GENERAL RELATIVITY & COSMOLOGY

A Quick Guide

Huan Bui

Colby College
Physics & Statistics
Class of 2021

December 26, 2018

Preface

Greetings,

General Relativity & Cosmology - A Quick Guide to is compiled from my PH335: General Relativity and Cosmology with professor Robert Bluhm at Colby College. The course is based on *A Short Course in General Relativity*, 3th Edition, by James Foster and J. David Nightingale.

One of my favorite things to do is to transfer my handwritten lecture notes into nice LaTeXdocuments. I started working on these documents last year with *Special Relativity - A Quick Guide to*, which was based on (part of) my PH241: Modern Physics I with professor Charles Conover. I am still working on it, along with another one on Linear Algebra (based on MA253 with Otto Bretscher), and of course General Relativity - the one you are reading now. All of these .pdf files can be found on my website, under the "A Quick Guide to..." tab. For the "raw" content, i.e., my handwritten lecture notes, feel free to visit the "Lecture Notes" tab.

Today is Dec 22, 2018. In the spring I will be doing an independent study on Classical Fields Theory with prof. Bluhm. I will also be taking Matrix Analysis with prof. Leo Livshits. I'm excited to be taking these advanced physics and mathematics courses, and I look forward to sharing my journey with you through my lecture notes.

Enjoy!

Contents

Preface . . . 1

1 Overview and Review **4**
1.1 Review of Special Relativity . . . 4
1.2 The Equivalence Principle . . . 6
1.3 Versions of the Equivalence Principle . . . 7
1.3.1 The Strong Equivalence Principle . . . 7
1.3.2 The Weak Equivalence Principle . . . 7

2 Review of Vector Calculus **8**
2.1 Operations & Theorems . . . 8
2.2 Coordinate Systems . . . 12

3 Flat 3-dimensional space **15**
3.1 Curvilinear coordinates . . . 15
3.2 Basis vectors . . . 15
3.3 Natural basis . . . 15
3.4 Dual basis . . . 17
3.5 Contravariant and covariant vectors . . . 20
3.5.1 The suffix notation (vector calculus) . . . 20
3.5.2 The Einstein summation notation . . . 20
3.5.3 Contravariant and covariant vectors . . . 21
3.6 Metric tensor . . . 25
3.6.1 Summations as matrix products . . . 27
3.7 Coordinate transformation . . . 30
3.7.1 Vectors . . . 30
3.7.2 Tensors . . . 38
3.8 Scalars . . . 40

4 Flat spacetime **41**
4.1 Special Relativity . . . 41
4.1.1 The Lorentz Transformations, revisited . . . 43
4.1.2 A curiosity about the Lorentz boosts . . . 46
4.1.3 The Poincaré transformations . . . 47
4.1.4 Velocity, momentum, and force . . . 47
4.2 Relativistic Electrodynamics . . . 47

5 Curved spaces **49**
5.1 2-dimensional curved spaces . . . 49
5.2 Manifolds . . . 49
5.3 Tensors on manifolds . . . 49

6 Gravitation and Curvature **50**
6.1 Geodesics and Affine connections $\Gamma^{\sigma}_{\mu\nu}$ 50
6.2 Parallel transport 50
6.3 Covariant differentiation 50
6.4 Newtonian limit 50

7 Einstein's field equations **51**
7.1 The stress-energy tensor $T^{\mu\nu}$ 51
7.2 Riemann curvature tensor $R^{\lambda}{}_{\mu\nu\sigma}$ 51
7.3 The Einstein equations 51
7.4 Schwarzschild solution 51

8 Predictions and tests of general relativity **52**
8.1 Gravitational redshift 52
8.2 Radar time-delay experiments 52
8.3 Black Holes 52

9 Cosmoslogy **53**
9.1 Large-scale geometry of the universe 53
9.1.1 Cosmological principle 53
9.1.2 Robertson-Walker (flat, open, closed) geometries 53
9.1.3 Expansion of the universe 53
9.1.4 Distances and speeds 53
9.1.5 Redshifts 53
9.2 Dynamical evolution of the universe 53
9.2.1 The Friedmann equations 53
9.2.2 The cosmological constant Λ 53
9.2.3 Equations of state 53
9.2.4 A matter-dominated universe ($\Lambda = 0$) [Friedmann models] 53
9.2.5 A flat, metter-dominated universe ($\Lambda = 0$) [old favorite model] 53
9.3 Observational cosmology 53
9.3.1 Hubble law 53
9.3.2 Acceleration of the universe 53
9.3.3 Matter densities and dark matter 53
9.3.4 The flatness and horizon problems 53
9.3.5 Cosmic Microwave Backgrond (CMB) anisotropy 53
9.4 Modern cosmology 53
9.4.1 Inflation 53
9.4.2 Dark energy (The cosmological constant problem) 53
9.4.3 Concordance model [new favorite model] 53
9.4.4 Open questions 53

1 Overview and Review

General relativity is a theory of gravity, replacing Newton's gravity law for heavy masses to give more precise predictions. However, we should keep in mind that general relativity is not yet compatible with quantum mechanics. There are numerous open problems in physics related to reconciling gravity and quantum mechanics.

1.1 Review of Special Relativity

Special relativity studies the kinematics and dynamics in relatively moving inertial reference frames. Most of special relativity and its consequences are encoded in the Lorentz transformations. A classic example of a Lorentz transformation that is often focused on in introductory special relativity is called the "Lorentz boost" in the x-direction. Note that there is nothing special about x or y or z. The x-Lorentz boost is essentially a coordinate transformation, i.e. given coordinates an event A in frame (S), we can calculate the coordinates of A in frame (S').

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix},$$

where $\beta = v/c$ and the Lorentz factor

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Since we are studying relativity, it is important to look at "invariants" - quantities that do not change under transformations. In special relativity, one such invariant is the spacetime invariant:

$$\begin{aligned} (\Delta S)^2 &= (c\Delta\tau)^2 \\ &= (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 \\ &= (c\Delta t')^2 - (\Delta x')^2 - (\Delta y')^2 - (\Delta z')^2 \\ &= (\Delta S')^2 = (c\Delta\tau')^2. \end{aligned}$$

This can be readily shown. In fact, we have verified this in *Special Relativity: A Quick Guide*. This quantity, roughly speaking, is a measure of **proper** distances and times. If we go to a rest frame of event A, such that $\Delta x' = \Delta y' = \Delta z' = 0$, then we get $\Delta t' = \Delta\tau$, where τ denotes the **proper time**, which is the time measured in the rest frame. This gives $(\Delta S)^2 = (c\Delta\tau)^2$.

In relativity in general and in Minkowski spacetime in particular, we are interested in two types of objects: **scalars** and **vectors**. Scalars are invariant

under general coordinate transformations. An example of a scalar is the spacetime invariant. Another scalar, which we probably will not see again in this text, is the metric signature ($\mathrm{sgn}([\eta_{\mu\nu}] = -2)$). Vectors, on the other hand, are a different type of objects, which transform in the same way under coordinate transformations, but are not invariant under general coordinate transformations in general, i.e., their components are not necessarily the same in different coordinate systems. In 4-dimensional spacetime, we work with 4-vectors, which simply means 4-component vectors.

In special relativity, we often talk about position vectors $\mathbf{x} = (ct, x, y, z)^\top = (ct, \vec{x})^\top$ and energy-momentum vectors $\mathbf{p} = (E/c, p_x, p_y, p_z)^\top = (E/c, \vec{p})^\top$. These vectors transform under the Lorentz transformation.

Example 1.1. The x-Lorentz boost applied to $\mathbf{p}$ gives

$$\begin{pmatrix} E'/c \\ p'_x \\ p'_y \\ p'_z \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} E/c \\ p_x \\ p_y \\ p_z \end{pmatrix}.$$

Notice that E/c transforms like ct (time) and $\vec{p}$ transforms like $\vec{x}$ (space). We also have the following invariant

$$\frac{E^2}{c^2} - p_x^2 - p_y^2 - p_z^2 = \frac{E^2}{c^2} - \vec{p} \cdot \vec{p}.$$

But recall that $(mc)^2 = E^2/c^2 - \vec{p} \cdot \vec{p}$, so

$$\frac{E^2}{c^2} - p_x^2 - p_y^2 - p_z^2 = (mc)^2.$$

If we go to a rest frame, such that $\vec{p} = \vec{0}$, then we obtain the famous rest mass-energy equivalence: $E = mc^2$.

Some quantities associated with a vector can be invariant (scalar), such as their norm. In fact, the spacetime invariant in Minkowski space is nothing but a dot product of a vector with itself:

$$\mathbf{a} \cdot \mathbf{a} = a^0 a^0 - a^1 a^1 - a^2 a^2 - a^3 a^3,$$

where $0, 1, 2, 3$ are indices, and $\mathbf{a}$ implies a 4-vectors, which should be distinguished from the 3-vector $\vec{a} = (a^1, a^2, a^3)^\top$. Notice that the dot product in Minkowski spacetime is defined differently from that in flat 3-dimensional Cartesian coordinate system. In the following chapters, we will explore how dot products are defined in general. Hint: the metric tensor plays an important role.

1.2 The Equivalence Principle

In 1907, Albert Einstein had the "happiest thought of his life" when he realized that in a freely falling frame (non rotating and/or accelerating), the effects of gravity go away, i.e., there is an equivalence between gravity and acceleration such that they can "undo" each other. For example, the following two situations are equivalent in terms of the acceleration experienced by the observer: (i) a person standing on Earth, and (ii) a person inside an elevator accelerating upwards at rate g in free space (no gravitational field). Likewise, the following two situations are also equivalent: (iii) a person floating in free space, and (iv) a person inside an airplane free falling towards Earth (this is commercially known as "zero-G flight"). we arrive at the statement of the Equivalence Principle:

"*A small, non-rotating, freely falling frame in a $\vec{g}$ field is an inertial frame.*"

The above statement is a direct result of Galileo's discovery that all objects have the same acceleration due to gravity. This result may seem a little bit *circular*, because it is actually a coincidence that the two roles of **mass**: (i) to cause gravitational force like charge in an electric field:

$$\vec{F} = \frac{GMm_G}{r^3}\vec{r} = m_G\vec{g}$$

and (ii) to measure inertia:

$$\vec{F} = m_I\vec{a}$$

are the same, i.e., $m_I = m_G$. It could have been that "gravitational mass" m_G and "inertial mass" m_I are not the same, in which case $\vec{a} \neq \vec{g}$, and the equivalence principle does not hold. However, the famous Eötvös experiment, which measured the correlation between inertial mass and gravitational mass, has shown that

$$\frac{|m_G - m_I|}{m_I} \leq 10^{-10}.$$

While the observations we have made so far can seem self-apparent, the consequence of the equivalence principle is the bending of light around massive objects. Consider a light ray going horizontally (left to right) pass an upward-accelerating elevator with an observer inside. For the observer, since he is accelerating upwards, he sees the light beam as bent, entering at the top-left of the elevator and exiting at the bottom-right. Now, according to the equivalence principle, because the upward-accelerating scenario is equivalent to the existence of a massive object (such that the observer experiences the same acceleration). This means that the observer, now no longer in the elevator, "sees" the light as bent around this massive object. Fig. 1. illustrates this postulate.

include figure here!!!

General relativity predicts that light going pass Earth's surface will "fall" by approximately 1Å, which is not observable. However, for a much more massive object like the Sun, general relativity predicts a bending of $1.75''$ (arc sec). This prediction was verified by the glorious experiment of Arthur Eddington.

Note that we could argue for the bending of light, using Newtonian physics. However, in order to get the correct predictions for the bending of light, we need general relativity. We will explore the reason behind this discrepancy in the following chapter. But roughly speaking, spacetime is assumed to be flat in Newtonian physics, while spacetime is curved by massive objects, according to general relativity.

One might ask: "How do we view falling objects on Earth as due to the curvature of spacetime?" The answer requires bringing back Minkowski spacetime diagrams.

insert figure here!!!

1.3 Versions of the Equivalence Principle

There are two versions of the equivalence principle, referred to as the *strong equivalence principle*, SEP, and the *weak equivalence principle*, WEP.

1.3.1 The Strong Equivalence Principle

The strong equivalence principle states that **all** of physics reduces to special relativity in a freely falling frame.

1.3.2 The Weak Equivalence Principle

The weak equivalence principle states that all point particles fall at the same rate in a gravitational field ($m_G = m_I$). We notice that WEP only applies to gravity, which makes it sufficient to develop general relativity, but not quantum mechanics. For this remaining of this note, we will rely on the WEP.

2 Review of Vector Calculus

2.1 Operations & Theorems

Recall that a path in 3-dimensional space can by parameterized a single variable, which we will call t, as

$$\vec{r}(t) = x(t)\hat{\imath} + y(t)\hat{\jmath} + z(t)\hat{k},$$

where the tangent to this path is given by

$$\dot{\vec{r}}(t) = \frac{d\vec{r}}{dt}.$$

The length of the underlying curve of this path is the integral over the line element $ds = ||\vec{r}|| = ||\dot{\vec{r}}||\, dt$:

$$L = \int_a^b dS = \int_a^b ||\dot{\vec{r}}||\, dt.$$

Example 2.1. Consider a path in $\mathbb{R}^2$ defined as $\vec{r}(t) = (t, \sin t)$, with $t \in [-2\pi, 2\pi]$. We instantly recognize that the underlying curve of this path is nothing but a sine curve in the xy-plane. The length of this curve is

$$L = \int_{-2\pi}^{2\pi} ||(1, \cos t)||\, dt = \int_{-2\pi}^{2\pi} \sqrt{1 + \cos^2 t}\, dt \approx 15.28.$$

Next, we consider a vector-valued function $\vec{F} : \mathbb{R}^3 \to \mathbb{R}^3$ defined as $\vec{F}(\vec{r}) = (F_x(x, y, z), F_y(x, y, z), F_z(x, y, z))$. Recall the two "del operators" from vector calculus: the **curl** and the **div**, whose definitions are

$$\text{curl}(\vec{F}) = \vec{\nabla} \times \vec{F} = \det \begin{pmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{pmatrix}$$

and

$$\text{div}(\vec{F}) = \vec{\nabla} \cdot \vec{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}.$$

where the "nabla" symbol represents the operator

$$\vec{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right).$$

We have also seen the **gradient** of a scalar field (or potential) $f : \mathbb{R}^3 \to \mathbb{R}$. The gradient is defined as

$$\vec{\nabla} f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right).$$

There are two kinds of potential functions in vector calculus: scalar potential and vector potential. While we are more familiar with the scalar potential (e.g. the gravitational potential or the electric potential in electromagnetism), there isn't as strong a connection between vector potentials and any physical meaning. However, we should still look at some examples.

Example 2.2. In electromagnetism, an electric field $\vec{E}$ is defined as the negative gradient of the electric potential ϕ, measured in Volts:

$$\vec{E} = -\vec{\nabla}\phi.$$

On the other hand, an magnetic field $\vec{B}$ is defined as the curl of the magnetic potential $\vec{A}$.

$$\vec{B} = \vec{\nabla} \times \vec{A}.$$

In this case, A is a vector potential. Note that we don't often talk about a vector potential of an electric field. This is simply because for a point charge, which produces an inverse-square field, the vector potential simply does not exist. The proof is straightforward. For curious readers, this proof can be an interesting exercise (the solution should be a one-liner). Or else, please refer to my vector calculus notes on my website. Hint: we need one of the two following fundamental theorems of vector calculus: Stokes' and Gauss' (or Ostrogradsky's).

Stokes' theorem states that, the circulation around a closed curved in $\mathbb{R}^3$ by a vector field $\vec{F}$ is the flux of $\vec{\nabla} \times \vec{F}$ through the surface S whose boundary $C = \partial S$ is the curve C.

$$\boxed{\oint_{C=\partial S} \vec{F} \cdot d\vec{s} = \iint_S \vec{\nabla} \times \vec{F} \cdot d\vec{S}.}$$

Gauss' theorem states that, if S is closed surface, oriented outward, and if $\vec{F}$ is defined throughout the solid region W enclosed by S, then the flux of $\vec{F}$ through S is the total divergence of S through the region W.

$$\boxed{\oiint_{S=\partial W} \vec{F} \cdot d\vec{S} = \iiint \vec{\nabla} \cdot \vec{F}\, dV.}$$

Line integrals are one of a few important operations in vector calculus. In the context of undergraduate or high school physics, we often use the line integral to calculate the work done by a force $\vec{F}$ over some path $\vec{r}$. The integral in general is simply an accumulation of the component of the vector field along the path.

$$\int_a^b \vec{F} \cdot d\vec{r}.$$

If $\vec{F}$ is a conservative field, then the line integral along $\vec{F}$ between two point A and B only depends on where A and B are, i.e., the line integral is path-independent. A good illustration of this fact is the way we compute the change in electric potential:

$$W = -\int \vec{E} \cdot d\vec{r} = \Delta\phi$$

To actually compute a line integral, we often look for symmetry first (or whether $\vec{F}$ is conservative), then parameterize $\vec{r} = \vec{r}(s)$ if we must. Then, the line integral becomes

$$\int_a^b \vec{F} \cdot d\vec{r} = \int_a^b \vec{F}(\vec{r}(s)) \cdot \frac{d\vec{r}}{ds}\, ds$$

Surface integrals give the flux of a vector field through a surface:

$$\int \vec{F} \cdot d\vec{S} = \int \vec{F} \cdot \vec{n}\, dS = \pm \int \vec{F} \cdot \vec{N}\, d\phi, d\theta,$$

where $\vec{F}$ is the vector field, and $\vec{n}$ is the unit normal vector to the surface. The last expression denotes the surface integral in terms of a parameterization $\vec{S}(\phi, \theta) = (x(\phi, \theta), y(\phi, \theta), z(\phi, \theta))$. $\vec{N}$ is the standard normal to the surface, which is not necessarily a unit vector (with norm of 1) and not necessarily in the same direction as $\vec{n}$, which is intrinsic to the surface.

Example 2.3. Consider the electric flux through a sphere:

$$\Phi_E = \oiint \vec{E} \cdot d\vec{a} = \frac{q}{\epsilon_0}.$$

The last expression comes from Gauss' (or Ostrogradsky's) theorem, equivalently the divergence theorem, where q is the charge enclosed.

We should introduce the Maxwell's equations as they "bring together" the concepts we have touched on so far. To see how, we consider the Maxwell's equations in integral form:

Gauss' law: $\boxed{\oiint \vec{E} \cdot d\vec{a} = \frac{q}{\epsilon_0}}$

No magnetic monopole: $\boxed{\oint \vec{B} \cdot d\vec{a} = 0}$

Faraday's law: $\boxed{\oint \vec{E} \cdot d\vec{s} = -\frac{d}{dt}\Phi_B = \frac{\partial}{\partial t}\int_A \vec{B} \cdot d\vec{a}}$

Ampere-Maxwell's law: $\boxed{\oint\!\!\!\oint \vec{B} \cdot d\vec{s} = \mu_0 I + \mu_0\epsilon_0 \frac{\partial}{\partial t}\int_A \vec{E} \cdot d\vec{a}}$

Using the theorems and facts we have discussed, we can convert the above equations into differential form. First, we can use Gauss' theorem on Gauss' law, with

$$q = \int_V \rho \, d^3r$$

where ρ is the charge volume density. So, Gauss' law becomes

$$\oint \vec{E} \cdot \, d\vec{a} = \int_V \vec{\nabla} \cdot \vec{E} \, d^3r = \frac{1}{\epsilon_0} \int_V \rho \, d^3r,$$

which implies that the divergence of any electric field $\vec{E}$ is proportional to the charge enclosed q.

$$\boxed{\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0}}$$

Applying Gauss' theorem to the second equation,

$$\oiint \vec{B} \cdot \, d\vec{a} = \int_V \vec{\nabla} \cdot \vec{B} \, d^3r = 0$$

we immediately see that the divergence of any magnetic field $\vec{B}$ is 0, i.e., there is no such thing as a "magnetic monopole."

$$\boxed{\vec{\nabla} \cdot \vec{B} = 0}$$

We can apply Stokes' theorem on the other two equations to turn them into differential form. Converting the third equation is as simple as bookkeeping:

$$\oint \vec{E} \cdot \, d\vec{s} = \int_A \vec{\nabla} \times \vec{E} \cdot \, d\vec{a} = -\frac{\partial}{\partial t} \int_A \vec{B} \cdot \, d\vec{a},$$

which says the curl of any electric field $\vec{E}$ is negatively proportional to the change in the magnetic field.

$$\boxed{\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}}$$

To convert the fourth equation into differential form, we define a new quantity, $\vec{J}$, as current (area) density, such that

$$I = \int_A \vec{J} \cdot \, d\vec{a}.$$

Again, applying Stokes' theorem to the last equation,

$$\oint \vec{B} \cdot \, d\vec{s} = \int_A \vec{\nabla} \times \vec{B} \cdot \, d\vec{a} = \mu_0 I + \mu_0 \epsilon_0 \frac{\partial}{\partial t} \int_A \vec{E} \cdot \, d\vec{a} = \mu_0 \int_A \left(\vec{J} + \epsilon_0 \frac{\partial \vec{E}}{\partial t} \right) \cdot \, d\vec{a}.$$

So, in differential form:

$$\boxed{\vec{\nabla} \times \vec{B} = \mu_0 \vec{J} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t}}$$

In a later chapter, we will see how to make these equations fully relativistic. For the curious reader wanting more information regarding this subsection, please refer to a standard textbook on vector calculus, or feel free to use my Vector Calculus lecture notes for a more formal treatment of this subject.

2.2 Coordinate Systems

There are numerous coordinate systems to describe 3-dimensional space. However, we are most familiar with the Cartesian coordinates (x, y, z), the Spherical coordinates (r, θ, ϕ), and the Cylindrical coordinates (ρ, ϕ, z). The Cartesian coordinate system is the most simple, elegant, and "nice" of the three systems, since the (x, y, z) coordinates can also be coefficients of the vector pointing from the origin to any point in space. We will see that this is not the case in general coordinate systems. One such coordinate system is the **spherical coordinate system** (illustrated in Fig. ???), where

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

and the inverse relations are

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2} \\ \theta &= \cos^{-1}\left(\frac{z}{r}\right) = \cos^{-1}\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) \\ \phi &= \tan^{-1}\left(\frac{y}{x}\right) \end{aligned}$$

where $0 \leq r$, $0 \leq \theta \leq \pi$, and $0 \leq \phi \leq 2\pi$.

insert figure here!!!!

The **volume element** is given by

$$dV = dx\, dy\, dz = r^2 \sin\theta\, dr\, d\theta\, d\phi.$$

Consider a sphere of radius R. The **area element** is given by

$$dA = R\, d\theta\, d\phi.$$

In **cylindrical coordinates**,

$$\begin{aligned} x &= \rho \cos\phi \\ y &= \rho \sin\phi \\ z &= z \end{aligned}$$

and the inversion relations are

$$\begin{aligned}
\rho &= \sqrt{x^2 + y^2} \\
z &= z \\
\phi &= \tan^{-1}\left(\frac{y}{x}\right)
\end{aligned}$$

The **volume element** is given by

$$dV = r\,dt\,dz\,d\phi.$$

Consider a cylinder of radius R, the **area element** is given by

$$dA = R\,dz\,d\phi.$$

Insert figure here!!!

The purpose of having the area and volume elements is so that we could integrate over regions, parameterized in spherical or cylindrical coordinates. While it seems like these quantities appear out of the blue, there is a general procedure to find them, given any arbitrary coordinate system. This "procedure" is called the **Jacobian**, which essentially gives the scaling factor as we move from one coordinate system into another. We know that the area element in polar coordinate is $dA = r\,dr\,d\theta$. We can see this is true, using the Jacobian. First, the Jacobian matrix, which is characteristic of a transformation from coordinate system u^i into $u^{j'}$ is defined as

$$\boxed{[U_i^{j'}] = \left[\frac{\partial u^{j'}}{\partial u^i}\right] = \begin{pmatrix} \frac{\partial u^1}{\partial u^{1'}} & \cdots & \frac{\partial u^1}{\partial u^{n'}} \\ \vdots & \ddots & \vdots \\ \frac{\partial u^n}{\partial u^{1'}} & \cdots & \frac{\partial u^n}{\partial u^{n'}} \end{pmatrix}}$$

Note that i and j' are indices, not powers. Now, the general **scaling factor** is defined by the following equality

$$\boxed{N = du^{1'}\,du^{2'}\ldots du^{n'} = \left|\det\left([U_i^j]\right)\right|\,du^1\,du^2\ldots du^n}$$

Example 2.4. The scaling factor as we move from 2-dimensional Cartesian coordinates into polar coordinates is

$$N = \det\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} = \det\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = r.$$

So, the area element is $dA = dx\,dy = r dr\,d\theta$, as expected.

Example 2.5. We can try repeating the same procedure to get the volume element for spherical coordinates, which we have said before to be $dV = r^2 \sin\theta \, dr \, d\theta \, d\phi$. First, define the Jacobian matrix:

$$[U_i^{j'}] = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \phi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \phi} \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix}$$

After some bookkeeping-type rearrangements and computation, we find that

$$N = \left|\det\left[U_i^{j'}\right]\right| = r^2 \sin\theta,$$

as expected. As a check, we can integrate this volume element over a region defined by $x^2 + y^2 + z^2 = R^2$ to find its volume

$$\int_0^R \int_0^\pi \int_0^{2\pi} r^2 \sin\theta \, dr \, d\theta \, d\phi = \frac{4}{3}\pi R^3,$$

which is not surprisingly the volume of the sphere of radius R.

3 Flat 3-dimensional space

Flat, or "Euclidean" space are spaces with **no curvature**. For instance, a flat sheet of paper in $\mathbb{R}^3$ is considered "flat" while the surface of a sphere embedded in $\mathbb{R}^3$ is considered curved space. There should also be a distinction between a sphere in $\mathbb{R}^3$ and a sphere **embedded** into $\mathbb{R}^3$. The former case is simply some subset of flat 3-dimensional space, so it is considered flat. But in the latter case, the sphere is a curved 2-dimensional space, although it lives in flat 3-dimensional space.

Our goal in this section is to see how we can use any arbitrary coordinate system, which specify points in 3-dimensions as an intersection of 3 surfaces. For instance, in Cartesian coordinates, a point (a^1, a^2, a^3) is specified by 3 planes: $x = a^1$, $y = a^2$, and $z = a^3$. In spherical coordinates, a point is specified by intersecting a sphere of radius r, a cone with azimuth angle θ, and a plane of constant ϕ. In cylindrical coordinates, a point is specified by intersecting a cylinder of radius ρ, vertical plane of constant ϕ to the $x-$axis, and the horizontal plane of constant z.

insert figure here!!!

3.1 Curvilinear coordinates

Curvilinear coordinates are an arbitrary coordinate system in Euclidean space where coordinate lines can be curved (hence curvilinear). Note that although coordinate lines can be curvy, the space is still flat. We can call (u, v, w) our arbitrary coordinates, which specify a point (a^1, a^2, a^3) by intersecting three surfaces: $u = a^1$, $v = a^2$, $w = a^3$. Let let relations between (u, v, w) and (x, y, z) be established as

$$\begin{aligned} u = u(x, y, z);\ x = x(u, v, w) \\ v = v(x, y, z);\ y = y(u, v, w) \\ w = w(x, y, z);\ z = z(u, v, w) \end{aligned}$$

Insert figure here!!!

3.2 Basis vectors

3.3 Natural basis

Next, since we want to be able to describe vectors using arbitrary curvilinear coordinates, we need to know a **basis set** of vectors that span the space. In Cartesian coordinates, such a basis set is $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$. Similarly, we want to find a set $\{\vec{e}_u, \vec{e}_v, \vec{e}_w\}$ that would give a basis for an arbitrary curvilinear coordinate system. The way we obtain this set is the same as how we obtain $\{\hat{\imath}, \hat{\jmath}\, \hat{k}\}$ from Cartesian coordinates. We notice that $\hat{\imath}$ is a vector that follows the change in

the x-direction with y and z fixed, i.e., $\hat{i}$ is a tangent vector to a vector $\vec{r}$ along the change in x. Hence

$$\hat{i} = \frac{\partial \vec{r}}{\partial x}.$$

Likewise, $\hat{j} = \frac{\partial \vec{r}}{\partial y}$, and $\hat{k} = \frac{\partial \vec{r}}{\partial z}$. By analogy, we can define the set $\{\vec{e}_u, \vec{e}_v, \vec{e}_w\}$ with

$$\vec{e}_u = \frac{\partial \vec{r}}{\partial u},\ \vec{e}_v = \frac{\partial \vec{r}}{\partial v},\ \vec{e}_w = \frac{\partial \vec{r}}{\partial w},$$

In general, given a coordinate system with $(u^1, u^2, \ldots, u^n)$, we can define a **natural basis set** by letting

$$\boxed{\vec{e}_{u^i} = \frac{\partial \vec{r}}{\partial u^i}}$$

where $i = 0, 1, 2, \ldots, n$ are indices. Note that the set $\{\vec{e}_{u^i}\}$ need not be orthogonal. The only requirement is that they have to be linearly independent and span the space. We also note that they need not be unit vectors. It is possible to make a basis set of only unit vectors by letting

$$\hat{e}_u = \frac{\vec{e}_u}{||\vec{e}_u||},$$

but we will see that this definition is not so useful in general relativity. A question one might ask is: "What is **natural** about this basis set?" The answer, unfortunately, will be provided in the following sections, where we will discuss how they give rise to the **metric tensor**.

We will often use $\{\hat{i}, \hat{j}, \hat{k}\}$ as a reference basis, i.e., we can express $\{\vec{e}_u, \vec{e}_v, \vec{e}_w\}$ in terms of these:

$$\vec{e_u} = (e_u)_x \hat{i} + (e_u)_y \hat{j} + (e_u)_z \hat{k}.$$

Example 3.1. We can look at the basis set for the spherical coordinate system, $\{\vec{e}_r, \vec{e}_\theta, \vec{e}_\phi\}$. First, let the coordinates (u, v, w) be (r, θ, ϕ), the spherical coordinates. A vector $\vec{r}$, expressed in spherical coordinates, has the form:

$$\vec{r} = \begin{pmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{pmatrix},$$

where $r \geq 0$, $0 \leq \theta \leq \pi$, and $0 \leq \phi \leq 2\pi$. To obtain the basis vectors, we simply take $\partial\vec{r}/\partial r$, $\partial\vec{r}/\partial\theta$, and $\partial\vec{r}/\partial\phi$. We expect that our resulting vectors should form the column vectors of the Jacobian matrix $[U_i^{j'}]$, which, as we have discussed previously, has the following form:

$$[U_i^{j'}] = \begin{pmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix}.$$

What we're doing now, in terms of linear algebra, is essentially constructing $[U_i^{j'}]$ - the change-of-basis matrix. Let's see if our described procedure works:

$$\vec{e}_r = \frac{\partial \vec{r}}{\partial r} = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}$$
$$\vec{e}_\theta = \frac{\partial \vec{r}}{\partial \theta} = \begin{pmatrix} r\cos\theta\cos\phi \\ r\cos\theta\sin\phi \\ -r\sin\theta \end{pmatrix}$$
$$\vec{e}_\phi = \frac{\partial \vec{r}}{\partial \phi} = \begin{pmatrix} -r\sin\theta\sin\phi \\ r\sin\theta\cos\phi \\ 0 \end{pmatrix}$$

Not surprisingly,

$$[U_i^{j'}] = \begin{pmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix} = (\vec{e}_r|\vec{e}_\theta|\vec{e}_\phi)\,.$$

Next, we can check if $\vec{e}_r, \vec{e}_\theta, \vec{e}_\phi$ is mutually-perpendicular set. (Note that I avoid the word "orthogonal" here, since it carries different meanings in physics and in linear algebra). We can also check whether the three vectors are unit vectors. It suffices to take the dot products of $\vec{e}_r, \vec{e}_\theta, \vec{e}_\phi$ among themselves:

$$\begin{aligned}
\vec{e}_r \cdot \vec{e}_r &= \sin^2\theta(\cos^2\phi + \sin^2\phi) + \cos^2\theta = 1 \\
\vec{e}_r \cdot \vec{e}_\theta &= r\cos^2\phi\sin\theta\cos\theta + r\sin^2\phi\sin\theta\cos\theta - r\sin\theta\cos\theta = 0 \\
\vec{e}_r \cdot \vec{e}_\phi &= -r\sin^2\theta\sin\phi\cos\phi + r\sin^2\theta\sin\phi\cos\phi = 0 \\
\vec{e}_\theta \cdot \vec{e}_\theta &= r^2\cos^2\phi\cos^2\theta + r^2\cos^2\phi\sin^2\theta + r^2\sin^2\theta = r^2 \\
\vec{e}_\theta \cdot \vec{e}_\phi &= 0 \\
\vec{e}_\phi \cdot \vec{e}_\phi &= r^2\sin^2\phi\sin^2\theta + r^2\sin^2\theta\cos^2\phi = r^2\sin^2\theta.
\end{aligned}$$

We notice that all dot products of cross-coordinates result in 0, which means the basis vectors we found are mutually perpendicular (we can also see this fat geometrically, based on how the spherical coordinates are defined). But we also find that these basis vectors are not unit vectors, since their norms are not necessarily 1:

$$\begin{aligned}
||\vec{e}_r|| &= 1 \\
||\vec{e}_\theta|| &= r \\
||\vec{e}_\phi|| &= r\sin\theta.
\end{aligned}$$

3.4 Dual basis

The procedure we have described in the previous subsection allows us to find a **natural basis** set of vectors for a coordinate system. However, basis sets are

not unique. In this section, we look at an alternative basis set of vectors, called the **dual basis** set, denoted $\{\vec{e}^u, \vec{e}^v, \vec{e}^w, \dots\}$. Instead of deriving these from the finding the tangent vectors, we find the normal of surfaces of constant $u, v, w, \dots$

Recall that the **gradient** operator $\vec{\nabla}$, when applied to a scalar field f, gives the normal vectors to surfaces of constant f. Since curvilinear coordinates are given by coordinate lines where u, v, or w is constant, we can let

$$\begin{aligned}\vec{e}^u &= \vec{\nabla} u \\ \vec{e}^v &= \vec{\nabla} v \\ \vec{e}^w &= \vec{\nabla} w,\end{aligned}$$

and so on, where $\vec{e}^u$ is a normal vector to the surface of constant u. Or more generally,

$$\boxed{\vec{e}^{u^i} = \vec{\nabla} u^i}$$

Now, we might wonder what the dual basis in Cartesian coordinates is. It turns out the dual and natural basis sets in Cartesian coordinates are the same:

$$\begin{aligned}\vec{e}^x &= \vec{\nabla} x = (1,0,0)^\top = \hat{\imath} = \vec{e}_x \\ \vec{e}^y &= \vec{\nabla} y = (0,1,0)^\top = \hat{\jmath} = \vec{e}_y \\ \vec{e}^z &= \vec{\nabla} z = (0,1,0)^\top = \hat{k} = \vec{e}_z.\end{aligned}$$

There's actually no magic behind this "coincidence." The reason the basis sets are the same is simply that the direction of increasing/decreasing x is the same as the normal vector to the surface of constant x, i.e., the yz-plane.

insert figure here!!!

However, we suspect that this is not the case in general curvilinear coordinates, since the direction of increasing u, for instance, should not necessarily be the same as the normal vector to the surface of constant u, as illustrated here:

insert figure here!!!

Example 3.2. We can look at the dual basis set for the spherical coordinate system. Based on what we have established so far, we can compute the dual basis set, using the inverse relations $(x, y, z) \to (r, \theta, \phi)$.

$$\vec{e}^r = \vec{\nabla} r = \vec{\nabla}(x^2+y^2+z^2)^{1/2} = (x^2+y^2+z^2)^{-1/2}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}.$$

Tedious and painstaking computations give us the other two vectors:

$$\vec{e}^{\theta} = \frac{1}{r}\begin{pmatrix} \cos\theta\cos\phi \\ \cos\theta\sin\phi \\ -\sin\theta \end{pmatrix}$$
$$\vec{e}^{\phi} = \frac{1}{r}\begin{pmatrix} -\frac{\sin\phi}{\sin\theta} \\ \frac{\cos\phi}{\sin\theta} \\ 0 \end{pmatrix}$$

We can readily compare $\{\vec{e}^{r}, \vec{e}^{\theta}, \vec{e}^{\phi}\}$ to $\{\vec{e}_r, \vec{e}_\theta, \vec{e}_\phi\}$ and find that while $\vec{e}^{r} = \vec{e}_r$, $\vec{e}^{\theta} \neq \vec{e}_\theta$ and $\vec{e}^{\phi} \neq \vec{e}_\phi$.

Example 3.3. Let's consider a non-familiar case of paraboloidal surfaces (u, v, w), a non-orthogonal set, where

$$\begin{aligned} x &= u + v \\ y &= u - v \\ z &= 2uv + w, \end{aligned}$$

which gives the inverse relations

$$\begin{aligned} u &= \frac{1}{2}(x + y) \\ v &= \frac{1}{2}(x - y) \\ w &= z - \frac{1}{2}(x^2 - y^2). \end{aligned}$$

Let $\vec{r} = (x, y, z)^\top = (u + v, u - v, 2uv + w)^\top$. We first obtain the natural basis set:

$$\begin{aligned} \vec{e}_u &= \frac{\partial \vec{r}}{\partial u} = (1, 1, 2v)^\top \\ \vec{e}_v &= \frac{\partial \vec{r}}{\partial v} = (1, 1, 2u)^\top \\ \vec{e}_w &= \frac{\partial \vec{r}}{\partial w} = (0, 0, 1)^\top, \end{aligned}$$

which is an mutually perpendicular set, by inspection (i.e., all cross-dot product terms are zero). Next, we obtain the dual basis set:

$$\begin{aligned} \vec{e}^{u} &= \vec{\nabla} u = \frac{1}{2}\vec{\nabla}(x + y) = \left(\frac{1}{2}, \frac{1}{2}, 0\right)^\top \\ \vec{e}^{v} &= \vec{\nabla} v = \frac{1}{2}\vec{\nabla}(x - y) = \left(\frac{1}{2}, -\frac{1}{2}, 0\right)^\top \\ \vec{e}^{w} &= \vec{\nabla} w = \vec{\nabla}\left(z - \frac{1}{2}(x^2 - y^2)\right) = (-x, y, 1)^\top = (-u - v, u - v, 1)^\top \end{aligned}$$

But notice that we get non-zero cross-dot product terms:

$$\begin{aligned}\vec{e}^u \cdot \vec{e}^w &= -v \\ \vec{e}^u \cdot \vec{e}^v &= 0 \\ \vec{e}^v \cdot \vec{e}^w &= -u.\end{aligned}$$

The reason for observing what value the dot products between basis vectors are will become clear when we talk about the **metric tensor** in the next subsections.

3.5 Contravariant and covariant vectors

3.5.1 The suffix notation (vector calculus)

In general coordinate systems, there can be many more than three coordinates, and it can be burdensome to keep track of different letters representing these coordinates, especially if they don't follow a convention. It is convenient to change our notation from "letter-based" to "index-based." For example, instead of using (u, v, w) as coordinates, we will now use (u^1, u^2, u^3), for $i = 1, 2, 3$. Note that we are using upper indices here. As we will discuss soon, upper and lower indices have different meanings.

Similarly, our notation for basis vectors also change. For the natural basis, we "code" $\{\vec{e}_u, \vec{e}_v, \vec{e}_w\}$ as $\{\vec{e}_i\}$, where $i = 1, 2, 3$. For the dual basis, we "code" $\{\vec{e}^u, \vec{e}^v, \vec{e}^w\}$ as $\{\vec{e}^i\}$ for $i = 1, 2, 3$. Our vector notation also changes as a result, but there are additional rules. Let vector $\vec{\lambda}$ be given. In the **natural basis**, the components of λ is written with **upper** indices,

$$\vec{\lambda} = \lambda^1 \vec{e}_1 + \lambda^2 \vec{e}_2 + \lambda^3 \vec{e}_3 = \sum_{i=1}^{3} \lambda^1 \vec{e}_i.$$

and in the **dual basis**, we use **lower** indices for the components:

$$\vec{\lambda} = \lambda_1 \vec{e}^1 + \lambda_2 \vec{e}^2 + \lambda_3 \vec{e}^3 = \sum_{i=1}^{3} \lambda_i \vec{e}^i$$

3.5.2 The Einstein summation notation

We shall make another "leap in notation" as we introduce the Einstein summation convention:

Any index that appears once up and once down is automatically summed

For example, from now on, we will write $\vec{\lambda}$ as:

$$\vec{\lambda} = \sum_{i=1}^{3} \lambda^i \vec{e}_i = \lambda^i \vec{e}_i$$

Note that this summation convention removes us from indicating the minimum and maximum values for our indices, which allows us to keep the mathematics as applicable to any n-dimensional system as possible. We should also notice that indices can be exchangeable, especially in summed situations. For example,

$$\boxed{a^i b_i = a^j b_j = \cdots = \sum_{n=1}^{N} a^n b_n}$$

Keep in mind that $a_i b_i$ does not make sense in the Einstein convention. If we want to denote a sum, we have to specify the range of i as follows:

$$\sum_{i=n}^{N} a_i b_i.$$

Likewise, $a_i b^i c^i$ is not defined either. Remember that only "1 up index, 1 down index" is allowed. Finally, it is widely agreed upon that certain letters are reserved for some specific cases. For instance, $i, j, k, l, \cdots = 1, 2, 3$ are reserved for 3-dimensional space. We use Greek characters $\mu, \nu, \alpha, \beta, \cdots = 0, 1, 2, 3$ for 4-dimensional spacetime, $A, B, C, \cdots = 1, 2$ for 2-dimensional spaces, and $a, b, c, \cdots = 1, 2, \ldots, N$ for N-dimensional manifolds.

3.5.3 Contravariant and covariant vectors

Earlier, we have established that any vector $\vec{\lambda}$ can be written in two "upper-lower index" ways, depending on whether we are using the natural or the dual basis set:

$$\vec{\lambda} = \lambda^i \vec{e}_i = \lambda_i \vec{e}^{\,i}.$$

We shall now call λ^i the **contravariant** components and λ_i the **covariant** components. (Remember: "co" is "low"). Contravariant vectors are simply vectors with contravariant components. Likewise, covariant vectors are vectors with covariant components. Later in this text, we will denote $\vec{\lambda}$ simply as λ^i (contravariant vector) or λ_i (covariant vector).

Putting together what we have established so far about index notation and Einstein summation convention, we can now look at the last interesting dot product: between a vector in the natural basis and one in the dual basis: $\vec{e}^{\,i} \cdot \vec{e}_j$. Before we proceed, keep in mind that this single expression $(\vec{e}^{\,i} \cdot \vec{e}_j)$ represents 9 different objects because there is not summation here (i is not necessarily j and $i, j = 1, 2, 3$). Now, recall the definitions of these vectors:

$$\vec{e}^{\,i} = \vec{\nabla} u^i = \frac{\partial u^i}{\partial x}\hat{\imath} + \frac{\partial u^i}{\partial y}\hat{\jmath} + \frac{\partial u^i}{\partial z}\hat{k}$$
$$\vec{e}_j = \frac{\partial \vec{r}}{\partial u^j} = \frac{\partial x}{\partial u^j}\hat{\imath} + \frac{\partial y}{\partial u^j}\hat{\jmath} + \frac{\partial z}{\partial u^j}\hat{k}$$

Therefore,

$$\vec{e}^{\,i} \cdot \vec{e}_j = \frac{\partial u^i}{\partial x}\frac{\partial x}{\partial u^j} + \frac{\partial u^i}{\partial y}\frac{\partial y}{\partial u^j} + \frac{\partial u^i}{\partial z}\frac{\partial z}{\partial u^j}$$

But notice that this is nothing but the chain rule, if we think about u^i as $u^i(x, y, z)$. So,

$$\boxed{\vec{e}^{\,i} \cdot \vec{e}_j = \frac{\partial u^i}{\partial u^j}}$$

But also notice that the u^i's are independent variables, i.e.,

$$\frac{\partial u^i}{\partial u^j} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

We can introduce a two-index object, the **Kronecker delta**, defined as

$$\boxed{\delta^i_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}}$$

We obtain an elegant relation:

$$\boxed{\vec{e}^{\,i} \cdot \vec{e}_j = \delta^i_j}$$

Again, keep in mind that this is not a single equation but rather 9. Six of which are $0 = 0$ and three $1 = 1$. We observe that for $u \neq v$, $\vec{e}^{\,u} \cdots \vec{e}_v = 0$, i.e., $\vec{e}^{\,u} \perp \vec{e}_v$, which makes sense, by definition:

insert figure here!!!

As a little aside, every time we see an two-index object, we should immediately think of a matrix. In matrix form, the Kronecker delta is our beloved identity matrix I:

$$[\delta^i_j] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

What about the dot (or inner) products of $\{\vec{e}^{\,i}\}$ and $\{\vec{e}_j\}$ among themselves? We have calculated the terms before, when we introduced spherical coordinates, but we have not thought about their significance. Let us define another two-index object, called the **metric tensor** by

$$\boxed{g_{ij} = \vec{e}_i \cdot \vec{e}_j}$$

which in matrix form is:

$$[g_{ij}] = \begin{pmatrix} \vec{e}_1 \cdot \vec{e}_1 & \vec{e}_1 \cdot \vec{e}_2 & \vec{e}_1 \cdot \vec{e}_3 \\ \vec{e}_2 \cdot \vec{e}_1 & \vec{e}_2 \cdot \vec{e}_2 & \vec{e}_2 \cdot \vec{e}_3 \\ \vec{e}_3 \cdot \vec{e}_1 & \vec{e}_3 \cdot \vec{e}_2 & \vec{e}_3 \cdot \vec{e}_3 \end{pmatrix}.$$

And likewise, we define the **inverse metric tensor** as

$$\boxed{g^{ij} = \vec{e}^{\,i} \cdot \vec{e}^{\,j}}$$

which in matrix form is:

$$[g^{ij}] = \begin{pmatrix} \vec{e}^{\,1} \cdot \vec{e}^{\,1} & \vec{e}^{\,1} \cdot \vec{e}^{\,2} & \vec{e}^{\,1} \cdot \vec{e}^{\,3} \\ \vec{e}^{\,2} \cdot \vec{e}^{\,1} & \vec{e}^{\,2} \cdot \vec{e}^{\,2} & \vec{e}^{\,2} \cdot \vec{e}^{\,3} \\ \vec{e}^{\,3} \cdot \vec{e}^{\,1} & \vec{e}^{\,3} \cdot \vec{e}^{\,2} & \vec{e}^{\,3} \cdot \vec{e}^{\,3} \end{pmatrix}.$$

Since dot products are commutative, i.e, $\vec{e}^{\,i} \cdot \vec{e}^{\,j} = \vec{e}^{\,j} \cdot \vec{e}^{\,i}$ and $\vec{e}_i \cdot \vec{e}_j = \vec{e}_j \cdot \vec{e}_i$, we get an interesting property:

$$\boxed{g_{ij} = g_{ji} \text{ and } g^{ij} = g^{ji}}$$

In other words, the matrices $[g_{ij}]$ and $[g^{ij}]$ are **symmetric**. Just for sanity check, we can look at $[g_{ij}]$ for Cartesian coordinates, which (not surprisingly), turns out to be the identity matrix, i.e., $[g_{ij}] = [g^{ij}] = [\delta^i_j] = I$, which is symmetric. However, $[g_{ij}]$ are not **diagonal**, in general. A counter-example is $[g_{ij}]$ in paraboloidal coordinates.

We are of course not limited to looking only at dot products of the basis vectors among themselves. Consider the dot product of two vectors $\vec{\mu} = \mu^i \vec{e}_i = \mu_i \vec{e}^{\,i}$ and $\vec{\lambda} = \lambda^i \vec{e}_i = \lambda_i \vec{e}^{\,i}$. We immediately see that there are more than one way to express $\vec{\mu} \cdot \vec{\lambda}$, all of which are equivalent. Now, we have to be careful with writing the dot product with indices, because $\lambda^i \vec{e}^{\,i} \mu^i \cdot \vec{e}^{\,i}$ is not defined in the Einstein summation convention. We have to realize that the indices of $\vec{\mu}$ and $\vec{\lambda}$ should be independent, so it suffices to replace i with j as indices for one of the vectors. We are now safe to proceed.

$$\boxed{\vec{\lambda} \cdot \vec{\mu} = \lambda^i \vec{e}_i \cdot \mu^j \vec{e}_j = \lambda_i \vec{e}^{\,i} \cdot \mu_j \vec{e}^{\,j} = \lambda_i \vec{e}^{\,i} \cdot \mu^j \vec{e}_j = \lambda^i \vec{e}_i \cdot \mu_j \vec{e}^{\,j}}$$

We have shown that $\vec{e}^{\,i} \cdot \vec{e}_j \delta^i_j$, $\vec{e}^{\,i} \cdot \vec{e}^{\,j} = g^{ij}$, and $\vec{e}_i \cdot \vec{e}_j = g_{ij}$, we obtain a number of relations:

$$\begin{aligned} \vec{\lambda} \cdot \vec{\mu} = \lambda^i \vec{e}_i \cdot \mu^j \vec{e}_j &= g_{ij} \lambda^i \mu^j \\ &= \lambda_i \vec{e}^{\,i} \cdot \mu_j \vec{e}^{\,j} = g^{ij} \lambda_i \mu_j \\ &= \lambda_i \vec{e}^{\,i} \cdot \mu^j \vec{e}_j = \delta^i_j \lambda_i \mu^j = \lambda_i \mu^i = \lambda_j \mu^j \\ &= \lambda^i \vec{e}_i \cdot \mu_j \vec{e}^{\,j} = \delta^j_i \lambda^i \mu_j = \lambda^i \mu_i = \lambda^j \mu_j \end{aligned}$$

We can summarize the statements above into five equivalent statements:

$$\boxed{\vec{\lambda} \cdot \vec{\mu} = g_{ij}\lambda^i\mu^j = g^{ij}\lambda_i\mu_j = \lambda_i\mu^i = \lambda^i\mu_i}$$

From the second and fourth expressions, we get

$$\boxed{g_{ij}\mu^j = \mu_i}$$

Likewise, if we look at the third and fifth expressions:

$$\boxed{g^{ij}\lambda_i = \lambda^j}$$

This means we can use the metric tensor g_{ij} and the inverse metric tensor g^{ij} to go back and forth between covariant and covariant vector components. Or, in a more "bookkeeping" sense, g^{ij} raises an index, while g_{ij} lowers an index. These operations are useful to remember, since they will become especially handy when we have objects with more than one upper and lower indices.

But we can extract even more information from the above two equalities by combining them:

$$\mu^i = g^{ij}\mu_j = g^{ij}\left(g_{jk}\mu^k\right).$$

Note that we need an extra index k here since the indices of the expansion of μ_j are independent of the indices of g^{ij}, only except for j. Now, since it is also true that $\mu^i = \delta^i_j\mu^j$, the following has to be true as well:

$$\boxed{g^{ij}g_{jk} = \delta^i_k}$$

We obtain a similar equality if we start with the covariant component μ_i:

$$\boxed{g_{ij}g^{jk} = \delta^k_i}$$

The above two equalities imply that the matrices $[g^{ij}]$ and $[g_{ij}]$ are **inverses** of each other. Hence, we call g_{ij} the **metric tensor** and g^{ij} the **inverse metric tensor**.

Example 3.4. Consider $\vec{a} = \vec{e}_\theta$ in spherical coordinates. Express $\vec{a}$ with contravariant components, covariant components, and find the norm of $\vec{a}$, with and without the metric g_{ij}.

Since $\vec{a} = \vec{e}_\theta$, a single natural basis vector, we can easily write $\vec{a}$ as a contravariant vector as

$$a^i = (a^1, a^2, a^3)^\top = (0, 1, 0)^\top.$$

We can calculate the covariant components of $\vec{a}$ using the metric tensor g_{ij}:

$$\begin{aligned} a_1 &= g_{1j}a^j = g_{11}a^1 = 0 \\ a_2 &= g_{2j}a^j = g_{22}a^2 = r^2 \\ a_3 &= g_{3j}a^j = g_{33}a^3 = 0. \end{aligned}$$

So, $\vec{a}$, with covariant components, is

$$a_i = (0, r^2, 0)^\top$$

Next, we can calculate $||\vec{a}||$ in two ways, and both will give the same answer:

$$\begin{aligned} ||\vec{a}||^2 &= a^i a_i = a^1 a_1 + a^2 a_2 + a^3 a_3 = r^2 \\ ||\vec{a}||^2 &= g_{ij}a^i a^j = g_{11}a^1 a^1 + g_{22}a^2 a^2 + g_{33}a^3 a^3 = r^2. \end{aligned}$$

So, $||\vec{a}|| = r$.

3.6 Metric tensor

In the previous subsection we have defined the metric tensor and the inverse metric tensor. In this subsection, we will discuss their physical significance and related properties.

To start, the metric tensor contains information about physical lengths and the geometry of the space. To illustrate this point, let us consider a path $\vec{r} = \vec{r}(t)$ in flat 3-dimensional space (Euclidean). For vector calculus, we know that the length of the underlying curve, denoted as L, is given by

$$L = \int_a^b ||\dot{\vec{r}}||\, dt.$$

Let this curve be expressed in a general curvilinear coordinate system with coordinates (u, v, w), i.e., $\vec{r}(t) = (u(t), v(t), w(t))$, it follows that

$$\begin{aligned} \dot{\vec{r}} &= \frac{\partial \vec{r}}{\partial u}\frac{du}{dt} + \frac{\partial \vec{r}}{\partial v}\frac{dv}{dt} + \frac{\partial \vec{r}}{\partial w}\frac{dw}{dt} \\ &= \vec{e}_u \frac{du}{dt} + \vec{e}_v \frac{dv}{dt} + \vec{e}_w \frac{dw}{dt} \\ &= \vec{e}_i \frac{du^i}{dt}, \end{aligned}$$

in Einstein's notation. So,

$$\left|\left|\dot{\vec{r}}\right|\right| = \sqrt{\dot{\vec{r}} \cdot \dot{\vec{r}}} = \sqrt{\vec{e}_i \frac{du^i}{dt} \cdot \vec{e}_j \frac{du^j}{dt}} = \sqrt{g_{ij}\frac{du^i}{dt}\frac{du^j}{dt}}$$

Putting everything (so far) together, we get

$$\boxed{L = \int_a^b \left|\left|\frac{d\vec{r}}{dt}\right|\right| dt = \int_a^b \sqrt{g_{ij}\frac{du^i}{dt}\frac{du^j}{dt}}\, dt}$$

Let ds the infinitesimal distance, we get

$$ds = \sqrt{g_{ij} \frac{du^i}{dt} \frac{du^j}{dt}}\, dt.$$

Bringing the dt into the square root, and squaring both sides, we get

$$\boxed{ds^2 = g_{ij}\, du^i\, du^j}$$

Even though ds^2 has units of length squared, we still shall refer to this quantity as the **line element**. Let us look at a few examples to see how the above equality manifests in geometries we are familiar with.

Example 3.5. In Cartesian coordinates, the natural basis vectors are $\{\vec{e}_i\}\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$. Observe that $g_{ij} = \vec{e}_i \cdot \vec{e}_j = \delta_{ij}$, since the $\vec{e}_i$'s are both unit vectors and mutually perpendicular. So, as a matrix:

$$[g_{ij}] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The line element, $ds^2 = g_{ij}\, du^i\, du^j$, is a sum of nine terms. However, we notice that $g_{ij} = 0$ if $i \neq j$ and 1 otherwise. Therefore the line element is:

$$ds^2 = g_{ij}\, du^i\, du^j = dx^2 + dy^2 + dz^2,$$

which we are very familiar with. Indeed, the moral of this example is to show that while the form of the line element can be derived geometrically, a rigorous way to obtain the line element is actually through the metric. For Cartesian coordinates, deriving the line element is more or less trivial, because we realize that the above equality is simply Pythagorean theorem in the three dimensions. However, as we will see in the next two examples, it is more difficult to obtain the line element by "inspection," and that realizing the line element from the metric is more reliable.

Example 3.6. We have shown that, in spherical coordinates:

$$\begin{aligned} \vec{e}_r \cdot \vec{e}_r &= 1 \\ \vec{e}_\theta \cdot \vec{e}_\theta &= r^2 \\ \vec{e}_\phi \cdot \vec{e}_\phi &= r^2 \sin^2\theta \\ \vec{e}_r \cdot \vec{e}_\theta = \vec{e}_r \cdot \vec{e}_\phi &= \vec{e}_\theta \cdot \vec{e}_\phi = 0, \end{aligned}$$

which give the metric, in matrix form, as:

$$[g_{ij}] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2\theta \end{pmatrix}.$$

Therefore, the line element in spherical coordinates is:

$$ds^2 = g_{ij}\, du^i\, du^j = dr^2 + r^2\, d\theta^2 + r^2 \sin^2\theta\, d\phi^2.$$

While one can certainly derive this expression geometrically, it is no doubt more challenging to do than in Cartesian coordinates.

Example 3.7. Find the length of a curve in spherical coordinates with the parameterization $\vec{r}(t) = (r(t), \theta(t), \phi(t)) = (t, \pi/4, 4t)$, where $0 \leq t \leq \pi$.

It is possible to proceed without knowing what the underlying curve of this path is. However, it certainly does not harm to know what length we are trying to find here. First, notice that θ is fixed, so we are limited to a cone, oriented vertically along the z-axis. Now, since $0 \leq t \leq \pi$, and that $\phi = 4t$, we know that the curve makes two revolutions around the z-axis. So we can imagine the curve looking something like this:

Insert figure here!!!

Find the length of a curve is only "mechanical" once we already know the line element ds^2, which we have derived in the previous example. Plus, we know that $d\theta = 0$ since θ is constant. So,

$$\begin{aligned}
L &= \int ds \\
&= \int_0^\pi \sqrt{\frac{dr^2}{dt^2} + r^2 \frac{d\theta^2}{dt^2} + r^2 \sin^2\theta \frac{d\phi^2}{dt^2}}\, dt \\
&= \int_0^\pi \sqrt{1 + 0 + \left(4t \sin\left(\frac{\pi}{4}\right)\right)^2}\, dt \\
&= \int_0^\pi \sqrt{1 + 8t^2}\, dt \\
&\approx 14.55.
\end{aligned}$$

As we have seen before in the earlier sections, the metric tensor does not just encode physical lengths. g_{ij} are also ubiquitous in calculating norms of vectors and inner products (or dot products). As a reminder:

$$\begin{aligned}
&\textbf{Norm: } ||\vec{\lambda}||^2 = \vec{\lambda} \cdot \vec{\lambda} = g_{ij} \lambda^i \lambda^j \\
&\textbf{Inner product: } \vec{\lambda} \cdot \vec{\mu} = g_{ij} \lambda^i \mu^j.
\end{aligned}$$

They can also be used to raise/lower indices, as we have seen.

3.6.1 Summations as matrix products

As a little "aside," writing vectors and 2-component tensors (such as g_{ij}) as matrices can be more convenient for us. But note that more general tensors

(with three or more indices) can no longer be written as matrices.

Suppose that we have two square matrices $\tilde{A} = [a_{ij}]$ and $\tilde{B} = [b_{ij}]$. Let matrix $\tilde{C} = \tilde{A}\tilde{B}$. In matrix multiplication,

$$\boxed{\tilde{C} = \begin{pmatrix} & \vdots & \\ \dots & c_{ij} & \dots \\ & \vdots & \end{pmatrix} = \begin{pmatrix} \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \end{pmatrix} \begin{pmatrix} \dots & b_{1j} & \dots \\ & \vdots & \\ \dots & b_{nj} & \dots \end{pmatrix}}$$

i.e., the ij^{th} element of $\tilde{C}$ is

$$\boxed{c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}}$$

where k on the a term indicates a column, while k on the b term indicates a row. Notice that the summed index (k) goes as "column-row," whereas $i - j$ go as row-column. Note that I do not have a restriction on n, because the above relations are true in any n-dimensional space.

Similarly, we can multiply vectors in matrix notations. Let two vectors $\vec{F}, \vec{G} \in \mathbb{R}^n$ be given. Assuming that the metric is the identity, their inner product is given by:

$$\boxed{\vec{F} \cdot \vec{G} = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} \cdot \begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix} = \vec{F}^\top \vec{G} = \sum_{k=1}^{n} f_k g_k.}$$

With a general metric g_{ij}, we want to express $\vec{\lambda} \cdot \vec{\mu} = g_{ij} \lambda^i \mu^j$ as a matrix multiplication. We have to make sure that the matrix multiplication is defined, i.e., the rows of the right-multiplying element is equal to the columns of the left-multiplying element. There are two things we can do: ordering the ij indices, and transposing. Let $\tilde{\lambda}$ denote $\vec{\lambda} = [\lambda^i]$, and $\tilde{\mu}$ denote $\vec{\mu} = [\mu^i]$.

For contravariant vectors:

Let $\tilde{G} = [g_{ij}]$, where i is the row indicator and j is the column indicator. This means that the index of $\vec{\lambda}$ has to be a column indicator, so we transpose $\vec{\lambda}$. The index of μ is a row indicator, which matches with j-the column indicator of $\tilde{G}$, so we are good to proceed. Again, note that n is not necessarily 3.

$$\boxed{\vec{\lambda} \cdot \vec{\mu} = \lambda^i g_{ij} \mu^j = \begin{pmatrix} \lambda_1 & \dots & \lambda_n \end{pmatrix} \begin{pmatrix} g_{11} & \dots & g_{1n} \\ \vdots & \ddots & \vdots \\ g_{n1} & \dots & g_{nn} \end{pmatrix} \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \tilde{\lambda}^\top \tilde{G} \tilde{\mu}}$$

For covariant vectors:

We can find express the inner product of covariant vectors in matrix multiplication, using the form we have found for contravariant products. Consider two covariant vectors $\tilde{\lambda}^* = [\lambda_i]$ and $\tilde{\mu}^* = [\mu_i]$ and the inverse metric tensor $\tilde{G}^{-1} = [g_{ij}]^{-1} = [g^{ij}]$ (recall that $g_{ij}g^{jk} = \delta_i^k$). Similar to the contravariant vector case, but now with $\vec{\lambda} \cdot \vec{\mu} = g^{ij}\lambda_i\mu_j$:

$$\boxed{\tilde{\lambda}^* \cdot \tilde{\mu}^* = \begin{pmatrix} \lambda_1 & \dots & \lambda_n \end{pmatrix} = \begin{pmatrix} g^{11} & \dots & g^{1n} \\ \vdots & \ddots & \vdots \\ g^{n1} & \dots & g^{nn} \end{pmatrix} \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \tilde{\lambda}^{*\top}\tilde{G}^{-1}\tilde{\mu}}$$

As an little aside, to write $\tilde{\lambda}^*$ in terms of the metric tensor matrix and the contravariant vector, we simply remember that $\lambda_i = g_{ij}\lambda^j$. So,

$$\tilde{\lambda}^* = \tilde{G}\tilde{\lambda}$$
$$\tilde{\lambda} = \tilde{G}^{-1}\tilde{\lambda}^*$$

In matrix form, the above two equalities can be written as:

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} g_{11} & \dots & g_{1n} \\ \vdots & \ddots & \vdots \\ g_{n1} & \dots & g_{nn} \end{pmatrix} \begin{pmatrix} \lambda^1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

and

$$\begin{pmatrix} \lambda^1 \\ \vdots \\ \lambda^n \end{pmatrix} = \begin{pmatrix} g^{11} & \dots & g^{1n} \\ \vdots & \ddots & \vdots \\ g^{n1} & \dots & g^{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

respectively.

Metric tensor and inverse metric tensor

As mentioned, $[g^{ij}] = [g_{ij}]^{-1}$. In matrix form:

$$\boxed{\tilde{G}^{-1}\tilde{G} = \begin{pmatrix} g^{11} & \dots & g^{1n} \\ \vdots & \ddots & \vdots \\ g^{n1} & \dots & g^{nn} \end{pmatrix} \begin{pmatrix} g_{11} & \dots & g_{1n} \\ \vdots & \ddots & \vdots \\ g_{n1} & \dots & g_{nn} \end{pmatrix} = \operatorname{diag}(1, \dots, 1) = \tilde{I}}$$

Example 3.8. Find the inverse metric tensor g^{ij} in spherical coordinates, given the metric tensor g_{ij} is:

$$[g_{ij}] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix}$$

By definition, $[g^{ij}] = [g_{ij}]^{-1}$. So taking the inverse of $[g_{ij}]$ will give us $[g^{ij}]$. But notice that $[g_{ij}]$ is diagonal, so g^{ij} is simply the reciprocal of g_{ij}, if $g_{ij} \neq 0$, and 0 otherwise. This gives

$$[g^{ij}] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^{-2} & 0 \\ 0 & 0 & r^{-2}\sin^{-2}\theta \end{pmatrix}$$

3.7 Coordinate transformation

We have already seen the usage of multiple (more or less conventional) coordinate systems in the three-dimensional space. In general n-dimensional spaces, the options for coordinate systems are almost unlimited. It is crucial for us, as relativists, to develop for rigorous procedure to transform vector and tensor components as we move from one coordinate system to another. Note that vector and tensor **components** change under general coordinate transformation. It is common in physics to associate a vector or a tensor with a physical quantity, and coordinate system, or reference frame, etc., with perspectives. It makes sense that what we see might differ at different perspectives, but the physical quantity itself is independent of how we view it. In other words, vectors and tensors do not change under general coordinate transformations, only their components do.

In this subsection, we will learn how to transform between arbitrary coordinate systems. We will also learn how vector and tensor components transform, as well as what vectors and tensors are.

3.7.1 Vectors

We as physicists often think of vectors as an object with magnitude and direction. But this definition is incomplete, as it disregards a reference frame in which the vector is viewed. As we have mentioned in the introduction to this section, under general coordinate transformations, vectors do not change, but their components do, since the basis set of the coordinate system changes. For instance,

insert figure here!!!

insert figure here!!!

In the above two figures, the vector $\vec{\lambda}$ is the same. However, one frame is rotated by θ from another. In frames (S) and (S'), $\vec{\lambda}$ will have different components, but the "physical object" $\vec{\lambda}$ does not change. For instance, we can calculate the norm of $\vec{\lambda}$, $||\vec{\lambda}||$, and find that it is the same in both frames. In fact, **scalars are invariant under general coordinate transformations**.

Next, let's talk **notations**. Let's assume, without loss of generality, that we are only dealing with contravariant vectors here (the rules we will establish

shortly will apply to covariant vectors as well). Consider $\vec{\lambda}$ in frame (S). In index-notation, we can write $\vec{\lambda} = \{\lambda^i\}$. Consider a different frame (a "primed" frame), denoted (S'). We will denote the components of $\vec{\lambda}$ in (S') as $\lambda^{i'}$. Just to get us familiar with the new notation, let's go through an example:

Example 3.9. Suppose that $\vec{\lambda}$ is described two coordinate systems: spherical and cylindrical. Let the "unprimed" coordinate system be the spherical coordinate system. Then,

$$\{u^i\} = (r, \theta, \phi)$$
$$\{u^{i'}\} = (\rho, \phi, z).$$

As simple as that!

Back to (S) and (S'). Since these are two different coordinate systems, the basis sets and metric tensors are also different:

$$(S) : \vec{e}_i = \frac{\partial \vec{r}}{\partial u^i},\ \vec{e}^{\,i} = \vec{\nabla} u^i,\ g_{ij} = \vec{e}_i \cdot \vec{e}_j$$
$$(S') : \vec{e}_{i'} = \frac{\partial \vec{r}}{u^{i'}},\ \vec{e}^{\,i'} = \vec{\nabla} u^{i'},\ g_{i'j'} = \vec{e}_{i'} \cdot \vec{e}_{j'}.$$

Now, since $\vec{\lambda}$ is the same in both (S) and (S'), it must be true that"

$$\vec{\lambda} = \lambda^i \vec{e}_i = \lambda^{i'} \vec{e}_{i'},$$

i.e., λ^i and $\vec{e}_i$ must transform in a way that leaves $\vec{\lambda}$ unchanged. Consider $\vec{r} = \vec{r}(u^{i'}) = \vec{r}(u^{i'}(u^j))$. By the chain rule:

$$\vec{e}_j = \frac{\partial \vec{r}}{\partial u^j} = \frac{\partial \vec{r}}{\partial u^{i'}} \frac{\partial u^{i'}}{\partial u^j} = \vec{e}^{\,i'} \frac{\partial u^{i'}}{\partial u^j}.$$

Let us define, for simplicity sake,

$$\boxed{U^{i'}_j = \frac{\partial u^{i'}}{\partial u^j}}$$

Note that there is nothing new here. $U^{i'}_j$ simply generalizes 9 partial derivatives (because we are in 3-dimensional space). Together as a matrix, $[U^{i'}_j]$ is nothing but the Jacobian matrix "representing" the transformation $(S) \to (S')$:

$$[U^{i'}_j] = \begin{pmatrix} \frac{\partial u^{1'}}{\partial u^1} & \frac{\partial u^{1'}}{\partial u^2} & \frac{\partial u^{1'}}{\partial u^3} \\ \frac{\partial u^{2'}}{\partial u^1} & \frac{\partial u^{2'}}{\partial u^2} & \frac{\partial u^{2'}}{\partial u^3} \\ \frac{\partial u^{3'}}{\partial u^1} & \frac{\partial u^{3'}}{\partial u^2} & \frac{\partial u^{3'}}{\partial u^3} \end{pmatrix}$$

So, the chain rule for $\vec{e}_j$ becomes the general coordinate transformation rule for the natural basis sets $\vec{e}_{i'} \to \vec{e}_i$.

$$\boxed{\vec{e}_j = U^{i'}_j \vec{e}_{i'}}$$

Next, since $\vec{\lambda} = \lambda^{i'}\vec{e}_{i'} = \lambda^j\vec{e}_j = \lambda^j U^{i'}_j \vec{e}_{i'}$, we obtain the general coordinate transformation rule for contravariant vector components $\lambda^i \to \lambda^{i'}$:

$$\boxed{\lambda^{i'} = U^{i'}_j \lambda^j}$$

We can, of course, conjecture that the transformation rule for contravariant vector components $\lambda^{i'} \to \lambda^i$ is:

$$\boxed{\lambda^{i'} = U^j_{i'} \lambda^{i'}}$$

and we would be correct. The "proof" follows here. First, we look at the Jacobian matrix representing the transformation $(S') \to (S)$:

$$[U^j_{i'}] = \left[\frac{\partial u^j}{\partial u^{i'}}\right],$$

which is simply the inverse of $\left[U^{i'}_j\right]$. So, it follows that:

$$[U^{j'}_k][U^k_{i'}] = \text{diag}(1, \ldots, 1) = [\delta^{j'}_{i'}] = [\delta^j_i] = \tilde{I},$$

and

$$[U^k_{i'}][U^{i'}_j] = \text{diag}(1, \ldots, 1) = [\delta^k_j] = \tilde{I},$$

because matrices that are inverses of each other are commutative. Element-wise:

$$\boxed{U^k_{i'} U^{i'}_j = \delta^k_j = \delta^{k'}_{j'} = U^{k'}_i U^i_{j'}}$$

So we can now verify our little "conjecture:"

$$\begin{aligned}
\lambda^{i'} &= U^{i'}_j \lambda^j \\
U^k_{i'} \lambda^{i'} &= U^{i'}_j U^k_{i'} \lambda^j \\
U^k_{i'} \lambda^{i'} &= \delta^k_j \lambda^j \\
U^k_{i'} \lambda^{i'} &= \lambda^k.
\end{aligned}$$

We notice that nothing makes $U^{i'}_j$ more special than $U^i_{j'}$. Also, while we use contravariant vectors more often than covariant vectors, nothing makes contravariant vectors more special than covariant vectors. So if transformation rules apply to contravariant vectors, we should also expect them to work for covariant vectors:

$$\boxed{\begin{aligned}
\lambda_{i'} &= U^j_{i'} \lambda_j \\
\lambda_i &= U^{j'}_i \lambda_{j'}
\end{aligned}}$$

At this point it looks like we are inventing new mathematics by playing a little game of swapping indices, and it is reasonable for us to grow a little bit suspicious

of what we are doing here. But we can show that our "index-swapping intuition" is correct. Let us consider (again) a vector $\vec{\lambda} = \lambda_j \vec{e}^{\,j} = \{\lambda_j\}$, where $\vec{e}^{\,j}$ is a vector in the dual basis:

$$\vec{e}^{\,j} = \vec{\nabla} u^j = \frac{\partial u^j}{\partial x}\hat{\imath} + \frac{\partial u^j}{\partial y}\hat{\jmath} + \frac{\partial u^j}{\partial z}\hat{k}.$$

Let us apply to chain rule:

$$\frac{\partial u^j}{\partial x} = \frac{\partial u^j}{\partial u^{i'}}\frac{\partial u^{i'}}{\partial x}.$$

Therefore,

$$\begin{aligned}
\vec{e}^{\,j} &= \vec{\nabla} u^j \\
&= \frac{\partial u^j}{\partial x}\hat{\imath} + \frac{\partial u^j}{\partial y}\hat{\jmath} + \frac{\partial u^j}{\partial z}\hat{k} \\
&= \frac{\partial u^j}{\partial u^{i'}}\frac{\partial u^{i'}}{\partial x}\hat{\imath} + \frac{\partial u^j}{\partial u^{i'}}\frac{\partial u^{i'}}{\partial y}\hat{\jmath} + \frac{\partial u^j}{\partial u^{i'}}\frac{\partial u^{i'}}{\partial z}\hat{k} \\
&= \sum_{i=1}^{3} \frac{\partial u^j}{\partial u^{i'}}\frac{\partial u^{i'}}{\partial x}\hat{\imath} + \frac{\partial u^j}{\partial u^{i'}}\frac{\partial u^{i'}}{\partial y}\hat{\jmath} + \frac{\partial u^j}{\partial u^{i'}}\frac{\partial u^{i'}}{\partial z}\hat{k} \\
&= \sum_{i=1}^{3} \frac{\partial u^j}{\partial u^{i'}}\left(\frac{\partial u^{i'}}{\partial x}\hat{\imath} + \frac{\partial u^{i'}}{\partial y}\hat{\jmath} + \frac{\partial u^{i'}}{\partial z}\hat{k}\right) \\
&= \sum_{i=1}^{3} \frac{\partial u^j}{\partial u^{i'}}\vec{\nabla} u^{i'} \\
&= \sum_{i=1}^{3} U^j_{i'}\vec{e}^{\,i'}.
\end{aligned}$$

Or, in Einstein's summation notation, the transformation rule for the dual basis vectors is:

$$\boxed{\vec{e}^{\,j} = U^j_{i'}\vec{e}^{\,i'}}$$

And the inverse transformation follows:

$$\boxed{\vec{e}^{\,j'} = U^j_{i'}\vec{e}^{\,i'}}$$

Note that, the above relation is in fact true for any n-dimensional space. Our "proof" here is restricted to only 3-dimensional space. However, it is not so difficult to fix this slight issue. If we redefine the $\vec{\nabla}$ operator for n-dimensional space, then we can easily see that the new proof is identical to the proof above, except for the maximal value of n.

We easily obtain the transformation rules for the covariant vector components using the results above:

$$\vec{\lambda} = \lambda_{i'}\vec{e}^{\,i'} = \lambda_j \vec{e}^{\,j} = \lambda_j U^j_{i'} \vec{e}^{\,i'} = \lambda_{i'} U^{i'}_j \vec{e}^{\,j}.$$

Therefore,

$$\boxed{\begin{aligned} \lambda_{i'} &= U^j_{i'} \lambda_j \\ \lambda_j &= U^{i'}_j \lambda_{i'} \end{aligned}} \tag{1}$$

Below is a summary of what we have learned so far about general coordinate transformations:

Natural basis:	$\vec{e}_j = U^{i'}_j \vec{e}_{i'}$	$\vec{e}_{j'} = U^i_{j'} \vec{e}_i$
Dual basis:	$\vec{e}^{\,j} = U^j_{i'} \vec{e}^{\,i'}$	$\vec{e}^{\,j'} = U^{j'}_i \vec{e}^{\,i}$
Contravariants:	$\vec{\lambda}^{j'} = U^{j'}_i \vec{\lambda}^i$	$\vec{\lambda}^j = U^j_{i'} \vec{\lambda}^{i'}$
Covariants:	$\vec{\lambda}_j = U^{i'}_j \vec{\lambda}_{i'}$	$\vec{\lambda}_{j'} = U^i_{j'} \vec{\lambda}^{j'}$

Since components of a vector must transform according to the rules above under general coordinate transformations, we can turn these rules into a more general definition of a vector.

Definition 3.1. A vector is a quantity whose components transform contravariantly as $\lambda^{i'} = U^{i'}_j \lambda^j$ under a general coordinate transformation $u^{i'} = u^{i'}(u^j)$.

Remark 3.1. We are often interested in vector fields, a collection of vectors at different points in space whose components depend on the coordinates: $\lambda^i = \lambda^i(u^j)$. Since vectors have to transform correctly, we require that at any point P, we require that $\lambda^{i'} = U^{i'}_j \lambda^j$ be true for a vector field (defined on λ^i) to be defined.

Remark 3.2. Not all 3-tuple of functions are vectors. Consider a 3-tuple of coordinates. Let $\lambda^i = u^i$ and $\lambda^{j'} = u^{j'}$, where $u^{j'} = u^{j'}(u^i)$. For λ^i to be make a vector field, it is required that $\lambda^{i'} = U^{i'}_j \lambda^j$. This implies that $u^{i'} = U^{i'}_j u^j$, with $U^{i'}_j = \partial u^{i'}/\partial u^j$. However, we immediately see that this is **not true in general**:

$$\boxed{u^{i'} \neq \frac{\partial u^{i'}}{\partial u^j} u^j}$$

But rather, $u^{i'} = u^{i'}(u^j) = f(u^j)$. So, coordinates themselves do not make vectors in general, because they, as components, do not transform correctly under general coordinate transformations. This is why we never use g_{ij} or g^{ij} to lower or raise the index of u^i or u_i, i.e., $u^i \neq g^{ij} u_j$, and $u_j \neq g_{ij} u^i$, in general.

There are exemptions to the "general rule," of course. For instance, consider **linear transformations**, where $u^{j'} = u^{j'}(u^i) = C^{j'}_i u^i$, where $C^{j'}_i$ are

constants. What this means is that the new coordinates are simply linear combinations of the old coordinates. It follows that

$$\frac{\partial u^{j'}}{\partial u^k} = C_i^{j'} \frac{\partial u^i}{\partial u^k} = C_i^{j'} \delta_k^i = C_k^{j'}.$$

If we let $k = i$ (we're just swapping dummy indices here), then $C_i^{j'} = \partial u^{j'}/\partial u^i = U_i^{j'}$, i.e., $u^{j'} = U_i^{j'} u^i$ under linear transformations. Therefore, u^i transform correctly, and in this case the coordinates u^i can form a vector, $\{u^i\}$.

Remark 3.3. Properly speaking we can define vectors with respect to a particular class of transformations. It is possible for something to be a vector with respect to a class of transformation, but not a vector with respect to another. Such a class of transformation (default) is the general coordinate transformations.

Example 3.10. Can differentials of coordinates (du^i) form a vector?

The solution is actually quite simple. Without loss of generality, let us consider a hypothetical vector $\{du^i\}$ in 3-dimensional space. $\{du^1, du^2, du^3\}$. Applying the chain rule:

$$du^{i'} = \frac{\partial u^{i'}}{\partial u^j} du^j = U_j^{i'} du^j.$$

We see that du^i transform correctly under general coordinate transformations. Therefore, du^i can make a vector.

Example 3.11. Find $[U_i^{j'}]$ for a coordinate transformation from Cartesian to spherical in flat 3-dimensional space. Let $\{u^j\} = \{x, y, z\}$ and $\{u^{j'}\} = \{r, \theta, \phi\}$. We find the Jacobian matrix:

$$[U_i^{j'}] = \begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} & \frac{\partial r}{\partial z} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} & \frac{\partial \theta}{\partial z} \\ \frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial y} & \frac{\partial \phi}{\partial z} \end{pmatrix} = \frac{1}{r} \begin{pmatrix} r\sin\theta\cos\phi & r\sin\theta\sin\phi & r\cos\theta \\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ \frac{-\sin\phi}{\sin\theta} & \frac{\cos\phi}{\sin\theta} & 0 \end{pmatrix}$$

Example 3.12. Consider $\vec{\lambda} = (1, 0, 0)^\top = \{\lambda^{i'}\}$ in Cartesian coordinates. What are the components of $\vec{\lambda}$ in spherical coordinates?

With our developed procedure, we should be able to obtain the answer with little thinking:

$$\vec{\lambda} = \{\lambda^{i'}\} = \begin{pmatrix} \lambda^{1'} \\ \lambda^{2'} \\ \lambda^{3'} \end{pmatrix} = [U_j^{i'} \lambda^j] = \frac{1}{r} \begin{pmatrix} r\sin\theta\cos\phi & r\sin\theta\sin\phi & r\cos\theta \\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ \frac{-\sin\phi}{\sin\theta} & \frac{\cos\phi}{\sin\theta} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

which gives

$$\begin{pmatrix} \lambda^{1'} \\ \lambda^{2'} \\ \lambda^{3'} \end{pmatrix} = \frac{1}{r} \begin{pmatrix} r \sin\theta \cos\phi \\ \cos\theta \cos\phi \\ \frac{-\sin\phi}{\sin\theta} \end{pmatrix}.$$

In the natural basis of spherical coordinates,

$$\vec{\lambda} = \sin\theta \cos\phi \, \vec{e}_r + \frac{1}{r} \cos\theta \cos\phi \, \vec{e}_\theta - \frac{\sin\phi}{r \sin\theta} \, \vec{e}_\phi.$$

We have said before that the "physical quantity" $\vec{\lambda}$ does not change. We can test this by finding $||\vec{\lambda}||$ in Cartesian and spherical coordinates and compare the results. In Cartesian, the metric tensor is $g_{ij} = \delta_i^j$, so

$$||\vec{\lambda}||_{\text{Cartesian}} = \sqrt{g_{ij}\lambda^i\lambda^j} = \sqrt{1^2 + 0^2 + 0^2} = 1,$$

as expected. The metric tensor $g_{i'j'}$ of spherical coordinate, in matrix form, is:

$$[g_{i'j'}] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2\theta \end{pmatrix}.$$

It follows that

$$\begin{aligned} ||\vec{\lambda}||_{\text{Spherical}} &= \sqrt{g_{i'j'}\lambda^{i'}\lambda^{j'}} = g_{1'1'}\left(\lambda^{1'}\right)^2 + g_{2'2'}\left(\lambda^{2'}\right)^2 + g_{3'3'}\left(\lambda^{3'}\right)^2 \\ &= \sqrt{(\sin\theta\cos\phi)^2 + r^2\left(\frac{1}{r}\cos\theta\cos\phi\right)^2 + r^2\sin^2\theta\left(-\frac{\sin\phi}{r\sin\theta}\right)^2} \\ &= \sqrt{(\sin\theta\cos\phi)^2 + (\cos\theta\cos\phi)^2 + \sin^2\phi} \\ &= \sqrt{\cos^2\phi + \sin^2\phi} \\ &= 1. \end{aligned}$$

Not surprisingly, $||\vec{\lambda}_{\text{Cartesian}}|| = ||\vec{\lambda}_{\text{Spherical}}|| = 1$.

Example 3.13. Find $U_i^{j'}$ for a rotation by an angle ϕ around the z-axis in Cartesian coordinates. Suppose $\vec{a} = \{a^i\} = (1, 1, 0)^\top$. What is $\vec{a} = \{a^{i'}\}$ after a rotation by ϕ.

insert figure here!!!

From linear algebra, we know that a rotation by ϕ around the z-axis is given by:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

So,

$$[U_i^{j'}] = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Given that $\vec{a} = \{a^i\} = (1,1,0)^\top$,

$$\begin{pmatrix} \lambda^{1'} \\ \lambda^{2'} \\ \lambda^{3'} \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\phi - \sin\phi \\ \sin\phi + \cos\phi \\ 0 \end{pmatrix}$$

Notice that because the coefficients of $[U_i^{j'}]$ are constants, the coordinates $\{x, y, z\}$ can make a vector. But note (again) that this is not true in general.

Next, we can find $||\vec{\lambda}||$ in both coordinate systems again and show that they are equal. In Cartesian coordinates, $\{\lambda^i\} = (1,1,0)^\top$. So, $||\vec{\lambda}|| = \sqrt{2}$. In the rotated frame:

$$||\vec{\lambda}|| = \sqrt{g_{i'j'}\lambda^{i'}\lambda^{j'}}.$$

We seem to be stuck here, since we have yet found the form of the new metric $g_{i'j'}$. However, we observe that $g_{i'j'} = e^{i'} \cdot \vec{e}_{j'}$, which is just either 1 (if $i = j$) or 0 (if $j \neq j$), because the new coordinate system is just the "tilted" version of the old one. So we expect that new metric tensor is the same as the old metric tensor, which is nothing but $g_{i'j'} = \mathrm{diag}(1,1,1)$. We can proceed to calculate the norm:

$$\begin{aligned} ||\vec{\lambda}|| &= \sqrt{(\cos\phi - \sin\phi)^2 + (\sin\phi + \cos\phi)^2 + 0^2} \\ &= \sqrt{2\left(\sin^2\phi + \cos^2\phi\right)} \\ &= \sqrt{2}, \end{aligned}$$

which is of course what we found earlier.

The above example motivates what we will discuss next. We have seen how vector components transform under general coordinate transformations, and how their norms (scalar) remain invariant. However, we start to question how the metric tensor transforms, in general. In the above example, the only reason we was able to calculate the norm of $\vec{\lambda}$ was that we could guess the form of the metric tensor, using some cleverness. In general, though, we will not be able to solve this problem without knowing exactly the form of the metric. Our next task, as apparent as it is now, is to find out how metric tensors in particular and tensors in general transform under general coordinate transformations.

3.7.2 Tensors

We should first ask ourselves what tensors are. While a rigorous definition of a tensor will come shortly in this section (hint: it will resemble that of a vector, i.e., based on how their components transform under general coordinate transformations), we can first think of tensors as generalizations of vectors, i.e., with more than one index and hence multi-dimensional. Consider a physical object like a balloon which stretches and compresses multi-dimensionally when some force is applied to it. We call this "response" the stress tensor, which has 9 components in 3-dimensional space: $F_{xx}, F_{xy}, F_{xz}, F_{yx}, F_{yy}, F_{yz}, F_{zx}, F_{zy}, F_{zz}$.

Definition 3.2. A tensor is a multi-component quantity whose components transform as contravariant or covariant vector components.

Example 3.14. Find the condition for $\tau^{ij}{}_{k}{}^{l}$ to be a tensor.

Well, a tensor has to transform correctly, so we can play a little game of index-replacing:

$$\boxed{\tau^{i'j'}{}_{k'}{}^{l'} = U^{i'}_{m} U^{j'}_{n} U^{p}_{k'} U^{l'}_{p} \tau^{mn}{}_{p}{}^{q}}$$

under a general coordinate transformation $u^{i'} \to u^{i'}(u^j)$.

Example 3.15. Show that g_{ij} is a tensor.

g_{ij} is a tensor, in fact, by definition:

$$\boxed{g_{i'j'} = U^k_{i'} \vec{e}_k \cdot U^l_{j'} \vec{e}_l = U^k_{i'} U^l_{j'} \vec{e}_k \cdot \vec{e}_l = U^k_{i'} U^l_{j'} g_{kl}}$$

Similarly, the inverse g^{ij} is also a tensor:

$$\boxed{g^{k'l'} = U^{k'}_{i} \vec{e}^{\,i} \cdot U^{l'}_{j} \vec{e}^{\,j} = U^{k'}_{i} U^{l'}_{j} \vec{e}^{\,i} \cdot \vec{e}^{\,j} = U^{k'}_{i} U^{l'}_{j} g^{ij}}$$

Definition 3.3. A tensor $\tau^{abc\ldots}{}_{def\ldots}$ is of type (r, s) when it has r contravariant components and s covariant components.

Example 3.16. g_{ij} is a type $(0, 2)$ tensor. g^{ij} is a type $(2, 0)$ tensor.

Vectors are just a special type of tensors. Contravariant vectors λ^i are type $(1, 0)$ tensors, while covariant vectors λ_i are type $(0, 1)$ tensors.

Remark 3.4. The transformation $U^{i'}_{j}$ is NOT a tensor. Indeed, it does not make sense to talk about how a transformation transforms.

Example 3.17. Write $g_{i'j'} = U^k_{i'} U^l_{j'} g_{kl}$ as a matrix equation.

We let $\tilde{G} = [g_{ij}]$, $\tilde{G}' = [g_{i'j'}]$, and $\hat{U} = \tilde{U}^{-1} = [\partial u^k / \partial u^{i'}]$. To create a matrix equation, we have to be aware of when an index indicate a row or column. The correct "row-column-sensitive" order expression is:

$$[g_{i'j'}] = [U^k_{i'}][g_{kl}][U^l_{j'}],$$

or equivalently,

$$\boxed{\tilde{G}' = \hat{U}^\top \tilde{G} \hat{U}}$$

Example 3.18. Using the transformation rules for tensors, obtain $g_{i'j'}$ for a rotation by ϕ around the z-axis in 3-dimensional space.

Recall that

$$[U^{j'}_i] = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and that $g_{ij} = \mathrm{diag}(1,1,1)$ in 3-dimensional Cartesian coordinate system. To find $\tilde{G}'$, we first have to find $U^k_{i'}$, the inverse transformation. We expect that $U^k_{i'}$ is simply a rotating by $-\phi$ (since this intuitively "reverses" what $U^{k'}_i$ does). But we, as always, find $U^k_{i'}$ mechanically:

$$\hat{U} = \left[\frac{\partial u^k}{\partial u^{i'}}\right] = \tilde{U}^{-1} = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

as anticipated. Using the result of the previous example,

$$\begin{aligned} \tilde{G}' &= \hat{U}^\top \tilde{G} \hat{U} \\ &= \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned}$$

just as guessed. So our intuition was correct after all. But we realize an important fact: the metric tensor can remain the same after a non-trivial transformation (in this case the transformation is a rotating around the z-axis). We also notice that the metric tensors are the same under a transformation if the transformation matrix is **orthogonal**, i.e., its inverse the equal to its transpose:

$$\hat{U} = \tilde{U}^{-1} = \tilde{U}^\top.$$

Remark 3.5. Only tensors of type (r, s) with $r + s \leq 2$ can be displayed as matrices and/or matrix multiplications. It is not possible to write $\tau^{ij}{}_{kl}$ as a matrix.

3.8 Scalars

As we have seen from time to time, scalars are invariant under general coordinate transformations. As tensors, scalars are of type $(0,0)$, i.e., they have open indices.

Example 3.19. Show that the magnitude of a vector is a scalar.

Consider $\vec{a} = a^i = a^{i'}$ in two different coordinate systems. The magnitude of $\vec{a} = ||\vec{a}||$ is a scalar if:

$$\vec{a} \cdot \vec{a} = a^i a_i = a^{i'} a_{i'}.$$

This can be readily shown, since we already know the transformation rules for vector components:

$$\begin{aligned} a^{i'} a_{i'} &= \left(U^{i'}_{j}\, a^j\right)\left(U^k_{i'} a_k\right) \\ &= U^{i'}_{j}\, a^j U^k_{i'} a_k a^j \\ &= \delta^k_j a_k a^j \\ &= a^k a_k \\ &= a^i a_i. \end{aligned}$$

Therefore, $||\vec{a}||$ is a scalar.

Example 3.20. Show that the line element $ds^2 = g_{ij}\, du^i\, du^j$ is a scalar.

Similar to the previous example, we have to show that:

$$g_{i'j'}\, du^{i'}\, du^{j'} = g_{ij}\, du^i\, du^j.$$

We simply apply the transformation rule to each factor and play a little game of index-arranging:

$$\begin{aligned} g_{i'j'}\, du^{i'}\, du^{j'} &= \left(U^k_{i'} U^l_{j'} g_{kl}\right)\left(U^{i'}_{m}\, du^m\right)\left(U^{j'}_{n}\, du^n\right) \\ &= U^k_{i'} U^l_{j'} U^{i'}_{m} U^{j'}_{n}\, g_{kl}\, du^m\, du^n \\ &= \delta^k_m \delta^l_n\, g_{kl}\, du^m\, du^n \\ &= g_{kl}\, du^k\, du^l \\ &= g_{ij}\, du^i\, du^j. \end{aligned}$$

Therefore, $ds^2 = g_{ij}\, du^i\, du^j$ is a scalar.

4 Flat spacetime

Recall that in space, we require 3 coordinates. It makes sense that in order to describe spacetime, we need an extra coordinate to describe time, in which case, we will switch from Roman-character indices $(i, j, k, \cdots = 1, 2, 3)$ to Greek-character indices $(\mu, \nu, \sigma, \dots) = 0, 1, 2, 3$. We also tend to (by an unwritten convention and of course not mandatory) write vectors as capital letters to distinguish 4-vectors from 3-vectors. For example, below are a number of ways we express a 4-vector:

$$X^\mu = \left\{X^0, X^1, X^2, X^3\right\} = \left(X^0, \vec{X}\right) = \left(X^0, X^i\right).$$

Notice that the "hat" notation is reserved for 3-vectors.

4.1 Special Relativity

Coordinate transformations in special relativity are Lorentz transformations under which the spacetime interval (in Cartesian coordinates)

$$ds^2 = c^2\, dt^2 + dx^2 + dy^2 + dz^2 = \left(dx^0\right)^2 + \left(dx^1\right)^2 + \left(dx^2\right)^2 + \left(dx^3\right)^2,$$

which gives "physical distances" in spacetime, is invariant. We also recall that

$$ds^2 = g_{ij}\, du^i\, du^j,$$

we can "read off" the metric from the spacetime interval as diag $= (1, -1, -1, -1)$. This is referred to as the Minkowski metric (flat 4-dimensional spacetime in Cartesian coordinates), denoted as $\eta_{\mu\nu}$:

$$[\eta_{\mu\nu}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Because any frame in special relativity can be connected to the original frame by a Lorentz transformation, i.e.,

$$ds^2 = ds'^2 = \left(dx^{0'}\right)^2 + \left(dx^{1'}\right)^2 + \left(dx^{2'}\right)^2 + \left(dx^{3'}\right)^2,$$

The Minkowski metric (in Cartesian coordinates) is the same for any reference frame:

$$[\eta_{\mu'\nu'}] = [\eta_{\mu\nu}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

This says that:

$$\boxed{ds^2 = \eta_{\mu\nu}\, dX^\mu\, dX^\nu = \eta_{\mu'\nu'}\, dX^{\mu'}\, dX^{\nu'}}$$

For general coordinate systems in three dimensions, the metric $g_{ij} \neq \mathrm{diag}(1,1,1)$. In these cases, the Minkowski metric has the following form:

$$[\eta_{\mu\nu}] = \begin{pmatrix} 1 & 0 \\ 0 & -[g_{ij}] \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & -[g_{ij}] & \\ 0 & & & \end{pmatrix}.$$

More generally, however, spacetime can be curved, and the metric tensor of the three-dimensional space can be non-Cartesian. In these cases, we simply replace $\eta_{\mu\nu}$ by $g_{\mu\nu}$. It follows that:

$$\boxed{ds^2 = g_{\mu\nu}\, du^\mu\, du^\nu}$$

in general. The "moral" here is that $\eta_{\mu\nu}$ is reserved for flat Minkowski spacetime.

Same as before, the metric tensor encodes all "physical properties" of the Minkowski spacetime and allows us to transform contravariant components into covariant components and vice versa. For instance, consider a contravariant vector $\lambda^\mu = (\lambda^0, \lambda^1, \lambda^2, \lambda^3) = (\lambda^0, \vec{\lambda})$. The covariant vector is given by simply "lowering" the index of the contravariant vector using the Minkowski metric tensor:

$$\lambda_\mu = \eta_{\mu\nu}\lambda^\nu = (\lambda_0, \lambda_1, \lambda_2, \lambda_3).$$

The inverse transformation has the same form:

$$\lambda^\mu = \eta^{\mu\nu}\lambda_\nu.$$

Note that if we use the Minkowski metric tensor in Cartesian coordinates, then

$$\lambda_\mu = \eta_{\mu\nu}\lambda^\nu = (\lambda_0, \lambda_1, \lambda_2, \lambda_3) = (\lambda^0, -\lambda^1, -\lambda^2, -\lambda^3).$$

We observe that

$$\begin{aligned} \lambda^0 &= \lambda_0 \\ \lambda^i &= -\lambda_i. \end{aligned}$$

We also observe that the inverse Minkowski metric tensor is Minkowski metric tensor itself:

$$\boxed{[\eta^{\mu\nu}] = [\eta_{\mu\nu}]^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}}$$

Inner products of vectors follow the same rule as before:

$$\boxed{\mathbf{a} \cdot \mathbf{b} = a^\mu b_\mu = a_\mu b^\mu = \eta_{\mu\nu} a^\mu b^\nu = \eta^{\mu\nu} a_\mu b_\nu}$$

On a more practical side of things, we have to be careful with the signs that go with the terms when we solve problems:

$$\begin{aligned}\mathbf{a} \cdot \mathbf{b} &= a^1 b_1 + a^2 b_2 + a^2 b_2 + a^3 b_3 \\ &= a_1 b_1 - a_2 b_2 - a_2 b_2 - a_3 b_3 \\ &= a^1 b^1 - a^2 b^2 - a^2 b^2 - a^3 b^3.\end{aligned}$$

At this point, we might wonder what the basis vectors $\vec{e}_\mu$ in the flat, 4-dimensional Minkowski spacetime in Cartesian coordinates are. We can, of course, find them, from the metric tensor $\eta_{\mu\nu}$. However, one thing for sure: they are not going to be the basis vectors $\hat{\imath}, \hat{\jmath}, \hat{k}$. In fact, the basis vectors that make up $\eta_{\mu\nu}$ will likely contain imaginary parts, since $\vec{e}_i \cdot \vec{e}_j = -1$ for $i = j$. The point is: we don't really need the basis vectors once we already have the metric, because after all, almost all of our computations involve the metric tensor.

4.1.1 The Lorentz Transformations, revisited

Recall that Lorentz transformations are (not necessarily strictly linear) transformations from one **inertial frame** (S) to another **inertial frame** (S').

The most general Lorentz transformations, also referred to collectively as "Poincaré transformations" include:

- Lorentz boost: inertial frames with relative velocity v.
- Translation: origins don't coincide at $t' = t = 0$.
- Spatial rotation
- Spatial inversion: or parity transformations ($x' = -x$)
- Time reversal: $t' = -t$.

We can also make further distinctions among the listed transformations:

- Inhomogeneous Lorentz transformations: have non-trivial translation.
- Homogeneous Lorentz transformations: no translations.
- Improper Lorentz transformations: with time reversal or parity.
- Proper Lorentz transformation: no parity or time reversal.

We can first look at **homogeneous, proper Lorentz transformations with no rotations**, i.e., Lorentz boosts. These are probably the simplest of Lorentz transformations that we encountered in introductory special relativity. A classic example is the Lorentz boost along the x-axis:

$$\begin{pmatrix} a^{0'} \\ a^{1'} \\ a^{2'} \\ a^{3'} \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a^0 \\ a^1 \\ a^2 \\ a^3 \end{pmatrix}$$

where the transformation matrix element $U^{i'}_j$ in 3-dimensional space is now generalized to 4-dimensional spacetime:

$$\boxed{X^{\mu'}_{\nu} = \frac{\partial a^{\mu'}}{\partial a^{\nu}}}$$

Specifically for Lorentz boosts, we let the general $X^{\mu'}_{\nu}$ be $\Lambda^{\mu'}_{\nu}$. This slight change in notation gives:

$$\boxed{[\Lambda^{\mu'}_{\nu}] = \left[\frac{\partial a^{\mu'}}{\partial a^{\nu}}\right] = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}$$

We notice immediately that $\Lambda^{\mu'}_{\nu}$ is constant for all μ, ν. This means (i) Lorentz transformations are linear transformations, and (ii) Cartesian coordinates X^{μ} can form the components of a vector under Lorentz transformations because

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

is obeyed. While this concept might seem foreign, we realize we have seen this many times before in special relativity, only in a slightly different form:

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix},$$

which gives us the very first Lorentz-Einstein equations that we are familiar with:

$$\begin{cases} ct' = \gamma(ct - \beta x) \\ x' = \gamma(x - c\beta t) \\ y' = y \\ z' = z. \end{cases}$$

Keep in mind that the equalities above are only true because $\Lambda^{\mu'}_{\nu}\Lambda^{\nu}_{\sigma'} = \delta^{\mu}_{\sigma}$, and are not true in general, because coordinates don't form vectors in general.

We can find the inverse of $\Lambda^{\mu'}_{\nu}$ in two ways: the intuitive way, and the mechanical way. Intuitively, $\Lambda^{\mu}_{\nu'}$, or the inverse transformation, undoes what $\Lambda^{\mu'}_{\nu}$, so it has to be a boost in the opposite direction ($v \to -v$ therefore $\beta \to -\beta$). So if $\Lambda^{\mu'}_{\nu}$ is an x-boost, then its inverse is simply:

$$[\Lambda^{\mu}_{\nu'}] = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

which not surprisingly gives us the inverse Lorentz-Einstein transformation equations:

$$\begin{cases} ct = \gamma(ct' + \beta x) \\ x = \gamma(x + c\beta t) \\ y = y \\ z = z. \end{cases}$$

A reasonable next step is to look at **proper homogeneous Lorentz transformation with rotation(s)**. Recall from linear algebra that such a transformation is simply a composition of multiple, more elementary transformations (pure boost and/or pure rotation). For example, a pure rotation an angle ϕ about the z-axis has the following matrix:

$$[\Lambda^{\mu'}_{\nu}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi & \sin\phi & 0 \\ 0 & -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

To think how to do about how to construct a transformation matrix has includes both a boost and a rotation, we can first do a gentle example.

Example 4.1. Find the transformation matrix for a Lorentz boost in the y-axis in two ways: intuition and composition of rotation by $\pi/2$ and boost along the x-direction.

We can easily predict that the matrix for a boost in y should be almost the same as that for a boost x, except for the locations of the terms in the matrix. We know this because the x-direction is no more special than the y-direction. So,

$$[\Lambda^{\mu'}_{\nu}] = \begin{pmatrix} \gamma & 0 & -\beta\gamma & 0 \\ 0 & 0 & 0 & 0 \\ -\beta\gamma & 0 & \gamma & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

We should get the same matrix by doing the "composition method." We can simply think of a boost in the y-direction as a rotating by $\pi/2$, followed by a boost along the x-direction, then followed by a rotation by $-\pi/2$. A final transformation matrix is a product of the list matrix, in right-to-left order:

$$
\begin{aligned}
[\Lambda^{\mu'}_{\nu}] &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \gamma & 0 & -\beta\gamma & 0 \\ 0 & 0 & 0 & 0 \\ -\beta\gamma & 0 & \gamma & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},
\end{aligned}
$$

as expected.

The moral of the above example is to show that any Lorentz boost along an arbitrary direction can be found as a combination of a boost along the x-direction and a spatial rotation:

$$
\boxed{[\Lambda^{\mu'}_{\nu}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi & 0 \\ 0 & \sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi & \sin\phi & 0 \\ 0 & -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}
$$

4.1.2 A curiosity about the Lorentz boosts

We can actually make Lorentz boosts look like rotation using hyperbolic functions:

$$
\begin{aligned}
\sinh(\alpha) &= \frac{e^{\alpha} - e^{-\alpha}}{2} \\
\cosh(\alpha) &= \frac{e^{\alpha} + e^{-\alpha}}{2} \\
\operatorname{sech}(\alpha) &= \frac{1}{\cosh(\alpha)} \\
\operatorname{csch}(\alpha) &= \frac{1}{\sinh(\alpha)} \\
\tanh(\alpha) &= \frac{\sinh(\alpha)}{\cosh(\alpha)} \\
\coth(\alpha) &= \frac{1}{\tanh(\alpha)},
\end{aligned}
$$

which have the following identities:

$$
\begin{aligned}
\cosh^2(\alpha) - \sinh^2(\alpha) &= 1 \\
1 - \tanh^2(\alpha) &= \operatorname{sech}^2(\alpha).
\end{aligned}
$$

If we let $\beta = \tanh(\phi)$, where the angle ϕ is called "rapidity," then it follows that:

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-\tanh^2(\phi)}} = (\operatorname{sech}(\phi))^{-1} = \cosh(\phi),$$

and

$$\gamma\beta = \sinh(\phi).$$

So the Lorentz boost $[\Lambda^{\mu'}_{\nu}]$ becomes:

$$[\Lambda^{\mu'}_{\nu}] = \begin{pmatrix} \cosh(\phi) & -\sinh(\phi) & 0 & 0 \\ -\sinh(\phi) & \cosh(\phi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

4.1.3 The Poincaré transformations

$$\boxed{a^{\mu'} = \Lambda^{\mu'}_{\nu} a^{\nu} + b^{\mu'}}$$

Example 4.2. stuff

Example 4.3. stuff

Example 4.4. stuff

4.1.4 Velocity, momentum, and force

4.2 Relativistic Electrodynamics

Gauss' law: $\boxed{\oiint \vec{E} \cdot d\vec{a} = \frac{q}{\epsilon_0}}$

No magnetic monopole: $\boxed{\oint \vec{B} \cdot d\vec{a} = 0}$

Faraday's law: $\boxed{\oint \vec{E} \cdot d\vec{s} = -\frac{d}{dt}\Phi_B = \frac{\partial}{\partial t}\int_A \vec{B} \cdot d\vec{a}}$

Ampere-Maxwell's law: $\boxed{\oint\!\!\!\oint \vec{B} \cdot d\vec{s} = \mu_0 I + \mu_0\epsilon_0 \frac{\partial}{\partial t}\int_A \vec{E} \cdot d\vec{a}}$

Definition 4.1. The electromagnetic field strength tensor $F^{\mu\nu}$ is given by:

$$[F^{\mu\nu}] = \begin{pmatrix} 0 & E^1/c & E^2/c & E^3/c \\ -E^1/c & 0 & B^3 & -B^2 \\ -E^2/c & -B^3 & 0 & B^1 \\ -E^3/c & B^2 & -B^1 & 0 \end{pmatrix}.$$

The Maxwell equations, in index notation become:

$$stuff$$

5 Curved spaces

5.1 2-dimensional curved spaces

5.2 Manifolds

5.3 Tensors on manifolds

6 Gravitation and Curvature

6.1 Geodesics and Affine connections $\Gamma^{\sigma}_{\mu\nu}$

6.2 Parallel transport

6.3 Covariant differentiation

6.4 Newtonian limit

7 Einstein's field equations

7.1 The stress-energy tensor $T^{\mu\nu}$

7.2 Riemann curvature tensor $R^{\lambda}{}_{\mu\nu\sigma}$

7.3 The Einstein equations

7.4 Schwarzschild solution

8 Predictions and tests of general relativity

8.1 Gravitational redshift

8.2 Radar time-delay experiments

8.3 Black Holes

9 Cosmoslogy

9.1 Large-scale geometry of the universe

9.1.1 Cosmological principle

9.1.2 Robertson-Walker (flat, open, closed) geometries

9.1.3 Expansion of the universe

9.1.4 Distances and speeds

9.1.5 Redshifts

9.2 Dynamical evolution of the universe

9.2.1 The Friedmann equations

9.2.2 The cosmological constant Λ

9.2.3 Equations of state

9.2.4 A matter-dominated universe ($\Lambda = 0$) [Friedmann models]

9.2.5 A flat, metter-dominated universe ($\Lambda = 0$) [old favorite model]

9.3 Observational cosmology

9.3.1 Hubble law

9.3.2 Acceleration of the universe

9.3.3 Matter densities and dark matter

9.3.4 The flatness and horizon problems

9.3.5 Cosmic Microwave Backgrond (CMB) anisotropy

9.4 Modern cosmology

9.4.1 Inflation

9.4.2 Dark energy (The cosmological constant problem)

9.4.3 Concordance model [new favorite model]

9.4.4 Open questions