit circle. This bundle may be thought as $S^1 \times \mathbb{R}$, which is a cylinder. Source: Alexandrov (2007).

As we defined, TM is just a set. However, one can define opens set in TM by requiring that the map $\pi : TM \to M$, $(p, v) \mapsto p \in M$, is continuous[4]. Since the n-dimensional smooth manifold M already

[4]This is made by the initial topology.

has a maximal atlas and we have managed to define open sets in TM, we can construct coordinate systems as follows:

1. Choose (U_α, ψ_α) be a coordinate system in M.
2. Take notice that the pre-image $\pi^{-1}(U_\alpha) \subseteq TM$ is an open set by definition.
3. Define a map $\Psi_\alpha : \pi^{-1}(U_\alpha) \to \mathbb{R}^{2n}$ by
$$(p, v) \mapsto (\psi_\alpha(p), v^1, \dots, v^n) \in \mathbb{R}^{2n},$$
where the n real numbers v^k are the components of v in the induced basis by U_α.
4. Repeat the previous steps for every coordinate system in the maximal atlas of M.

The collection
$$\mathcal{A}^{TM} = \{(\pi^{-1}(U_\alpha), \Psi_\alpha)\}_\alpha$$
is an atlas for TM. By extending it to the maximal atlas, we have constructed a smooth structure in TM and it is now regarded as a $2n$-dimensional smooth manifold. The triple (TM, π, M) is a kind of smooth vector bundle.

We are now ready to define the following:

Definition B.5. *Let M be a smooth manifold. A **vector field on** M is a continuous map $X : M \to TM$, such that $\pi \circ X = \mathbb{1}_M$.*

Smooth vector fields are defined analogously. The set of all smooth vector fields on a manifold is denoted either by $\Gamma(TM)$ or $\mathfrak{X}(M)$. It clearly is a real vector space, but it is also a $C^\infty(M)$-module.

Besides the very physical definition of vector fields given above, one could also identify these fields as maps from $C^\infty(M)$ into itself via the following prescription:

$$Xf : U \to \mathbb{R}, \ (Xf)(p) \doteq X_p(f).$$

Lastly, given smooth vector fields $X, Y \in \mathfrak{X}(M)$, one defines the **Lie bracket** as a map which send these fields into another smooth vector field, denoted by $[X, Y]$. This new vector field is defined as:

$$[X, Y]f \doteq X(Yf) - Y(Xf).$$

This simple product turns $\mathfrak{X}(M)$ into a **real Lie algebra**; in other words, the Lie bracket satisfies the following:

Proposition B.3. *Let M be a smooth manifold and $X, Y, Z \in \mathfrak{X}(M)$. Then*

(i) *For all $a, b \in \mathbb{R}$:*
$$[aX + bY, Z] = a[X, Z] + b[Y, Z],$$
$$[Z, aX + bY] = a[Z, X] + b[Z, Y].$$

(ii) $[X, Y] = -[Y, X]$.

(iii) $[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0$.

B.3 Cotangent Space

To any vector space V one can associate the vector space of all linear functionals on V, named as the (algebraic) **dual of V**, and denoted by V^*. In view of this fact, given $p \in M$, we define the **cotangent space to** M **at p** as:

$$T_p^*M \doteq (T_pM)^*.$$

Its elements are **covectors**, which I imagine is short for cotangent vector.

Since everything is finite dimensional here, $\dim T_pM = \dim T_p^*M$. Naturally, due to a chart induced basis in T_pM, we can define a basis in T_p^*M via the covectors $\left.dx^\mu\right|_p : T_p^*M \to \mathbb{R}$:

$$\left.dx^\mu\right|_p \left(\left.\frac{\partial}{\partial x^\nu}\right|_p \right) \doteq \delta^\mu_\nu.$$

Because of their nature, if $\omega \in T_p^*M$ e $\nu \in T_pM$, then:

$$\omega(\nu) = \omega\left(\nu^\mu \left.\frac{\partial}{\partial x^\mu}\right|_p\right) = \nu^\mu \omega\left(\left.\frac{\partial}{\partial x^\mu}\right|_p\right) \equiv \omega_\mu \nu^\mu.$$

On the other hand:

$$dx^\mu(\nu) = \nu^\nu \left.dx^\mu\right|_p \left(\left.\frac{\partial}{\partial x^\nu}\right|_p\right) = \nu^\mu.$$

Hence,

$$\omega(\nu) = \omega_\mu \left.dx^\mu\right|_p (\nu) \iff \omega = \omega_\mu \left.dx^\mu\right|_p .$$

Example B.6. *One can define the cotangent bundle* T^*M *in the an analogous fashion as done for* TM. *The real vector space of smooth covector fields is often denote as* $\mathfrak{X}^*(M)$, $\Omega(M)$ *or* $\Gamma(T^*M)$.

Example B.7. *Let* $f : M \to \mathbb{R}$ *be smooth. Its differential* df *is the covector field defined by*

$$df_p(\nu) \doteq \nu f, \text{ for } \nu \in T_pM.$$

In a chart induced basis this means that

$$df = \frac{\partial f}{\partial x^\mu} dx^\mu,$$

where $\partial f/\partial x^\mu$ *is the function* $p \in M \mapsto \left.\frac{\partial f}{\partial x^\mu}\right|_p \in \mathbb{R}$

Exercise B.3. *Let* (U, φ) *and* (V, ψ) *be overlapping charts in a smooth manifold* M. *For a covector* $\omega \in T_p^*M$, *show that*

$$\tilde{\omega}_\nu = \frac{\partial x^\mu}{\partial \tilde{x}^\nu}(p)\omega_\mu, \tag{B.2}$$

and

$$\left.d\tilde{x}^\nu\right|_p = \frac{\partial \tilde{x}^\nu}{\partial x^\mu}(p) \left.dx^\mu\right|_p ,$$

where objects with ~ are being described by (V, ψ).

B.4 Tensor Fields on Manifolds and Differential Forms

Main Bibliography

John M Lee. *Introduction to Riemannian manifolds*. Springer, 2018.

John M. Lee. *Introduction to Smooth Manifolds*. Springer-Verlag New York, 2012.

Roger Penrose. *Techniques in Differential Topology in Relativity*. SIAM, 1987.

Robert M. Wald. *General Relativity*. University of Chicago Press, 2010.

Norbert Straumann. *General relativity*. Springer Science & Business Media, 2013.

Sean M. Carroll. *Spacetime and Geometry: An Introduction to General Relativity*. Cambridge University Press, 2019. doi: 10.1017/9781108770385.

Stomatapoll. Two coordinate charts on a manifold, 2012. URL https://commons.wikimedia.org/wiki/File:Two_coordinate_charts_on_a_manifold.svg. Accessed on 11/08/2021.

Loring Wuliang Tu. *An Introduction to Manifolds*. Universitext. Springer-Verlag, 2011.

Oleg Alexandrov. Tangent bundle, 2007. URL https://commons.wikimedia.org/wiki/File:Tangent_bundle.svg. Accessed on 11/08/2021.