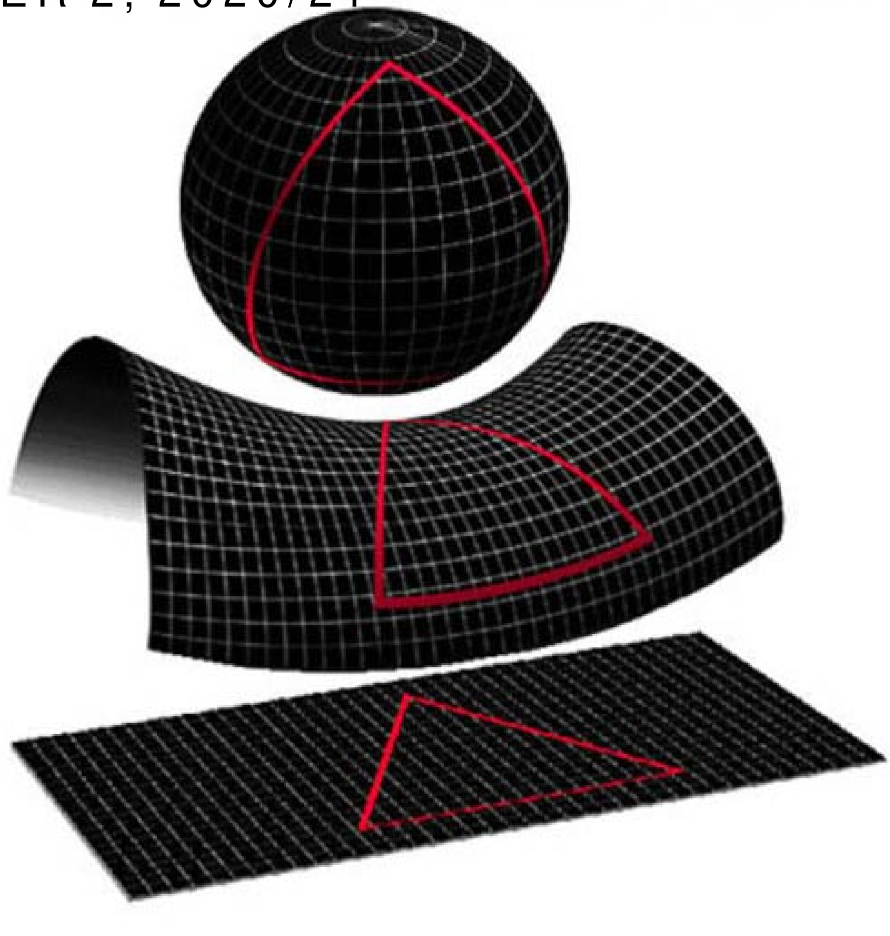
D. TSANG & V. SCOWCROFT

PHYS 40112:

GENERAL RELATIVITY & COSMOLOGY

SEMESTER 2, 2020/21



DEPARTMENT OF PHYSICS, UNIVERSITY OF BATH

PH40112: General Relativity and Cosmology Course Syllabus Semester II, 2020/21

Lecturers:

- Dr. David Tsang, D.Tsang@bath.ac.uk; 3W2.2A.
- Dr. Victoria Scowcroft, V.Scowcroft@bath.ac.uk; 3W2.2E.

Goals: This unit aims to develop a mathematically rigorous description of general relativity and cosmology, including derivation of the Einstein field equations and the exploration of the Friedmann-Robertson-Walker solution, which forms the basis for modern cosmological models. It will also provide an overview of observational cosmology, and the techniques and observations that have constrained current cosmological models. The material covered in this unit will provide a strong foundation for further work in theoretical astrophysics and cosmology.

Requisites: While taking this unit you should be taking or have taken PH30101 and PH30111 or equivalents, or have obtained permission from the Lecturers and your Director of Studies.

Credit and Assessment: 6 Credits. This course is 100% exam.

Online Course Information: The course Moodle `https://moodle.bath.ac.uk/course/view.php?id=58189` is where readings, documents, notes, problem sets and solutions will be posted regularly.

Contact Hours: There will be approximately 100 minutes of Lectures per week, split up by subtopic, along with a 50 minute LOIL/Office hours, where topics will be briefly reviewed and student questions are answered. Each week there will also be a mandatory Worksheet session over zoom. Dr. David Tsang will be presenting the first 6 weeks of material, focusing on General Relativity, while Dr. Victoria Scowcroft will present the following 5 weeks, focusing on Cosmology. Extra office hours, revision sessions, and lecturer availability during the revision week will be posted to the Moodle and discussion forum.

Course Schedule

	Week Starts	Monday (15:15) Zoom	Thursday(14:15) Zoom	Topics pre-reading
1	1 Feb	LOIL 1 / OH (DT)	WS 1 (DT)	Ch.1 Vector Spaces
2	8 Feb	LOIL 2 / OH (DT)	WS 2 (DT)	Ch.2 Tensors
3	15 Feb	LOIL 3 / OH (DT)	WS 2 (DT)	Ch.3 Manifolds
4	22 Feb	LOIL 4 / OH (DT)	WS 3 (DT)	Ch.4 Physics on Mfds.
5	1 Mar	LOIL 5 / OH (DT)	WS 3 (DT)	Ch.5 Physics of Curv.
6	8 Mar	LOIL 6 / OH (DT)	WS 4 (DT)	Ch 6. Rel. Cosmology
7	15 Mar	TBD (VS)	TBD (VS)	TBD
8	22 Mar	TBD (VS)	TBD (VS)	TBD
9	12 Apr	TBD (VS)	TBD (VS)	TBD
10	19 Apr	TBD (VS)	TBD (VS)	TBD
11	26 Apr	TBD (VS)	TBD (VS)	TBD

Learning Outcomes

After taking this unit the student should be able to:

- show proficiency in calculations involving tensors;
- understand the basics of geometric/coordinate-free vector and tensor analysis;
- understand the basic principles of differential geometry;
- perform covariant differentiation of tensor quantities in curved space;
- provide a physical explanation of the Einstein field equations;
- show how the Friedmann-Robertson-Walker metric is an exact solution to the Einstein equations;
- describe the key ideas behind cosmology and the expanding universe;
- describe the observations and techniques used to establish the accelerating expansion of the universe, including systematics and sources of uncertainty;
- describe the observation and analysis of the Cosmic Microwave Background Radiation, including systematics and sources of uncertainty;
- describe current state of the art cosmology experiments, how their results compare to each other, and current open questions that they address.

Recommended Textbooks

(Library List: `http://www.bath.ac.uk/library/gen/leganto/index.php?course=PH40112`)

1. ***Carroll, SM.***, **Spacetime and Geometry: an Introduction to General Relativity** Pearson; 2014. —- This book covers most of the basics, focusing on a direct index approach.
2. ***Misner, C.; Thorne, K; and Wheeler, J.***, **Gravitation** (Thorne KS, Wheeler JA, eds.); 1973 —- This enormous book contains everything you would need to know, in a coordinate-free approach, but it is more useful as a reference rather than a way to first learn the subject.
3. ***Wald, RM.***, **General Relativity**, University of Chicago Press; 1984 —- This classic uses abstract index notation, which is very useful for calculations, but makes it hard to develop geometric intuition.
4. ***D'Inverno, R.***, **Introducing Einstein's Relativity**, Oxford University Press; 1992 —- An excellent introductory text, using abstract index notation.
5. ***Jeevanjee, N.***, **An Introduction to Tensors and Group Theory for Physicists**, 2nd ed, Springer, 2015 —- This is a very well written introduction to tensors and group theory useful for aspiring theorists.
6. ***Moore, T. A.***, **A General Relativity Workbook**, University Science Books, 2013 —- The tutorial worksheets have been developed using this book as a template. The material covered isn't exactly the same as our class, but it is a good introduction, particularly for building your ability to calculate.

During the Semester:

- **Pre-reading:** This unit will cover extremely challenging material, approaching things in a more mathematically rigorous way than many of you will be used to. We will often provide pre-reading in the form of detailed notes, selected textbook chapters, and articles for you on the Moodle, to be read *before* watching certain lectures. It is ***extremely important*** that you at least skim these before the indicated lectures, as the discussion will assume you have at least heard of the background/concepts from the readings. While the lectures will provide an overview of this material, much more detail will be available in the readings, and you will be rewarded for revisiting them in depth.
- **Lectures:** There will be approximately 100 minutes of lectures a week which we will try to post before the end of the previous week, which will be recorded via the Panopto/ReView service. These lectures will serve to provide an active discussion of the material introduced by the assigned background reading. Lecture notes will be captured/scanned regularly, and posted to

the moodle, though any board work (particularly for tutorials and problem classes) may not be captured well.

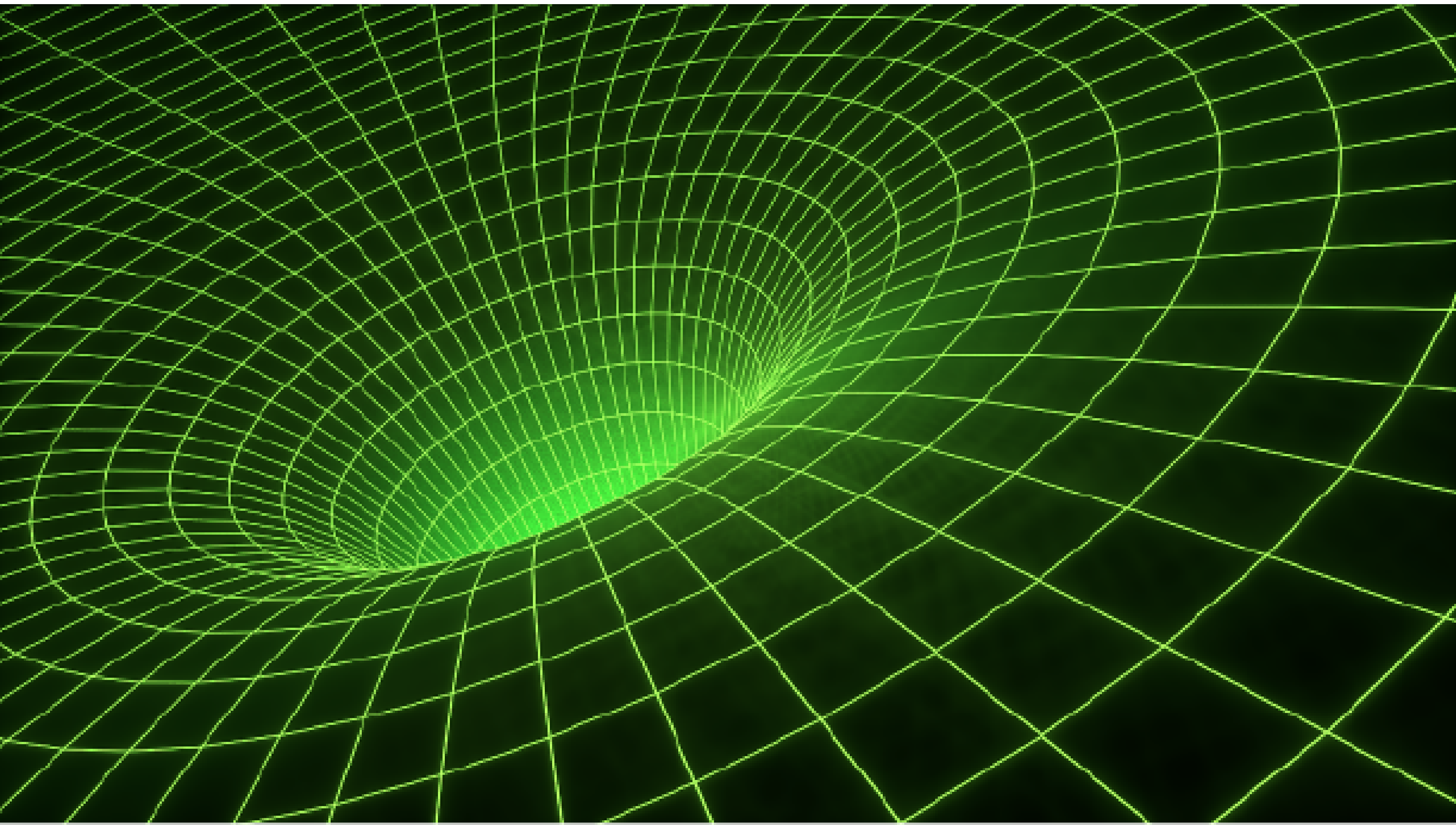
- **Worksheet Tutorials:** Each week on Thursdays, there will be a tutorial, or a problem class. We have developed interactive group worksheets that we will work on during these sessions. You should definitely read over the problems before hand and think about how you would tackle each one while you are watching the lectures and reading the notes, as there is not enough time during the worksheet sessions to solve them cold. Some worksheets will be completed over the course of 2 sessions.

University of Bath Principles of Academic Integrity: The University's principles of academic integrity are set out in http://www.bath.ac.uk/quality/documents/QA53.pdf. Any violation of these principles will be referred directly to the appropriate Director of Studies for investigation.

Congratulations! You made it to the end of the syllabus! To demonstrate that you have read this document, please email Dr. Tsang a nice picture of a dog (or animal of your choice) before WS 1.

PART I:

GENERAL RELATIVITY



Introduction to General Relativity

In this course we will develop, in a mathematically rigorous way, the theoretical foundations of General Relativity and Cosmology. We will also cover the observations and observational techniques that provide the empirical evidence for modern cosmology. It is assumed that the student is familiar with relativistic kinematics in curved spacetime as covered in PHYS 30101: Intro to General Relativity. This material will be reviewed where necessary, in order to cast it in the appropriate mathematical context. To study the source and dynamics of curved spacetime, we must first develop the geometric framework to be able to discuss curvature, and how it is generated.

We will be covering the physics of General Relativity (GR), i.e. the physics of gravity as we understand it. GR does not view gravity as a force, like other forces of nature, but rather as a consequence of spacetime geometry. In order to be able to mathematically describe gravity in GR, we need a way of describing the geometry of 4-dimensional spacetime.

This can seem difficult, however it isn't that scary if you develop your geometric formalism and intuition carefully. GR is viewed by the public consciousness as an incredibly complex topic, however it isn't, really, provided we develop the mathematical language necessary to express it. The first several lectures/weeks in this class will be spent climbing this mathematical summit, so that we can appreciate the grand vista of Einstein's theory.

We can begin with a little philosophy. Why is gravity so different from the other forces? Simply put, you don't "feel" the force of gravity. What you feel (i.e. what produces internal forces) is the chair pushing up on your butt. Why is this so?

In Newtonian gravity we know that

$$F = m_{\text{inertial}} \times a = -\frac{GMm_{\text{gravity}}}{r^2} \tag{1}$$

where the gravitational "charge" m_{gravity} is no more than 1 part in $\sim 10^{11}$ different than the inertial mass m_{inertial}. Since there is no differential acceleration felt across your body there is no net internal force.

Einstein's key insight was that if you are freely falling, you won't notice the force of gravity, i.e. you can't distinguish whether you are freely falling in a gravitational field or just floating fa from anything else in deep space.

The "morals" (as in Aesop, not morality) of GR can be summed up as

1. Every inertial observer measures the same laws of physics locally (Galilean Relativity)

2. The speed of light in a vacuum is a law of physics (Special Relativity)

3. Freely falling observers are inertial observers (General Relativity)

We usually think of inertial observers as those moving in "straight lines". Here we understand that inertial observers are those that measure physics "as it is", rather than induced by the way they are moving. This leads to a somewhat starting revelation for those of us in the classroom: In GR we are not considered inertial observers (not just because of the rotating reference frame of the Earth's surface) because we are not freely falling. On the Earth's surface the inertial observer is the elevator rider whose cable has just been cut. This is a fundamental change in how we view physics. In some sense inertial observers are still travelling in "straight" lines, and it is us that are on a curved path. Obviously, we don't see our classroom as a curved path in space, but it is indeed curved in the sense of space and time combined.

The mathematical ability to describe this curvature, to describe "straight" (i.e. geodesic) paths in curved spacetime, and to describe the laws of physics in the presence of curvature will be the focus of the first half of the unit. In order to express the laws of physics we want to be able to describe how things like vectors can be described and how they change while moving through curved spacetime.
In order to develop the language necessary to express Einstein's theory quantitatively, we have to develop the mathematical notion of a "tensor", and the easiest way to do this is to go back and make sure we know what a "vector" really is.

1
Mathematical Preliminaries

1.1 Linear Algebra of Vectors and Dual Vectors

In the first part of this course we will develop the basic geometrical formalism necessary to study the differential geometry of curved spacetimes. For more references you can look at Schutz, Carroll, or Jeevanjee. These notes will borrow from the approach of all of these, and I will try to note where the notation is different.

We will start with a careful consideration of what is meant by vectors and vector spaces, using these to define dual-vectors, also known as one-forms, co-vectors or "transpose" vectors. Tensors then follow as natural generalisations of these geometric objects, independent of a coordinate basis.

We will then define the concept of manifolds, tangent spaces, cotangent spaces, and metric spaces. We will demonstrate how to take derivatives on curved manifolds. We can then show how we can characterise the curvature of a manifold, with and without a metric. We will review geodesics on Reimmannian manifolds, and show how geodesic deviation also captures curvature.

1.1.1 Vectors and Vector Spaces

By this point in your physics education, you should be intimately familiar with the concept of vectors, and how they are used to describe physics. One of the goals of this course is to understand tensors, which act as multi-linear (scalar valued) functions on a ***vector space*** (and its dual). In carefully defining abstract vector spaces, we will provide a generalisation of the vectors you are familiar with, and show how these tie together several other mathematical objects familiar to us from physics.

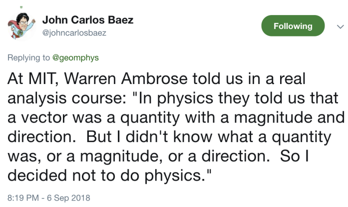


Figure 1.1: How mathematicians really feel about vectors.

The idea behind this kind of "abstract" definition is to try to distill the nature of vectors down to the basic essentials, such that it applies to the widest class of objects, but still maintains a "vector"-like nature. By providing a definition, we will clarify the nature of vectors and vector spaces, allowing us the mathematical precision and flexibility to understand the geometry they describe, no matter the context.

The notion of a vector as a quantity that has both magnitude and direction is not sufficiently general, as seen from quantum mechanics where complex-valued wave functions represent states of a particle. So how do we define a vector then? We don't. Instead we define a vector space! Once we have a vector space, whatever belongs to the vector space is called a vector.

Definition: A C-Vector Space

AN (ABSTRACT) VECTOR SPACE (V, C) consists of:

- A set V (whose elements are called vectors)
- A set of scalars C, e.g. $\mathbb{R}$ or $\mathbb{C}$ (technically C can be any "algebraic field", a set of numbers equipped with addition and multiplication operations)
- The definition of some addition $(+)$ and scalar multiplication $(\cdot)$ operators under which the vector space is ***closed*** (meaning that these operations performed on elements of V result in another element of V), which also satisfy the axioms:
  - **(C)** $v + w = w + v \quad \forall v, w \in V,$ (**C**ommutativity of $V+$)
  - **(A)** $v + (w + x) = (v + w) + x \quad \forall v, w, x \in V,$ (**A**ssociativity of $V+$)
  - **(N)** $\exists$ a vector $0 \in V$ such that $v + 0 = v \quad \forall v \in V,$ (**N**eutral element (identity) under $V+$)
  - **(I)** $\forall v \in V, \exists$ a vector $-v \in V$ such that $v + (-v) = 0,$ (**I**nverse under $V+$)

  - **(A)** $(c_1 \cdot c_2) \cdot v = c_1 \cdot (c_2 \cdot v) \quad \forall c_1, c_2 \in C, v \in V,$ (**A**ssociativity of $C\cdot$)
  - **(D)** $c \cdot (v + w) = c \cdot v + c \cdot w \quad \forall v, w \in V, c \in C,$ (**D**istributivity of $C\cdot$ over $V+$)
  - **(D)** $(c_1 + c_2) \cdot v = c_1 \cdot v + c_2 \cdot v \quad \forall c_1, c_2 \in C, v \in V,$ (**D**istributivity of $C\cdot$ over $C+$)
  - **(U)** $\exists 1 \in C$ such that $1 \cdot v = v \quad \forall v \in V,$ (**U**nit element in C)
- We often refer to V (equipped with $+$ and $\cdot$) as a C-vector space.

In the above definition of an abstract vector space, most of the axioms required may seem pedantic or redundant, but they mostly exist to make sure the notion of addition and scalar multiplication behave as expected (CANI), and are consistent across the vectors and the scalars (ADDU)[1].

You may have noticed we haven't included any notion of things like dot products. We can introduce the idea of an inner product later, but it isn't necessary for the minimal notion of a vector space. As we will see later, equipping a vector space with an inner product introduces additional non-trivial mathematical structure to the space.

An algebraic field is a abstract "group" with certain well behaved addition and multiplication (and their inverse) operations. To be a considered a vector space, the scalars C must technically be an algebraic field, though there are "modules" that are almost like vector spaces but with a more general "algebraic ring" replacing the scalars. Understanding the details of these algebraic structures is beyond the scope of this course, but I will try to point them out when we encounter them just for your cultural interest.

[1] In practice, given a reasonable definition of the addition and scalar multiplication operations, the only nontrivial vector space properties that one needs to check are ensuring the vector space is indeed closed under both $+$ and $\cdot$, and that the set of vectors V includes both the 0 vector **(N)**, and the negative (additive inverse) of each vector **(I)**.

Example 1.1: Real vector space $\mathbb{R}^n$

The vector spaces (V, C) that you are most used to considering are of course the n-dimensional cartesian space $V = \mathbb{R}^n$, with real scalars $C = \mathbb{R}$. Addition and scalar multiplication are defined as normal for $v, w \in \mathbb{R}$:

$$(v^1, \ldots, v^n) + (w^1, \ldots, w^n) \equiv (v^1 + w^1, \ldots, v^n + w^n) \quad (1.1)$$

$$c(v^1, \ldots, v^n) \equiv (cv^1, \ldots, cv^n). \quad (1.2)$$

You can check for yourself that all the axioms above are satisfied, but it should be fairly intuitive. Such spaces are, of course, the ones we are most used to thinking about, $\mathbb{R}^3$ is just regular 3-D cartesian space, $\mathbb{R}^4$ can represent a 4-D spacetime, $\mathbb{R}^6$ can represent position-velocity configuration space for a single particle in classical mechanics.

Why is $\mathbb{R}^n$ not a* complex *vector space?
I.e why is $(V, C) = (\mathbb{R}^n, \mathbb{C})$ with the same addition and scalar multiplication as above NOT a vector space?

$(\mathbb{R}^n, \mathbb{C})$ is not a vector space because for any $v \neq 0 \in \mathbb{R}^n$ and $c \in \mathbb{C}$ with $\text{Im}(c) \neq 0$, we have $cv \notin \mathbb{R}^n$. Thus $\mathbb{R}^n$ is ***not closed*** under complex scalar multiplication, and cannot be a complex vector space.

Exercise 1.1: $\mathbb{C}^n$

Show that $\mathbb{C}^n$ can be considered both a real and a complex vector space, i.e. show that $(V, C) = (\mathbb{C}^n, \mathbb{R})$ and $(\mathbb{C}^n, \mathbb{C})$ are both valid vector spaces under the standard definitions of component-wise vector addition and complex scalar multiplication.

Exercise 1.2: $M_n(\mathbb{R})$ the set of $n \times n$ matrices

Show that $M_n(\mathbb{R})$ the set of $n \times n$ real valued matrices can be considered a real vector space under component-wise addition and scalar multiplication.

Exercise 1.3: The 2-Sphere

Consider the set of vectors in $\mathbb{R}^3$ that describe points on the unit 2-sphere, i.e. $S^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$. Show that S^2 cannot form a (real) vector space under the standard definitions of vector addition and scalar multiplication.

1.1.2 *Linear Independence and Vector Bases*

YOU SHOULD ALREADY BE INTUITIVELY FAMILIAR with the notion of a "basis" for a vector space: it's a set of vectors from which we can construct all other vectors in that vector space. In order to discuss basis vectors we first need the notion of linear independence.

Let us consider a set of distinct vectors $S = \{v_1, v_2, \ldots, v_n\}$. These vectors are ***linearly dependant*** if there exists some set of scalars $c^1, c^2, \ldots, c^n \in C$, ***not all of which are*** 0, such that

Definition of linear dependence and independence

$$c^1 v_1 + c^2 v_2 + \ldots + c^n v_n = 0. \tag{1.3}$$

In other words there is at least one vector in S that can be written as a linear combination of the others. If this is ***not*** the case, then we say that the set S is ***linearly independent***, and none of its elements can be written as linear combinations of the other elements.

Definition: A Vector Basis

A BASIS FOR A VECTOR SPACE V is then defined to be an ***ordered*** linearly independent set $\mathcal{B} = \{e_1, \ldots, e_n\} \subseteq V$ which ***spans*** V, i.e. the set of all vectors of the form

$$c^1 e_1 + c^2 e_2 + \ldots + c^n e_n \qquad \text{with each } c^i \in C, \tag{1.4}$$

is the same as V. From here on, we will mostly choose to label bases with the symbol e_i rather than a generic v, since they are important in choosing how we represent quantities related to a vector space.

One can show that all finite basis sets of a given vector space must have the same number of elements, so the ***dimension*** of a vector space V, (i.e. $\dim V$), is the number of elements of any finite basis. Some vector spaces (like function spaces in Quantum Mechanics) which have no finite bases, are called ***infinite dimensional***, with dimensionality given by the cardinality of the infinite basis.

Dimension of a vector space: $\dim V$

Above we have surreptitiously introduced some notation by choosing to label the basis vectors e_i with label i down. This choice is arbitrary convention, but will inform all of the index gymnastics on which much of the calculations within general relativity are based.

Basis vectors, e_i, are labeled with ***down*** indices, by convention.

Definition: Vector Component Representation

CHOOSING THESE BASIS VECTORS we can introduce the extremely useful notion of ***vector components*** $\{v^i\}$ such that any vector v in a n-dimensional vector space (V, C) can be expressed in terms of the basis vectors $\{e_i\} \in V$,

$$v = v^1 e_1 + v^2 e_2 + \ldots + v^n e_n = \sum_{i=1}^{n} A^i e_i, \tag{1.5}$$

where each $v^i \in C$ such that we can write all elements of the vector space as n-tuples with respect to $\mathcal{B}$. We often write these components as a ***column vector***,

$$[v]_{\mathcal{B}} = \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix} \tag{1.6}$$

just as we are used to doing in regular n-dimensional Cartesian space. When we want to be explicit about what basis

we have adopted we will use the notation above, with the subscript $\mathcal{B}$, though when it is obvious we will often drop the subscript.
With a choice of vector basis, any n-dimensional vector space can be made to "look" like $\mathbb{R}^n$ or $\mathbb{C}^n$, however it is important to remember that ***vectors exist independently of any chosen basis!*** It should be obvious that the components that represent a given vector $\boldsymbol{v}$ will be different depending on the basis you choose to represent it in.

Notice above that we have carefully labeled the vector components with raised indices, so that they are paired with the lowered indices of the basis vectors. This is so that we can adopt the ***Einstein summation*** convention, which is that whenever an index is repeated in an expression, once as a superscript, and once as a subscript, then summation over that index is implied, e.g.

Einstein Summation Convention

$$\boldsymbol{v} = v^i \boldsymbol{e}_i \equiv \sum_i v^i \boldsymbol{e}_i \tag{1.7}$$

in order to save us the ink of writing all the summation signs! The repeated (up/down) pairs are called ***dummy indices***, as opposed to unpaired indices which are called ***free indices***, which should match across equalities. You are always free to rename the dummy index labels (i, j, k etc), making sure to keep them paired correctly. Careful usage of such ***index juggling*** is a useful skill that will become second nature with practice.

dummy indices

1.1.3 Linear Mappings and Linear Operators

MAPPINGS CAN BE THOUGHT OF AS FUNCTIONS that map elements between some domain set V to some range set W (typically written as $V \to W$). A mapping $f : V \to W$ takes each $V \ni v \mapsto w \in W$.

The mapping $f : V \to W$ is ***invertible*** if and only if the mapping f is ***one-to-one*** and ***onto*** (i.e. it is bijective). You should convince yourself that this makes sense. If a mapping is invertible then there exists another mapping $f^{-1} : W \to V$, called the ***inverse***, such that $ff^{-1} = \mathrm{id}_W$ and $f^{-1}f = \mathrm{id}_V$, where id_V and id_W are the identity operators on each set.

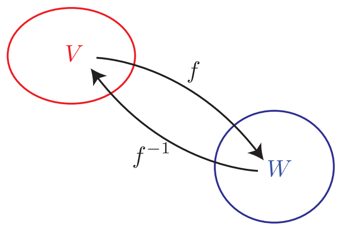


Figure 1.2: A mapping $f : V \to W$ takes elements of V to elements of W. If f is bijective (one-to-one and onto), then the inverse f^{-1} exists and maps $f^{-1} : W \to V$.

If (V, C) and (W, C) are both vector spaces which share the same set of scalars, then the mapping $f : V \to W$ is considered a ***linear mapping*** or ***linear transformation*** [2] if it also obeys the linearity condition,

$$f(c\boldsymbol{v}_1 + \boldsymbol{v}_2) = cf(\boldsymbol{v}_1) + f(\boldsymbol{v}_2), \qquad \forall \boldsymbol{v}_1, \boldsymbol{v}_2 \in V, \forall c \in C \tag{1.8}$$

where $F(\boldsymbol{v}_1)$ are $F(\boldsymbol{v}_2)$ vectors in W. Note that in the equation above the $+$ symbol on the left hand side is the addition operation associated with V, while the $+$ symbol on the right hand side is the addition operation associated with W.

[2] These maps are also called ***vector space homomorphisms***, i.e. they are mappings (morphisms - from *morphe*: shape in Greek) that preserve (*homo*: same) a vector space's linear structure under addition and scalar multiplication. If the linear map is also invertible (i.e. bijective) then we can call it a ***vector space isomorphism*** (from *iso*: equal). Any two vector spaces with an isomorphism between them are called ***isomorphic***, and can often be considered equivalent, (written as $V \cong W$).

For convenience we will denote ***linear mappings*** with the notation $f : V \xrightarrow{\text{lin}} W$, and ***invertible linear mappings*** with the notation $f : V \xrightarrow{\sim} W$ (since these are vector space isomorphisms).

IF A LINEAR MAPPING Λ acts between a V and itself, we call this mapping $\Lambda : V \xrightarrow{\text{lin}} V$ a ***linear operator***. Linear operators can be completely defined by how they act on a set of basis vectors $\mathcal{B} = \{\boldsymbol{e}_i\}$. Since $\Lambda(\boldsymbol{e}_i)$ is some other vector in V, we can describe it in terms of the basis vectors, i.e. (using the Einstein summation convention)

Linear operators are vector space homomorphisms that map V to itself, i.e. $V \to V$. These are sometimes called vector space ***endomorphisms*** (*endo*: within). Endomorphisms that are also isomorphisms (bijective) are called ***automorphisms*** (*auto*: self).

$$\Lambda(\boldsymbol{e}_i) = \Lambda_i{}^j \boldsymbol{e}_j \tag{1.9}$$

where $\Lambda_i{}^j$ is the jth component of the column vector $[\Lambda(\boldsymbol{e}_i)]_{\mathcal{B}}$. The coefficients $\Lambda_i{}^j$ can also be taken to be components of the matrix representation of the linear transformation in basis $\mathcal{B}$ which we can denote by

Matrix representation of a linear operator

$$[\Lambda]_{\mathcal{B}} = \begin{bmatrix} \Lambda_1{}^1 & \Lambda_1{}^2 & \dots & \Lambda_1{}^n \\ \Lambda_2{}^1 & \Lambda_2{}^2 & \dots & \Lambda_2{}^n \\ \vdots & \vdots & \ddots & \vdots \\ \Lambda_n{}^1 & \Lambda_n{}^2 & \dots & \Lambda_n{}^n \end{bmatrix}. \tag{1.10}$$

The action of Λ on an arbitrary vector $\boldsymbol{v}$ can then be reconstructed through linearity, giving

$$\Lambda(\boldsymbol{v}) = \Lambda(v^i \boldsymbol{e}_i) = v^i \Lambda(\boldsymbol{e}_i) = v^i \Lambda_i{}^j \boldsymbol{e}_j \tag{1.11}$$

There are several notations for applying mapping Λ to vector $\boldsymbol{v}$ which we will use interchangeably: the "functional" form $\Lambda(\boldsymbol{v})$, the operator form $\Lambda \boldsymbol{v}$, and the "bracket" form $\langle \Lambda, \boldsymbol{v} \rangle$. These all mean the same thing, but one may be slightly more convenient than the others depending on the context. You should at least be aware of these so you don't get confused when reading other texts.

An important thing to keep in mind is that a linear operator is ***not*** the same thing as its matrix representation, just as we know a vector is not the same thing as its column vector representation. Choosing a different set of basis vectors gives a completely different set of matrix components.

THE SET OF ALL LINEAR OPERATORS on an n-dimensional vector space (V, C) is often written as $\mathcal{L}(V) = \{\Lambda \mid \Lambda : V \xrightarrow{\text{lin}} V\}$. (Note that $\mathcal{L}(V)$ can itself be considered an $n \times n$ dimensional C-vector space! See if you can show this, perhaps by assuming an arbitrary basis of V). There is something special about this set of mappings; since we do not need to introduce any new mathematical "structure" to define the domain or range of the mappings, we can consider this set of all linear mappings $\mathcal{L}(V)$ as given to us "for free" as soon as we are given the vector space V. As we will see below, this is not the only geometric notion that we get automatically when considering a vector space in isolation.

The set of all linear operators $\mathcal{L}(V)$ is sometimes also written $\mathrm{End}(V)$ for the set of endomorphisms on V. Similarly the set of all linear transformations from vector space V to W is often called $\mathrm{Hom}(V, W)$, for homomorphism.

Every vector space comes automatically "equipped" with a set of linear operators, $\mathcal{L}(V)$.

IT CAN BE SHOWN that a ***linear operator***, Λ, is ***invertible***, $\Lambda : V \xrightarrow{\sim} V$, if and only if the only vector it maps to the zero vector is the zero vector itself. This is a useful fact, if somewhat unobvious. See if you can show this is the case (i.e. use linearity to show that such an operator must be one-to-one and onto and vice versa).

invertibility of linear operators

1.1.4 A Change of Basis

One important example of a linear operator is simply switching which basis vectors we are using to describe a vector space. Consider an n-dimensional vector space (V, C) from which we have chosen two different arbitrary bases $\mathcal{B} = \{e_i\}$ and $\tilde{\mathcal{B}} = \{\tilde{e}_i\}$.

Let Λ be the linear operator that maps the basis vector $\boldsymbol{e}_i$ to basis vector $\tilde{\boldsymbol{e}}_i$, such that

Change of Basis Operator

$$\Lambda(\boldsymbol{e}_i) = \tilde{\boldsymbol{e}}_i = \Lambda_i{}^j \boldsymbol{e}_j. \tag{1.12}$$

i.e. the coefficient $\Lambda_i{}^j$ is the jth component of the column vector $[\tilde{\boldsymbol{e}}_i]_{\mathcal{B}}$.

As long as none of the basis vectors in $\tilde{\mathcal{B}}$ are the zero vector (you should see that this is impossible due to linear independence of the basis vectors in $\tilde{\mathcal{B}}$), then this transformation should be invertible, and we can write the inverse transformation Λ^{-1} which converts from $\tilde{\mathcal{B}}$ back to $\mathcal{B}$

Inverse Change of Basis

$$\Lambda^{-1}(\tilde{\boldsymbol{e}}_i) = \boldsymbol{e}_i = (\Lambda^{-1})_i{}^j \tilde{\boldsymbol{e}}_j \tag{1.13}$$

where the coefficient $(\Lambda^{-1})_i{}^j$ is the jth component of the column vector $[\boldsymbol{e}_i]_{\tilde{\mathcal{B}}}$.

Passive Transformation of Basis Vectors and Components

We can use either vector basis in order to express an arbitrary vector $\boldsymbol{v}$. If we are merely changing how we choose to ***represent*** a vector, we are not changing the fundamental vector itself, and it should remain unchanged such that

$$\boldsymbol{v} = \tilde{v}^i \tilde{\boldsymbol{e}}_i = v^i \boldsymbol{e}_i. \tag{1.14}$$

We have above that $\boldsymbol{e}_i$ can be transformed by Λ to get $\tilde{\boldsymbol{e}}_i$. ***In order for vector to remain fundamentally unchanged*** we require that the components v^i must transform to $\tilde{v}^i$ in a way that ***exactly compensates*** for the change in basis vectors. Using the inverse change of basis above to write $\boldsymbol{e}_i$ in terms of $\tilde{\boldsymbol{e}}_i$, we see

$$\begin{aligned} \tilde{v}^i \tilde{\boldsymbol{e}}_i &= v^i \boldsymbol{e}_i = v^i \Lambda^{-1}(\tilde{\boldsymbol{e}}_i) \\ &= v^j (\Lambda^{-1})_j{}^i \tilde{\boldsymbol{e}}_i \\ \Rightarrow \tilde{v}^i &= (\Lambda^{-1})_j{}^i v^j \end{aligned} \tag{1.15}$$

such that

$$\begin{aligned} \boldsymbol{v} = \tilde{v}^i \tilde{\boldsymbol{e}}_i &= v^j (\Lambda^{-1})_j{}^k \Lambda_k{}^i \boldsymbol{e}_i \\ &= v^j \delta_j i \boldsymbol{e}_i \\ &= v^i \boldsymbol{e}_i \end{aligned} \tag{1.16}$$

Thus we see that for passive transformations of basis representation $\mathcal{B} \xrightarrow{\sim} \tilde{\mathcal{B}}$ the ***basis vector transforms*** with Λ, while the ***vector component transforms*** with the inverse Λ^{-1}.

Such transformations are called ***passive transformations***, since they leave the fundamental geometric object unchanged. ***Active transformations***, like physical rotations, actually change the physical or geometric properties of the object, whereas passive transformations, like rotations of the coordinate system, only change the basis with which we describe the objects in the vector space.

Things that transform like basis vectors (with Λ) are said to ***transform covariantly***, while things that transform like the vector components (with Λ^{-1}) are said to ***transform contravariantly***.

Contravariant and covariant transformations

In many textbooks, vectors are often only represented by these ***contravariant*** components (with index up), without the implicit basis vectors. ***It is important to remember that while these vector components are indeed transformed contravariantly, through Λ^{-1}, the geometric object, the vector itself, remains unchanged due to a passive transformation of basis!***

Part of the power of the ***index notation*** we use in Relativity comes from the automated bookkeeping that occurs when we always write things that transform covariantly with free index ***down***, while quantities that transform covariantly are always written with free ***index up***. If we encounter an object that has no free indices (such as for a vector $\boldsymbol{v} = v^i \boldsymbol{e}_i$) we know that the object is ***invariant*** under a passive basis transformation.

1.1.5 *Examples of Linear Operators: Rotations, Lorentz Transformations*

- rotations of 3-vectors
- Lorentz transformations of vectors
- Adjoint representation of a linear operator acting on the vector space of linear operators.

1.2 *Dual Vectors and Dual Spaces*

Given a C-vector space V, we saw above how we automatically have access to the set of all linear operators $\mathcal{L}(V)$, without requiring any additional mathematical structure since these are all possible linear mappings from $V \xrightarrow{\text{lin}} V$.

It may have occurred to you already that we also already have access to another complete set of linear mappings, namely the ones that take $V \xrightarrow{\text{lin}} C$, that is the set of ***all scalar valued linear mappings***. This set of ***linear*** mappings $V^* = \{\boldsymbol{\alpha} \mid \boldsymbol{\alpha} : V \xrightarrow{\text{lin}} C\}$, is called the ***dual space*** of the vector space V.

V^* is sometimes written as $\mathrm{Hom}(V, C)$.

The elements of V^* are called ***dual vectors***, (sometimes ***covectors*** or ***bra*** vectors in quantum mechanics), and they are ***objects that eat vectors and spit out scalars***. Notationally we will usually write this with either bracket or function notation

$$\langle \boldsymbol{\alpha}, \boldsymbol{v} \rangle \equiv \boldsymbol{\alpha}(\boldsymbol{v}) \in C, \qquad \forall \boldsymbol{\alpha} \in V^*, \forall \boldsymbol{v} \in V \tag{1.17}$$

For now, we will try to write dual vectors using bolded greek letters $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \ldots)$, while vectors will be written with bolded roman letters $(\boldsymbol{v}, \boldsymbol{w}, \ldots)$.

THE SET OF DUAL VECTORS V^* IS ITSELF A VECTOR SPACE! To see this, we just need to show that the "obvious" addition and scalar multiplication operations indeed obey the vector space axioms. What do we mean by the "obvious" operations? Just those that they inherit from the scalars C. Since a dual vector acting on a vector produces a scalar, we can define the dual-vector addition and scalar multiplication for an arbitrary vector $\boldsymbol{v}$, i.e. $\forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in V^*, \forall \boldsymbol{v} \in V$, and $\forall c \in C$,

$$[\boldsymbol{\alpha} + \boldsymbol{\beta}](\boldsymbol{v}) = \boldsymbol{\alpha}(\boldsymbol{v}) + \boldsymbol{\beta}(\boldsymbol{v}) \tag{1.18}$$

$$[c \cdot \boldsymbol{\alpha}](\boldsymbol{v}) = c \cdot \boldsymbol{\alpha}(\boldsymbol{v}). \tag{1.19}$$

where the right hand side operations are just the ones between scalars. Thus the dual vectors satisfy the vector space axioms (CANI ADDU) by inheriting them from the scalar operations, where the neutral element (N) under addition is just the dual vector that maps all vectors to $0 \in C$, i.e. $V^* \ni \mathbf{0} : \boldsymbol{v} \mapsto 0$; and the additive inverse (I) of a dual vector $\boldsymbol{\alpha}$ is just $-\boldsymbol{\alpha} : \boldsymbol{v} \mapsto -[\boldsymbol{\alpha}(\boldsymbol{v})]$.

It is tempting, based on our experience in flat Cartesian space, to associate a particular dual vector with a given vector, however without a metric (i.e. a non-degenerate inner product) there generally isn't such an association, as any dual vector can be applied to any vector. Since we haven't yet introduced a metric structure onto the vector space, this correspondence isn't yet established.

For finite dimensional vector spaces, however, we can use a particular choice of vector basis $\mathcal{B} = \{\boldsymbol{e}_i\}$ in order to induce a corresponding basis on the dual vector space, by arbitrarily choosing to use the Kroenecker delta symbol[3],

$$\delta_i^j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}, \tag{1.20}$$

[3] Here, in general, "0" and "1" are the additive and multiplicative identity elements, respectively, for the C scalars.

such that we define the ***dual basis*** $\mathcal{B}^* = \{\boldsymbol{e}^j\}$ that spans V^* defined by

$$\boldsymbol{e}^j(\boldsymbol{e}_i) = \delta_i^j. \tag{1.21}$$

The ***dual basis*** $\mathcal{B}^*$ is sometimes called the ***reciprocal basis*** of $\mathcal{B}$.

Remember that this "association" is purely arbitrary, there is nothing "canonical" about it beyond pure convention.

Using the functional notation, and this dual basis, we see that the key property of any dual vector $\boldsymbol{\alpha} \in V^*$ is that it is entirely determined by how it acts on the basis vectors we've chosen for V, since by the linearity property we have, $\forall \boldsymbol{v} \in V$,

$$\boldsymbol{\alpha}(\boldsymbol{v}) = \boldsymbol{\alpha}(v^i \boldsymbol{e}_i) \tag{1.22}$$

$$= v^i \boldsymbol{\alpha}(\boldsymbol{e}_i) \tag{1.23}$$

where we've used the Einstein summation convention. We can define the "components" of a dual vector to be

$$\alpha_i = \boldsymbol{\alpha}(\boldsymbol{e}_i) \tag{1.24}$$

such that by the definition of the dual basis we have

$$\boldsymbol{\alpha} = \alpha_i \boldsymbol{e}^i \tag{1.25}$$

which is why this is a convenient basis to assume for the dual space.

The bracket notation for applying dual-vectors to vectors is useful here, because it establishes an equivalence that exists between (finite) vector spaces and dual spaces. Writing,

$$\langle \boldsymbol{\alpha}, \boldsymbol{v} \rangle \in C \tag{1.26}$$

we can clearly associate the scalar valued linear map $\boldsymbol{\alpha}$ with the scalar valued linear map $\langle \boldsymbol{\alpha}, \cdot \rangle$, which maps $V \xrightarrow{\text{lin}} C$,

$$\boldsymbol{\alpha}(\boldsymbol{v}) = \langle \boldsymbol{\alpha}, \boldsymbol{v} \rangle, \tag{1.27}$$

$\forall \boldsymbol{v} \in V$, since these are just alternate ways of writing the same map.

Similarly[4], we can establish an equivalence (isomorphism) between the vectors and $\text{Hom}(V^*, C)$, i.e. the set of maps $V^* \xrightarrow{\text{lin}} C$, by defining the action of a vector $\boldsymbol{v} \in V$, on a dual-vector $\boldsymbol{\alpha} \in V^*$ as

$$\boldsymbol{v}(\alpha) \equiv \langle \boldsymbol{\alpha}, \boldsymbol{v} \rangle \tag{1.28}$$

which we can write as $\boldsymbol{v} \cong \langle \cdot, \boldsymbol{v} \rangle$.

It is not that difficult to show that such an equivalence is one-to-one and onto for a finite dimensional vector space, and this means that we can consider the original vector space to be dual to its dual space, i.e.,

$$V^{**} = V. \tag{1.29}$$

[4] Caution: this is only true for a finite dimensional ($\dim V < \infty$) vector space.

1.3 *Metrics and Metric Duals*

Notice that the above bracket notation looks a lot like a "dot product" or an "inner product", between two vectors. While this may look similar, we have ***not*** yet defined a notion of such a product between two vectors. If we introduce such a notion, we add additional structure onto a vector space, turning it into an inner product space, or a metric space.

We can define a non-degenerate bilinear form[5] $(\cdot|\cdot)$ as function that takes in an ordered pair of vectors from V and produces a scalar from C. This map must be ***symmetric and linear*** in both arguments and satisfy the non-degeneracy condition:

$$\forall \boldsymbol{v} \neq 0 \in V, \exists \boldsymbol{w} \in V, \text{ such that } (\boldsymbol{v}|\boldsymbol{w}) \neq 0, \tag{1.30}$$

i.e., $(\cdot|\cdot)$ can tell the difference between $\boldsymbol{v}$ and $\boldsymbol{v} + \boldsymbol{v}'$ for any $\boldsymbol{v}' \neq 0$.

Additionally, if this bilinear map is positive definite, i.e.

$$(\boldsymbol{v}|\boldsymbol{v}) > 0, \qquad \forall \boldsymbol{v} \neq 0 \in V \tag{1.31}$$

then we call this map a "(definite) inner product". A vector space equipped with an inner product has a notion of angles, distances, and a norm, and is called an "inner product space". Note that in Relativity the metric does not satisfy the positive-definite condition, it is known as an "indefinite inner product", but it possesses many of the same properties which we will exploit.[6]

[5] Note that for complex vector spaces we would instead examine non-degenerate Hermitian maps, which are linear in the second argument and become the complex conjugate when interchanging the arguments.) Such products are used on the infinite dimensional Hilbert spaces of Quantum Mechanics.

[6] Formally, a metric space is more general than an inner product space as "a metric space" is any set with a well-defined notion of distance (a metric) between any pair of elements. An (indefinite) inner product space induces a norm on a vector space, which in turn induces a metric on the base set of vectors.

The Metric Dual

Given a metric $(\cdot|\cdot)$, the ***metric dual*** of a vector $\boldsymbol{v} \in V$ is the one form $\boldsymbol{\alpha} \in V^*$ such that

$$\boldsymbol{\alpha} = (\boldsymbol{v}|\cdot) \tag{1.32}$$

so that $\forall \boldsymbol{w} \in V$,

$$\boldsymbol{\alpha}(\boldsymbol{w}) = (\boldsymbol{v}|\boldsymbol{w}) \tag{1.33}$$

Note that only after a metric is specified can we identify a ***particular covector to a particular vector***, through the metric dual.

2

Tensors

We are now ready to move on to discussing one of the main topics of this unit, tensors. Tensors have a (relatively unearned) reputation of being very complicated, but based on our understanding of vector spaces and linear maps, they are actually quite simple.

Given a vector space (V, C), we define the set of (p, q) tensors on V,

$$\mathcal{T}^p_q V \equiv \underbrace{V \otimes \ldots \otimes V}_{p \text{ times}} \otimes \underbrace{V^* \otimes \ldots \otimes V^*}_{q \text{ times}} \tag{2.1}$$

For now we will just take the right hand side to be another name or label for $\mathcal{T}^p_q V$. We will define the tensor product $\otimes$ below.

to be the set of all ***multilinear*** maps that eat p co-vectors, and q vectors, and spit out a scalar.

$$\mathcal{T}^p_q V \equiv \{T | T : \underbrace{V^* \times \ldots \times V^*}_{p \text{ times}} \times \underbrace{V \times \ldots \times V}_{q \text{ times}} \xrightarrow{\text{lin}} C\} \tag{2.2}$$

where the $\times$ represents the Cartesian product,

The Cartesian Product

The Cartesian product $A \times B$, of two sets A and B, is defined as the set of all ordered pairs (a, b), where $a \in A$, and $b \in B$,

$$A \times B \equiv \{(a, b) | a \in A \text{ and } b \in B\}$$

This can be extended to multiple Cartesian products, yielding the set of all ordered n-tuples,

$$A_1 \times \ldots \times A_n = \{(a_1, \ldots, a_n) | a_1 \in A_1, \ldots, a_n \in A_n\}.$$

If this is applied to multiple copies of the same set, we often use the notation,

$$A^n \equiv \underbrace{A \times \ldots \times A}_{n \text{ times}},$$

e.g. $\mathbb{R}^4$ or $\mathbb{C}^n$.

Here ***multilinearity*** means that a tensor obeys the linearity condition ***in each argument***, such that $\forall T \in \mathcal{T}^p_q V$, $\forall \boldsymbol{\alpha}_n, \boldsymbol{\beta} \in V^*, \forall \boldsymbol{v}_n, \boldsymbol{w} \in$

V, and $\forall c \in C$,

$$T(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n + c\boldsymbol{\beta}, \ldots, \boldsymbol{\alpha}_p, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_q) = T(\boldsymbol{\alpha}_1, \ldots \boldsymbol{\alpha}_p, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_q) + cT(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\beta}, \ldots, \boldsymbol{\alpha}_p, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_q) \quad (2.3)$$

and

$$T(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_p, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n + c\boldsymbol{w}, \ldots, \boldsymbol{v}_q) = T(\boldsymbol{\alpha}_1, \ldots \boldsymbol{\alpha}_p, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_q) + cT(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_p, \boldsymbol{v}_1, \ldots, \boldsymbol{w}, \ldots, \boldsymbol{v}_q) \quad (2.4)$$

and so on for each of the p co-vector arguments, and q vector arguments.

2.0.1 *Vectors and Dual Vectors as Tensors*

We can view the dual vectors V^* as just the set of all $(0, 1)$ tensors, i.e.,

$$\mathcal{T}^0_1 V \cong V^* \quad (2.5)$$

Here the $\cong$ symbol means "is isomorphic to", i.e. there exists a vector space isomorphism between the two sets, so that they can be viewed as equivalent, in terms of linear algebra.

since 1-forms/covectors eat vectors and return scalars. Similarly we can view the $(1, 0)$ tensors as isomorphic to the vectors V,

$$\mathcal{T}^1_0 V \cong V \quad (2.6)$$

since we can define

You have to be a bit more careful here for infinite dimensional vector spaces

$$T_{\boldsymbol{v}}(\boldsymbol{\alpha}) \mapsto \boldsymbol{\alpha}(\boldsymbol{v}), \forall \boldsymbol{v} \in V, \text{ and } \boldsymbol{\alpha} \in V^*. \quad (2.7)$$

Indeed, we can also view the linear operators as $(1, 1)$ tensors, $\mathcal{T}^1_1 V \cong \mathcal{L}(V)$, by associating a tensor $T_\Lambda \in \mathcal{T}^1_1 V$ with each linear operator $\Lambda \in \mathcal{L}(V)$,

$$T_\Lambda(\boldsymbol{v}, \boldsymbol{\alpha}) \mapsto \boldsymbol{\alpha}(\Lambda(\boldsymbol{v})), \forall \boldsymbol{\alpha} \in V^*, \forall \boldsymbol{v} \in V. \quad (2.8)$$

Note that the ***scalars*** can be considered $(0, 0)$ tensors, since they take 0 covectors, 0, vectors and return a scalar!

The Transpose of a Linear Operator

For a given linear transformation $\Lambda \in \mathcal{L}(V)$, we can consider an associated member of the endomorphisms of V^*, i.e.

$$\mathcal{L}(V^*) = \{\Phi | \Phi : V^* \xrightarrow{\text{lin}} V^*\}. \quad (2.9)$$

We can define $\Lambda^T \in \mathcal{L}(V^*)$ such that

$$\underbrace{(\Lambda^T(\boldsymbol{\alpha}))}_{\in V^*}(\boldsymbol{v}) \mapsto \boldsymbol{\alpha}(\Lambda(\boldsymbol{v})), \quad (2.10)$$

which we can call the "transpose" of Λ.

The inner product of a vector space is a particular example of a (0,2) tensor, since it eats two vectors and returns a scalar. $(\cdot|\cdot) \in \mathcal{T}^0_2 V$

Note that while a linear operator, being a (1,1) tensor can be written as a matrix in any particular basis, since it takes in a vector and returns a vector, for $\Lambda \in \mathrm{End}(V)$ and $\boldsymbol{v} \in V$,

$$\Lambda(\boldsymbol{v}) \in V. \tag{2.11}$$

The metric $(\cdot|\cdot)$ is a particular $(0,2)$ tensor, and hence, if it takes in a vector, returns an object that wants to eat another vector, i.e. it returns a ***dual vector***.

$$(v|\cdot) \in V^*. \tag{2.12}$$

This is why it is a little bit of a cheat to write a spacetime metric in matrix form. Really it doesn't act like a matrix (linear operator) at all.

The Tensor Product $\otimes$

It is convenient to define the tensor product between two tensors that may come from ***different*** tensor spaces. Consider $T \in \mathcal{T}^p_q V$, and $S \in \mathcal{T}^r_s V$. The tensor product of T and S,

$$(T \otimes S) \in \mathcal{T}^{p+r}_{q+s} V$$

is defined, point-wise, as

$$\begin{aligned}&(T \otimes S)(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_{p+r}, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{q+s}) = \\ &\qquad T(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_p, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_q) \cdot S(\boldsymbol{\alpha}_{p+1}, \ldots, \boldsymbol{\alpha}_{p+r}, \boldsymbol{v}_{q+1}, \ldots, \boldsymbol{v}_{q+s})\end{aligned}$$

Note that we don't actually need the tensors to have the same base spaces for the tensor product to be applied!

For example, we can construct a (1, 1) tensor $T \in \mathcal{T}^1_1 V$ by taking $\boldsymbol{v} \in V \cong \mathcal{T}^1_0 V$ and $\boldsymbol{\alpha} \in V^* \cong \mathcal{T}^0_1 V$, such that

$$\begin{aligned}T &= \boldsymbol{v} \otimes \boldsymbol{\alpha} \\ T(\boldsymbol{\beta}, \boldsymbol{w}) &= \boldsymbol{v}(\boldsymbol{\beta}) \cdot \boldsymbol{\alpha}(\boldsymbol{w}) = \boldsymbol{\beta}(\boldsymbol{v}) \cdot \boldsymbol{w}(\boldsymbol{\alpha}),\end{aligned}$$

$\forall \boldsymbol{\beta} \in V^*$ and $\forall \boldsymbol{w} \in V$.
It is notationally convenient to "overload" the tensor product "function" to work on the spaces themselves, to just mean the spaces of all such tensors,

$$\mathcal{T}^p_q V \equiv \underbrace{V \otimes \ldots \otimes V}_{p \text{ times}} \otimes \underbrace{V^* \otimes \ldots \otimes V^*}_{q \text{ times}}$$

which we used as another name of the set of all (p,q) tensors at the beginning of this section.

2.0.2 *Tensor Components*

Let us assume we have a basis $\mathcal{B} = \{e_i\}$ for V, and a dual basis $\mathcal{B}^* = \{e^i\}$ for V^*. We can define the "components" of a tensor $T \in \mathcal{T}^p_q V$ to be

$$T^{a_1 \ldots a_p}{}_{b_1 \ldots b_q} \equiv T(\boldsymbol{e}^{a_1}, \ldots, \boldsymbol{e}^{a_p}, \boldsymbol{e}_{b_1}, \ldots, \boldsymbol{e}_{b_q}) \in C \tag{2.13}$$

Note that we don't really need to use the dual basis here, but it makes things easier to write down.

To reconstruct a tensor T from its components we can use the tensor product,

$$T = T^{a_1 \ldots a_p}{}_{b_1 \ldots b_q} \cdot \boldsymbol{e}_{a_1} \otimes \ldots \otimes \boldsymbol{e}_{a_p} \otimes \boldsymbol{e}^{b_1} \otimes \ldots \otimes \boldsymbol{e}^{b_q}. \tag{2.14}$$

On tutorial worksheet 2, you will show explicitly, for arbitrary basis sets of vectors and dual vectors, that this does indeed reconstruct the tensor T.

If a vector space, V, has a basis $\mathcal{B} = \{\boldsymbol{e}_i\}$, then $\mathcal{T}^p_q V$ naturally inherits a basis made up of the tensor product of basis vectors. Using this it is easy to show that if the base vector space has (vector space) dimension $\dim V$, then the tensor space $\mathcal{T}^p_q V$ has (vector space) dimension $(\dim V)^{p+q}$, where we often call $p+q$ the ***rank*** of a tensor.

Rank and dimension of a tensor.

Change of Basis for Tensors

Previously, we examined how the vector components changed due to a passive change of basis. Let us examine a similar change of basis, given by a linear transformation/(1,1) tensor, Λ, and its inverse, to transform between two sets of bases $\mathcal{B} = \{\boldsymbol{e}_i\}$ and $\tilde{\mathcal{B}} = \{\tilde{\boldsymbol{e}}_i\}$

$$\begin{aligned} \tilde{\boldsymbol{e}}_a &= \Lambda(\boldsymbol{e}_a) \\ &= \Lambda^b{}_c \boldsymbol{e}_b \underbrace{\otimes \boldsymbol{e}^c(\boldsymbol{e}_a)}_{\cdot \delta^c_a} \\ &= \Lambda^b{}_a \boldsymbol{e}_b \end{aligned}$$

Similarly, we see that the components of the inverse transformation are merely the elements of the matrix inverse of Λ.

$$\boldsymbol{e}_b = (\Lambda^{-1})_b{}^a \tilde{\boldsymbol{e}}_a.$$

How do the components of covectors and tensors change under such a basis transformation?
We have the components of $\boldsymbol{\alpha} \in V^* \cong \mathcal{T}^0_1 V$

$$\begin{aligned} \alpha_a \equiv \boldsymbol{\alpha}(\boldsymbol{e}_a) &= \boldsymbol{\alpha}((\Lambda^{-1})_a{}^b \tilde{\boldsymbol{e}}_b) \\ &= (\Lambda^{-1})_a{}^b \boldsymbol{\alpha}(\tilde{\boldsymbol{e}}_b) \\ &= (\Lambda^{-1})_a{}^b \tilde{\alpha}_b \end{aligned}$$

where the second equality makes use of the linearity of $\boldsymbol{\alpha}$.
We see that the components of a co-vector must transform in the ***same way as the basis vectors***, i.e. they transform ***covariantly*** (and the dual basis covectors must transform contravariantly).
We can revisit the vectors by viewing them also $(1,0)$ tensors that eat covectors and give us numbers. For $\boldsymbol{v} \in \mathcal{T}^1_0 V \cong V$

we have,

$$\begin{aligned} v^a \equiv \boldsymbol{v}(\boldsymbol{e}^a) &= \boldsymbol{e}^a(\boldsymbol{v}) \\ &= \boldsymbol{e}^a(\tilde{v}^b \tilde{\boldsymbol{e}}_b) \\ &= \tilde{v}^b \boldsymbol{e}^a(\Lambda^c{}_b \boldsymbol{e}_c) \\ &= \tilde{v}^b \Lambda^c{}_b \underbrace{\boldsymbol{e}^a(\boldsymbol{e}_c)}_{=\delta^a_c} \\ &= \Lambda^a{}_b \tilde{v}^b \end{aligned}$$

In changing from basis $\mathcal{B}$ to $\tilde{\mathcal{B}}$ an arbitrary (p,q) tensor T has components that are related by,

$$\begin{aligned} T^{i_1 \dots i_p}{}_{j_1 \dots j_q} &= T(\underbrace{\boldsymbol{e}^{i_1}, \dots, \boldsymbol{e}^{i_p}}_{\dots(\Lambda^{-1})_k{}^i \tilde{\boldsymbol{e}}^k \dots}, \underbrace{\boldsymbol{e}_{j_1}, \dots, \boldsymbol{e}_{j_q}}_{\dots \Lambda^\ell{}_j \tilde{\boldsymbol{e}}_\ell \dots}) \\ &= (\Lambda^{-1})_{k_1}{}^{i_1} \dots (\Lambda^{-1})_{k_p}{}^{i_p} \Lambda^{\ell_1}{}_{j_1} \dots \Lambda^{\ell_q}{}_{j_q} T(\tilde{\boldsymbol{e}}^k, \dots \tilde{\boldsymbol{e}}_\ell, \dots) \\ &= (\Lambda^{-1})_{k_1}{}^{i_1} \dots (\Lambda^{-1})_{k_p}{}^{i_p} \Lambda^{\ell_1}{}_{j_1} \dots \Lambda^{\ell_q}{}_{j_q} \tilde{T}^{k_1 \dots k_p}{}_{\ell_1 \dots \ell_q} \end{aligned}$$

Note that many (most!) textbooks on relativity start by treating the ***components*** of tensors (and vectors, and one-forms) as the objects themselves, rather than just a set of numbers that represent the abstract mathematical object in a particular basis.

As such, you'll often see the components v^i referred to as the ***vector*** v^i. We, however, are bolder and wiser than others, and understand that this is only correct if we implicitly include the basis vectors (and dual basis covectors, and tensors products of the bases, etc) along with the components. By remembering this we understand that while these mathematical objects themselves are ***invariant*** under a choice of basis, their ***components*** depend on the basis in which they are represented.

Many of these same textbooks will also define tensors to be things that (have components that) ***transform like a tensor*** (see box above), rather than the multilinear map to the scalars that we have used. These same texts will define vectors as things that transform as vectors (contravariantly), and dual vectors to be things that transform as dual vectors (covariantly). As long as you keep in mind those texts are really only talking about the components, it turns out that these relatively unintuitive and ungeometric definitions of vectors, dual vectors and tensors are equivalent to the ones we have used (i.e. an object satisfies one definition, if and only if it satisfies the other), however I find these component-transformation definitions particularly uninformative.

3 Manifolds

In the previous chapter we established the basic concepts of vector spaces, and saw how, through the concept of linear mappings, each vector space comes automatically accompanied by an infinite menagerie of dual vectors, linear operators, and tensors. We saw how equipping a vector space with an inner product (the metric tensor), provides a way to measure lengths of vectors, and allows us to associate a particular dual vector to each vector.

This understanding of vector spaces certainly goes a long way to helping us describe relativity, in particular, and physics in general, however, in this section we will see that vector (and tensor) fields in general spacetimes don't live in just a single vector space, they actually exist spread across an infinity of vector spaces which are attached to each point of a mathematical "surface" called a manifold.

To study manifolds we will begin by developing a few key concepts from Topology[1]. We will learn about the concepts of charts and atlases, and define what we mean by a "smooth" multi-dimensional manifold.

[1] Specifically Point-Set Topology

We will then learn how to think about the differentiation of scalar, vector, and tensor fields, and how we can use the "locally flat" nature of a manifold to bootstrap our knowledge of calculus in Euclidean $\mathbb{R}^n$ to work in curved spaces.

3.1 Point-Set Topology

In this section we will explore some basic concepts in topology, in order to be able to define a manifold precisely. While this will be by no means an exhaustive tour of a vast field, it will give you a flavour of topological spaces. Three key concepts for the study of topological spaces are continuity, compactness and connectedness. Intuitively:

- Continuous functions, take "nearby" points to "nearby" points.
- Compact sets are those that can be "covered" by finitely many sets of arbitrarily "small size".
- Connected sets are sets that cannot be divided into two pieces that are "far apart".

We will define these concepts precisely below.

In its most basic form, a ***topological space*** is a pair $(X, \mathcal{O}_X)$, where X is a set of objects[2] and $\mathcal{O}_X$ is called ***a topology***, and is a set that contains all ***open sets*** of X.

[2] we can call each object a "point" in the set X, hence point-set topology

Open, Closed, and Clopen Sets

What is an open set for a topological space $(X, \mathcal{O}_X)$? It's just something in $\mathcal{O}_X$! Note that the topology $\mathcal{O}_X$ is not a ***completely*** arbitrary set of subsets of X; We require that $\mathcal{O}_X$

- contains the empty set, $\varnothing$ (also called the null set),
- contains X itself,
- contains the collection of the (possibly infinite) union of any open sets $\bigcup_i U_i \in \mathcal{O}_X$,
- also contains collection of finite intersection of any open sets $\bigcap_i U_i \in \mathcal{O}_X$.

A set is called ***closed*** if its ***complement*** in X is open, i.e. S is called closed iff $S' = \{p | p \in X \text{ and } p \notin S\} \in \mathcal{O}_X$.
If a set is both open and closed it is sometimes called "clopen", that is S is considered clopen if $S \in \mathcal{O}_X$ ***and*** $S' \in \mathcal{O}_X$.

We only include the intersection of a finite number of open sets, since allowing infinite intersections can lead to closed sets in most standard topologies.

It can be shown that a set is clopen iff its boundary is empty, and that a topological space $(X, \mathcal{O})$ is ***connected*** iff the only clopen sets are $\varnothing$ and X.

On the real line we have the "usual" or "standard" topology, defined by the set of open 1D intervals and their unions and (finite) intersections, that is,

$$(\mathbb{R}, \mathcal{O}_s), \quad \text{where} \quad \forall a < b \in \mathbb{R} \text{ the set } \{x | a < x < b, x \in \mathbb{R}\} \in \mathcal{O}_S.$$

$\mathcal{O}_s$ stands for "standard" topology

$\mathcal{O}_D$ stands for "discrete" topology

The integers $(\mathbb{Z}, \mathcal{O}_D)$ have a "discrete" topology, where every element of Z by itself is a member of $\mathcal{O}_D$.

Continuity and open sets

Consider a mapping, $\phi : X \to Y$, from a topological space $(X, \mathcal{O}_X)$ to a topological space $(Y, \mathcal{O}_Y)$, The mapping ϕ is called ***continuous*** if and only if

$$\forall U \in \mathcal{O}_Y, \qquad \text{preimage}_\phi(U) \in \mathcal{O}_X,$$

where the ***preimage*** is defined as

$$\text{preimage}_\phi(U) \equiv \{x | x \in X, \phi(x) \in U\}.$$

Note that if X and Y are metric spaces, this definition of continuity is equivalent to the standard ϵ-δ definition of continuity from real/complex analysis.

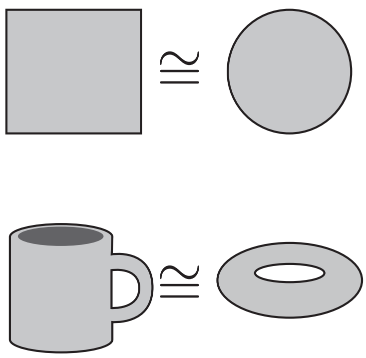


Figure 3.1: Colloquially: topological spaces are geometric objects and homeomorphisms are mappings that continuously deform (bending/stretching but not tearing, so that continuity is preserved) an object into a new shape. (Here the $\cong$ symbol indicates they are topological isomorphic, or homeomorphic.)

The above notion of ***continuity*** is the kind of mapping that preserves the ***structure*** (openness) of a topological space. If we have a mapping ϕ that is a bijection and is ***bicontinuous***, (i.e. ϕ and its inverse ϕ^{-1} are both continuous) then it is a topological isomorphism, which is confusingly called a ***homeomorphism***.

Spaces which have a homeomorphism between them are called ***homeomorphic*** and are considered topologically equivalent.

Compact Topological Spaces

A topological space $(X, \mathcal{O}_X)$ is called ***compact*** if each of its ***open covers*** has a ***finite subcover***.
An ***open cover*** of X is a collection of open subsets of X, e.g. $\{U_i\}_{i\in A}$, with each $U_i \in \mathcal{O}_X$, such that

Definition of an open cover

$$X = \bigcup_{i\in A} U_i.$$

If for ***any*** open cover, $\{U_i\}_{i\in A}$, there is a ***finite*** subset $B \subseteq A$, such that

$$X = \bigcup_{i\in B} U_i$$

Definition of a finite subcover

then the topological space $(X, \mathcal{O}_X)$ is called ***compact***.

Compactness of a topological space generalises the idea of an interval on the real line being closed, i.e. that ***it contains all its own limit points***. For example, a subspace of $\mathbb{R}^n$ is compact if and only if it is both closed and bounded[3].

Compactness is often a useful property to keep track of since it can be shown that every homeomorphic mapping of a compact space is compact. Additionally, if we are given a compact metric space (a compact topological space with distances defined), then every sequence of points has a convergent subsequence. This fact is often used as the definition of compactness for real and complex analysis.

[3] This is called the Heine-Borel theorem.

Connected Topological Spaces

A topological space $(X, \mathcal{O}_X)$ is ***disconnected*** if it is the union of two (or more) disjoint non-null open sets, i.e., $\exists U_1, U_2 (\neq \varnothing) \in \mathcal{O}_X$, such that

$$X = U_1 \cup U_2 \quad \text{and} \quad U_1 \cap U_2 = \varnothing.$$

Otherwise it is called ***connected***.
A topological space is called ***totally disconnected*** if all of the points are considered open one-point sets, e.g. $(\mathbb{Z}, \mathcal{O}_D)$.

Example 3.1: The Standard Topology of $\mathbb{R}^n$

The topological space $(\mathbb{R}^n, \mathcal{O}_s)$ uses the ***standard*** topology, which consists of the set of all open "balls" of the form

$B_r(\boldsymbol{x}_o)$

$$B_r(\boldsymbol{x}_o) = \{\boldsymbol{x} | \boldsymbol{x} \in \mathbb{R}^n, (\boldsymbol{x} - \boldsymbol{x}_o)^2 < r^2\}$$

$\forall r \in \mathbb{R}, \forall \boldsymbol{x}_o \in \mathbb{R}^n$, as well as all unions and finite intersections of the open balls.

Example 3.2: The Composition Map

$f \circ g$ is pronounced "f after g", "f composed with g", or "f circ g"

We take a brief aside here to discuss the composition map $\circ$. If we have mappings $f : U \to V$, and $g : W \to U$, for some sets U, V, W, the composition map $f \circ g$ is defined as

$$\begin{aligned} f \circ g : W &\to V \\ w &\mapsto f(g(w)), \qquad \forall w \in W. \end{aligned}$$

Let us consider three topological sets, $(A, \mathcal{O}_A)$, $(B, \mathcal{O}_B)$, and $(C, \mathcal{O}_C)$. Let $\phi : A \to B$ be a homeomorphism.
Show that a mapping $(\psi \circ \phi) : A \to C$ is continuous ***if and only if*** the mapping $\psi : B \to C$ is continuous.
Proof: We have to prove both the "if" and the "only if".

- If ψ is a continuous map, then any open subset $U_C \in \mathcal{O}_C$ has the preimage$_\psi(U_c) = U_B \in \mathcal{O}_B$. Since ϕ is a continuous map, we have preimage$_\phi(U_B) = U_A \in \mathcal{O}_A$, and hence

$$\begin{aligned} \text{preimage}_{\psi\circ\phi}(U_C) &= \text{preimage}_\phi(\text{preimage}_\psi(U_C)) \\ &= \text{preimage}_\phi(U_B) \\ &= U_A \in \mathcal{O}_A, \end{aligned}$$

  proving the "if".

- Suppose that $\psi \circ \phi : A \to C$ is continuous ***but*** $\psi : B \to C$ is ***not***. Then there must exist an open subset of $U_C \in \mathcal{O}_C$, such that

$$\text{preimage}_\psi(U_C) \notin \mathcal{O}_B.$$

  However, since $\psi \circ \phi$ is continuous we have that preimage$_{\psi\circ\phi}(U_C) = U_A \in \mathcal{O}_A$ must be open. Since $\phi : A \to B$ is a ***homeomorphism*** is must be bicontinuous, so that $\phi^{-1} : B \to A$ is also continuous, hence we must have

$$\text{preimage}_{\phi^{-1}}(U_A) = U_B \in \mathcal{O}_B$$

  be open. This contradicts the result above, hence ψ must be continuous if $\psi \circ \phi$ is continuous, proving the "only if".
  □

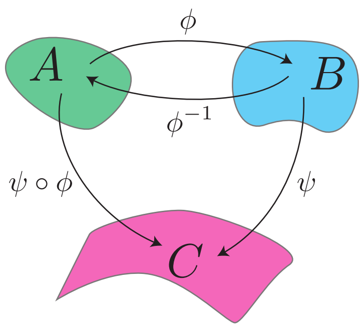


Figure 3.2: If ϕ is a homeomorphism, the composition map $\psi \circ \phi$ is continuous iff ψ is a continuous map.

3.2 Topological Manifolds

A ***topological manifold*** is a generalisation of an n-dimensional Euclidean space $\mathbb{R}^n$.

More precisely a manifold is a topological space that is ***locally*** homeomorphic to $\mathbb{R}$, and is given by the triple

$$(M, \mathcal{O}_M, \mathcal{A}_M),$$

where M is the base space set, $\mathcal{O}_M$ is the topology, and $\mathcal{A}_M$ is called the ***atlas*** of M. The atlas of M is a set of (overlapping) "***charts***" which homeomorphically map subsets of M to subsets of the Euclidean space $(\mathbb{R}^n, \mathcal{O}_s)$ with the standard topology.

Figure 3.3: You can think of an Atlas like a collection of OS maps that cover the globe (a manifold). We know we can't really map the entire earth to a 2D flat, continuous chart, but we can use many such small OS maps as charts that overlap to cover the entire globe.

Charts

Formally, chart is a pair (U, ϕ),

- an open ***chart domain***, $U \in \mathcal{O}_M$, which is a subset of M,
- and a homeomorphism called a ***chart map***, ϕ, i.e.

$$\phi : U \to V$$

where $V \in \mathcal{O}_s$ is an open subset of $\mathbb{R}^n$, and ϕ is both bi-continuous and invertible.

Atlases and dimensionality

An ***atlas*** $\mathcal{A}_M$ for a topological space M is an indexed set of charts $\mathcal{A}_M = \{(U_i, \phi_i) : i \in I\}$ (where I is just the set of chart indices for this particular atlas, which may or may not be countable), such that the chart domains of $\mathcal{A}_M$, cover M, i.e.

$$M = \bigcup_{i \in I} U_i.$$

If each chart in $\mathcal{A}_M$ has a range in n-dimensional Euclidean space, $\phi_i(U_i) \subset \mathbb{R}^n$, then $(M, \mathcal{O}_M, \mathcal{A}_M)$ is called an ***n-dimensional manifold***, and we can write $\dim M = n$.

Note that there are other technical requirements for a topological space to be a manifold that we don't have time to get into, but that you can look up. Specifically, any topological space that is

- ***a Hausdorff space*** (any two points can be separated with open sets),
- ***second countable*** (any open set can be written as the union of some ***countable*** "base" of open sets),
- and ***para-compact*** (you don't need "too many" open sets in order to cover the base space),

that is equipped with an atlas is a manifold.

So we have defined each part of a manifold, and seen how the charts in the atlas of a manifold let us link an n-dimensional manifold directly to the Euclidean space of $\mathbb{R}^n$. Thus using these charts, we can bootstrap our extensive knowledge of mathematics on $\mathbb{R}^n$ to the more general notion of a manifold.

3.2.1 C^k-Manifolds

So far we have dealt with continuous maps, in this abstract definition of continuity, rather than the notion of continuity we have from dealing with functions in Euclidean space.

Given two overlapping charts (U_1, ϕ_1) and (U_2, ϕ_2), we can now talk about the $\mathbb{R}^n \to \mathbb{R}^n$ functions

$$\phi_2 \circ \phi_1^{-1} : \phi_1(U_1 \cap U_2) \to \phi_2(U_1 \cap U_2), \quad (3.1)$$

$$\phi_1 \circ \phi_2^{-1} : \phi_2(U_1 \cap U_2) \to \phi_1(U_1 \cap U_2) \quad (3.2)$$

where $\phi_1(U_1 \cap U_2) \subseteq \mathbb{R}^n$ and $\phi_2(U_1 \cap U_2) \subseteq \mathbb{R}^n$ are the ***images*** of the overlap in the domains. These functions allow you to compare two charts and are sometimes called ***transition maps*** between those charts.

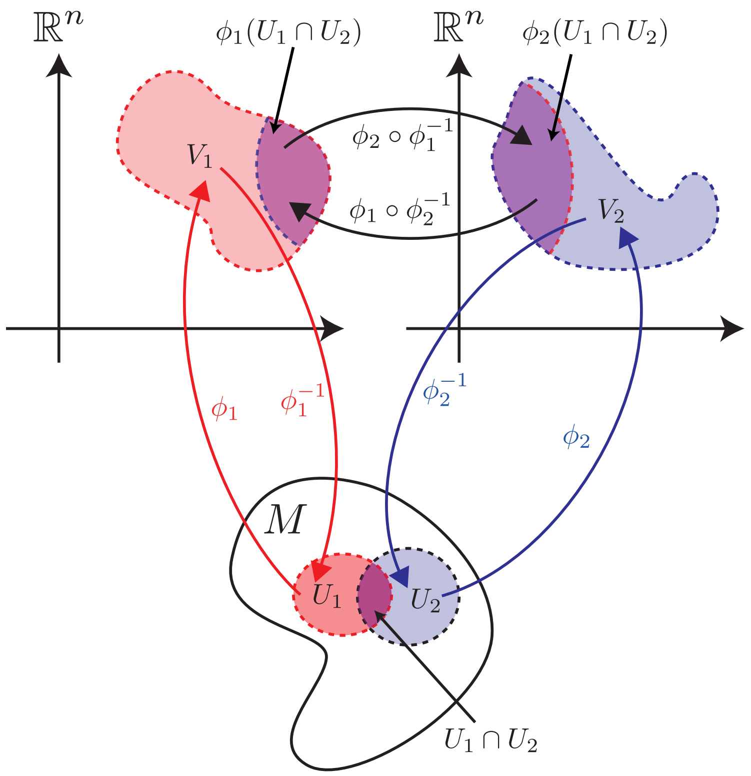


Figure 3.4: For a manifold M any two overlapping charts (U_1, ϕ_1) and (U_2, ϕ_2), the transition maps $\phi_2 \circ \phi_1^{-1}$ and $\phi_1 \circ \phi_2^{-1}$ are only defined in the regions $\phi_1(U_1 \cap U_2)$ and $\phi_2(U_1 \cap U_2)$ respectively.

These transition functions take Euclidean coordinates to Euclidean coordinates, and are therefore truly functions in the sense that we are used to. In studying these functions, it is useful to introduce a notion of a C^k-function

C^k-functions

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is called a C^k function if and only if the kth derivatives of f are continuous.
Thus $C^0(\mathbb{R}^n, \mathbb{R}^m)$ is the set of continuous functions that map $\mathbb{R}^n \to \mathbb{R}^m$. $C^1(\mathbb{R}^n, \mathbb{R}^m)$ is the set of all such functions that have continuous first derivatives, and so on. The set $C^\infty(\mathbb{R}^n, \mathbb{R}^m)$ is the set of all such functions that are ***infinitely differentiable***. We often call C^∞-functions ***smooth*** functions. Note that any C^k-function is automatically also a C^{k-1}-function.

Two charts of an n-dimensional manifold are said to be C^k-compatible if their transition maps are members of $C^k(\mathbb{R}^n, \mathbb{R}^n)$. A C^k-atlas is an atlas that contains only C^k-compatible charts, that is any two charts in the atlas are C^k-compatible.

A C^k-atlas that all possible C^k-compatible charts called a ***maximal atlas***. Note that there may be multiple distinct maximal atlases, as there may be several equivalent but C^k-incompatible groups of charts.

A manifold is called a C^k-manifold if it is equipped with a C^k-atlas, and a manifold is called a ***smooth*** manifold if it is equipped with a C^∞-atlas, that is all possible transition maps are smooth.

All manifolds are automatically C^0-manifolds, since topological continuity guarantees that the charts (since they are homeomorphisms) must be C^0- compatible.

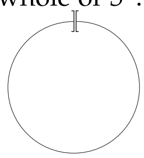
The choice of an atlas, which provides overlapping charts that map $M \to \mathbb{R}^n$ and have chart domains that cover M, is what we formally mean when we say that a manifold "locally looks like $\mathbb{R}^n$".

Most manifolds can't be covered with a single chart. Consider the circle, sometimes called S^1.

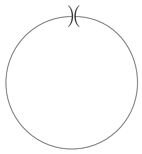
The manifold S^n usually refers to the n-dimensional surface of a $(n+1)$-dimensional sphere. Here S_1 is the 1D surface of a 2D sphere (circle).

Example 3.3: The Circle: S^1

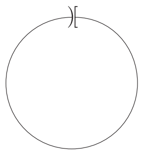
The usual way to give a coordinate system (chart) relating a circle to $\mathbb{R}$, would be $\theta : S^1 \to \mathbb{R}$, where $\theta = 0$ at one point, and it wraps around to $\theta = 2\pi$.
For θ to be a chart map, it needs its domain U_θ to be ***open***.
We need to show that an open chart domain can't cover the whole of S^1.



The closed interval with $\theta = [0, 2\pi]$ that includes both 0 and 2π is clearly not open.

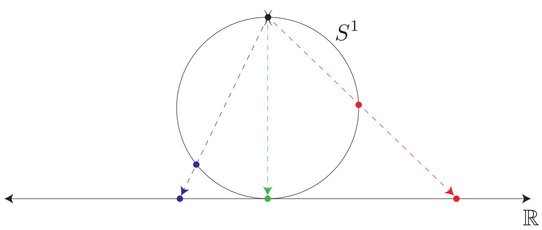


The open interval with $\theta = (0, 2\pi)$ that does not include 0 or 2π, so it doesn't cover the point at the top of the circle.

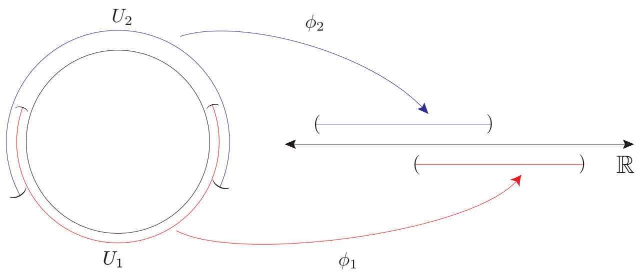


The half-open/half-closed interval with $\theta = (0, 2\pi]$ does include this top point, but isn't open.
The same is true for $\theta = [0, 2\pi)$.

We could try to map the circle to the ***entire*** real line by, for example, mapping each point in the circle to the point on the real line made by projecting a line from the top of the circle through a point on S^1 onto a line tangent to the bottom of the circle.



However, we see that this chart domain does not include the top of the circle itself, since that would map to the same point as the bottom.
So we need two charts (U_1, ϕ_1), and (U_2, ϕ_2) that have open domains which ***overlap*** and cover the entire circle.



This makes sense because the circle is compact (a topological property preserved by homeomorphisms) and $\mathbb{R}$ is not, therefore S^1 ***cannot*** be homeomorphic to the real line,

3.2.2 *Differentiable Maps and Differmorphisms*

Now we can talk about differentiable maps ***between*** manifolds by stealing this notion from $\mathbb{R}^n$.

Let's consider two manifolds, $(M, \mathcal{O}_M, \mathcal{A}_M)$ and $(N, \mathcal{O}_N, \mathcal{A}_N)$ that have some map ϕ between them,

$$\phi : M \supseteq U \to V \subseteq N. \tag{3.3}$$

Let us assume that the manifold M has chart maps x and $\tilde{x}$ with domains overlapping so that they each contain a domain of interest, U, such that $(U, x) \in \mathcal{A}_M$ and $(U, \tilde{x}) \in \mathcal{A}_M$. Similarly let us assume the charts (V, y) and $(V, \tilde{y})$ are both in atlas $\mathcal{A}_N$.

$$\begin{tikzcd}
\mathbb{R}^{\dim M} \supseteq \tilde{x}(U) \arrow[r, "\tilde{y}\circ\phi\circ\tilde{x}^{-1}"] & \tilde{y}(V) \subseteq \mathbb{R}^{\dim N} \\
M \supseteq U \arrow[u, "\tilde{x}"'] \arrow[d, "x"] \arrow[r, "\phi"] & V \subseteq N \arrow[u, "\tilde{y}"'] \arrow[d, "y"] \\
\mathbb{R}^{\dim M} \supseteq x(U) \arrow[r, "y\circ\phi\circ x^{-1}"] \arrow[uu, bend left, "\tilde{x}\circ x^{-1}"] & y(V) \subseteq \mathbb{R}^{\dim N} \arrow[uu, bend right, "\tilde{y}\circ y^{-1}"']
\end{tikzcd}$$

Figure 3.5: The "commutative" diagram for looking at differentiability of map $\phi : M \to N$

The map ϕ is called k-differentiable, if the $\mathbb{R}^{\dim M} \to \mathbb{R}^{\dim N}$ function

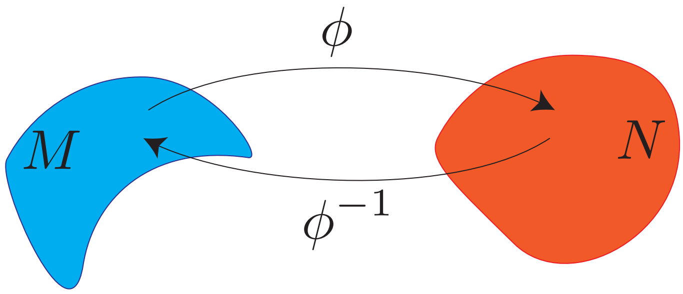
$$y \circ \phi \circ x^{-1} : x(U) \to y(V) \tag{3.4}$$

is continuously differentiable k-times, that is, it is a member of $C^k(\mathbb{R}^{\dim M}, \mathbb{R}^{\dim N})$. The set of k-differentiable maps between M and N is called $C^k(M, N)$.

If we say that map ϕ is ***smooth*** or simply "differentiable" we usually mean $\phi \in C^{\infty}(M, N)$.

Diffeomorphisms

A ***diffeomorphism***, ϕ, between manifolds M and N, is an infinitely differentiable (smooth) map between the manifolds, that also has a smooth inverse map, ϕ^{-1} (i.e. it is a bijection with an infinitely differentiable inverse).



The diffeomorphisms are the ***structure preserving maps of smooth manifolds***, and as such they are often called isomorphisms of smooth manifolds.

Using the notion of calculus on $\mathbb{R}^n$, we have developed a notion of differentiability of maps. We can use this to construct some additional useful concepts on manifolds by considering differentiable maps to specific manifolds.

3.2.3 *Functions on a Manifold*

Let's consider the set of ***smooth scalar functions on*** M, that is

$$C^{\infty}(M, \mathbb{R}) = \{f | f : M \to \mathbb{R}, f \circ \phi \in C^{\infty}(\mathbb{R}^{\dim M}, \mathbb{R})\}, \tag{3.5}$$

for all chart maps ϕ, where $f \in C^{\infty}(M, \mathbb{R})$ assigns a scalar value to each point in M in an infinitely differentiable way. We often call such scalar functions on a manifold ***scalar fields***.

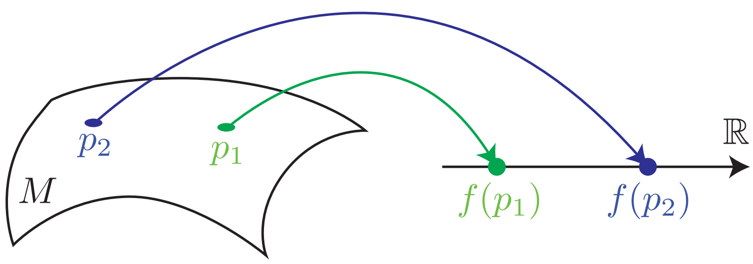


Figure 3.6: A scalar function $f \in C^{\infty}(M, \mathbb{R})$ smoothly maps points on M to scalar values in $\mathbb{R}$.

In fact, we can make $C^{\infty}(M, \mathbb{R})$ a vector space by adopting ***pointwise*** $+$ and $\cdot$ operations inherited from the real numbers. I.e.

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

$$[f+g](p) = f(p) + f(g), \qquad \forall p \in M. \tag{3.6}$$
$$[c \cdot f](p) = c \cdot f(p), \qquad \forall p \in M, \forall c \in \mathbb{R}. \tag{3.7}$$

3.2.4 *Curves on Manifolds*

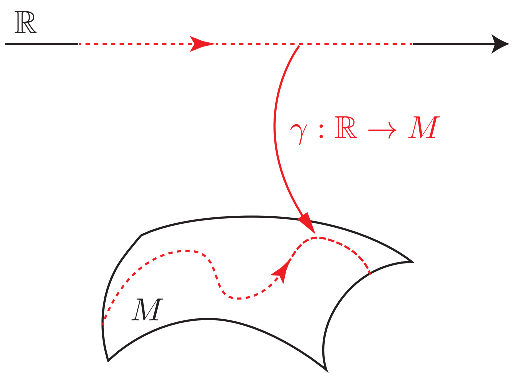


Figure 3.7: A smooth curve $\gamma : \mathbb{R} \to M$ is an infinitely differentiable map that describes a path on M.

We can also define ***smooth (1-dimensional) curves*** on a manifold by looking at a mappings of the form

$$\gamma : \mathbb{R} \to M \tag{3.8}$$

where γ is a C^∞ map with respect to the charts in $\mathcal{A}_M$.

The obvious example of a curve on a manifold are the inverse maps of the the coordinate grids for each chart (see Figure 3.8).

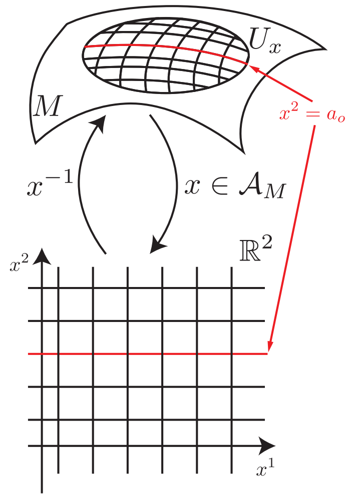


Figure 3.8: An example of a curve $\gamma = x^{-1}(\cdot, a_o)$ on the 2D manifold M is the the line where, for a chart $(\boldsymbol{x}, U_x)$, the chart coordinate $x^2 = a_o$ is constant.

3.3 *Tangent Spaces*

Establishing what we mean by scalar fields on a manifold, it would be nice to know how to define a vector field (or a tensor field) on an arbitrary manifold. However, we don't yet have a notion of what a vector on a manifold is!

Consider all possible vectors you could have located at a point on the manifold of S^2, i.e. the surface of a sphere.

If we allow ourselves the luxury of embedding this manifold in $\mathbb{R}^3$, that is we consider S^2 as a subset of $\mathbb{R}^3$, then we can talk about vectors at a point in the ***tangent plane***.

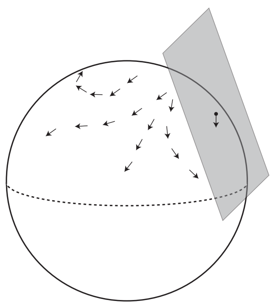


Figure 3.9: Vectors on the surface of the sphere don't really live on the curved surface of the sphere, rather they exist on planes that are ***tangent*** to the point where they are anchored.

But it would be better to describe these vectors in a way that doesn't require embedding in a higher Euclidean space, since we don't always have access to this higher space.

The key, it turns out, is the word "tangent". The notion that vectors are ***tangent*** to the manifold is intuitive and correct, but how do we define "tangent" using the tools that we have developed?

In differential geometry $X_{\gamma,p}$ is usually called the ***tangent vector*** to the curve γ at $p \in M$. You can think of $X_{\gamma,p}$ (by itself, rather than acting on a function) as something like the ***velocity*** of a particle at a particular point p, moving along the curve γ. When it acts on a function f, it just gives you the derivative of f that the particle sees as it moves along γ.

The Directional Derivative Operator

$$\begin{tikzcd} \mathbb{R} \arrow[r, "\gamma"] \arrow[rr, bend left, "f\circ\gamma"] & M \arrow[r, "f"] & \mathbb{R} \end{tikzcd}$$

Let $\gamma : \mathbb{R} \to M$ be a smooth (C^∞) 1-dimensional curve through a point $p \in M$, and let f be an arbitrary scalar function. Without loss of generality we may assume that $p = \gamma(0)$.

The ***directional derivative operator at*** p ***along*** γ is defined as the linear map

$$X_{\gamma,p} : C^\infty(M, \mathbb{R}) \xrightarrow{\text{lin}} \mathbb{R} \tag{3.9}$$

$$f \mapsto \underbrace{\underbrace{(f \circ \gamma)}_{\mathbb{R}\to\mathbb{R}}{}'(0)}_{\in\mathbb{R}} \tag{3.10}$$

where $'$ denotes the usual derivative of a function. This operator takes in scalar functions (scalar fields) on M, and gives you back the derivative of the function along the curve γ at the point p.

Example 3.4: Twice as fast

Let $\delta : \mathbb{R} \to M$, and $\gamma : \mathbb{R} \to M$ be two curves parameterised by $\lambda \in \mathbb{R}$, such that

$$\delta(\lambda) = \gamma(2\lambda).$$

Then the directional derivatives of an arbitrary scalar function $f \in C^\infty(M \to \mathbb{R})$ along δ and γ at $p \in M$ are related by

$$\begin{aligned} X_{\delta,p}(f) &= (f \circ \delta)'(0) \\ &= 2(f \circ \gamma)'(0) \\ &= 2X_{\gamma,p}(f), \end{aligned}$$

so that the directional derivative operators are related by

$$X_{\delta,p} = 2X_{\gamma,p},$$

i.e. the curve δ has twice the "velocity" of the curve γ.

The Tangent (Vector) Space at $p \in M$

For a manifold M, the ***Tangent Space at point*** $p \in M$ is defined as

$$T_pM \equiv \{X_{\gamma,p} | \gamma \text{ is a smooth curve through } p\},$$

i.e. it is the set of directional derivatives that can be constructed out of ***any curve*** that passes through point p.

Can this set, T_pM, be made into a (real) vector space? In order to do this we need to define appropriate addition and scalar multiplication operations,

$$+ : T_pM \times T_pM \to T_pM \tag{3.11}$$
$$\cdot : \mathbb{R} \times T_pM \to T_pM \tag{3.12}$$

that obey the vector space axioms (CANI ADDU).

We can define the $+$ and $\cdot$ operations on arbitrary $X_{\gamma,p}, X_{\delta,p} \in T_pM$ ***pointwise***, that is by ***defining how the result applies*** to any $f \in C^1(M, \mathbb{R})$,

$$(X_{\gamma,p} + X_{\delta,p})(f) \equiv X_{\gamma,p}(f) + X_{\delta,p}(f) \tag{3.13}$$
$$(c \cdot X_{\gamma,p})(f) \equiv c \cdot X_{\gamma,p}(f) \tag{3.14}$$

where the operations on the right are operations on real numbers.

Most of ***CANI ADDU*** are satisfied by virtue of the properties inherited from these scalar operations. We can define the additive **N**eutral element to be the directional derivative operator of the curve that maps ***any*** real number to point p, and hence will return a 0 for a directional derivative of any function. We can define an additive **I**nverse to $X_{\gamma,p}$ by choosing

$$-X_{\gamma,p} \equiv X_{\delta,p}, \tag{3.15}$$
$$\text{where } \forall\lambda \in \mathbb{R}, \qquad \delta(\lambda) \equiv \gamma(-\lambda). \tag{3.16}$$

However we are not done, we need to prove that the result of these pointwise definitions operations are ***also in*** T_pM (closure). This isn't necessarily obvious!

Closure of T_pM Under Scalar Multiplication

Above, we defined pointwise scalar multiplication, for arbtirary $c \in \mathbb{R}$,

$$(c \cdot X_{\gamma,p})(f) \equiv c \cdot X_{\gamma,p}(f), \qquad \forall f \in C^\infty(M, \mathbb{R}).$$

However we still need to show that $c \cdot X_{\gamma,p}(f)$ is an element of T_pM, i.e. ***we need to find a curve, σ, such that***

$$X_{\sigma,p} = c \cdot X_{\gamma,p}.$$

Luckily, we already looked at something very similar in Example 3.4

We can just choose curve $\sigma \in C^\infty(R, M)$ such that

$$\sigma(\lambda) = \gamma(c \cdot \lambda), \qquad \forall \lambda \in \mathbb{R},$$

so that for all functions f,

$$\begin{aligned} X_{\sigma,p}(f) &= (f \circ \sigma)'(0) \\ &= c \cdot (f \circ \gamma)'(0) \\ &= c \cdot X_{\gamma,p}(f), \end{aligned}$$

and hence, $c \cdot X_{\gamma,p} = X_{\sigma,p} \in T_pM$. $\square$

Closure of T_pM Under Addition

If we have two curves $\gamma : \mathbb{R} \to M$ and $\delta : \mathbb{R} \to M$ (with $\gamma(0) = \delta(0) = p \in M$), we need to show that there exists ***another curve***, $\sigma : \mathbb{R} \to M$, such that

$$X_{\sigma,p} = X_{\gamma,p} + X_{\delta,p}.$$

We will do this by explicitly ***constructing*** this curve, with the aid of an arbitrary coordinate chart $(U_x, x) \in \mathcal{A}_M$, (with $p \in U_x$) by choosing a curve σ defined by,

$$\sigma = \underbrace{x^{-1}}_{\mathbb{R}^n \to M} \circ \underbrace{(x \circ \gamma + x \circ \delta - x(p))}_{\mathbb{R} \to \mathbb{R}^n}.$$

Using this definition we can calculate how the tangent vector $X_{\sigma,p}$ acts on an arbitrary function $f \in C^\infty(M, \mathbb{R})$:

$$\begin{aligned} X_{\sigma,p} f &\equiv (f \circ \sigma)'(0) \\ &= [f \circ x^{-1} \circ (x \circ \gamma + x \circ \delta - x(p))]'(0) \\ &= \partial_i [\underbrace{f \circ x^{-1}}_{\mathbb{R}^n \to \mathbb{R}}]|_{x(p)} \cdot [(\underbrace{x \circ \gamma + x \circ \delta - x(p)}_{\mathbb{R} \to \mathbb{R}^n})^i]'(0) \\ &= \partial_i [f \circ x^{-1}]|_{x(p)} (x^i \circ \gamma)'(0) \\ &\quad + \partial_i [f \circ x^{-1}]|_{x(p)} (x^i \circ \delta)'(0) \\ &\quad - \partial_i [f \circ x^{-1}]|_{x(p)} (\underbrace{x^i(p)}_{\text{constant}})'(0) \end{aligned}$$

where we are using the multivariable chain rule and the Einstein sum convention to take the derivative, and x^i is the ith coordinate of the x chart. Recombining these we get,

$$\begin{aligned} X_{\sigma,p} f &= [\underbrace{f \circ x^{-1} \circ x \circ \gamma}_{f \circ \gamma} + \underbrace{f \circ x^{-1} \circ x \circ \delta}_{f \circ \delta} - \underbrace{f \circ x^{-1} \circ x(p)}_{\text{constant}}]'(0), \\ &= [f \circ \gamma]'(0) + [f \circ \delta]'(0), \\ &= X_{\gamma,p} f + X_{\delta,p} f. \end{aligned}$$

Thus we see that $X_{\sigma,p} = X_{\gamma,p} + X_{\delta,p}$ and T_pM is indeed closed under pointwise addition. $\square$

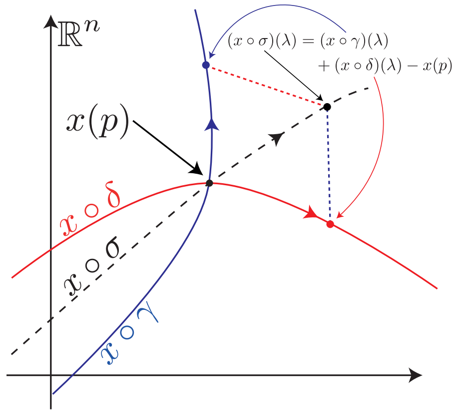


Figure 3.10: We define curve σ with the help of a chart x, such that $x \circ \sigma = x \circ \gamma + x \circ \delta - x(p)$. While this ***does*** define a different curve σ for every choice of chart x, each of those σ actually has the same directional derivative operator (tangent vector) $X_{\sigma,p}$.

3.4 *A Coordinate basis for T_pM*

In the previous section we showed that T_pM was indeed a vector space, when equipped with pointwise addition and scalar multiplication. We also had the notion of manifold ***dimension***, with $\dim M$ given by the dimension of the Euclidean space it is locally homeomorphic to.

Really, here by $\dim M$, we mean the dimension of the manifold at point p. Note that we could make a weird manifold, for example, by taking the union of two manifolds with different dimensions. Such a a manifold obviously has different dimension at different points. In GR we really only talk about manifolds with a single dimension, often called ***pure manifolds***.

But we also have the ***vector space dimension*** of the tangent vector space T_pM. Even though these two notions of dimensionality are completely different, it turns out that,

$$\dim T_pM = \dim M. \tag{3.17}$$

To prove this we can begin by constructing a vector space basis for T_pM.

The Coordinate Basis Vectors

Given an arbitrary chart $(U_x, x) \in \mathcal{A}_M$, where $p \in U_x$. (without loss of generality we may assume that $x(p) = (0, \ldots, 0)$, since we can just subtract constants from any x to get this to be true), we want to choose a set of curves

$$\gamma_i : \mathbb{R} \to U_x, \tag{3.18}$$

where the index $i = 1 \ldots \dim M$ such that

$$(x^j \circ \gamma_i)(\lambda) = \delta^j{}_i \cdot \lambda. \tag{3.19}$$

Here $\lambda \in \mathbb{R}$, and $(x^j \circ \gamma_i)$ is the jth component (projection) of the $\mathbb{R} \to \mathbb{R}^{\dim M}$ function $(x \circ \gamma_i)$, which are the coordinate curves that pass through point p.
We then construct the tangent vector to $\gamma_i(0) = p$ and define this as the basis vector

$$\boldsymbol{e}_i \equiv X_{\gamma_i, p} \in T_pM, \tag{3.20}$$

such that for an arbitrary function $f \in C^\infty(M, \mathbb{R})$,

$$\boldsymbol{e}_i f = (\underbrace{f \circ \gamma_i}_{\mathbb{R}\to\mathbb{R}})'(0) = (\underbrace{f \circ x^{-1}}_{\mathbb{R}^n\to\mathbb{R}} \circ \underbrace{x \circ \gamma_i}_{\mathbb{R}\to\mathbb{R}^n})'(0).$$

Applying the multidimensional chain rule (as usual) we get

$$\begin{aligned} \boldsymbol{e}_i f &= \partial_j (f \circ x^{-1})|_{x(p)} \cdot \underbrace{(x^j \circ \gamma_i)'(0)}_{\delta^j{}_i} \\ &= \partial_i (f \circ x^{-1})|_{x(p)} \equiv \left(\frac{\partial}{\partial x^i}\right)_p f. \end{aligned}$$

Thus we can ***define*** the symbol

$$\left(\frac{\partial}{\partial x^i}\right)_p : C^\infty(M, \mathbb{R}) \to \mathbb{R}, \qquad \left(\frac{\partial}{\partial x^i}\right)_p \equiv \boldsymbol{e}_i \equiv X_{\gamma_i, p} \tag{3.21}$$

to mean the ***coordinate basis vector*** associated with the ith component of the x-coordinate chart. This basis set is also sometimes called the ***chart-induced vector basis***.

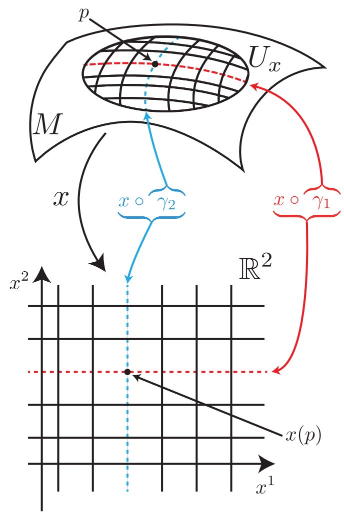


Figure 3.11: We construct a coordinate basis for the tangent space T_pM by using the coordinate curves $x \circ \gamma_i$ for a particular chart x.

$$\begin{tikzcd} \mathbb{R} \arrow[r, "\gamma_i"] \arrow[rr, bend left, "f\circ\gamma_i"] \arrow[dr, bend right, "x\circ\gamma_i"'] & M \arrow[r, "f"] \arrow[d, "x"] & \mathbb{R} \\ & \mathbb{R}^n \arrow[ur, bend right, "f\circ x^{-1}"'] & \end{tikzcd}$$

Figure 3.12: Commutative diagram for the coordinate curves.

The above is in contrast to the usual partial derivative operator which we wrote above as

$$\partial_i : C^\infty(\mathbb{R}^n, \mathbb{R}) \to \mathbb{R} \tag{3.22}$$

which is the normal partial derivative with respect to the ith slot of some function on $\mathbb{R}^n$.

In order to show that this "basis" $\{(\partial/\partial x^i)_p\}_i$ is actually a vector space basis for T_pM we still need to prove that it ***spans*** T_pM and it is ***linearly independent.***

- $\{(\partial/\partial x^i)_p\}_i$ ***spans*** T_pM: we need to show that any $X \in T_pM$ can be written as

$$X = \underbrace{X^i}_{\in\mathbb{R}} \left(\frac{\partial}{\partial x^i}\right)_p . \tag{3.23}$$

  To do this we choose arbitrary curve $\sigma : \mathbb{R} \to M$ smooth through $\sigma(0) = p$, so that $X = X_{\sigma,p}$. We then have, $\forall f \in C^\infty(M, \mathbb{R})$,

$$\begin{aligned} X_{\sigma,p} f &= (f \circ \sigma)'(0), \\ &= (f \circ \underbrace{x^{-1} \circ x}_{\mathrm{id}_M} \circ \sigma)'(0), \\ &= \underbrace{\partial_i (f \circ x^{-1}(x(p)))}_{\left(\frac{\partial}{\partial x^i} f\right)_p} \cdot (x^i \circ \sigma)'(0), \\ &= \underbrace{\left[(x^i \circ \sigma)'(0) \cdot \left(\frac{\partial}{\partial x^i}\right)_p \right]}_{X^i \cdot \left(\frac{\partial}{\partial x^i}\right)_p} f, \\ \Rightarrow X_{\sigma,p} &= X^i \left(\frac{\partial}{\partial x^i}\right)_p . \end{aligned} \tag{3.24}$$

  Thus $\{(\partial/\partial x^i)_p\}$ spans T_pM.

$$\left(\frac{\partial}{\partial x^i}\right)_p f = \partial_i (f \circ x^{-1})|_{x(p)}$$

$$f : M \to \mathbb{R}$$

$$(f \circ x^{-1}) : \mathbb{R}^{\dim M} \to \mathbb{R}$$

$$\left(\frac{\partial}{\partial x^i}\right)_p : C^1(M, \mathbb{R}) \to \mathbb{R}$$

$$\partial_i : C^1(\mathbb{R}^{\dim M}, \mathbb{R}) \to C^0(\mathbb{R}^{\dim M}, \mathbb{R})$$

Figure 3.13: How to translate between the ith directional derivative basis vector and the partial derivative of a $\mathbb{R}^{\dim M} \to \mathbb{R}$ function with respect to the ith slot.

- ***Linear independence:*** Let $f \in C^\infty(M, \mathbb{R})$ be an arbitrary smooth function. Consider the sum, and assume we have c^i such that,

$$\left[c^i \left(\frac{\partial}{\partial x^i}\right)_p \right] f = 0. \tag{3.25}$$

  If this implies that ***all*** scalar c^i must be zero, then the basis must be linearly independent.

  To show that all c^i must be zero, let us choose a particular function $f = x^j : M \to \mathbb{R}$. That is, we choose the jth component of the x-coordinate map.

$$c^i \left(\frac{\partial}{\partial x^i}\right)_p x^j = 0 \tag{3.26}$$

$$c^i \partial_i (\underbrace{x^j \circ x^{-1}}_{\text{proj. oper.}})(x(p)) = 0 \tag{3.27}$$

$$c^i \delta^j{}_i = 0 c^j = 0. \tag{3.28}$$

  Thus we have shown that the tangent space coordinate basis is also linearly independent.

The expression $(x^i \circ x^{-1})$ is just the projection to the ith component, i.e.

$$(x^i \circ x^{-1})[(a_1, \ldots, a_{\dim M})] = a_i.$$

Since $\{(\partial/\partial x^i)_p\}_{i=1}^{\dim M}$ acts as a vector basis for T_pM, we must have

$$\dim T_pM = \dim M,$$

as we set out to prove. $\square$

3.5 *The Cotangent Space*

Given a the vector space T_pM, we know that we can immediately construct its dual space. The ***cotangent space*** of a manifold M, at $p \in M$ is just

$$T_p^*M \equiv (T_pM)^*, \tag{3.29}$$

i.e. the dual space of T_pM. As we know from our discussion of vector spaces, it must have the same dimension as T_pM.

The Gradient Operator d_p

We can also construct a coordinate basis for the cotangent space, called the ***cotangent coordinate basis*** by first defining the ***gradient operator at point*** p,

$$d_p : C^\infty(M, \mathbb{R}) \to T_p^*M \tag{3.30}$$

$$f \mapsto d_pf \tag{3.31}$$

where for any ***tangent vector*** $X \in T_pM$

$$(d_pf)(X) \equiv Xf. \tag{3.32}$$

When this operator acts on a function f, it returns a covector that, when combined with a vector X, returns the derivative of f along the direction of X, hence we call it the ***gradient operator***.

This is why the gradient is naturally considered a covector!

The Cotangent Coordinate Basis Covectors

We can use the gradient operator to construct a basis for the cotangent space T_p^*M. Given an arbitrary chart $(U_x, x) \in \mathcal{A}_M$, with $p \in U_x$, we can apply the gradient operator, d_p to the coordinate component functions x^i, to get

$$\{\underbrace{d_px^1}_{\in T_p^*M}, \ldots, d_px^{\dim M}\}. \tag{3.33}$$

This covector basis is called the ***coordinate cotangent basis***, or the ***chart-induced covector basis***.

For each index $i = 1 \dots \dim M$, we can apply the coordinate basis covector $d_p x^i$ to a coordinate basis vector $(\partial/\partial x^j)_p$,

$$(d_p x^i)\left(\frac{\partial}{\partial x^j}\right)_p = \left(\frac{\partial}{\partial x^j}\right)_p x^i \tag{3.34}$$

$$= \partial_j(\underbrace{x^i \circ x^{-1}}_{\text{proj. map}})(x(p)) \tag{3.35}$$

$$= \delta^i{}_j \tag{3.36}$$

Hence we see that the coordinate cotangent basis $\{d_p x^i\}$ is actually the dual/reciprocal basis to $\{(\partial/\partial x^i)_p\}$!

Now that we have a notions of a vector space T_pM, and a covector space T^*_pM at each point p, we can similarly construct a tensor space at p,

$$\mathcal{T}^r_s(T_pM) \equiv \{T | T : \underbrace{T^*_pM \times \dots \times T^*_pM}_{r\text{ times}} \times \underbrace{T_pM \times \dots T_pM}_{s\text{ times}} \xrightarrow{\text{lin}} \mathbb{R}\} \tag{3.37}$$

which has basis induced by coordinate map x,

$$\left\{\left(\frac{\partial}{\partial x^{i_1}}\right)_p \otimes \dots \otimes \left(\frac{\partial}{\partial x^{i_r}}\right)_p \otimes d_p x^{j_1} \otimes \dots \otimes d_p x^{j_s}\right\}. \tag{3.38}$$

3.6 Vector, Covector, and Tensor Fields on Manifolds

We can now take an arbitrary smooth manifold, examine an open set U_x with some chart, and using the chart coordinates, x, establish bases in the tangent and cotangent spaces for each point in U_x.

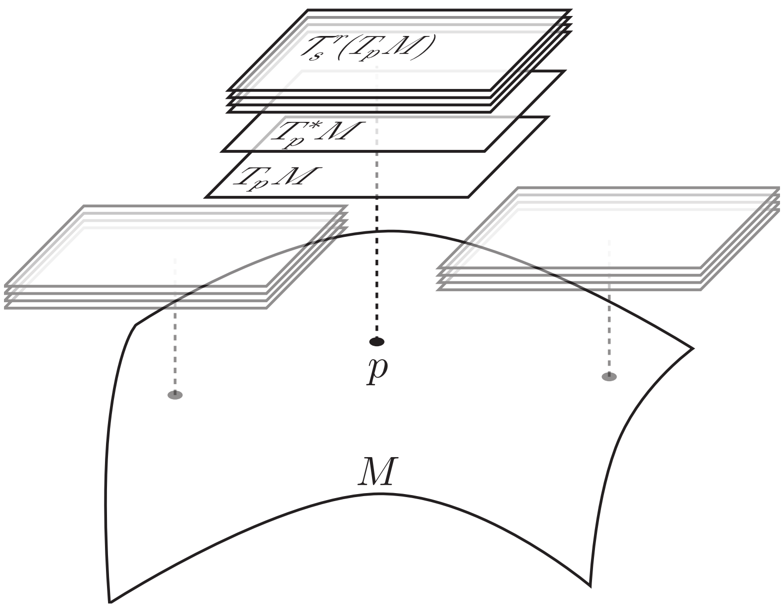


Figure 3.14: The tangent bundle, cotangent bundle, and tensor bundles on a manifold M are made up of all the tangent spaces, cotangent spaces, and tensor spaces attached at each point $p \in M$. The "fibres" of these associated tensor spaces are attached to each base point in M.

That is for each point p in the manifold, there is a copy of each of these tangent, cotangent and tensor spaces! They are completely

separate spaces and we have no a-priori way of knowing how to take e.g. a vector in one, and compare it with a vector in another.

To study the geometric properties of this collection of vector/tensor spaces formally, we would need to discuss the notion of a tangent bundle TM, and Fibre bundles, which are beyond the scope of this course. But you can imagine the vector/tensor spaces as fibres anchored to the manifold at each point.

Figure 3.15: Fibre bundles can be viewed as analogous to the collection of hairbrush fibres (tensor spaces) attached to the cylindrical handle (base manifold).

Vector fields can be thought of as "sections" of the bundle, i.e. choosing one element of the vector space for every single point in the manifold. Just as we defined ***smooth*** functions, we can similarly define smooth sections to give smooth vector/covector/tensor fields.

A ***vector field*** $\boldsymbol{X} : M \to TM$ is a mapping from the manifold M to the tangent bundle TM, defined by the set of all possible pairs of base points and vectors,

$$TM \equiv \{(p, V_p) | p \in M, V_p \in T_pM\},$$

such that for each $p \in M$,

$$\boldsymbol{X}(p) = (p, X_p) \tag{3.39}$$

assigns a single tangent vector,

$$X_p \in T_pM. \tag{3.40}$$

Formally, a tangent bundle is always defined with a projection operator $\pi : TM \to M$ defined by $\pi((p, X_p)) = p$ that relates an element in the tangent bundle to its root point in the base space. A vector field $\boldsymbol{X}$ is a ***section*** of the tangent bundle, i.e. it is a mapping from $M \to TM$ such that $\pi(X(p)) = p, \forall p \in M$. You don't need to worry about the details of bundles for this unit.

For any smooth $C^\infty(M, \mathbb{R})$ function $f : M \to \mathbb{R}$, we can define the action of a vector field $\boldsymbol{X}$ on f as

$$\boldsymbol{X}f : M \to \mathbb{R} \tag{3.41}$$

$$p \mapsto X_p f, \qquad \forall p \in M. \tag{3.42}$$

If $\boldsymbol{X}f \in C^\infty(M, \mathbb{R})$, then we say that the vector field $\boldsymbol{X}$ is ***a smooth vector field***.

In this way we can say that $\boldsymbol{X}$ also represents a mapping from $C^\infty(M, \mathbb{R}) \to C^\infty(M, \mathbb{R})$, since it takes a scalar function, and maps it to a directional derivative of the function at each point.

You may suspect that the vector fields on M make up a vector space, given our love of vector spaces. However, it turns out that vector fields are only ***almost*** a vector space. Vector fields are mathematical objects called ***modules***, which are like vector spaces, but instead of having associated algebraic fields as the scalars, they are associated with ***algebraic rings*** which do not require multiplicative inverses. For a vector field on M, the associated ring is the set of scalar valued functions $C^\infty(M, \mathbb{R})$ by which you can multiply each vector at each point. A function $f \in C^\infty(M, \mathbb{R})$ does ***not*** have a multiplicative inverse if $f(p) = 0$ at any point $p \in M$.

Similarly ***a covector field*** $\boldsymbol{\alpha}$ is a section of the cotangent bundle T^*M

$$\boldsymbol{\alpha} : M \to T^*M, \tag{3.43}$$

$$\text{where } T^*M \equiv \{(p, \alpha_p) \ | \ p \in M, \alpha_p \in T_p^*M\}, \tag{3.44}$$

which acts on a vector field $\boldsymbol{X}$ to produce the map

$$\boldsymbol{\alpha X} : M \to \mathbb{R} \tag{3.45}$$

$$p \mapsto \boldsymbol{\alpha}_p(X_p) \tag{3.46}$$

A smooth covector field is a covector field that produces a smooth $\mathbb{M} \to \mathbb{R}$ function if it acts on a smooth vector field.

Similarly, ***a smooth*** (r, s)***-tensor field***, is a tensor field that produces a smooth $\mathbb{M} \to \mathbb{R}$ function if it is fed r smooth covector fields, and s smooth vector fields.

3.7 (Psuedo-)Riemannian Manifolds

For a manifold M, we have a notion of vector fields, covector fields, and tensor fields that can exist attached to points on M. We may want an additional structure that lets us promote each tangent space into an ***inner product space*** by providing an inner product at each point on the manifold. Since this inner product is a linear map between two vectors and a number we recognize it as a $(0, 2)$ tensor. If we have such an inner product for each point on the manifold we have a $(0, 2)$ tensor field, which we call ***the metric***. If a metric is ***positive definite*** at all points, then the manifold is called ***a Reimannian manifold***. If the metric is merely ***non-degenerate*** at all points, then it is called ***a pseudo-Reimannian manifold***. Regular Eulerian space is particularly simple case of a Reimannian manifold. Spacetime is an example of a psuedo-Reimannian manifold. We call this inner product tensor a metric because it allows us to measure "lengths" of curves on the manifold.

We usually write the metric tensor $g = g_{\mu\nu} dx^\mu \otimes dx^\nu$ (for a particular choice of coordinates). The metric tensor can be used to "***lower the index***" of a vector to obtain the covector that is its metric dual, which has components $v_\mu = g_{\mu\nu} v^\nu$. Similarly the inverse metric $g^{-1} = g^{\mu\nu} \frac{\partial}{\partial x^\mu} \otimes \frac{\partial}{\partial x^\nu}$ can be used to "***raise the index***" of one-forms to their associated vectors. Note that we often omit the $^{-1}$ when writing the components of the inverse metric since the index placement makes it clear what we mean. The metric and inverse metric components are related by $g^{\mu\alpha} g_{\alpha\nu} = \delta^\mu{}_\nu$.

3.8 Pullbacks and Pushforwards

Consider two smooth manifolds M and N and a smooth map between them, $\phi : M \to N$.

If we have a function defined on the manifold N, $f : N \to \mathbb{R}$, how does this naturally map to the manifold M? Using the composition map.

The Pullback of a Function

The ***pullback*** of the function f by ϕ is the smooth function defined by

$$\phi^* f : M \to \mathbb{R}$$
$$(\phi^* f)(p) = [f \circ \phi](p) = f(\phi(p))$$

at each point $p \in M$.

$$\begin{tikzcd} & \mathbb{R} & \\ M \arrow[ur, "\phi^* f = f\circ\phi"] \arrow[rr, "\phi"] & & N \arrow[ul, "f"'] \end{tikzcd}$$

$$C^\infty(M, \mathbb{R}) \xleftarrow{\ \phi^*\ } C^\infty(N, \mathbb{R})$$

We call ϕ^* the pullback, since it pulls a function defined on the range of ϕ back to its domain. We can't push this function forward since we aren't guaranteed that the inverse map ϕ^{-1} exists and is well defined.

We can use this notion of the pullback of functions in order to construct an analogous notion for vectors, however, we will see that this must be defined in the ***opposite direction***.

The Pushforward of a Vector

Consider a vector $\boldsymbol{X}_p$ tangent to some point on a manifold $p \in M$. Remember that tangent vector $\boldsymbol{X}_p : C^\infty(M, \mathbb{R}) \to \mathbb{R}$ is a directional derivative operator at p, and can be applied to any function in $C^\infty(M, \mathbb{R})$.
Using the pullback of an arbitrary function $f \in C^\infty(N, \mathbb{R})$, we can then define the ***pushforward***, $(\phi_*)_p \boldsymbol{X}_p$ of the vector $\boldsymbol{X}_p$, such that,

$$\underbrace{(\phi_*)_p \boldsymbol{X}_p}_{\in T_{\phi(p)}N} f = \boldsymbol{X}_p \underbrace{\phi^* f}_{\in C^\infty(M,\mathbb{R})} = \boldsymbol{X}_p(f \circ \phi)$$

Let $\gamma : M \to \mathbb{R}$ be a curve such that $X_{\gamma,p} = \boldsymbol{X}_p$, with $\gamma(0) = p$. The pushforward can then be written as

$$\begin{aligned}(\phi_*)_p \boldsymbol{X}_p f &= X_{\gamma,p}(f \circ \phi) \\ &= (f \circ \phi \circ \gamma)'(0)\end{aligned}$$

What does this look like in terms of particular coordinate charts, i.e. $(U_x, \boldsymbol{x}) \in \mathcal{A}_M$ and $(\phi(U_x), \boldsymbol{y}) \in \mathcal{A}_N$?

$$\begin{aligned}(\phi_*)_p X_{\gamma,p} f &= (\underbrace{f \circ \phi \circ \gamma}_{\mathbb{R} \leftarrow N \leftarrow M \leftarrow \mathbb{R}})'(0) \\ &= (\underbrace{f \circ y^{-1}}_{\mathbb{R} \leftarrow \mathbb{R}^{\dim N}} \circ \underbrace{(y \circ \phi \circ \gamma}_{\mathbb{R}^{\dim N} \leftarrow \mathbb{R}}))'(0) \\ &= \partial_i (f \circ y^{-1})|_{y(\phi(p))} (y^i \circ \phi \circ \gamma)'(0) \\ &= \partial_i (f \circ y^{-1})|_{y(\phi(p))} (\underbrace{y^i \circ \phi \circ x^{-1}}_{\mathbb{R} \leftarrow \mathbb{R}^{\dim M}} \circ \underbrace{x \circ \gamma}_{\mathbb{R}^{\dim M} \leftarrow \mathbb{R}})'(0) \\ &= \left[\partial_j (y^i \circ \phi \circ x^{-1})|_{x(p)} \underbrace{(x^j \circ \gamma)'(0)}_{X_p^j} \right] \left(\frac{\partial}{\partial y^i} \right)_{\phi(p)} f\end{aligned}$$

where X_p^j is the jth component of the vector $\boldsymbol{X}_p$ using the $\boldsymbol{x}$-coordinate induced basis on T_pM. Let us assume that $\mathcal{B}_x$ and $\mathcal{B}_y$ are the coordinate induced bases in T_pM and $T_{\phi(p)}N$ respectively. Then we can write,

$$\boldsymbol{Y}_{\phi(p)} \equiv (\phi_*)_p \boldsymbol{X}_p = \underbrace{\partial_j (y^i \circ \phi \circ x^{-1})|_{x(p)} X_p^i}_{Y^i_{\phi(p)}} \underbrace{\left(\frac{\partial}{\partial y^i} \right)_{\phi(p)}}_{y\text{-coord basis}}$$

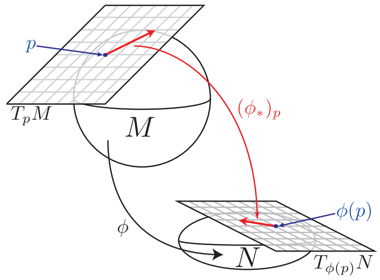


Figure 3.16: The pushfoward of a vector via a map $\phi : M \to N$ at point $p \in M$, is a map $(\phi_*)_p : T_pM \to T_{\phi(p)}N$. Notationally, it can also be written as $D\phi_p$, $\phi'(p)$ or $T_p\phi$. It can be thought of as a linear approximation of ϕ near p, and for particular sets of coordinates on each manifold, it is given by the Jacobian matrix of the map, i.e. the total derivative of ϕ at point p.

where $Y^i_{\phi(p)}$ is the ith component of $\boldsymbol{Y}_{\phi(p)}$ in the $\mathcal{B}_y$ basis. Writing this out pedantically, we see

$$[\boldsymbol{Y}_{\phi(p)}]^i_{\mathcal{B}_y} = [(\phi_*)_p]^i{}_j [\boldsymbol{X}_p]^j_{\mathcal{B}_x}.$$

Here we recognise $[(\phi_*)_p]^i{}_j \equiv \partial_j(y^i \circ \phi \circ x^{-1})|_{x(p)}$ as the $(\dim N \times \dim M)$ Jacobian matrix of partial derivatives, which represents the linear transformation $(\phi_*)_p : T_pM \overset{\text{lin}}{\longrightarrow} T_{\phi(p)}N$.

The Pushforward of a Vector Field

Consider a vector field $\boldsymbol{X}$ on M, (that is a section of the tangent bundle TM), and an injective (1-to-1) smooth map $\phi : M \to N$.
We can then define $\phi_*\boldsymbol{X}$, ***the pushforward of the vector field*** $\boldsymbol{X}$, to be the pushforward at each point $p \in M$,

$$(\phi_*\boldsymbol{X})(p) = (\phi(p), (\phi_*)_p\boldsymbol{X}_p) \in T_{\phi(M)}N$$

$$\begin{tikzcd} M \arrow[r, "\phi"] \arrow[d, "\boldsymbol{X}"'] & N \arrow[d, "\phi_*\boldsymbol{X}"] \\ TM \arrow[r, "\phi_*"] & T_{\phi(M)}N \end{tikzcd}$$

where $T_{\phi(M)}N$ denotes the tangent bundle of N, with base space restricted to the image $\phi(M)$.

The lack of a guaranteed inverse for ϕ prevents us from being able to push functions forward, and therefore pull vectors back. (If the ϕ is actually a diffeomorphism, these "opposite" mappings can be defined using the pullbacks and pushforwards of the inverse map ϕ^{-1}.)

Note that if the map ϕ is ***not 1-to-1***, i.e. if more than one point in M is mapped to any one point in N, then we ***cannot*** really define the pushforward of a vector field to an image set that includes this point, since it would be multi-valued there.

Similarly, we can then go ***back again*** to define the ***pullback of a covector field***.

The Pullback of a Covector Field

Let's consider a covector field $\boldsymbol{\alpha} : N \to T^*N$ defined on N, such that at each point $p \in M$, $\phi(p) \in N$

$$\boldsymbol{\alpha}(\phi(p)) = (\phi(p), \boldsymbol{\alpha}_{\phi(p)}),$$
$$\text{where } \boldsymbol{\alpha}_{\phi(p)} \in T^*_{\phi(p)}N$$

is a covector defined in the cotangent space at $\phi(p)$. We can then define, $\phi^*\boldsymbol{\alpha}$, ***the pullback of the covector field*** $\boldsymbol{\alpha}$, such that $\forall p \in M$,

$$(\phi^*\boldsymbol{\alpha})(p) = (p, (\phi^*\boldsymbol{\alpha})_p)$$

In contrast to the pushforward of a vector field (which fails for non-injective maps), the pullback of a covector field can always be used, since it doesn't depend on the existence of a vector field, only arbitrary vectors at each point, and there is no issue if a covector from some point $\phi(p_1) = \phi(p_2)$ is pulled back to two different points p_1 and p_2.

where

$$(\phi^* \boldsymbol{\alpha})_p \boldsymbol{X}_p = \boldsymbol{\alpha}_{\phi(p)}(\phi_* \boldsymbol{X}_p), \qquad \forall \boldsymbol{X}_p \in T_p M$$

That is, the scalar value returned by the pulled-back one-form, $(\phi^* \alpha)_p$ acting on $\boldsymbol{X}_p$ is the same as the scalar value of the original one-form acting on the pushfoward of $\boldsymbol{X}_p$.

$$\begin{tikzcd} M \arrow[r, "\phi"] \arrow[d, "\phi^* \boldsymbol{\alpha}"'] & N \arrow[d, "\boldsymbol{\alpha}"] \\ T^*M & T^*_{\phi(M)} N \arrow[l, "\phi^*"] \end{tikzcd}$$

What does this look like in terms of coordinates? For coordinate charts $(U, x) \in \mathcal{A}_M$ and $(\phi(U), y) \in \mathcal{A}_N$, with $p \in U$, we can

$$\begin{aligned} \underbrace{(\phi^* \boldsymbol{\alpha})_p \boldsymbol{X}_p}_{\text{on } M} &= \underbrace{\boldsymbol{\alpha}_{\phi(p)}(\phi_* \boldsymbol{X}_p)}_{\text{on } N} \\ [(\phi^* \boldsymbol{\alpha})_p]_i^{\mathcal{B}_x^*} [\boldsymbol{X}_p]^i_{\mathcal{B}_x} &= [\boldsymbol{\alpha}_{\phi(p)}]_j^{\mathcal{B}_y^*} [(\phi_*)_p]^j{}_i [\boldsymbol{X}_p]^i_{\mathcal{B}_x} \\ \Rightarrow [(\phi^* \boldsymbol{\alpha})_p]_i^{\mathcal{B}_x^*} &= [(\phi_*)_p]^j{}_i [\boldsymbol{\alpha}_{\phi(p)}]_j^{\mathcal{B}_y^*} \end{aligned}$$

where $[(\phi_*)_p]^j{}_i \equiv \partial_i (y^j \circ \phi \circ x^{-1})|_{x(p)}$ is the Jacobian matrix from the pushforward.

The vector pushforward and the covector pullback rely on using the Jacobian in opposite senses.

If $\tilde{\boldsymbol{\alpha}}$ is a one-form in $T_{\phi(p)} N$, and $\boldsymbol{\alpha}$ is it's pullback into $T_p M$, their components are related by

$$[(\phi_*)_p]^i_j [\tilde{\boldsymbol{\alpha}}]_i^{\mathcal{B}_y^*} = [\boldsymbol{\alpha}]_j^{\mathcal{B}_x^*}.$$

If $\boldsymbol{X}$ is a vector in $T_p M$ and $\tilde{\boldsymbol{X}}$ is its pushfoward into $T_{\phi(p)} N$, then the components are related by

$$[\tilde{\boldsymbol{X}}]^i_{\mathcal{B}_y} = [(\phi_*)_p]^i_j [\boldsymbol{X}]^j_{\mathcal{B}_x}$$

Using the above it is straightforward to show that if ϕ is a diffeomorphism, then the Jacobians $[(\phi_*)_p]^i{}_j$ and $[(\phi_*^{-1})_{\phi(p)}]^i{}_j$ must related by the matrix inverse.

The Pullback of a $(0, r)$ Tensor Field

The pullback of a $(0, r)$ tensor field $T \in \mathcal{T}_r^0 TN$ can be defined similarly, at each point $p \in M$,

$$\underbrace{(\phi^* T)_p}_{\in \mathcal{T}_r^0 T_p M}(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_r) = T_{\phi(p)}((\phi_*)_p \boldsymbol{X}_1, \ldots, (\phi_*)_p \boldsymbol{X}_r)$$

for arbitrary vectors $\boldsymbol{X}_i \in T_p M$.
In components this has

$$[(\phi^* T)_p]_{i_1 \ldots i_r} = [(\phi_*)_p]^{j_1}{}_{i_1} \ldots [(\phi_*)_p]^{j_r}{}_{i_r} [T_{\phi(p)}]_{j_1 \ldots j_r}$$

The Pushforward of a $(r, 0)$ Tensor Field

The pushfoward of a tensor $S_p \in \mathcal{T}_0^r T_p M$ follows as

$$\underbrace{[(\phi_*)_p S_p]}_{\in \mathcal{T}_0^r T_{\phi(p)} N}(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r) = S_p((\phi^*)_{\phi(p)} \boldsymbol{\alpha}_1, \ldots (\phi^*)_{\phi(p)} \boldsymbol{\alpha}_r)$$

for arbitrary 1-forms $\boldsymbol{\alpha}_i \in T_{\phi(p)} N$. In components we can write this as

$$[(\phi_*)_p S_p]^{j_1 \dots j_r} = [(\phi_*)_p]^{j_1}{}_{i_1} \dots [(\phi_*)_p]^{j_r}{}_{i_r} S_p^{i_1 \dots i_r}$$

As before if ϕ is a $1 - to - 1$ map, then this can be extended to the pushforward $(r,0)$ tensor field in TM

3.8.1 *Coordinate Transformations*

Let's take a look back at transformations between different coordinates for vectors and tensors. Previously, we considered the transformation of bases and coordinates of a single vector space V using elements of the endomorphisms of V, $\mathcal{L}(V) = \mathrm{End}(V)$. Now we are equipped to see what happens in terms of our coordinate basis at each of the tangent (tensor) spaces attached to our manifold.

Consider a m-dimensional smooth manifold $(M, \mathcal{O}_M, \mathcal{A}_M)$ with coordinate charts $(U, \boldsymbol{x}) \in \mathcal{A}_M$ and $(U, \boldsymbol{y}) \in \mathcal{A}_M$, and a function $f \in C^\infty(M \to \mathbb{R}))$

$$\begin{tikzcd}
 & \mathbb{R}^m \arrow[dl, bend right, "f\circ \boldsymbol{x}^{-1}"'] \arrow[dd, bend left=60, "\boldsymbol{y}\circ \boldsymbol{x}^{-1}"] \\
\mathbb{R} & U \subseteq M \arrow[l, "f"'] \arrow[u, "\boldsymbol{x}"] \arrow[d, "\boldsymbol{y}"'] \\
 & \mathbb{R}^m \arrow[ul, bend left, "f\circ \boldsymbol{y}^{-1}"]
\end{tikzcd}$$

A $\boldsymbol{x}$-coordinate basis vector for T_pM acting as the directional derivative of f is given by

$$\begin{aligned}
\left(\frac{\partial}{\partial x^\mu}\right)_p f &= \partial_\mu (f \circ \boldsymbol{x}^{-1})(\boldsymbol{x}(p)) \\
&= \partial_\mu (\underbrace{f \circ \boldsymbol{y}^{-1}}_{\mathbb{R}\leftarrow\mathbb{R}^m} \circ \underbrace{\boldsymbol{y} \circ \boldsymbol{x}^{-1}}_{\mathbb{R}^m\leftarrow\mathbb{R}^m})(\boldsymbol{x}(p)) \\
&= \partial_{\mu'}(f \circ \boldsymbol{y}^{-1})|_{\boldsymbol{y}(p)} \cdot \partial_\mu (y^{\mu'} \circ \boldsymbol{x}^{-1})|_{x(p)} \\
&= \underbrace{\left(\frac{\partial y^{\mu'}}{\partial x^\mu}\right)_p}_{\partial_\mu (y^{\mu'} \circ \boldsymbol{x}^{-1})(x(p))} \left(\frac{\partial}{\partial y^{\mu'}}\right)_p f.
\end{aligned} \tag{3.47}$$

Note that not all linear transformations are coordinate transformations. Only transformations of the form $[\Lambda]^{\mu'}{}_\nu = \partial y^{\mu'} / \partial x^\nu$ qualify.

The above defines how the basis vector field $(\partial/\partial x^\mu)$ transforms to the basis vector field $(\partial/\partial y^{\mu'})$.

So a ***vector field*** $V = V^\alpha (\partial/\partial x^\alpha)$ has components that transforms as

$$V^{\alpha'} = \frac{\partial y^{\alpha'}}{\partial x^\alpha} V^\alpha .$$

Similarly, 1-forms and tensors transform according to the index placement we have chosen for $\boldsymbol{\alpha} \in T*M$, the components and basis covector fields transform as

$$\alpha_{\mu'} = \frac{\partial x^{\mu}}{\partial y^{\mu'}} \alpha_{\mu} \tag{3.48}$$

$$dy^{\mu'} = \frac{\partial y^{\mu'}}{\partial x^{\mu}} dx^{\mu}. \tag{3.49}$$

Tensor Densities and the Jacobian determinant

We haven't had time to talk about integration on curved manifolds, but integration over a volume occurs via integration of a volume-form, e.g. $dV = dx^1 \wedge dx^2 \wedge dx^3$, in 3 dimensions, where as we know from the worksheet, the wedge product $\wedge$ is the antisymmetric tensor product. When this volume form is mapped to a different coordinate tensor basis, e.g. $dV' = dy^1 \wedge dy^2 \wedge dy^3$ it gains a factor of the determinant of the Jacobian, $|\det(\partial y^{i'}/\partial x^j)|$, so that $dV = |\det(\partial x^{i'}/\partial x^j)| dV'$, due to the way coordinate basis 1-forms transform and the antisymmetry of the wedge product combining to yield the determinant. If a tensor is something that should be integrated over a volume, you must treat it as a ***tensor density***, and this change in the volume form must be taken into account when it transforms. Generalising this to things that may be integrated W times, a (p, q) ***tensor density*** $\mathcal{T}$ of weight W transforms via,

$$\mathcal{T}^{i'_1 \ldots i'_p}{}_{j'_1 \ldots j'_q} = \left| \det\left(\frac{\partial y^{i'}}{\partial x^j} \right) \right|^W \frac{\partial y^{i'_1}}{\partial x^{i_1}} \cdots \frac{\partial y^{i'_p}}{\partial x^{i_p}} \frac{\partial x^{j_1}}{\partial y^{j'_1}} \cdots \frac{\partial x^{j_q}}{\partial y^{j'_q}} \mathcal{T}^{i_1 \ldots i_p}{}_{j_1 \ldots j_q}. \tag{3.50}$$

It is straightforward to show (using the expression for how the metric tensor transforms, and properties of determinants) that the Jacobian determinant can be conveniently written as,

$$\left| \det\left(\frac{\partial y^{\mu'}}{\partial x^{\nu}} \right) \right| = \sqrt{-\det(g_{\mu\nu})}, \tag{3.51}$$

which is often written as $\sqrt{-g}$.

We don't have time to cover this in this unit, however I'm including it for your reference.

4
Physics on Curved Manifolds

We are now ready to apply what we have learned about differential geometry to physics ***on*** curved spacetimes. Here we will assume you are familiar with special relativity, which takes place on a 4-dimensional psuedo-Riemannian manifold with metric given by the Minkowski metric at all points.

4.1 Local Flatness and the Einstein Equivalence Principle (EEP)

The first physical principle we want to start with is the ***Weak Equivalence Principle***. This dates back to Galileo and Newton and states that the ***inertial mass*** of an object is the same as the ***gravitational mass***, as we discussed at the start of this class.

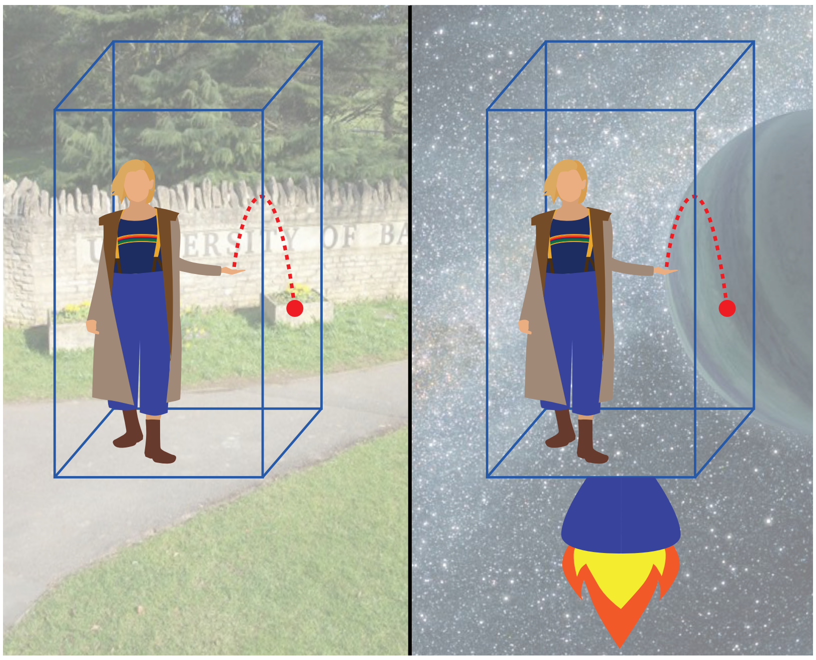


Figure 4.1: The Weak Equivalence Principle says that a Doctor of Physics cannot perform any experiment with the motion of test particles to tell if her police box is uniformly accelerating or in a gravitational field. The Einstein Equivalence Principle says that she cannot perform ***any*** (non-gravitational) local experiment that can tell the difference.

As a result of the universality of the Weak Equivalence Principle, we can formulate a thought expierment. We consider a Doctor of Physics in a small (police) box, unable to see out. The Doctor can do experiments inside the box, measuring the motion of test particles. The results of these experiments would be different if the box was sitting on Earth, or in orbit. However, it would also be different if the box was being accelerated. The point of the Weak Equivalence Principle is that there is ***no way to distinguish between gravitational effects from equivalent accelerations using the laws of motion***: if the box was accelerated upwards at 9.8 $\mathrm{m/s^2}$, for example, any mechanical experiment done inside the box would be the same as if it were sitting on the surface of the Earth. (This would not be the case if the Doctor was trying to e.g. disentangle electromagnetic interactions from accelerations. The "charge" in this case is different to the inertial mass, and thus things with different charges but the same mass would accelerate differently.)

The ***Strong Equivalence Principle*** (SEP) is the notion that any local gravitational or non-gravitational experiment behaves the same in a freely falling laboratory as in a flat spacetime. GR actually obeys this stronger equivalence principle, but some theories of gravity obey the EEP, but not the SEP.

Of course this universality and indistinguishability is only over small regions. If the box was big enough, there would be gravitational effects that vary over the box: there would be essentially tidal effects that would be different from uniform accelerations. So we must restrict ourselves to "small enough" regions of spacetime, just as we have local homeomorphisms from a manifold to Euclidean space, if we zoom in to a small enough region.

Einstein decided it was reasonable to strengthen the Weak Equivalence Principle into something that applied to ***any kind of experiment***. The Einstein equivalence principle (EEP), states that ***any non-gravitational physics in a freely falling laboratory appears* locally *the same as physics in flat spacetime*** (i.e. in special relativity).

Local Flatness

One of the reasons the EEP works is that in General Relativity the spacetime manifold is ***locally flat***, meaning that there exists ***some*** local coordinate transformation that takes the metric $g_{\mu\nu}$ to the Minkowski metric of flat space,

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{4.1}$$

in a small open neighbourhood around a particular point on the spacetime manifold $p \in M$. More explicitly we can show that we can always choose a Lorentz frame with a set of ***Locally Flat Coordinates*** at any point p, such that $g_{\mu'\nu'}(p) = \eta_{\mu'\nu'}$ and $\partial g_{\mu'\nu'}/\partial x^{\gamma'}|_p = 0$.

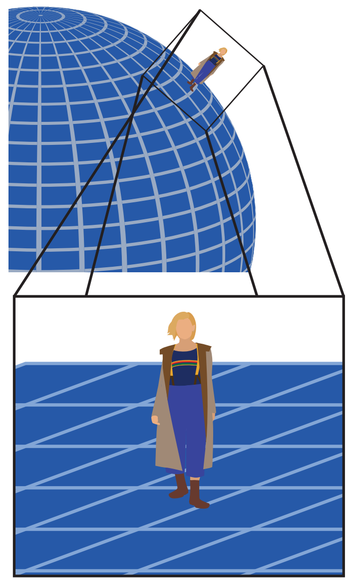


Figure 4.2: Spacetime is locally flat, i.e. if you zoom in far enough it looks like Minkowski space. This is analogous to zooming in far enough on a sphere so that it locally looks like a simple plane.

Counting free parameters

Since the $(0,2)$ tensor $g_{\mu\nu}$ transforms via a coordinate change as

$$g_{\mu'\nu'} = \Lambda^{\mu}{}_{\mu'}\Lambda^{\nu}{}_{\nu'}g_{\mu\nu}$$

we can compare this with an arbitrary coordinate transformation $\Lambda^{\mu}{}_{\mu'} = \partial x^{\mu}/\partial x^{\mu'}$ (where the $'$ index denotes a different set of coordinates). Taylor expanding the above equation at point p, we see that the metric and its derivatives in the x' coordinates are given by

$$g_{\mu'\nu'}|_p = \left[\Lambda^{\mu}{}_{\mu'}\Lambda^{\nu}{}_{\nu'}g_{\mu\nu}\right]_p$$

$$\left.\frac{\partial g_{\mu'\nu'}}{\partial x^{\gamma'}}\right|_p = \left[\Lambda^{\mu}{}_{\mu'}\Lambda^{\nu}{}_{\mu'}\frac{\partial g_{\mu\nu}}{\partial x^{\gamma'}} + 2\frac{\partial \Lambda^{\mu}{}_{\mu'}}{\partial x^{\gamma'}}\Lambda^{\nu}{}_{\nu'}g_{\mu\nu}\right]_p$$

$$\left.\frac{\partial^2 g_{\mu'\nu'}}{\partial x^{\gamma'}\partial x^{\delta'}}\right|_p = \Bigg[\Lambda^{\mu}{}_{\mu'}\Lambda^{\nu}{}_{\mu'}\frac{\partial^2 g_{\mu\nu}}{\partial x^{\gamma'}\partial x^{\delta'}} + 4\frac{\partial \Lambda^{\mu}{}_{\mu'}}{\partial x^{\gamma'}}\Lambda^{\nu}{}_{\nu'}\frac{\partial g_{\mu\nu}}{\partial x^{\delta'}}$$
$$+ 2\frac{\partial \Lambda^{\mu}{}_{\mu'}}{\partial x^{\gamma'}}\frac{\partial \Lambda^{\nu}{}_{\nu'}}{\partial x^{\delta'}}g_{\mu\nu} + 2\frac{\partial^2 \Lambda^{\mu}{}_{\mu'}}{\partial x^{\gamma'}\partial x^{\delta'}}\Lambda^{\nu}{}_{\nu'}g_{\mu\nu}\Bigg]_p$$

where we've used the fact that both the metrics and the mixed partial derivatives must be symmetric under interchange of indices.
Do we have enough freedom in the coordinate change to choose ***Locally Flat Coordinates*** at point p? We can check this by counting the free parameters in the metric, the coordinate transformation, and their derivatives.

- The tensor $g_{\mu\nu}$ has 10 independent components (due to symmetry) while the arbitrary transformation $\Lambda^{\mu}{}_{\mu'} \equiv \frac{\partial x^{\mu}}{\partial x^{\mu'}}$ has 16 free parameters. Thus in order for $g_{\mu'\nu'}|_p = \eta_{\mu'\nu'}$, we only need to constrain 10 of those 16, leaving 6 free (which corresponds to the 6 degrees of freedom of rotations/Lorentz boosts).

- Fixing these above values, the first derivative of the transformation $\frac{\partial \Lambda^{\mu}{}_{\mu'}}{\partial x^{\gamma'}} = \frac{\partial^2 x^{\mu}}{\partial x^{\mu'}\partial x^{\gamma'}}$ has $4 \times 10 = 40$ free parameters (because of symmetry of the mixed partial derivatives), while $\frac{\partial g_{\mu\nu}}{\partial x^{\gamma'}}$ also has $10 \times 4 = 40$ free parameters, which gives us just enough freedom to eliminate the first derivative of the metric at point p in the Locally Flat Coordinates.

- Fixing all the values above again, we see that in order to eliminate $\frac{\partial^2 g_{\mu\nu}}{\partial x^{\gamma'}\partial x^{\delta'}}$ we would need $10 \times 10 = 100$ free parameters in $\frac{\partial^2 \Lambda^{\mu}{}_{\mu'}}{\partial x^{\gamma'}\partial x^{\delta'}} = \frac{\partial^3 x^{\mu}}{\partial x^{\mu'}\partial x^{\gamma'}\partial x^{\delta'}}$, of which we only have $4 \times 20 = 80$. These 20 "missing" terms means we cannot eliminate the 2nd derivative of the metric by changing frames. In fact these terms describe the ***curvature*** of spacetime!

$$\begin{pmatrix} g_{00} & g_{01} & g_{02} & g_{03} \\ g_{01} & g_{11} & g_{12} & g_{13} \\ g_{02} & g_{12} & g_{22} & g_{23} \\ g_{03} & g_{13} & g_{23} & g_{33} \end{pmatrix} \; 4 + 6 = 10$$

Figure 4.3: A symmetric rank 2 tensor in 4D spacetime has 10 free parameters.

The symmetry of partial derivatives means that we don't have 4^4 independent components of $\frac{\partial^3 x^{\mu}}{\partial x^{\mu'}\partial x^{\gamma'}\partial x^{\delta'}}$. The 3 lower indices can either be

- ***All The Same:*** 4 possible choices for this index.
- ***Two the Same:*** $\binom{4}{1} \times \binom{3}{1} = 4 \times 3 = 12$ possible choices for the indices
- ***All different:*** $\binom{4}{3} = 4$ possible choices for the indices

So there are $4 + 12 + 4 = 20$ possible choices for the lower indices. There are 4 possible choices of the upper index, giving us $20 \times 4 = 80$ free components.

How can we find a set of ***Locally Flat Coordinates***? The EEP tells us that a freely falling physicist in an elevator/police box can't tell ***locally*** that they aren't just in the flat space of Special

Relativity. This means that we could always use the coordinate system (of cartesian rulers and a clock) they carry around with them as the Locally Flat Coordinates. While this coordinate grid is not necessarily locally flat everywhere else, along the curve she traces out as her police box freely falls they are indeed the Locally Flat Coordinates. Thus at any point p, we can just adopt the local (cartesian) coordinate system of a freely falling observer passing that point.

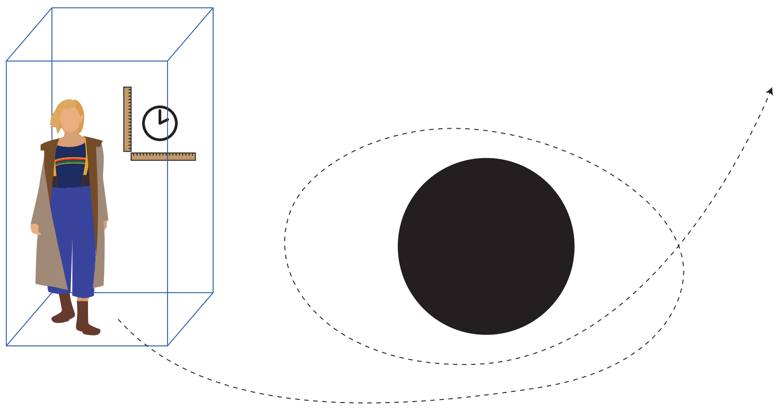


Figure 4.4: The Locally Flat Coordinates can be chosen by using the rulers and clock carried by a freely falling Doctor of Physics.

The Locally Flat Coordinates we have defined here are often called ***Fermi Normal Coordinates***. One can also construct ***Riemann Normal Coordinates*** by using geodesics launched by the Doctor of Physics in 4 perpendicular directions to define coordinates at each point.

The Geodesic Equation from the EEP

Let's consider a freely falling astronaut whose path on the Manifold is described by the curve $\gamma(\tau) : \mathbb{R} \to M$, where the parameter τ is the proper time measured by the astronaut. Let the Locally Flat Coordinates of the spacetime traveller be $\xi : M \to \mathbb{R}^4$. We will contrast this by some arbitrary coordinate chart $x : M \to \mathbb{R}^4$ that covers the area we are interested in.

In the Locally Flat Coordinates of the astronaut she feels no acceleration, i.e. her position is such that

$$(\xi^\mu \circ \gamma)''(\tau) = 0$$

$$\begin{tikzcd}
 & \mathbb{R}^m \arrow[dd, bend left, "\xi\circ x^{-1}"] \\
\mathbb{R} \arrow[ur, bend left, "x\circ\gamma"] \arrow[r, "\gamma"] \arrow[dr, bend right, "\xi\circ\gamma"'] & M \arrow[u, "x"'] \arrow[d, "\xi"] \\
 & \mathbb{R}^m
\end{tikzcd}$$

We can insert the identity $x^{-1} \circ x$,

$$(\xi^\mu \circ x^{-1} \circ x \circ \gamma)''(\tau) = 0$$

and apply the multi-d chain rule twice to get

$$\underbrace{(x^\nu \circ \gamma)''}_{\equiv \frac{d^2 x^\nu}{d\tau^2}\big|_{x(\gamma)}} \cdot \underbrace{\partial_\nu(\xi^\mu \circ x^{-1})|_{x(\gamma)}}_{\equiv \frac{\partial \xi^\mu}{\partial x^\nu}\big|_\gamma}$$

$$+ \underbrace{(x^\nu \circ \gamma)'}_{\equiv \frac{dx^\nu}{d\tau}\big|_{x(\gamma)}} \cdot \underbrace{\partial_\nu \partial_\lambda (\xi^\mu \circ x^{-1})|_{x(\gamma)}}_{\equiv \frac{\partial^2 \xi^\mu}{\partial x^\nu \partial x^\lambda}\big|_\gamma} \cdot \underbrace{(x^\lambda \circ \gamma)'}_{\equiv \frac{dx^\lambda}{d\tau}\big|_{x(\gamma)}} = 0$$

Multiplying both sides by $\partial x^\sigma / \partial \xi^\mu$ we get

$$\left[\frac{d^2 x^\sigma}{d^2\tau^2} + \underbrace{\frac{\partial x^\sigma}{\partial \xi^\mu}\frac{\partial^2 \xi^\mu}{\partial x^\nu \partial x^\lambda}}_{\Gamma^\sigma_{\nu\lambda}} \cdot \frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau}\right]_{x(\gamma)} = 0, \tag{4.2}$$

which is the usual geodesic equation, noting that $\Gamma^\sigma_{\nu\lambda}$ is the Christoffel connection we will encounter later.
Note that the geodesic equation involves only 2nd derivatives of the Locally Flat Coordinates ξ. We know that curvature only shows up in the 3rd derivative of a coordinate (i.e. the 2nd derivative of the coordinate transformation). This tells us that the curvature is captured not by the geodesic equation, but the geodesic ***deviation*** equation, which tells us how nearby geodesics diverge.

Note that symmetry of the mixed partial in the expression for the connection shows that it is in fact the only the ***torsion free*** part of the connection that contributes to the geodesic. For a more general connection we would identify this term as the torsion free part, $\Gamma^\sigma_{(\nu\lambda)}$. We will discuss what this means later in the notes.

4.2 *Connections and Covariant Derivatives*

For C^1 functions on a manifold (e.g. $f : M \to \mathbb{R}$), the derivative as we move along a curve on a manifold is well defined. However, if we want to take ***derivatives of vector fields, or tensor fields*** in general, things get a bit more complicated. Fundamentally this is because the tangent spaces (and their associated tensor spaces) at different points are completely separate from one another.

In order to take a derivative in the traditional sense, we need to be able to compare a vector from one tangent space with a vector in a ***nearby*** tangent space. However, we haven't yet constructed a way to do this on these disconnected vector spaces[1].

A manifestation of this issue is the fact that ***partial derivatives of vector or tensor field components do not transform like tensors*** and thus ***partial derivatives of tensors are not tensors (i.e. they are coordinate dependent quantities)***.

Consider the vector field $\boldsymbol{V}$ and two different coordinate systems x and x', taking the partial derivatives of the components

$$V^{\nu'} = \Lambda^{\nu'}{}_\nu V^\nu = \frac{\partial x^{\nu'}}{\partial x^\nu} V^\nu \tag{4.3}$$

$$\frac{\partial V^{\nu'}}{\partial x^{\mu'}} = \frac{\partial}{\partial x^{\mu'}}\left(\frac{\partial x^{\nu'}}{\partial x^\nu} V^\nu\right)$$

$$= \frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial^2 x^{\nu'}}{\partial x^\nu \partial x^\mu} V^\nu + \frac{\partial x^{\nu'}}{\partial x_\nu}\frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial V^\nu}{\partial x^\mu}$$

$$\frac{\partial}{\partial x^{\mu'}} V^{\nu'} = \left[\frac{\partial x^{\nu'}}{\partial x^\nu}\frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial}{\partial x^\mu} + \frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial^2 x^{\nu'}}{\partial x^\mu \partial x^\nu}\right] V^\nu \tag{4.4}$$

The second term in the square brackets prevents the partial derivative from transforming like a tensor.

To "fix" this we can introduce a non-tensor linear transformation, ***the connection***, $\Gamma^\mu_{\nu\lambda}$, which transforms in exactly the right way to

[1] One of the ways we could do this is to use another vector field as a generator of a flow along which we map a vector or tensor field using the pushforward. This is called ***Lie Dragging***, and is analogous to letting yourself be carried by a current. This process uses the ***exponential map*** at point p, $\exp_p : T_pM \to M$, which takes a vector field $\boldsymbol{X} \mapsto \gamma_{X,p}(1)$ mapping it to a point in the manifold one unit along the curve $\gamma_{X,p}(t)$ that passes through point p and is tangent to the vector field $\boldsymbol{X}$ at each point it flows through. We can then define the ***Lie Derivative*** of tensor T along X at p, as

$$(\mathcal{L}_X T)_p = \left.\frac{d}{dt}\right|_{t=0} ((\phi_t)^*_p T_{\phi_t(p)})$$

where $t \in \mathbb{R}$ and $(\phi_t)_p = \exp_p(tX) = \gamma_{X,p}(t)$, is a one-parameter family of diffeomorphisms defined by the exponential map.

Unfortunately we don't have time to cover the Lie derivative, but it is one of the 3 important derivatives in differential geometry (the Lie derivative, $\mathcal{L}_X$, the exterior derivative, d, and the covariant derivative, ∇).

cancel out the misbehaving term. Once this connection is ***specified*** for our manifold, we have a way to link our tangent spaces together (this is why it is called the connection).

The Covariant Derivative of a Vector Field

We can define the ***covariant derivative of a vector field V*** as

$$\nabla_\mu V^\nu \equiv \frac{\partial}{\partial x^\mu} V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda \tag{4.5}$$

We need this connection to obey the non-tensorial transformation law

$$\Gamma^{\nu'}_{\mu'\lambda'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\lambda}{\partial x^{\lambda'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \Gamma^\nu_{\mu\lambda} - \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\lambda}{\partial x^{\lambda'}} \frac{\partial^2 x^{\nu'}}{\partial x^\mu \partial x^\lambda}. \tag{4.6}$$

such that (you should be able to check this yourselves)

$$\nabla_{\mu'} V^{\nu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \nabla_\mu V^\nu, \tag{4.7}$$

transforms as a tensor.

In coordinate-free notation, we can define the covariant derivative in terms of the directional derivative of vector field $\boldsymbol{W}$ in the direction of vector field $\boldsymbol{V}$, $(D_{\boldsymbol{V}} \boldsymbol{W})_\varphi \equiv (V^\mu \partial_\mu \boldsymbol{W})_\varphi$, and how it changes for a particular coordinate chart map $\varphi : M \subseteq U_\varphi \to \mathbb{R}^{\dim M}$ that covers the domain of interest. The covariant derivative of $\boldsymbol{W}$ in the direction of $\boldsymbol{V}$ is given by

$$\nabla_{\boldsymbol{V}} \boldsymbol{W} = (D_{\boldsymbol{V}} \boldsymbol{W})_\varphi + \Gamma[\varphi](\boldsymbol{V}, \boldsymbol{W})$$

where the connection, $\Gamma[\varphi](\boldsymbol{V}, \boldsymbol{W})$, is bilinear in the vector arguments, $\boldsymbol{V}$ and $\boldsymbol{W}$, but depends on the choice of coordinate chart φ in a non-tensorial way, i.e. it depends not just on the first derivative of a transition map, but also the second derivative. This is chosen so that the resulting covariant derivative does ***not*** depend on the choice of φ.

Note that while the object $\Gamma[\cdot](\cdot,\cdot)$ is not a tensor, the object $\Gamma[\varphi](\cdot,\cdot)$, for some fixed chart map φ, is a perfectly good tensor.

In general we want a covariant derivative to have several properties. A covariant derivative needs to:

- Be ***linear***, such that

$$\nabla(T + S) = \nabla T + \nabla S \tag{4.8}$$

- It should also obey the ***Leibnitz rule*** (i.e. the product rule for a wider definition of multiplication), such that

$$\nabla(T \otimes S) = (\nabla T) \otimes S + T \otimes (\nabla S), \tag{4.9}$$

  where T and S are arbitrary rank tensors.

- The covariant derivative takes any (p, q) tensor field, and maps it to a $(p, q + 1)$ tensor field.

- We also want a ***covariant derivative to commute with index contractions***,

$$\nabla_\mu (T^\lambda{}_{\lambda\rho}) = (\nabla T)_\mu{}^\lambda{}_{\lambda\rho}. \tag{4.10}$$

- We want the ***covariant derivative of a scalar field, ϕ to just be the regular partial derivative***,

$$\nabla_\mu \phi = \frac{\partial \phi}{\partial x^\mu}. \tag{4.11}$$

A manifold can be equipped with ***any*** connection that satisfies these properties, in fact one can view the connection as something that must be separately specified in order to compare vectors in different tangent spaces. In GR we typically choose the Levi-Civita connection, which is uniquely determined when we insist on a connection that satisfies certain convenient properties.

The Covariant Derivative of Covectors

With the above properties, and the covariant derivative of vectors defined by Equation (4.5), we can show that the covariant derivatives of a one-form ω must be,

$$\nabla_\mu \omega_\nu = \frac{\partial}{\partial x^\mu}\omega_\nu - \Gamma^\lambda_{\mu\nu}\omega_\lambda. \tag{4.12}$$

The Covariant Derivative of Tensors

With the covariant derivatives of vectors and one forms we can use the Leibniz rule to construct the covariant derivative of a general tensor

$$\begin{aligned}\nabla_\sigma T^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_q} \equiv\ & \partial_\sigma T^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_q} \\ & + \Gamma^{\mu_1}_{\sigma\lambda} T^{\lambda\mu_2\ldots\mu_p}{}_{\nu_1\ldots\nu_q} + \Gamma^{\mu_2}_{\sigma\lambda} T^{\mu_1\lambda\mu_3\ldots\mu_p}{}_{\nu_1\ldots\nu_q} + \ldots \\ & - \Gamma^{\lambda}_{\sigma\nu_1} T^{\mu_1\ldots\mu_p}{}_{\lambda\nu_2\ldots\nu_q} - \Gamma^{\lambda}_{\sigma\nu_2} T^{\mu_1\ldots\mu_p}{}_{\nu_1\lambda\nu_3\ldots\nu_q} - \ldots\end{aligned} \tag{4.13}$$

The covariant derivative of a tensor ***density*** of weight W is given by

$$\nabla_\gamma \mathcal{T}^{\mu\ldots}{}_{\nu\ldots} = (\text{usual terms}) - W\Gamma^\sigma_{\sigma\gamma}\mathcal{T}^{\mu\ldots}{}_{\nu\ldots}.$$

We have no requirements about symmetry of the connection under interchange of indices, however we note that the antisymmetric part of the connection, known as the ***torsion***, is in fact a tensor, as can be seen directly from the transformation of the connection in Equation (4.6),

$$T_{\mu\nu}{}^\lambda \equiv \Gamma^\lambda_{\mu\nu} - \Gamma^\lambda_{\nu\mu} \equiv 2\Gamma^\lambda_{[\mu\nu]}. \tag{4.14}$$

The square brackets in the superscript or subscript are shorthand for antisymmetrisation of the expression via the surrounded indices. For example, with 2 upper indices,

$$T^{[ab]} = \frac{1}{2}(T^{ab} - T^{ba}).$$

As we noted above, geodesic motion depends only on the ***torsion-free*** component of the connection,

Another (optional) nice property we might want in a connection, because it makes calculation a bit easier, is ***metric compatibility***, i.e.

$$\nabla_\rho g_{\mu\nu} = 0 \tag{4.15}$$

In some books the partial and covariant derivatives are denoted within the indices using commas and semicolons, respectively:

$$T^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_q,\sigma_1\ldots\sigma_r} \equiv \partial_{\sigma_1}\ldots\partial_{\sigma_r} T^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_q}$$
$$T^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_q;\sigma_1\ldots\sigma_r} \equiv \nabla_{\sigma_1}\ldots\nabla_{\sigma_r} T^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_q}.$$

which implies the same for the inverse metric, $\nabla_\rho g^{\mu\nu} = 0$ and that such a covariant derivative ***commutes with raising and lowering of indices***, e.g. $g_{\mu\lambda}\nabla_\rho V^\lambda = \nabla_\rho V_\mu$.

The Levi-Civitae Connection

If we require the connection to be ***torsion free***, $\Gamma^\lambda_{\mu\nu} = \Gamma^\lambda_{(\mu\nu)}$, and ***metric compatible***, then it turns out there is only one choice of connection, ***the Levi-Civitae connection, (also called the Christoffel connection, or the Riemannian connection)***.

We can determine the expression for the Levi-Civitae connection in terms of the metric by rearranging the 64 equations of the metric compatibility condition above with three different

Round brackets ( ) in the superscript or subscript, denote symmetrisation over the surrounded indices, e.g. for two indices,

$$T_{(ab)} = \frac{1}{2}(T_{ab} + T_{ba}).$$

combinations of indices,

$$
\begin{aligned}
\nabla_\rho g_{\mu\nu} &= \partial_\rho g_{\mu\nu} - \Gamma^\lambda_{\rho\mu}\, g_{\lambda\nu} - \Gamma^\lambda_{\rho\nu}\, g_{\lambda\mu} = 0 \\
\nabla_\mu g_{\nu\rho} &= \partial_\mu g_{\nu\rho} - \Gamma^\lambda_{\mu\nu}\, g_{\lambda\rho} - \Gamma^\lambda_{\mu\rho}\, g_{\lambda\nu} = 0 \qquad (4.16) \\
\nabla_\nu g_{\rho\mu} &= \partial_\nu g_{\rho\mu} - \Gamma^\lambda_{\nu\rho}\, g_{\lambda\mu} - \Gamma^\lambda_{\nu\mu}\, g_{\lambda\rho} = 0
\end{aligned}
$$

where we've just written the same 64 equations again 3 times each with a slight permutation of the bookkeeping of the indices.
Remembering that the metric and the torsion free connection must both be symmetric under interchange of the lower indices, we can subtract the the latter two expressions from the first to get,

$$
\begin{aligned}
&\partial_\rho g_{\mu\nu} - \partial_\mu g_{\nu\rho} - \partial_\nu g_{\rho\mu} + \Gamma^\lambda_{\mu\nu} g_{\lambda\rho} + \Gamma^\lambda_{\nu\mu} g_{\rho\lambda} \\
&\quad - (\cancelto{0}{\Gamma^\lambda_{\rho\mu} g_{\lambda\nu} - \Gamma^\lambda_{\mu\rho} g_{\nu\lambda}}) - (\cancelto{0}{\Gamma^\lambda_{\rho\nu} g_{\mu\lambda} - \Gamma^\lambda_{\nu\rho} g_{\lambda\mu}}) = 0 \qquad (4.17)
\end{aligned}
$$

The 3 index components of the Levi-Civitae connection $\Gamma^\lambda_{\mu\nu}$ are called the Christoffel symbols (of the second kind). They are sometimes written as $\left\{ {\lambda \atop \mu\nu} \right\}$. Note that mathematicians sometimes refer to the covariant derivative itself as "the connection" and the connection coefficients (Christoffel symbols) as the "difference between connections".

which gives,

$$\partial_\rho g_{\mu\nu} - \partial_\mu g_{\nu\rho} - \partial_\nu g_{\rho\mu} + 2\Gamma^\lambda_{\mu\nu} g_{\lambda\rho} = 0. \qquad (4.18)$$

Applying the inverse metric, $g^{\sigma\rho}$, and solving for $\Gamma^\sigma_{\mu\nu}$ we get an expression for the Levi-Civitae connection in terms of the metric and its derivatives,

$$\Gamma^\sigma_{\mu\nu} = \frac{1}{2} g^{\sigma\rho} (\partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu}). \qquad (4.19)$$

For differential geometry on general manifolds we don't have to use the Levi-Civitae connection, but for spacetime in GR, where we have a metric, this is the connection (and covariant derivative) we use. Note that nothing but parsimony prevents us from using a connection with torsion. In fact alternate theories of gravity sometimes include torsion terms.

The connection only vanishes for flat-space cartesian coordinates. Spherical coordinates in flat space, for example, have non-vanishing connection, which is why extra terms appear in e.g. the formulae for the divergence of vector fields in spherical coordinates.

4.3 *The Principle of General Covariance*

Another consequence of the Einstein Equivalence Principle, is ***the Principle of General Covariance***, which states that the general laws of physics must be expressed in identical (tensor) equations relative to all other systems whichever way they are moving.

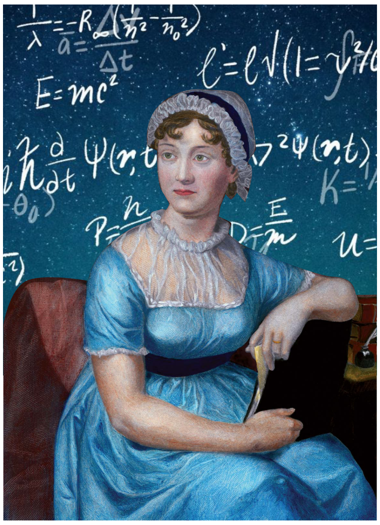
The reason that we want a way to express differential equations covariantly, i.e. in a way that transforms as a tensor, is that such equations are ***coordinate independent***. If the components of a tensor transform as they should, then the real object (the component and the tensor basis together) is invariant under coordinate transformations. Thus if we can express something ***as a proper tensor equation it should hold for any choice of coordinate systems.***

Practically, what this means is that if we can state a (local) law of physics in flat space/special relativity, using a good tensor equation, then this ***same*** tensor equation holds in curved spacetime of GR. For most laws of physics (except gravitational) we are used to, this involves replacing the partial derivative with a covariant derivative to obtain a well-behaved tensor expression.

$$\partial_i \rightarrow \nabla_\mu$$

One can, for example, do semi-classical quantum mechanics on the background of a curved spacetime by simply replacing the partial derivatives in Schroedinger's equation with the covariant derivative. This doesn't tell us how ***gravity*** works in a quantum way though, just the effect of background curvature on quantum mechanics.

How can we describe gravity then? We can't just promote Newton's gravity to a tensor expression, as general covariance only holds for local, non-gravitational physics.



It is a truth universally acknowledged that a physicist, in possession of a physical problem, must be in want of a differential equation.
— Jane Austen (probably)

5
Physics of Curved Manifolds

We now have all the machinery we need to talk about the ***curvature of spacetime***. The curvature of a manifold is defined by the Riemann tensor which is derived from the connection on that manifold. In this chapter we will finally discuss this spacetime curvature and how it is generated in Einstein's gravity.

5.1 Curvature and the Riemann Tensor

Curvature intrinsic to a manifold manifests in several ways that are all related to the same object, ***the Riemann curvature tensor***.

As we discussed earlier, initially parallel geodesics do not remain parallel in curved space (geodesic deviation). Additionally, covariant derivatives of tensors ***do not commute*** in curved space. We also have that ***parallel transport*** of a vector around closed loops in curved space leads to a change in that vector. Each of these ends up depending on the same object that describes the intrinsic curvature of the manifold.

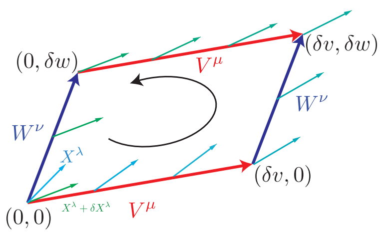
The Reimann Tensor

Here we will define this ***Riemann tensor*** by considering the commutator of covariant derivatives, acting on a vector field.

$$\begin{aligned}
[\nabla_\mu, \nabla_\nu]V^\rho &\equiv \nabla_\mu \nabla_\nu V^\rho - \nabla_\nu \nabla_\mu V^\rho \\
&= \partial_\mu(\nabla_\nu V^\rho) - \Gamma^\lambda_{\mu\nu}\nabla_\lambda V^\rho + \Gamma^\rho_{\mu\sigma}\nabla_\nu V^\sigma \\
&\qquad - \ldots (\mu \leftrightarrow \nu) \ldots \\
&= \partial_\mu\partial_\nu V^\rho + (\partial_\mu \Gamma^\rho_{\sigma\nu})V^\sigma + \Gamma^\rho_{\nu\sigma}\partial_\mu V^\sigma \\
&\qquad - \Gamma^\lambda_{\mu\nu}\partial_\lambda V^\rho - \Gamma^\lambda_{\mu\nu}\Gamma^\rho_{\lambda\sigma}V^\sigma \\
&\qquad + \Gamma^\rho_{\nu\sigma}\partial_\nu V^\sigma + \Gamma^\rho_{\mu\sigma}\Gamma^\sigma_{\nu\lambda}V^\lambda \\
&\qquad - \ldots (\mu \leftrightarrow \nu) \ldots \\
&= \underbrace{\left(\partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\nu \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma}\right)}_{\equiv R^\rho{}_{\sigma\mu\nu}} V^\sigma \\
&\qquad - 2\Gamma^\lambda_{[\mu\nu]}\nabla_\lambda V^\rho
\end{aligned}$$

We see that the left hand side is a good tensor, being constructed of covariant derivatives, while the term on the

Imagine we ***parallel transport*** a vector $\boldsymbol{X} = X^\sigma \boldsymbol{e}_\sigma$ around some infinitesimal loop defined by two other vectors $\boldsymbol{V} = V^\mu \boldsymbol{e}_\mu$ and $\boldsymbol{W} = W^\nu \boldsymbol{e}_\nu$ with infinitesimal side lengths δv and δw respectively.



The action of parallel transport is independent of the coordinate system, so we know it should be described by a tensor. It is a linear transformation on a vector so it must take a vector, and return a vector. It must also depend on the path, described by the two vectors $\boldsymbol{V}$ $\boldsymbol{W}$. It must also be antisymmetric with respect the loop vectors since going the other way around the loop must give the inverse infinitesimal change. If the change in the ρth component of $\boldsymbol{X}$ is δX^ρ, this should be given by

$$\delta X^\rho = (\delta v)(\delta w) V^\mu W^\nu R^\rho{}_{\sigma\mu\nu} X^\sigma$$

where the effect of curvature is captured in

$$R = R^\rho{}_{\sigma\mu\nu} \boldsymbol{e}_\rho \otimes \boldsymbol{e}^\sigma \otimes \boldsymbol{e}^\mu \otimes \boldsymbol{e}^\nu$$

a (1,3) tensor with components antisymmetric in the last two indices, called the Riemann curvature tensor.

last line that contains the torsion tensor, $2\Gamma^{\lambda}_{[\mu\nu]}$, is also a good tensor. Therefore the remaining piece must also be a good tensor, so we define the term in brackets to be the components of ***the Riemann tensor***,

$$R^{\rho}{}_{\sigma\mu\nu} \equiv (\partial_\mu \Gamma^{\rho}_{\nu\sigma} - \partial_\nu \Gamma^{\rho}_{\mu\sigma} + \Gamma^{\rho}_{\mu\lambda}\Gamma^{\lambda}_{\nu\sigma} - \Gamma^{\rho}_{\nu\lambda}\Gamma^{\lambda}_{\mu\sigma}). \tag{5.1}$$

Note that this definition of the Riemann tensor only depends on the connection not the metric, as it holds for ***any manifold with a connection***, not just one with the Levi-Civitae connection.

We haven't shown that this Riemann tensor we derived by the commutator of the covariant derivatives is the same as the curvature tensor object we find by parallel transport (or for that matter in geodesic deviation), but it turns out it is indeed the same.

Given the metric, assuming the Levi-Civitae connection and using Local Flatness, let's consider what the Riemann tensor looks like in Locally Flat Cartesian coordinates (e.g. Fermi normal coordinates or Riemann normal coordinates) at some particular point p, such that the derivatives of the metric (i.e. the Christoffel symbols $\Gamma^{\lambda}_{\mu\nu}$) vanish, while the second derivatives of the metric (i.e. the first derivatives of $\Gamma^{\lambda}_{\mu\nu}$) do not.

We have that in these coordinates,

$$R^{\rho}{}_{\sigma\mu\nu} = \partial_\mu \Gamma^{\rho}_{\nu\sigma} - \partial_\nu \Gamma^{\rho}_{\mu\nu}.$$

Using the metric to lower the index, and substituting the expression for the Christoffel symbol in terms of the metric, we can find,

$$R_{\rho\sigma\mu\nu} = \frac{1}{2}(\partial_\mu\partial_\sigma g_{\rho\nu} - \partial_\mu\partial_\rho g_{\nu\sigma} - \partial_\nu\partial_\sigma g_{\rho\mu} + \partial_\nu\partial_\rho g_{\mu\sigma}).$$

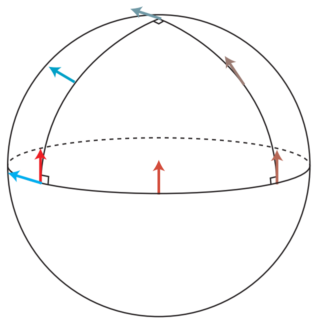


Figure 5.1: Parallel transport of a vector around a sphere results in a significant change in the vector orientation.

The expression for general coordinates is quite messy, but from the Locally Flat Coordinates version, we can extract its symmetries that should hold in all coordinates (since the below equations are tensor equations):

Symmetries of the Riemann Tensor

- The Riemann Tensor is antisymmetric in the first two indices

$$R_{\rho\sigma\mu\nu} = -R_{\sigma\rho\mu\nu}$$

- It is symmetric under the interchange of the first pair and the second pair of indices

$$R_{\rho\sigma\mu\nu} = R_{\mu\nu\rho\sigma}$$

- And finally the cyclic permutation of the last three indices vanishes

$$R_{\rho\sigma\mu\nu} + R_{\rho\mu\nu\sigma} + R_{\rho\nu\sigma\mu} = 0$$

which is equivalent to the vanishing of the antisymmetric piece of the last three indices,

$$R_{\rho[\sigma\mu\nu]} = 0$$

Considering these interdependent symmetries carefully, one can show that there are 20 independent components of the Riemann tensor, which are precisely the 20 components of the 2nd derivative of the metric that we couldn't get to vanish with a coordinate transform!

5.2 Riemann, and Ricci, and Bianchi, and Einstein

Using the Riemann tensor we can now construct several useful tensors and tensor identities.

The Ricci Tensor and Ricci Scalar

For a given Riemann tensor we can contract the raised index with one of the lower ones. If the connection is the Christoffel connection, with its (anti)symmetries there is only one independent contraction which we define as ***the Ricci tensor*** $R_{\mu\nu}$,

$$R_{\lambda\sigma} \equiv R^{\mu}{}_{\lambda\mu\sigma} \tag{5.2}$$
$$= \partial_\mu \Gamma^{\mu}_{\lambda\sigma} - \partial_\sigma \Gamma^{\mu}_{\lambda\mu} + \Gamma^{\nu}_{\lambda\sigma}\Gamma^{\mu}_{\mu\nu} - \Gamma^{\nu}_{\lambda\mu}\Gamma^{\mu}_{\sigma\nu}. \tag{5.3}$$

With the metric to raise one of the indices we can further contract to get the ***Ricci scalar***

$$R \equiv R^{\lambda}{}_{\lambda} \tag{5.4}$$
$$= g^{\lambda\sigma}(\partial_\mu \Gamma^{\mu}_{\lambda\sigma} - \partial_\sigma \Gamma^{\mu}_{\lambda\mu} + \Gamma^{\nu}_{\lambda\sigma}\Gamma^{\mu}_{\mu\nu} - \Gamma^{\nu}_{\lambda\mu}\Gamma^{\mu}_{\sigma\nu}). \tag{5.5}$$

The Ricci scalar (field) is often called ***the scalar curvature***.

The Bianchi Identities

In Locally Flat cartesian coordinates we can look at the components of the covariant derivative of the Riemann tensor at a point

$$\nabla_\lambda R_{\alpha\beta\mu\nu} = \partial_\lambda R_{\alpha\beta\mu\nu}$$
$$= \frac{1}{2}(\partial_\lambda\partial_\beta\partial_\mu g_{\alpha\nu} - \partial_\lambda\partial_\beta\partial_\nu g_{\alpha\mu} + \partial_\lambda\partial_\alpha\partial_\nu g_{\beta\mu} - \partial_\lambda\partial_\alpha\partial_\mu g_{\beta\nu}) \tag{5.6}$$

By symmetry of the metric and the order of partial derivatives we can do a little algebra and get the ***Bianchi identity***:

$$\nabla_\lambda R_{\alpha\beta\mu\nu} + \nabla_\nu R_{\alpha\beta\lambda\mu} + \nabla_\mu R_{\alpha\beta\nu\lambda} = 0$$

The Bianchi identity is true independent of coordinates as this is a tensor equation.
We can raise and contract the α and μ indices in the Bianchi identity using the metric to obtain

$$g^{\alpha\underline{\mu}}[\nabla_\lambda R_{\alpha\beta\underline{\mu}\nu} + \nabla_\nu R_{\alpha\beta\lambda\underline{\mu}} + \nabla_{\underline{\mu}} R_{\alpha\beta\nu\lambda}] = 0$$
$$\nabla_\lambda R_{\beta\nu} - \nabla_\nu R_{\beta\lambda} + \nabla_\mu R^{\mu}{}_{\beta\nu\lambda} = 0 \tag{5.7}$$

where we've used the antisymmetry of the second two indices of the Riemann tensor and the definition of the Ricci tensor. This is called the ***contracted Bianchi identity***.

Contracting again on β and ν we get

$$\begin{aligned} g^{\beta\nu}[\nabla_\lambda R_{\beta\nu} - \nabla_\nu R_{\beta\lambda} + \nabla_\mu R^\mu{}_{\beta\nu\lambda}] &= 0 \\ \nabla_\lambda R - \nabla_\nu R^\nu{}_\lambda - \nabla_\mu R^\mu{}_\lambda &= 0 \\ \nabla_\mu (2R^\mu{}_\lambda - \delta^\mu{}_\lambda R) &= 0 \end{aligned} \tag{5.8}$$

which is called ***the twice contracted Bianchi identity***.

The Einstein Tensor

Raising the indices of the terms inside the covariant derivative in the twice contracted Bianchi Identity, and dividing by 2, we can define the ***Einstein Tensor***

$$G^{\alpha\beta} \equiv R^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} R$$

such that the twice contracted Bianchi identity can be expressed as

$$\nabla_\beta G^{\alpha\beta} = 0$$

5.3 *Gravitation*

We want to connect curvature of spacetime to the Newtonian concept of a gravitational field.

In Newtonian gravity we have a gravitational field Φ, which obeys

$$\nabla^2 \Phi = 4\pi G \rho$$

where here we use ∇^2 to mean the regular Laplacian, G is Newton's constant and ρ is the mass density, which acts as the "source" term for the gravitational field.

5.3.1 *The Stress-Energy Tensor*

In a relativistic treatment we know mass, energy, and momentum are all frame dependent. How do we choose the "source" of gravity? Maybe we would choose the "source" to be the co-moving rest mass energy of a fluid, however this would only work for a one particular inertial observer, as someone moving in a different Lorentz frame would see a different amount of energy.

Spacetime contains a flowing "river" of 4-momentum. Each particle carries its 4-momentum vector with itself along its world line. Many particles, on many world lines, viewed in a smeared-out manner (continuum approximation), produce a continuum flow– a river of 4-momentum. Electromagnetic fields, neutrino fields, meson fields etc: they too contribute to this 4-momentum flow.

The Stress-Energy tensor is also sometimes called the ***Stress-Energy-Momentum tensor***, or the ***Energy-Momentum Tensor***, and it is the Lorentzian spacetime generalisation of the stress tensor from continuum mechanics.

It turns out that the right object to consider is the ***Stress-Energy tensor*** T at each event (point) p in spacetime, which is a linear machine that contains knowledge of the energy density, momentum density, and stress as measured by any and all observers at a particular point $p \in M$. Included are energy, momentum, and stress associated with all forms of matter and all non-gravitational fields.

The Stress-Energy Tensor

The stress-energy tensor is a linear machine with two vector slots (we can consider it as a machine with two ***covector*** slots instead by applying the spacetime metric to the input covectors to obtain their metric duals). We are always free to "raise" and "lower" indices, i.e. switch vector and covector arguments. (Note that one cannot really define a ***symmetric*** (1, 1) tensor under the interchange of a vector and a covector since these arguments are from completely different spaces!) Consider a small three-dimensional (oriented) parellelpiped in spacetime with vectors $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ for edges. ***How much 4-momentum crosses this volume in its positive sense*** (i.e. from its negative side to its positive side)? To calculate the answer we first have to construct a ***(3-)volume 1-form***

$$\Sigma_\mu = \varepsilon_{\mu\alpha\beta\gamma} A^\alpha B^\beta C^\gamma, \tag{5.9}$$

which provides a kind of "normal" 1-form to the 3D volume in 4D spacetime.
If we insert this volume 1-form into the the second slot† of the (2,0) version of the stress energy tensor we get

$$T(\cdot, \boldsymbol{\Sigma}) = \boldsymbol{p}(\cdot) = \begin{pmatrix} \text{Momentum vector crossing from} \\ \text{negative side toward the positive side.} \end{pmatrix}$$

To get the "projection" of the 4-momentum along a one-form $\boldsymbol{\alpha}$,

$$T(\boldsymbol{\alpha}, \boldsymbol{\Sigma}) = \boldsymbol{p}(\boldsymbol{\alpha}),$$

which has components,

$$p^\mu = T^{\mu\nu}\Sigma_\nu$$

This defines the stress energy tensor.

The totally antisymmetric tensor ε has components,

$$\varepsilon_{\mu\alpha\beta\gamma} = \begin{cases} 0 & \text{any repeated indices;} \\ +1 & \text{even permutations of indices;} \\ -1 & \text{odd permutations of indices;} \end{cases}$$

and is called the Levi-Civitae tensor. It can be thought of as the 4-dimensional equivalent to a cross product. (It naturally arises when we take wedge products.) It is often used to find the antisymmetric parts of an expression.

†The stress-energy tensor turns out to be symmetric, so we can actually use either slot. But this must be shown first.

Like all other tensors, the stress-energy tensor is a machine whose definition and significance transcend coordinate systems and reference frames. But any one observer, locked into one Lorentz frame, pays more attention to the components of T than to T itself. To each component he ascribes a specific physical significance. Of greatest interest, perhaps is the ***time-time*** component. It is the total density of mass-energy as measured in the observer's Lorentz frame:

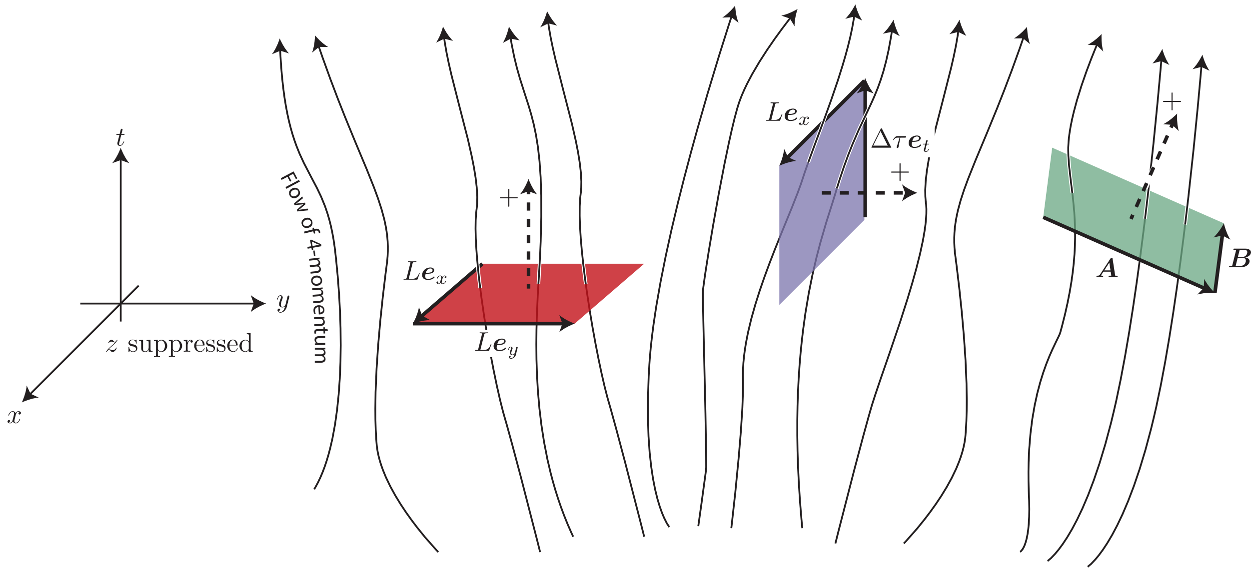


Figure 5.2: The "river" of 4-momentum flowing through spacetime, and three different 3-volumes across which it flows. (One dimension is suppressed in the picture so a 3-volume looks like a 2-volume.) The first 3-volume is the interior of a cubical soap box momentarily at rest in the depicted Lorentz frame. it's edges are Le_x, Le_y, Le_z and its volume 1-form with positive orientation $(+)$ towards the future, is $\mathbf{\Sigma} = L^3 \boldsymbol{dt}$. The second 3-volume is the "world sheet" swept out in time $\Delta\tau$ by the top of a square surface. The square's edges are Le_x and Le_z, and it's volume 1-form with "positive" sense is in the y-direction, is $\mathbf{\Sigma} = L^2 \Delta\tau \boldsymbol{dy}$. The third 3-volume is an arbitrary one, with edges $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$, with volume 1-form $\Sigma_\mu = \varepsilon_{\mu\alpha\beta\gamma} A^\alpha B^\beta C^\gamma$.

$$T_{00} = -T^0{}_0 = T^{00} = T(\boldsymbol{e}_0, \boldsymbol{e}_0) = \text{ density of mass-energy.} \tag{5.10}$$

(since in the observer's frame at this point $g_{\mu\nu} = \eta_{\mu\nu}$ that describe the flux of α-momentum across a surface of constant coordinate x^β, The ***spacetime*** componentw T^{j0} can be interpreted by considering the interior of a soap box at rest in the observer's frame. If its volume is V, then its volume 1-form is $\mathbf{\Sigma} = -V\boldsymbol{u} = V\boldsymbol{dt}$; and the μ-th component of 4-momentum inside it is

$$p^\mu = T(dx^\mu, \mathbf{\Sigma}) = VT(dx^\mu, dt) = VT^{\mu 0},$$

thus the four momentum per unit volume is $p^\mu / V = T^{\mu 0}$, i.e.

$$T^{00} = \text{density of mass-energy} \tag{5.11}$$

$$T^{j0} = \text{density of } j\text{-component of momentum} \tag{5.12}$$

The components $T^{\mu,k}$ can be interpreted using a two-dimensional surface area ΔS at rest in the observer's frame with a positive normal pointing in the k-direction. During the lapse of time Δt this 2-surface sweeps out a 3-volume with volume 1-form $\mathbf{\Sigma} = \Delta S \Delta t dx^k$. The μ-component of the 4-momentum that crosses the 2-surface in time Δt is

$$p^\mu = T(dx^\mu, \mathbf{\Sigma}) = \Delta S \delta t T(dx^\mu, dx^k) = \Delta S \delta t T^{\mu k}. \tag{5.13}$$

Thus the flux of 4-momentum is

$$\left(\frac{p^\mu}{\Delta S \Delta t}\right)_{\text{crossing surface} \perp \text{ to } \boldsymbol{e}_k} = T^{\mu k}, \tag{5.14}$$

or, equivalently,

$$
\begin{aligned}
T^{0k} &= k\text{-component of energy flux} && (5.15)\\
T^{jk} &= (j,k)\text{-component of "stress"} && (5.16)\\
&\equiv k\text{th component of flux of } j\text{-component of momentum}\\
&\equiv j\text{th component of force at } x^k - \epsilon \text{ acting on } x^k + \epsilon\\
&\quad\ \text{across the unit surface.}
\end{aligned}
$$

Symmetry of the Stress-Energy Tensor

Let's consider a particular Lorentz frame. Consider first the momentum density, (components T^{j0}) and the energy flux (components T^{0j}). They must be equal because energy = mass (i.e. "$E = Mc^2 = M$"),

$$
\begin{aligned}
T^{0j} &= (\text{energy flux})\\
&= (\text{energy density}) \times (\text{mean velocity of energy flow})^j\\
&= (\text{mass density}) \times (\text{mean velocity of mass flow})^j\\
&= (\text{momentum density}) = T^{j0}.
\end{aligned}
$$

Only the stress tensor T^{jk} remains. For it, one uses the same standard argument as in Newtonian theory of continua. Consider a very small cube, of side L, mass-energy $T^{00}L^3$, and moment of inertia $\sim\ T^{00}L^5$. With the space coordinates centred at the cube the expression for the z component of torque exerted on the cube by its surroundings is

$$
\begin{aligned}
\tau^z &= \underbrace{(-T^{yx}L^2)}_{(y\text{-force on } +x\text{-face})}\overbrace{(L/2)}^{(\text{lever arm at } +x \text{ face})} + (T^{yx}L^2)(-L/2)\\
&\qquad - (-T^{xy}L^2)(L/2) - (T^{xy}L^2)(-L/2)\\
&= (T^{xy} - T^{yx})L^3.
\end{aligned}
$$

Since the torque decreases as L^3, while the moment of inertia decreases as L^5, the torque will set an arbitrarily small cube into an arbitrarily great angular acceleration, which is clearly non-physical. To avoid this, the stresses must distribute themselves (with a timescale comparable to the shear velocity$/L$), so that the torque vanishes,

$$
T^{yx} = T^{xy} \xRightarrow{\text{all pairs}} T^{jk} = T^{kj}.
$$

Thus, the stress energy tensor T is symmetric.

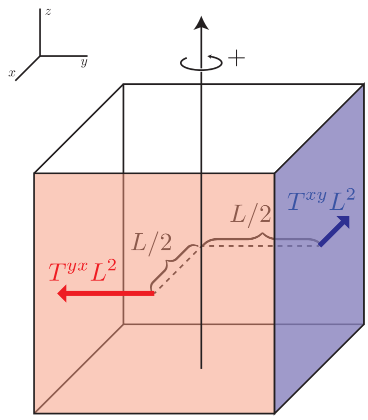


Figure 5.3: The torques in the z direction due to the stress on the $+x$ and $+y$ faces of a cube of material of size L. Note that the T^{yx} is the y component of the force that the block applies on its neighbour in the $+x$-direction, so the force it receives from its neighbour is equal and opposite.

Interpretations of the Stress-Energy Tensor

Let us consider the (0,2)-tensor version of the Stress-Energy Tensor, with components, $T_{\mu\nu}$ at some point (event) in spacetime, observed by an arbitrary observer also at that point.

- Insert the 4-velocity $\boldsymbol{u}$ of the observer into one of the slots; leave the other slot empty. The output is

$$T(\boldsymbol{u},\cdot) = T(\cdot,\boldsymbol{u}) = -\begin{pmatrix}\text{density of 4-momentum (covector),}\\ \text{"}d\boldsymbol{p}/dV\text{",}\\ \text{as measured in observer's}\\ \text{Lorentz frame at that point}\end{pmatrix}$$

- Insert the 4-velocity of the observer into one slot; insert an arbitrary unit vector $\boldsymbol{n}$ into the other slot. The output is then

$$T(\boldsymbol{u},\boldsymbol{n}) = T(\boldsymbol{n},\boldsymbol{u}) = -\begin{pmatrix}\text{component, "}\langle d\boldsymbol{p}/dV,\boldsymbol{n}\rangle\text{",}\\ \text{of 4-momentum density along the}\\ \boldsymbol{n}\text{-direction measured}\\ \text{in the observer's Lorentz frame}\end{pmatrix}$$

- Insert the 4-velocity of the observer into both slots. The output is the density of mass-energy as she measures in her Lorentz frame:

$$T(\boldsymbol{u},\boldsymbol{u}) = \begin{pmatrix}\text{mass-energy per unit volume}\\ \text{as measured in frame with}\\ \text{4-velocity } \boldsymbol{u}\end{pmatrix}$$

- Take two spacelike basis vectors of the observer's Locally Flat Coordinates, $\boldsymbol{e}_j$, $\boldsymbol{e}_k$. Insert $\boldsymbol{e}_j$ and $\boldsymbol{e}_k$ into the slots of T. The output is the j,k component of the stress as measured by that observer:

$$\begin{aligned}T_{jk} = T(\boldsymbol{e}_j,\boldsymbol{e}_k) &= T_{kj} = T(\boldsymbol{e}_k,\boldsymbol{e}_j)\\ &= \begin{pmatrix}j\text{-component of force acting}\\ \text{from side } x^k-\epsilon \text{ to side } x^k+\epsilon,\\ \text{across a unit surface area } \perp \boldsymbol{e}_k\end{pmatrix}\\ &= \begin{pmatrix}k\text{-component of force acting}\\ \text{from side } x^j-\epsilon \text{ to side } x^j+\epsilon,\\ \text{across a unit surface area } \perp \boldsymbol{e}_j.\end{pmatrix}\end{aligned}$$

Conservation of Energy and Momentum

In electrodynamics, the conservation of charge can be expressed by the differential equation

$$\partial_t(\text{charge density}) + \nabla\cdot(\text{current density}) = 0$$

i.e. $\partial_0 J^0 + \nabla \cdot J = 0$, or $\nabla_\mu J^\mu = 0$. Similarly the ***conservation of energy-momentum*** can be expressed by the fundamental geometric law,

$$\nabla_\nu T^{\mu\nu} = 0.$$

This is a tensor equation, and must hold even in curved spacetime by the principle of general covariance. This conservation law plays an important role in the general relativity.

5.3.2 *The Einstein Field Equations*

If we want the stress-energy tensor to act as our "source" term for the curvature of spacetime,

$$\text{Something to do with curvature} = T^{\mu\nu} \tag{5.17}$$

we will need some representation of that curvature that also have zero divergence so that the energy and momentum conservation remains true.

The "curvaturey" term that we need must be a $(2, 0)$ tensor that should be a combination of ***up to*** 2 partial derivatives of $g_{\mu\nu}$ since we know that the 2nd derivatives of the metric contain the "curvature" information. The term should also be symmetric, since $T^{\mu\nu}$ is symmetric.

One possible candidate is the Ricci tensor (with indices raised by the metric), which is of course all about curvature, is a second rank tensor and is symmetric.

$$R^{\mu\nu} \overset{?}{=} kT^{\mu\nu} \tag{5.18}$$

However we know from the second contracted Bianchi identity that

$$\nabla_\nu R^{\mu\nu} = \frac{1}{2} g^{\mu\nu} \nabla_\nu R \neq 0 = k \nabla_\nu T^{\mu\nu} \tag{5.19}$$

So we must have extra bits in the Ricci tensor that prevents it from having zero divergence.

However, the Einstein tensor

$$G^{\mu\nu} \equiv R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R \tag{5.20}$$

is precisely the Ricci tensor with these bits removed, as by the twice contracted Bianchi identity we have

$$\nabla_\nu G^{\mu\nu} = 0. \tag{5.21}$$

Thus we can try

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R = kT^{\mu\nu} \tag{5.22}$$

Note that we can also add a term proportional to $g^{\mu\nu}$ to $G^{\mu\nu}$ without changing the divergence (since we have a metric compatible covariant derivative) such that

$$G^{\mu\nu} + \Lambda g^{\mu\nu} = kT^{\mu\nu} \tag{5.23}$$

where Λ is known as ***the cosmological constant***.

In order to reproduce Newtonian gravity it turns out we have to choose† $k = 8\pi G/c^4$.

† Note that this choice of constant can be derived from the weak field limit of GR, by considering a metric of the form

$$\begin{aligned} ds^2 = &-c^2(1+2\Phi)dt^2 \\ &+ (1-2\Phi)(dx^2+dy^2+dz^2) \end{aligned}$$

where $\Phi << 1$.

If you work out the geodesic equations (using the Christoffel connection) for this metric, you would find that the spatial components gives (to lowest order in velocity)

$$\frac{d^2x^i}{dt^2} = -c^2\frac{\partial\Phi}{\partial x^j}\delta^{ij}$$

which is just the equation of motion in a Newtonian potential, while the 00 component of the Einstein field equations (with this particular choice of constants) gives

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\Phi = \nabla^2\Phi = 4\pi G\rho$$

which correctly recovers Newtonian gravity in the weak field limit.

The Einstein Field Equations

The ***Einstein Field Equations of General Relativity*** are

$$R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R \equiv G^{\mu\nu} = \frac{8\pi G}{c^4}T^{\mu\nu}. \tag{5.24}$$

where we can add a cosmological constant if desired. It tells us how spacetime curvature is sourced by matter and energy, which in turn describes spacetime background on which that matter and energy propagates.

Einstein's field equations are 2nd order non-linear differential equations for the metric tensor field $g_{\mu\nu}$. There are 10 independent equations in 4 dimensions (since both sides are symmetric 2-index tensors), which matches the 10 components of the metric. However, the Bianchi identity $\nabla_\mu G^{\mu\nu} = 0$ places 4 constraints on $R^{\mu\nu}$, so there are really only 6 independent equations. This makes sense: if the metric is a solution to Einstein's equations in one coordinate system, it must be a solution in all coordinate systems. This means that there are 4 unphysical degrees of freedom in $g_{\mu\nu}$, represented by the the the function $x^{\mu'}(x)$, and Einstein's equations should only constrain the remaining 6 coordinate-independent degrees of freedom.

These are extremely complicated partial differential equations: The Ricci scalar and tensor are contractions of the Riemann tensor, which involve both derivatives and products of the Christoffel symbols, which themselves involve the inverse metric and derivatives of the metric. The equations are ***non-linear***, so you can't superimpose two known solutions to get a third. Therefore, general solutions to Einstein's equations are very difficult to find. Even in vacuum, the equation

$$R^{\mu\nu} = 0 \tag{5.25}$$

is difficult to solve. Typically, the way to proceed analytically is to make simlyfing assumptions involving symmetry of the metric.

The nonlinearity of Einsteins equations is interesting. In Newtonian gravity, the potential due to two point masses is simply the sum of the potentials of two masses individually. This does not apply to general relativity (aside from in the weak field limit, where we can ignore terms $\sim \Gamma^2$), but there's a good reason for that: gravity couples with itself! This is a consequence of the equivalence principle. If gravity did not couple to itself, a gravitational atom (a bound state held together by gravity) would have different inertial and gravitational mass, due to negative binding energy only applying to the inertial mass.

Extra Material: The Einstein-Hilbert Action

Einstein's equations can also be derived by constructing an action principle for the behaviour of the metric, known as the Einstein-Hilbert action. As for any classical field theory, we want an integral over spacetime of a Lagrangian density:

$$S_{\text{EH}} = \int \mathcal{L}_{\text{EH}} d^4x. \tag{5.26}$$

The Lagrangian density is a tensor density (since the action must return a scalar value), so it is a scalar multiplied by the determinant of the Jacobian, $\sqrt{-g}$. What scalar can we make from the metric?
We know that by choosing Locally Flat coordinates, we can make the first derivatives of the metric vanish at any point, so any non-trivial scalar constructed from the metric (and independent of coordinates) must be dependent on the ***second*** derivative. The Riemann tensor is constructed from the 2nd derivatives, and the only independent scalar we can construct from the Riemann tensor is the Ricci scalar R. We didn't show (but it is true), that any tensor constructed from the metric and its 1st and 2nd derivatives can be expressed in terms of the metric and the Riemann tensor. Therefore, the only independent scalar constructed form the metric, which is no higher than second order in its derivatives, is the Ricci scalar. Hilbert figured out that this was therefore the simplest possible choice form the Lagrangian and proposed the ***Einstein-Hilbert*** action

$$S_{\text{EH}} = \frac{c^4}{16\pi G} \int \sqrt{-g} R d^4x. \tag{5.27}$$

Extremising the Einstein-Hilbert action over the metric tensor field results in the Einstein field equations for vacuum.

Including a term in the action to describe the physics of matter and non-gravitational fields $\mathcal{L}_M$ the action becomes

$$S = \int \left[\frac{c^4}{16\pi G} R + \mathcal{L}_M \right] \sqrt{-g} d^4x. \tag{5.28}$$

when varied with respect to the (inverse)metric, the first term gives the left hand side of the Einstein field equations, while the variation of 2nd term $\delta\mathcal{L}_M / \delta g^{\mu\nu}$ turns out to be

$$\frac{\delta \mathcal{L}_M}{\delta g^{\mu\nu}} = -\frac{1}{2} \left(T_{\mu\nu} + g_{\mu\nu} \mathcal{L}_M \right) \tag{5.29}$$

such that the Euler-Lagrange equations of the action are the Einstein field equations.

It's also worth noting that the coordinate freedom we use so liberally is an example of a symmetry called ***diffeomorphism invariance***, i.e. the action looks the same no matter if you transform coordinates through a diffemorphism, and pushforward all geometric quantities. This is an example of a ***gauge-symmetry***, and leads to the twice contracted Bianchi identity, and the conservation of energy and momentum in the form of, $\nabla_\nu T^{\mu\nu}$; in other words, $T^{\mu\nu}$ is the ***Noether current*** that is conserved because of the diffeomorphism invariance of the action.

6
Relativistic Cosmology

6.1 Why Relativistic?

Now we are ready to talk about the foundations for the 2nd half of the course: Cosmology, the study of the structure of the universe.

Why do we need GR to study cosmology?

- It should be obvious that we need GR when $M/R \gtrsim 1$ (in units of $G = c = 1$). For black holes, neutron stars and other such compact objects, this occurs due to small R. Far from these objects, Newton is fine and $M/R \ll 1$.

- Another way to get M/R closer to 1 is if M gets larger faster than R. This is the case for Cosmology. If we have a (roughly) uniform density on larger scales, this should have a mass increasing as R^3. Thus M/R scales with R^2, so even if M/R is small on the scales of our solar system, eventually it will become important.

- The solar system is nowhere near relativistic, with $M/R \sim 10^{-6}$. The galaxy is about 15kpc big and contains about 10^{11} stars and has $M/R \sim 10^{-6}$ as well. This is why Galactic structure doesn't really need relativity. Clusters are Mpc in scale, and have $M/R \sim 10^{-4}$.

- Assuming a mass density of $\rho = 10^{-26}$ kg/m^3 we would need to go out to ~ 6 Gpc for $M/R \sim 1$, which is well within the observable universe. To understand the large scale structure of the universe at these scales and beyond, we ***must use GR***.

Note that in Newtonian Gravity $\nabla^2\Phi = 4\pi G\rho$, so much of the solution depends on the boundary conditions, thus it depends on the ***rest of the universe***. However in GR, as long as we look at a small enough region (but far enough away so that $M/R \ll 1$, things can be treated as ***asymptotically flat***, even for small regions of curved space. Thus we can separate local (e.g. BHs) from global (cosmology) using local flatness.

6.2 Cosmological Principles in a Relativistic Framework

To build a relativistic description of cosmology, we need to make certain assumptions about the nature of the universe on large scales. These ***cosmological principles***, motivated by observations of the universe are:

Cosmological Principles

1. The Universe is homogeneous (on large scales), i.e., on large scales it looks the same in all directions.
2. It is isotropic – there is no consistently defined special direction (e.g. no global B-field)
3. There is a Hubble flow – galaxies, on average, recede from us at a speed proportional to their distance, $v = Hd$

In the cosmology of our neighbourhood of the universe, we can choose ***a preferred time*** whose hypersurfaces are homogeneous and isotropic, that expand with the Hubble flow as observed by any observer at rest relative to any location in those hypersurfaces. Why can we do this? It seems to violate everything we learned about relativity, where no frames are special. In this case we only have 1 universe. It ***is*** unique in terms of the symmetries of this special frame, and it would be foolish not to take advantage of the simplifications and symmetries it offered. Note that this is not to say we couldn't formulate just as good a relativistic cosmology using some other space-time slicing or frames, it would just be harder.

6.3 *Metrics for Cosmology*

Spacetime can be sliced into hypersurfaces of constant time that are homogeneous and isotropic. The mean rest frame of all galaxies agrees with this notion of simultaneity (we can ignore random velocity for now and assume it is all Hubble flow).

At each time slice $t = t_o$ the hypersurface has a line element

$$d\ell^2(t_o) = h_{ij}(t_o)dx^i dx^j. \tag{6.1}$$

You can think of this as the line element using the spatial metric induced from the ***pullback*** of the spacetime metric along the inclusion map for the hypersurface (i.e. the map that just identifies which elements of the manifold are on the submanifold).

For a different time slice, $t = t_1$, we can require the line element on that hypersurface to have the form

$$d\ell^2(t_1) = f(t_1, t_o)\, h_{ij}(t_o)\, dx^i dx^j = h_{ij}(t_1)\, dx^i dx^j, \tag{6.2}$$

such that it guarantees all h_{ij} terms grow at the same rate isotropically. Without loss of generality we can scale everything to time t_o and have

$$d\ell^2(t) = R^2(t)h_{ij}dx^i dx^j \tag{6.3}$$

where $R(t_o) = 1$ and R (which is not the Ricci scalar) grows in time.

The full spacetime must have

$$ds^2 = -dt^2 + g_{0i}dtdx^i + R^2(t)h_{ij}dx^i dx^j \tag{6.4}$$

where we've chosen $g_{00} = -1$ because t is the proper time at any spatial point.

Since each galaxy has a Locally Flat Coordinates that has the same idea of time, the coordinate basis vector in the time direction $\boldsymbol{e}_0$ must be orthogonal to the spatial coordinate basis vectors $\boldsymbol{e}_i$ such that

$$g_{0i} = \boldsymbol{e}_0 \cdot \boldsymbol{e}_i = 0 \tag{6.5}$$

(since the metric defines the inner product that determines orthogonality in each tangent space), which means we can drop those terms in the metric giving

$$ds^2 = -dt^2 + R^2(t) h_{ij} dx^i dx^j \tag{6.6}$$

we also need h_{ij} to be spherical symmetric (so the Universe looks isotropic and it should have the right signature. To do this it is traditional to rewrite h_{ij} such that

$$d\ell^2 = e^{2\Lambda(r)} dr^2 + r^2 d\Omega^2 \tag{6.7}$$

where the $\Lambda(r)$ term is just exponentiated to guarantee that the term is positive, and $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$ is the usual term in spherical coordinates in order to make it isotropic around $r = 0$.

Such a metric is, however, only necessarily isotropic around $r = 0$. How do we choose the function $e^{2\Lambda}$ to make it also homogeneous around any point? It turns out all we need to do is make sure the Ricci scalar curvature, $R^i{}_i$, of the 3D spatial metric of each hypersurface has the same value at all points (independent of position), i.e. spacetime is curved in the same way at the same time no matter where you are. ***This is equivalent to fixing the spatial (3D) trace of the (pulled-back) Einstein tensor $G^{ij} g_{ij}$ for each point.***

The value of the spatial scalar curvature is related to the ratio of the volume of geodesic balls on the spatial slice to that of an equivalent ball in Euclidean flat space.

The spatial components of the Einstein tensor using the above metric is

$$G_{rr} = -\frac{1}{r^2} e^{2\Lambda}(1 - e^{-2\Lambda}) \tag{6.8}$$

$$G_{\theta\theta} = -re^{-2\Lambda}\Lambda'(r) \tag{6.9}$$

$$G_{\phi\phi} = \sin^2\theta\, G_{\theta\theta} \tag{6.10}$$

We need the trace of G_{ij} to be constant, let's call it κ,

$$\begin{aligned}\kappa &= G_{ij} g^{ij} \\ &= -\frac{1}{r^2} e^{2\Lambda}(1 - e^{-2\Lambda}) e^{-2\Lambda} - 2re^{-2\Lambda}\Lambda' r^{-2} \\ &= -\frac{1}{r^2}[1 - (re^{-2\Lambda})']\end{aligned} \tag{6.11}$$

where we need κ to be constant independent of position. Integrating this allows us to solve for $e^{2\Lambda}$,

$$g_{rr} = e^{2\Lambda} = \frac{1}{1 + \frac{1}{3}\kappa r^2 - A/r} \tag{6.12}$$

where A is a constant of integration.

We need Local Flatness when $r = 0$, which requires $g_{rr}(r = 0) = 1$ (if this were not the case you could always tell things were in curved spacetimes by drawing tiny circles around $r = 0$ and measuring their circumferences to see they do not equal 2π times their radius, hence violating the EEP). This gives us $A = 0$. Usually people use the constant $k = -\kappa/3$ such that

$$g_{rr} = \frac{1}{1-kr^2} \tag{6.13}$$

$$d\ell^2 = \frac{dr^2}{1-kr^2} + r^2 d\Omega^2 \tag{6.14}$$

The Friedman-Robertson-Walker Metric

We find the full cosmological metric for a homogeneous and isotropic universe must take the form

$$ds^2 = -dt^2 + R^2(t)\left[\frac{dr^2}{1-kr^2} + r^2 d\Omega^2\right] \tag{6.15}$$

this is called the ***Friedman-Robertson-Walker metric***. Note that without loss of generality we can rescale r to make either $k = +1, 0,$ or -1.

Note that $\tilde{r} = R(t)r$ is the spherical radius (also called the areal radius, or the curvature radius), i.e. spherical surfaces of constant r have area ***defined to be*** $4\pi\tilde{r}^2$. That is the only sense in which it is akin to our usual r, i.e. the distance from the centre of its sphere to a point on the surface. In curved space these notions of radius are different, because of the non-zero scalar curvature!

6.4 *Cosmological Stress-Energy*

In our cosmological frame the stress energy tensor is taken to be of the form of a perfect fluid (no heat flow, no viscosity, and all forces are perpendicular to the surfaces) such that

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

where the components of the tensor has been written in the preferred[†] comoving coordinates of the expanding universe.

[†]Other, non-preferred frames, are available (and will have different forms of the stress-energy tensor).

Homogeneity tells us that all fluid properties depend only on the time $\rho = \rho(t)$, and $p = p(t)$ etc. We have from the Bianchi identity that $\nabla_\nu T^{\mu\nu} = 0$, and isotropy guarantees that the spatial components of this must vanish. Therefore, we only have the $\mu = 0$ components being nontrivial. Written out this is

$$\frac{d}{dt}(\rho R^3) = -p\frac{d}{dt}(R^3).$$

The term on the left is the rate of change of total mass-energy, while the term on the right is the power lost due to pdV work.

In a matter dominated universe we take $p = 0$ so that the matter (stars and dark matter) behave like pressure-less dust.

$$\frac{d}{dt}(\rho R^3) = 0$$

If radiation is dominating, then we have $p = \rho/3$, and

$$\begin{aligned} \frac{d}{dt}(\rho R^3) &= -\frac{\rho}{3}\frac{d}{dt}(R)^3 \\ \frac{d}{dt}(\rho R^4) &= 0. \end{aligned}$$