

GENERAL RELATIVITY & COSMOLOGY

A Quick Guide

Huan Bui

Colby College
Physics & Statistics
Class of 2021

December 20, 2018

Contents

1	Overview and Review	3
1.1	Review of Special Relativity	3
1.2	The Equivalence Principle	4
1.3	Versions of the Equivalence Principle	6
1.3.1	The Strong Equivalence Principle	6
1.3.2	The Weak Equivalence Principle	6
2	Review of Vector Calculus	7
2.1	Operations & Theorems	7
2.2	Coordinate Systems	11
3	Flat 3-dimensional space	14
3.1	Curvilinear coordinates	14
3.2	Basis vectors	14
3.3	Natural basis	14
3.4	Dual basis	16
3.5	Contravariant and covariant vectors	19
3.5.1	The suffix notation (vector calculus)	19
3.5.2	The Einstein summation notation	19
3.5.3	Contravariant and covariant vectors	19
3.6	Metric tensor	19
3.7	Coordinate transformation	19
3.8	Tensors	19
4	Flat spacetime	20
4.1	Special Relativity	20
4.2	Relativistic Electrodynamics	20
5	Curved spaces	21
5.1	2-dimensional curved spaces	21
5.2	Manifolds	21
5.3	Tensors on manifolds	21
6	Gravitation and Curvature	22
6.1	Geodesics and Affine connections $\Gamma_{\mu\nu}^{\sigma}$	22
6.2	Parallel transport	22
6.3	Covariant differentiation	22
6.4	Newtonian limit	22
7	Einstein's field equations	23
7.1	The stress-energy tensor $T^{\mu\nu}$	23
7.2	Riemann curvature tensor $R_{\mu\nu\sigma}^{\lambda}$	23
7.3	The Einstein equations	23
7.4	Schwarzschild solution	23

8	Predictions and tests of general relativity	24
8.1	Gravitational redshift	24
8.2	Radar time-delay experiments	24
8.3	Black Holes	24
9	Cosmoslogy	25
9.1	Large-scale geometry of the universe	25
9.1.1	Cosmological principle	25
9.1.2	Robertson-Walker (flat, open, closed) geometries	25
9.1.3	Expansion of the universe	25
9.1.4	Distances and speeds	25
9.1.5	Redshifts	25
9.2	Dynamical evolution of the universe	25
9.2.1	The Friedmann equations	25
9.2.2	The cosmological constant Λ	25
9.2.3	Equations of state	25
9.2.4	A matter-dominated universe ($\Lambda = 0$) [Friedmann models]	25
9.2.5	A flat, matter-dominated universe ($\Lambda = 0$) [old favorite model]	25
9.3	Observational cosmology	25
9.3.1	Hubble law	25
9.3.2	Acceleration of the universe	25
9.3.3	Matter densities and dark matter	25
9.3.4	The flatness and horizon problems	25
9.3.5	Cosmic Microwave Background (CMB) anisotropy	25
9.4	Modern cosmology	25
9.4.1	Inflation	25
9.4.2	Dark energy (The cosmological constant problem)	25
9.4.3	Concordance model [new favorite model]	25
9.4.4	Open questions	25

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 space-time invariant. Another scalar, which we probably will not see again in this text, is the metric signature ($\text{sgn}([\eta_{\mu\nu}] = -2)$). 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. For example, for 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 1\AA , 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 be parameterized by a single variable, which we will call t , as

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

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 $ds = ||\vec{r}'|| = ||\dot{\vec{r}}|| dt$:

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

For example, 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 $\vec{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ 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

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

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) $f : \mathbb{R}^3 \rightarrow \mathbb{R}$. 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. 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 .

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

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 .

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

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. An example of surface integrals we often see in physics is 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: $\oiint \vec{E} \cdot d\vec{a} = \frac{q}{\epsilon_0}$

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

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

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

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,

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

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:

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

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. For example, 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

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

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

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

For example, 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, we get $dA = dx \, dy = r \, dr \, d\theta$, as expected. 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

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

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 $\{\hat{i}, \hat{j}, \hat{k}\}$. 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 $\{\hat{i}, \hat{j}, \hat{k}\}$ from Cartesian coordinates. We notice that \hat{i} 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

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

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

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

where $i = 0, 1, 2, \dots, n$ 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:

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

For example, 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:

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

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:

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

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!!!

For example, 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 $(x, y, z) \rightarrow (r, \theta, \phi)$.

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

Tedious and painstaking computations give us the other two vectors:

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

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 $\vec{e}^\phi \neq \vec{e}_\phi$.

Let's consider a non-familiar example with 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 the dot products are the basis vectors when 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)

3.5.2 The Einstein summation notation

3.5.3 Contravariant and covariant vectors

3.6 Metric tensor

3.7 Coordinate transformation

3.8 Tensors

4 Flat spacetime

4.1 Special Relativity

4.2 Relativistic Electrodynamics

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 $\Gamma_{\mu\nu}^{\sigma}$

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, matter-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 Background (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