Advanced Principles of Mechanics I

Notes on USC Spring 2026 Course PHYS-161

Jiarui Liu

High School Student at Nanjing Foreign Language School
Website
LinkedIn
GitHub

jliu1978@usc.edu

Contents

Special Relativity

Dimensions and Dimensional Analysis

This will be an experimental course: starting with the concept of space-time, going back to kinematics, then relativistic dynamics, lastly energy.

Definition of physics:

Definition 1 | Physics is the natural quantitative science that deals with the fundamental interactions between matter and energy.

What arguably sets physics apart from the other sciences would be the reductionist nature. At the base of the "pyramid of the natural sciences" lies mathematics, the fundamental subjects. Up a level is Physics, then Chemistry, then Biology. Because it's quantitative, when we study the world around us with measurements and standardizations. This introduces us to *dimensions and unit*:

Claim 1 | Nearly all properties in classical physics can be described with one or more of the three dimensions: length, time, and mass.

The notation would be $[x]$ as the dimension of some variable or constant, for example

$$[x] = L. \tag{1.1}$$

The SI unit is the international unit. Remember that the standard for mass in SI is kilograms. Now we introduce the natural unit.

Definition 2 | The natural unit is defined by setting the values of the three constants:

$$c = \hbar = G = 1. \tag{1.2}$$

Here they still have dimensions, for example:

$$[c] = \frac{L}{T} = 1 \quad \Rightarrow \quad L = T. \tag{1.3}$$

Dimensional Analysis

A simple pendulum on Earth has length L, mass m. Given

$$[l] = L, [m] = M, \text{ and } [g] = \frac{L}{T^2} \tag{1.4}$$

determine, up to dimensionless factors, the period of the pendulum.

Solution. We have, apparently, the period τ with dimension T. We would assume the

powers with greek letters:

$$[\tau] = T = [l^\alpha m^\beta g^\gamma] = L^\alpha M^\beta \left(\frac{L}{T^2}\right)^\gamma. \tag{1.5}$$

This is solved simply as:

$$\begin{cases} \alpha + \gamma = 0, \\ \beta = 0, \\ -2\gamma = 1 \end{cases} \Rightarrow \begin{cases} \alpha = \frac{1}{2}, \\ \beta = 0, \\ \gamma = -\frac{1}{2} \end{cases}. \tag{1.6}$$

Which gives us

$$\tau \propto l^{\frac{1}{2}} g^{-\frac{1}{2}} = \sqrt{\frac{l}{g}} \tag{1.7}$$

Exercise 1 Look at the dimensions of $c, \hbar,$ and G. Create the Planck units- the Planck length, time, and mass, by combining the three constants in a unique way that they create a constant for length, time, and mass.

SECTION 2
Definitions

We start by defining the basics of kinematics- the study of "how things move". The simplest case of motion is none- an object being stationary. Now there will be a question: where is it. Now we define the frame of reference and point particles.

Definition 3 A frame of reference for an observer is a sufficiently precise laboratory of measurement approaches, i.e. clocks, meter sticks. Here, in classical physics, we assume a precise measurement without quantum fluctuation, and it can be arbitrarily precise.

And we define the point particle:

Definition 4 A point particle is an approximation whereby an object is treated as a zero-dimensional, mathematical point. It is valid when the size and shape of the object are irrelevant for the question under study.

In this frame of reference, we may set up Cartesian coordinates for space (x, y, z) to uniquely specify the position of a particle. The Cartesian coordinate is defined by three everywhere mutually perpendicular axes intersecting at the same point, defined as the origin.

Now it's the problem of incorporating time: the sequence of locations a particle passes can be described as time. It cannot, to our best knowledge be reversed, so it's the monotonically increasing parameter denoting the particle's position. We treat it as another label on the event: write (ct, x, y, z). Thus, to locate an event in space, we have four dimensions.

Remark 1 Notice that the time axis is not t but ct, because the dimension of it should be in length.

Thus we have the more formal definition or general statement:

Claim 2

> To specify an event in the universe, we provide four labels:
>
> $$(ct, x, y, z) \tag{2.1}$$
>
> where x, y and z are for space or location, and ct as time. Here c is introduced as a speed constant (not "light speed" in its literal sense), which we will cover in due course.

This encourages us to think in the perspective that the "arena" of the universe is a four-dimensional mathematical space that we will call spacetime. A teaser for the mathematical space: "manifolds". This will be covered in higher division courses.

Now we have the empirical facts about space.

Claim 3

> Here are the empirical facts about space:
>
> 1. It is in three dimensions, with an extra dimension of time,
>
> 2. It is isotropic, which means it looks the same in every direction,
>
> 3. It is homogenous, as in it is the same at every point.

In the larger context of general relativity, which includes gravity, there are many space-times. To distinguish ours, the one relevant for special relativity, we will call it Minkowski Spacetime. Thus to summarize, ***The universe is a four dimensional Minkowski spacetime, with points and events labeled*** (ct, x, y, z).

Observers and inertial frames. There is a special class of frames of reference are known as inertial frames. It is the result of Newton's first law, which we will cover later in the course.

Consider a free particle— a particle subject to no influences. (we assume that this is true in the context of classical mechanics). Then, an inertial frame is a frame in which a free particle moves uniformly in a straight line, or not at all. This implies that if one inertial frame exists, by moving the origin or setting another frame of reference that is moving at constant speed relative to our first, or even performing static rotations, these will all define inertial frames equally.

A straight line in this sense can be defined as the shortest distance between two points in space— when we talk about geometric spaces, we are talking about the geometry of Euclid. Another remark is that coordinates are entirely artificial, and the choice of coordinates is arbitrary. Thus the laws of physics *cannot* depend on coordinates. This actually has profound consequences: the laws that we write down are not arbitrary and are subject to the principle.

Take the vector for example. It is defined as a "quantity with a magnitude and a direction", independent of coordinate systems— that is the motivation for vectors, scalars, tensors... and other ideas in physics.

Coordinate Systems

Distance. To express distances in coordinate systems, we have Pythagorean's Theorem:

Theorem 1   For two points in space, in a Euclidean coordinate system, we have the distance, $\Delta\ell$ is defined as:
$$\Delta\ell^2 = \Delta x^2 + \Delta y^2 + \Delta z^2. \tag{3.1}$$

If the quantity is infinitesimal:

$$\mathrm{d}\ell^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2. \tag{3.2}$$

Remark 2   Note that using the calculus of variations would reveal the straight line as the shortest distance between two arbitrary points.

Henceforth— and unless otherwise stated— we will look in inertial frames. The first we will introduce, of course, is the Cartesian coordinates, where we choose three mutually perpendicular axes. The other two coordinate systems, spherical / cylindrical coordinates are defined well enough in Calculus, with the exception of using (ct, r, θ, φ) where θ is the azimuthal angle and φ is the polar coordinate, different from the mathematical definition.

Now we introduce the shorthand notation for the axes:

$$(ct, x, y, z) \;\Rightarrow\; (x^0, x^1, x^2, x^3). \tag{3.3}$$

Or, when in spacetime coordinates we have:

$$x^\mu, x^\nu, x^\lambda, x^\sigma, \; \mu, \nu, \lambda, \sigma \in [0, 3], \tag{3.4}$$

or, when we are in space,

$$x^i, x^j, x^m, x^k, \; i, j, m, k \in [1, 3]. \tag{3.5}$$

There will be an abuse of notation that is quite standard:

$$x^\mu = (x^0, x^1, x^2, x^3). \tag{3.6}$$

Defining time. The natural choice of parameterization is to use coordinate time:

$$t \equiv \frac{x^0}{c}. \tag{3.7}$$

Then, we would have

$$x^0 = x^0(t) = ct \quad \Rightarrow \quad x^\mu = x^\mu(t) = (ct, x^1(t), x^2(t), x^3(t)). \tag{3.8}$$

When we use Cartesian coordinates, or $x^1 = x, x^2 = y, x^3 = z$, we would have the positions defined, when we suppress the time factor for simplicity, would be

$$(x(t), y(t), z(t)). \tag{3.9}$$

Now we move on to the definition of a "frame", or a "slice" of spacetime, as:

Definition 5 | A frame is a spatial slice of spacetime where every point on the slice is simultaneous with every other.

We can consider this as a "stack of slices", where the progression sequence is the direction of time. But this is inaccurate: it has the underlying assumption that space and time are independent of observers: what is considered simultaneous in the perspective of one observer is not necessarily simultaneous in another. We will come back to this notion in time dilation.

Displacement. As a first attempt at motion, define the displacement Δx^i as:

$$\Delta x^i \equiv x_{\text{f}}^i - x_{\text{i}}^i. \tag{3.10}$$

Where $x_{\text{f}}^i = x^i(t_{\text{f}})$ and $x_{\text{i}}^i = x^i(t_{\text{i}})$, where the duration is $\Delta t = t_{\text{f}} - t_{\text{i}}$. We can immediately see that this is a crude definition of motion because it considers only the endpoint and starting point, providing just the "Line of Sight" distance between two points. So, to further clarify, we will define also its magnitude:

$$\left|\Delta x^i\right| \equiv \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}. \tag{3.11}$$

We notice that this is an extremely crude way still, because it does not count for any motion of the particle in between the time of motion. However if we take the limit, this becomes extremely useful. Let's first obtain this from the Euclidean metric.

PROOF | We first introduce a parameter, λ, where the line is parameterized by

$$x(\lambda) = a\lambda, \ y(\lambda) = b\lambda, \ z(\lambda) = c\lambda. \tag{3.12}$$

With $0 \leq \lambda le 1$. Let

$$L = \left|\Delta x^i\right|, \tag{3.13}$$

This, when we start with a line assuming $c = 0$, we get $y = \frac{b}{a}x$. When we take the length differential,

$$d\ell = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}, \tag{3.14}$$

the length would be:

$$L = \int d\ell = \int_0^1 \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2} \, d\lambda = \sqrt{a^2 + b^2 + c^2}. \tag{3.15}$$

For this context is is a bit of an overkill, but for spacetime the incorporation of time would lead to more counterintuitive answers. $\qquad \square$

SECTION 4

Vectors in Three-Dimensional Euclidean Space

From the definition of displacement we notice that it is correlated with direction- it has both a magnitude, and a direction. Thus we introduce the idea of vectors:

Definition 6 | A quantity with only magnitude is a scalar. A quantity with magnitude and direction is a vector. A vector can thus be denoted as a directed line segment, i.e. an arrow. Note that vectors are independent of coordinates.

Now we go on to define the algebra, transformation, and calculus of vectors. We will start with notation. Vectors will be noted as $\vec{A}$ or $\mathbf{A}$, with its magnitude $\left|\vec{A}\right|$ or A. For the direction of the vector we wil denote as the unit vector $\hat{\mathbf{A}}$.

In Cartesian coordinates this course will prefer $\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$. But there is a better and more general way to note this:

$$\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}. \tag{4.1}$$

This extends to the claim:

Claim 4 | Any arbitrary vector $\hat{\mathbf{A}}$ can be expressed as a linear combination of the basis vectors $\{\hat{e}_i\}$.

Vector addition. Geometrically we define the diagonal of the parallelogram formed by the two vectors to be added. We also define multiplication by a scalar:

Definition 7 | Let $\alpha \in \mathbb{R}$ and $\hat{\mathbf{e}}_i$ as the basis vectors. Define αe^i by:

$$|\alpha\hat{\mathbf{e}}_i| = |\alpha|. \tag{4.2}$$

If $\alpha > 0$ the result is parallel to the original vector, if $\alpha < 0$ is antiparallel to the original vector, if $\alpha = 0$ the result is a zero vector.

We also define, formally, vector addition as the diagonal of the parallelogram formed by the two vector components. This is important because we can now express any vector using these three base vectors:

$$\mathbf{A} = A^1\hat{\mathbf{e}}_1 + A^2\hat{\mathbf{e}}_2 + A^3\hat{\mathbf{e}}_3. \tag{4.3}$$

Where the components of $\mathbf{A}$ are $\{A^1, A^2, A^3\}$, and they can be zero or nonzero. Here the components are in superscript because in abstract algebra (and by extension four-dimensional space), these will represent component vectors. The magnitude of $\mathbf{A}$ is given by

$$|\mathbf{A}| \equiv A = \sqrt{(A^1)^2 + (A^2)^2 + (A^3)^2}. \tag{4.4}$$

We can now write, for two arbitrary vectors $\mathbf{A}$ and $\mathbf{B}$, and scalars, α and β, we can write the following:

Definition 8 | Linear combinations:

$$\alpha\mathbf{A} + \beta\mathbf{B} = (\alpha A^1 + \beta B^1)\hat{\mathbf{e}}_1 + (\alpha A^2 + \beta B^2)\hat{\mathbf{e}}_2 + (\alpha A^3 + \beta B^3)\hat{\mathbf{e}}_3. \tag{4.5}$$

*Note that we do not need to provide definitions for subtraction because $\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B})$. There is, though, a useful way of drawing the geometric- the final vector minus the initial vector, linking the tips of $\mathbf{A}$ and $\mathbf{B}$.

This sounds like a mouthful and is complicated to write. Usually we write out the sigma in a compact form, with i from 1 through 3, but it is still too difficult. Note, though, that i is a *dummy variable*, meaning that within the summation the variables can be reassigned for other letters. They are essentially placeholders for numbers, so in any sum we can freely change the letter i. But, if the variables are not summed (when they are *free variables*), they are not interchangeable. So. with that in mind, we introduce

The Einstein Summation Convention. We are writing the following:

Claim 5

We write the vector $\mathbf{A}$ as:

$$\mathbf{A} = A^1\hat{\mathbf{e}}_1 + A^2\hat{\mathbf{e}}_2 + A^3\hat{\mathbf{e}}_3 = \sum_{i=1}^{3} A^i\hat{\mathbf{e}}_i \equiv A^i\hat{\mathbf{e}}_i. \tag{4.6}$$

Essentially we omit the sigma. Is should be one superscript and one subscript, however in Euclidean space ($\mathbb{E}^3$) this does not matter; thus we could write $A_i\hat{\mathbf{e}}_i$. However in the Minkowski space this is important. Now we introduce an example:

Example 1

Write the following using the summation convention:

1. $\mathbf{A} \pm \mathbf{B}$,

2. $\alpha\mathbf{A}$.

Solution. (a)The first one is easily deducted:

$$\alpha\mathbf{A} + \beta\mathbf{B} = (\alpha A^1 + \beta B^1)\hat{\mathbf{e}}_1 + (\alpha A^2 + \beta B^2)\hat{\mathbf{e}}_2 + (\alpha A^3 + \beta B^3)\hat{\mathbf{e}}_3 \tag{4.7}$$

Writing out the summation and dropping the sigma:

$$\mathbf{A} \pm \mathbf{B} = (A^i \pm B^i)\hat{\mathbf{e}}_i. \tag{4.8}$$

(b) the second one we already did:

$$\alpha\mathbf{A} = \alpha A^i\hat{\mathbf{e}}_i. \tag{4.9}$$

Now an exercise question:

Exercise 2

Write A and $\hat{\mathbf{A}}$ using the summation convention.

Solution. We first write out the intermediate equation:

$$A = \sqrt{\left(A^1\right)^2 + \left(A^2\right)^2 + \left(A^3\right)^2}, \tag{4.10}$$

and it would change into (for clarity we don't put extra superscripts):

$$A = \sqrt{A^i A^i}. \tag{4.11}$$

And the unit vector $\hat{\mathbf{A}}$ would be easy:

$$\hat{\mathbf{A}} = \frac{A^i\hat{\mathbf{e}}_i}{\sqrt{A^i A^i}}. \tag{4.12}$$

Scalar products / Dot products. We define an operation (the "product") that multiplies vectors and gives scalars. This is noticed as $\mathbf{A} \cdot \mathbf{B}$. It is defined as:

Definition 9

The dot product is defined as:

$$\mathbf{A} \cdot \mathbf{B} = AB\cos\theta, \tag{4.13}$$

where θ is the smallest angle between A and B. Note that it is symmetric and communicative, where $\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}$.

This introduces us to the Kronecker Delta:

$$\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \delta_{ij}, \tag{4.14}$$

where the delta is a compact form of the identity matrix.

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \tag{4.15}$$

This means we have an orthonormal basis. Now we prove that $\mathbf{A} \cdot \mathbf{B} = A^i B^i$.

PROOF | We use the Kronecker Delta:

$$\mathbf{A} \cdot \mathbf{B} = (A^i \hat{\mathbf{e}}_i) \cdot (B^j \hat{\mathbf{e}}_j) = A^i B^j \delta_{ij}$$

$$= \sum_{i=1}^{3} \sum_{j=1} 3 A^i B^j \delta_{ij} \tag{4.16}$$

$$= A^i B^i.$$

$\square$

Now we go for an exercise:

Exercise 3 | Show that:

1. $\mathbf{A} \cdot \mathbf{A} = A^i A^i = A^2$,

2. $\hat{\mathbf{e}}_i \cdot \mathbf{A} = A^i$.

Solution. (a) For the first equation, we simply plug in $\mathbf{A}$ for the general definition above and have:

$$\mathbf{A} \cdot \mathbf{A} = A^i A^i. \tag{4.17}$$

And, because we know, from Equation 4.11,

$$A = \sqrt{A^i A^i} \quad \Rightarrow \quad A^2 = A^i A^i, \tag{4.18}$$

we can prove the original equation.

(b) We expand the equation:

$$\hat{\mathbf{e}}_i \cdot \mathbf{A} = \hat{\mathbf{e}}_i \cdot (A^j \hat{\mathbf{e}}_j) = A^j \delta_{ij} = A^i. \tag{4.19}$$

Another example here: we know that the Euclidean line element

$$d\ell^2 = dx^2 + dy^2 + dz^2. \tag{4.20}$$

Example 2 | Show that:

$$d\ell^2 = \delta_{ij} dx^i dx^j. \tag{4.21}$$

Solution. We have:

$$d\ell^2 = dx^2 + dy^2 + dz^2$$
$$= (dx^1)^2 + (dx^2)^2 + (dx^3)^2$$
$$= \sum_{i=1}^{3}(dx^i)^2 = \sum_{i=1}^{3}dx^i dx^i \qquad (4.22)$$
$$= \sum_{i=1}^{3}\sum_{j=1}^{3}\delta_{ij}dx^i dx^j.$$

Drop the summation and we get the result. This is why δ_{ij} is called the Euclidean metric.

An extension of this example is this: Suppose

$$ds^2 = g_{\mu\nu}(x)dx^\mu dx^\nu, \qquad (4.23)$$

solving for g gives you the function for gravity, which is the definition for "geometry of space". Also when we define

$$ds^2 \equiv \left(1 - \frac{2M}{r}\right)c\,dt^2 + \left(1 - \frac{2M}{r}\right)^{-1}dr^2 + r^2\sin^2\theta\,d\Omega^2, \qquad (4.24)$$

we can determine the event horizon of a black hole because when $r = 2M$, the second differential gets to infinity.

We now define the **Cross product**.

Definition 10

Multiply two vectors and get a vector through $\mathbf{A} \times \mathbf{B}$. We have:

$$\mathbf{A} \times \mathbf{B} = AB\sin\theta\,\hat{\mathbf{n}}, \qquad (4.25)$$

where the direction of $\hat{\mathbf{n}}$ is defined by the right hand rule. Note that this is anti-symmetric product (thus evidently not associative), because by the right hand rule,

$$\mathbf{A} \times \mathbf{B} = -\mathbf{B} \times \mathbf{A}. \qquad (4.26)$$

Now we need to write down the cross product of the basis vectors. We have:

$$\hat{\mathbf{e}}_i \times \hat{\mathbf{e}}_j = \epsilon_{ijk}\hat{\mathbf{e}}_k. \qquad (4.27)$$

The ϵ is defined as the Levi-Civita symbol, where

$$\epsilon_{ijk} = \begin{cases} 1, & \text{if } \{i,j,k\} \text{ is an even permutation of } \{1,2,3\}, \\ 1, & \text{if } \{i,j,k\} \text{ is an odd permutation of } \{1,2,3\}, \\ 0, & \text{otherwise.} \end{cases} \qquad (4.28)$$

Note that k is a free variable. The exercise now is to expand the vector multiplication of the base vectors.

Exercise 4

The first one is

$$\hat{\mathbf{e}}_1 \times \hat{\mathbf{e}}_2 = \epsilon_{121}\hat{\mathbf{e}}_1 + \epsilon_{122}\hat{\mathbf{e}}_2 + \epsilon_{123}\hat{\mathbf{e}}_3 = \epsilon_{123}\hat{\mathbf{e}}_3 = \hat{\mathbf{e}}_3, \qquad (4.29)$$

The second one is:

$$\hat{\mathbf{e}}_2 \times \hat{\mathbf{e}}_3 = \epsilon_{231}\hat{\mathbf{e}}_1 + \epsilon_{232}\hat{\mathbf{e}}_2 + \epsilon_{233}\hat{\mathbf{e}}_3 = \epsilon_{231}\hat{\mathbf{e}}_1 = \hat{\mathbf{e}}_1, \tag{4.30}$$

And the last one is:

$$\hat{\mathbf{e}}_1 \times \hat{\mathbf{e}}_3 = \epsilon_{132}\hat{\mathbf{e}}_2 = -\hat{\mathbf{e}}_2. \tag{4.31}$$

Now consider two arbitrary vectors $\mathbf{A}$ and $\mathbf{B}$. We have:

$$\begin{aligned}
\mathbf{A} \times \mathbf{B} &= (A^i\hat{\mathbf{e}}_i) \times (B^i\hat{\mathbf{e}}_j) \\
&= A^i B^j (\hat{\mathbf{e}}_i \times \hat{\mathbf{e}}_j) = A^i b^j (\epsilon_{ijk}\hat{\mathbf{e}}_k) \\
&= \epsilon_{ijk} A^i B^j \hat{\mathbf{e}}_k \\
&= \epsilon_{123} A^1 B^2 \hat{\mathbf{e}}_3 + \epsilon_{132} A^1 B^3 e^2 + \epsilon_{213} A^2 B^1 \hat{\mathbf{e}}_3 \\
&\quad + \epsilon_{231} A^2 B^3 \hat{\mathbf{e}}_1 + \epsilon_{312} A^3 B^1 \hat{\mathbf{e}}_2 + \epsilon_{321} A^3 B^2 \hat{\mathbf{e}}_1 \\
&= (A^2 B^3 - A^3 B^2)\hat{\mathbf{e}}_1 + (A^3 B^1 - A^1 B^3)\hat{\mathbf{e}}_2 + (A^1 b^2 - A^2 B^1)\hat{\mathbf{e}}_3.
\end{aligned} \tag{4.32}$$

This is cleaner than the traditional determinant format because it is reproducible in higher dimensions. Note though, that only the Levi-Civita symbol can be defined, and in four dimensions or more, we cannot define the cross product comfortably because there is one more dimension of ambiguity. Now we have an example.

Example 3 | Show that the identity holds:

$$\epsilon_{ijk}\epsilon_{kmn} = \delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}. \tag{4.33}$$

We know that for this the single free variable is k.

Also: Prove the identity

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C}. \tag{4.34}$$

Change of basis. Our basis vectors are coordinate basis vectors, derived from the coordinates we choose. Hence, if we change the coordinate, it implies a change of basis. The components would change after changes of basis, however, the dimensions of the vector itself never changes. We now discuss the change of coordinates through (a) a static rotation, and changing Cartesian to polar coordinates, both in 2D.

We start with rotation. Suppose we have the basic 2D Cartesian coordinates $\mathcal{S}$. Now we rotate it by an angle θ and get $\bar{\mathcal{S}}$. What we need is to get how are the coordinates $(\bar{x}, \bar{y})$ related to (x, y)? First, from the figure, we have:

$$\begin{aligned}
\bar{x} &= x\cos\theta + y\sin\theta, \\
\bar{y} &= -x\sin\theta + y\cos\theta.
\end{aligned} \tag{4.35}$$

Sidenote: this raises an immediate question: since distance is preserved, what are the coordinate transformations that preserve the Euclidean line element— that is,

$$\mathrm{d}\bar{l}^2 = \mathrm{d}l^2 \text{ ?} \tag{4.36}$$

Clearly, any $(\bar{x}, \bar{y}) = (x + a, y + b)$ where a and b are constants preserves $\mathrm{d}l^2$. This

restricts us to linear transformations:

$$\bar{x} = Ax + By,$$
$$\bar{y} = Cx + Dy.$$

(4.37)

To prove the uniformity of $\mathrm{d}\bar{l}^2$:

$$\begin{aligned}
\mathrm{d}\bar{l}^2 &= \mathrm{d}\bar{x}^2 + \mathrm{d}\bar{y}^2 \\
&= (A^2 + C^2)\mathrm{d}x^2 + (B^2 + D^2)\mathrm{d}y^2 + 2(AB + CD)\mathrm{d}x\mathrm{d}y \\
&= \mathrm{d}x^2 + \mathrm{d}y^2.
\end{aligned}$$

(4.38)

This gives us:

$$\begin{aligned}
A^2 + C^2 &= 1, \\
B^2 + D^2 &= 1, \\
AB + CD &= 0.
\end{aligned}$$

(4.39)

Solve the equation and we can get the matrix in the equation below (4.40):

$$\begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} = x \cdot \begin{pmatrix} \cos\theta \\ -\sin\theta \end{pmatrix} + y \cdot \begin{pmatrix} \sin\theta \\ \cos\theta \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

(4.40)

The basis vectors will change as any vector in the coordinate. Thus we can write a shorthand of:

$$\hat{\mathbf{e}}_{\bar{i}} = \hat{\mathbf{e}}_j R_{j\bar{i}}.$$

(4.41)

For any vector $\mathbf{A}$, we must have $\bar{\mathbf{A}} = \mathbf{A}$. Hence, we would need:

$$\bar{\mathbf{A}} = \mathbf{A}^{\bar{i}}\hat{\mathbf{e}}_{\bar{i}} = \mathbf{A}^{\bar{i}}(\hat{\mathbf{e}}_j R_{j\bar{i}})\hat{\mathbf{e}}_j \equiv \mathbf{A}^j\hat{\mathbf{e}}_j = \mathbf{A}.$$

(4.42)

This implies that:

$$\mathbf{A}^{\bar{i}} R_{j\bar{i}} = \mathbf{A}^j.$$

(4.43)

If we would like to get $\mathbf{A}^{\bar{i}}$ from $\mathbf{A}^j$ we would like an inverse that has the property

$$R_{\bar{k}j} R_{j\bar{i}} = \delta_{\bar{k}\bar{l}}.$$

(4.44)

This would imply that

$$\mathbf{A}^{\bar{i}} R_{\bar{k}j} R_{j\hat{i}} = \mathbf{A}^{\bar{i}}\delta_{\bar{k}\bar{i}} = \mathbf{A}^{\bar{k}} = R_{\bar{k}j}\mathbf{A}^j.$$

(4.45)

This introduces the rotation matrix, $\mathcal{R}$, where

$$\begin{aligned}
(\mathcal{R})_{kj} = R_{\bar{k}j} &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \\
(\mathcal{R}^{-1})_{kj} = R_{\bar{j}\bar{k}} &= \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.
\end{aligned}$$

(4.46)

Note that $\mathcal{R}^{-1} = \mathcal{R}^T$ and thus $\mathcal{R}^{-1}\mathcal{R} = \mathcal{R}^T\mathcal{R} = \mathbb{I}$, and that $\det(\mathcal{R}) = 1$. The set of 2×2 matrices for which $\mathcal{R}^{-1} = \mathcal{R}^T$ and $\det(\mathcal{R}) = 1$ is known as the Special Orthogonal Group, $SO(2)$. An element of the set is called a rotation matrix because its operation on vectors is to rotate them through a fixed angle θ. Now an example:

Exercise 5

Consider $\mathrm{d}x^i$ as the components of a vector $\mathrm{d}\mathbf{x}$. Show, using rotational matrices $(\mathcal{R}_{j\bar{i}})$, that $\mathrm{d}\bar{\ell}^2 = \mathrm{d}\ell^2$.

Solution. We expand $\mathrm{d}\bar{\ell}^2$:

$$\begin{aligned}
\mathrm{d}\bar{\ell}^2 &= \delta_{\bar{i}\bar{j}}\mathrm{d}x^{\bar{i}}\mathrm{d}x^{\bar{j}} \\
&= \delta_{\bar{i}\bar{j}}\left(\mathcal{R}_{\bar{i}m}\mathrm{d}x^m\right)\left(\mathcal{R}_{\bar{j}n}\mathrm{d}x^n\right) \\
&= \delta_{\bar{i}\bar{j}}\mathcal{R}_{\bar{i}m}\mathcal{R}_{\bar{j}n}\mathrm{d}x^m\mathrm{d}x^n \\
&= \mathcal{R}_{\bar{i}m}\delta_{\bar{i}\bar{j}}\mathcal{R}_{\bar{j}n}\mathrm{d}x^m\mathrm{d}x^n \quad\quad (4.47) \\
&= \mathcal{R}_{\bar{i}m}\mathcal{R}_{\hat{i}n}\mathrm{d}x^m\mathrm{d}x^n \\
&= \left(\mathcal{R}^T\right)_{m\bar{i}}(\mathcal{R})_{\bar{i}n}\,\mathrm{d}x^m\mathrm{d}x^n \\
&= \delta_{mn}\mathrm{d}x^m\mathrm{d}x^n = \mathrm{d}\ell^2.
\end{aligned}$$

Polar coordinates. We express polar coordinates with (r, θ). The definitions are well-known, but note that there is a *coordinate singularity* at the origin: when $r = 0$, θ can be any value. This, when we change back to Cartesian coordinates, will disappear, and we know that this is not a physical singularity. This will be studied more in general relativity (4.24).

The polar coordinates are effectively rotating θ throughout $\theta \in [0, 2\pi]$. This results in the definitions of the basis vectors to be related to the value of θ:

$$\begin{aligned}
\hat{\mathbf{e}}_r(\theta) &= \hat{\mathbf{e}}_x\cos\theta + \hat{\mathbf{e}}_y\sin\theta, \\
\hat{\mathbf{e}}_\theta(\theta) &= -\hat{\mathbf{e}}_x\sin\theta + \hat{\mathbf{e}}_y\cos\theta.
\end{aligned} \quad\quad (4.48)$$

While calculating the derivatives- we need to keep in mind when we take derivatives.

SECTION 5

Vector Calculus

We define a time dependent vector, $\mathbf{A} = \mathbf{A}(t)$. Now we derive its derivative with respect to time. We claim that:

$$\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} = \lim_{\Delta t \to 0}\frac{\mathbf{A}(t + \Delta t) - \mathbf{A}(t)}{\Delta t}. \quad\quad (5.1)$$