Notes for Lecture 1

Introduction

Let us make sure that we understand, comfortably, some really basic and really important expressions such as particle, physical dimension, vector, transformation, scalar etc. These expressions are fundamental to this course and to the whole physics, and will occur again and again.

1.1 General Overview

You may have learned a lot about Newtonian mechanics in lower level courses. In some minimal sense, you have learned everything already: you know Newton's laws and you can use free body diagrams and conservation principles to solve various problems. What more is there to learn? A lot more. We will see that it is not just a more sophisticated calculation of the same old Newton's laws.

Symmetry and conservation We will learn the true meaning (the "symmetry principle") of the momentum conservation, the energy conservation, and the angular momentum conservation. This is the very one thing that every student of mechanics should learn and appreciate.

More sophisticated treatment of Newton's law problems We will also learn how to treat Newton's law problems more in depth—the perturbation theory, coupled oscillators, and rigid body etc. It is not just more difficult problems but more conceptually sophisticated frameworks that we will learn.

Complexity We will see that simple ingredients can give quite complex outcome due to non-linear interactions or many particle behaviors. These are the topics of the chaos, coupled-oscillators, and waves.

New formalisms We learn the Lagrangian mechanics and the Hamiltonian mechanics, which, while equivalent to the Newtonian mechanics, are of greater value, in the sense that they can be more easily generalized to other advanced topics, such as quantum mechanics or quantum field theory.

Perturbation The perturbation theory has been mentioned already, but you cannot overemphasize it, and it is of general value far beyond Newtonian mechanics. This technique, while simple, is of great general importance outside mechanics and even outside physics, and should be introduced importantly at this stage.

1.2 Particle, body, degrees of freedom, dimension

Paricle, Body The "particle" is a fundamental concept in classical mechanics. By "particle" we mean an object whose size can be ignored. It is a mathematical point, assigned with a mass value. This is an approximation, of course, and what we call "particle" in classical mechanics is a sizable object, while it can vary greatly in size. For instance, when we consider the motion of the Earth around the Sun, then we might call the Earth a "particle." Or a "body," since we are talking about a celestial body. But, when we consider an apple falling from a tree, then the Earth is definitely not a particle, but we might call the apple a "particle" to a good approximation. But, when the apple hits the ground and breaks into hundreds of pieces, then . . . (I hope you get the idea). In any case, note that "particle" in classical mechanics should not be confused with "fundamental particle" such as the electron, the neutrino, the proton etc. The mechanical law governing these fundamental particles is quantum mechanics, not classical mechanics. In general, a particle in classical mechanics is a composite object consisting of a very large number of fundamental particles. For instance, a cat can be taken as a classical particle, in many circumstances, but it can hardly be taken as a quantum particle. Indeed, Newton's law should be thought of as an emerging law when a large number of quantum particles are coalesced together. As such any "particle" in classical mechanics have a large internal degrees of freedom (see below). When the internal degrees of freedom are important, the object can no longer be considered as a point, and the term "body" would be much more appropriate. Even then, when we refer to the average motion of the body, we may still use the term "particle."

Dimension The spatial dimension is the number of coordinates to specify the position of one particle. In classical mechanics, time is just a parameter, and so the spatial dimension is the only dimension that we care about. The spatial dimension of our world is 3. Consider an airplane flying in the sky, we need 3 coordinates. These could be the Cartesian coordinates, (x, y, z), the spherical coordinates, (r, θ, ϕ), the cylindrical coordinates (ρ, ϕ, z), or most likely in real life, (longitude, latitude, altitude). Whatever coordinate system we use, the number of coordinates is 3. Thus, the (spatial) dimension is 3. This is true for any motion. However, if a particle is constrained to go through a linear motion only, then effectively we need only one coordinate, say x, to specify its position. Thus, the effective dimension is 1. I will use the symbol, D, for dimension. $D = 3$ for 3 dimensions, and $D = 1$ for 1 dimension, etc. Also, I will use the short-hand 1D, 2D, or 3D, to mean one-dimensional, two-dimensional, or three-dimensional, respectively.

Degrees of freedom For a given mechanical system, the degrees of freedom refers to the number of coordinates necessary to specify the positions of all particles. For one particle system, the degrees of freedom is thus equal to D. For a many particle system consisting of N particles, the degrees of freedom is ND.

Dimension A more general concept than the spatial dimension is the dimension of a physical quantity. Recall that the seven base units of the SI unit system are m, kg, s, A, K, cd, and mol. For mechanical problems, we are concerned with the first three only. Let us say that a mechanical quantity has the SI unit $\mathrm{m}^\alpha \mathrm{kg}^\beta \mathrm{s}^\gamma$. Then, the **(physical) dimension** of that quantity is defined as $L^\alpha M^\beta T^\gamma$, where L means length, M means mass, and T means time. For example, the angular frequency ω has the dimension T^{-1}. We say that it has the dimension of inverse time. The energy has the dimension $L^2 M T^{-2}$. In thermodynamics, we learn that the heat has the same dimension as the energy. **No two physical quantities can be equal to each other, if their dimensions are different.** In other words, two quantities can be compared, added or subtracted, only if their dimensions are identical. For this reason, if you solve a problem using symbols, then the first thing that you must check is the dimension. This is because, if the dimension is incorrect, then there must have been a mistake that you need to correct. In short, when you do any physics problem, checking dimensions should be part of your second nature. There are some more mistake-proofing guidelines summarized in a Homer box below—we will use them whenever we do problems in this course.

1.3 Vectors and matrices

Solving a mechanical problem usually requires setting up a coordinate system. In doing so, we are free to choose a coordinate system that is the most useful and

How to avoid getting things wrong...

As you undoubtedly know by now, you should do almost all physics problems **symbolically**. That is, you should use symbols for variables, obtain a symbolic expression for your answer first, before finally plugging in numbers to get a numerical answer, if required. Next, you should double check your answer. The better you are, the more you know about your weakness as well as your strength, and double checking should be pretty much an instinctive routine while you are doing problems. Here, I offer some guidelines how to **double check your answers**, and be sure about your answer, before anybody makes that potentially unpleasant judgement about your answer. You should apply these guidelines to your symbolic answer first and foremost, if applicable, and then to your numeric answer. These guidelines are mighty important to prevent any potential embarrassment, not to mention a deep negative impact on your scores. The potential negative impact that you will be avoiding by heeding these rules can be immense for the top item of the list, and goes down gradually in its severity as the list goes down.

- Does the **physical dimension** of my answer make sense?
- Does the **sign** or **(scaling) trend** make sense?
- Does the answer make sense in **known limits**, if any?
- Does the **order of magnitude** of my answer make sense?
- D'oh! Did I drop 2, π, ... somewhere?

the most elegant *for us*. **The eventual physics answer is independent of the coordinate system**, so it does not matter what coordinate system we use. The coordinate system is something that we draw in space, out of the blue, just to make it easy to calculate things. It is an essential device for us, but physics does not, and should not, depend on our choice of the coordinate system.

Consider the Cartesian coordinate system. The coordinates consist of D numbers, and in three dimensions they are written as x, y, z. The position corresponding to these coordinates are usually denoted with a symbol such as $\vec{r}$ or $\vec{x}$. For the reason that will become clear below, we will write coordinates **column-wise**, as in

$$\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \qquad \text{if } D = 3 \tag{1.1}$$

Now, imagine rotating the coordinate system, or reflecting the coordinate system (like reflecting the world in a mirror), or inverting the coordinate system (reflecting all coordinates), while, in all cases, the origin of the coordinate system remains fixed. These are examples of **coordinate transformations**. We consider the particle position as fixed, as we instantaneously make the coordinate transformation. Namely, physics is one and the same, but our description can be different depending on the coordinate system. In the new coordinate system, the same position is now represented by different numbers, x', y', z'.

$$\vec{r}\,' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \qquad \text{if } D = 3 \tag{1.2}$$

Coordinate transformations considered above, namely rotation, reflection, and inversion, are **orthogonal transformations**.

Coordinate transformation $T : \vec{r} \to \vec{r}\,' = T(\vec{r})$

In general, $T(\vec{r})$ can be any function. Correspondingly, the transformation can be linear (see below) or non-linear.

Linear transformation $L : \vec{r} \to \vec{r}\,' = L(\vec{r}) = \overleftrightarrow{L}\vec{r}$

Here, $\overleftrightarrow{L}$ is a "**square matrix**." In general, a **matrix** means a rectangular array of numbers, consisting of m rows and n columns. A square matrix means $m = n$.

Any coordinate transformation that displaces the origin is an example of a non-linear transformation.

Orthogonal transformation $O : \vec{r} \to \vec{r}\,' = \overleftrightarrow{O}\vec{r}$

Here, $\overleftrightarrow{O}$ is an **orthogonal matrix**, which can be defined as a square matrix satisfying any one of the following four equivalent properties. (1) The column vectors of $\overleftrightarrow{O}$ are orthonormal. That is, each column vector is a unit length vector, and perpendicular to one another. (2) So are the row vectors of $\overleftrightarrow{O}$. (3) $\overleftrightarrow{O}\overleftrightarrow{O}^t = \overleftrightarrow{O}^t\overleftrightarrow{O} = \overleftrightarrow{1}$ (t = transpose). (4) $\overleftrightarrow{O}^{-1} = \overleftrightarrow{O}^t$.

So, an orthogonal transformation[1] is a special kind of linear transformation. Also, note that expressions such as $\overleftrightarrow{L}\vec{r}$ and $\overleftrightarrow{O}\vec{r}$ make sense as matrix multiplication, only if $\vec{r}$ is a *column* vector, which is our convention here. Finally, note that I am using a bi-directional arrow over a symbol to mean a matrix quantity, as in Gibbs' "dyadic notation." In particular, $\overleftrightarrow{1}$ means the identity matrix, whose diagonal elements are all 1's and whose non-diagonal elements are all 0's. A quick and dirty but very often used short-hand notation for $\overleftrightarrow{1}$ is 1. That is, if you see an expression such as $\overleftrightarrow{O}\overleftrightarrow{O}^t = 1$, you should automatically upgrade 1 on the right hand side to an identity matrix of correct dimensions. This applies not only to 1, but also to any number, which, if equated with a matrix, should be interpreted as that number times the identity matrix.

It is worth noting the **fundamental definition of a linear transformation**:

$$L(a\vec{r}_1 + b\vec{r}_2) = aL(\vec{r}_1) + bL(\vec{r}_2) \tag{1.3}$$

for any positions $\vec{r}_1$ and $\vec{r}_2$ and any numbers a, b. This definition is completely equivalent to the above definition: L is a linear transformation if L can be represented by a matrix multiplication $L(\vec{r}) = \overleftrightarrow{L}\vec{r}$.

Just like the position, any physical quantity can be represented as a set of numbers, given a coordinate system. Let's take an arbitrary physical quantity. We measure it in one coordinate system ($\vec{r}$), and call it Q. The same quantity can be measured in the

[1]Please note that in the first few lectures, we will not consider complex numbers; we will confine ourselves to real numbers only. For example, discussion in this chapter is limited to real matrices. Thus, an orthogonal matrix must be a real matrix. In case you need to generalize to complex numbers (not commonly in classical mechanics, but absolutely from square one in quantum mechanics), you must generalize many concepts to those appropriate for complex numbers. For example, an orthogonal matrix must be generalized to a **unitary matrix**. A unitary matrix is one whose inverse is its "Hermitian conjugate" (transpose followed by complex conjugate). The Hermitian conjuration is usually denoted by a † (dagger) symbol, as in $\overleftrightarrow{U}^\dagger$. So, a complex matrix is unitary if and only if $\overleftrightarrow{U}\overleftrightarrow{U}^\dagger = 1$. It follows that any orthogonal matrix is also a unitary matrix, but the converse is not true in general.

transformed coordinate system ($\vec{r}\,'$), and we will call it Q'. How the transformation from Q to Q' is related to, or not related to, the coordinate transformation itself, is an important characteristic of the physical quantity. For one, that is how we define a vector quantity and a scalar quantity. The following is a more precise definition than a vague definition that one encounters in elementary physics courses.

Vector Any physical quantity whose representation, $\vec{V}$, transforms just like the position $\vec{r}$, for an arbitrary orthogonal coordinate transformation, is called a vector quantity. Namely, $\overleftrightarrow{O}\vec{V} = \vec{V}\,'$.

Examples of vector quantities include position (by definition!), velocity, momentum, force, angular momentum, and acceleration.

Scalar Any physical quantity whose representation, S, remains unchanged by an arbitrary orthogonal coordinate transformation is called a scalar quantity. Namely, $S = S'$.

For instance, time is independent of coordinate systems in classical mechanics, and thus it is a scalar quantity.[2] Mass is another example. For given vector quantities, scalar quantities can be derived from them also. E.g., the magnitude of a vector and the angle between two vectors are scalar quantities, according to the following property.

[2]In the relativistic mechanics of Einstein, time is no longer absolute when speeds close to the speed of light are involved. Instead, time should be considered as another axis of the coordinate system, part of the four dimensional "space-time" vector. It is no longer a scalar.

Scalar product is scalar, indeed.

The **scalar product** of two vectors, $\vec{A}$ and $\vec{B}$, is defined as

$$\vec{A} \cdot \vec{B} = \sum_i A_i B_i$$

where A_i's (B_i's) are the Cartesian components of $\vec{A}$ ($\vec{B}$). It is invariant under an orthogonal transformation. In other words, the scalar product is indeed **a scalar quantity for orthogonal transformations**.

We will see how this property arises, in a little bit, but let us discuss some key facts, first.

The scalar product of two vectors can also be understood as a matrix multiplication (recall that a vector is also a matrix after all).

$$\begin{aligned} \vec{A} \cdot \vec{B} &= \vec{A}^t \vec{B} \\ &= \vec{B}^t \vec{A} \end{aligned} \tag{1.4}$$

On the right-hand side, the matrix multiplication of a row vector ($\vec{A}^t$ or $\vec{B}^t$) and a column vector ($\vec{B}$ or $\vec{A}$, respectively) is seen to result in a number[3].

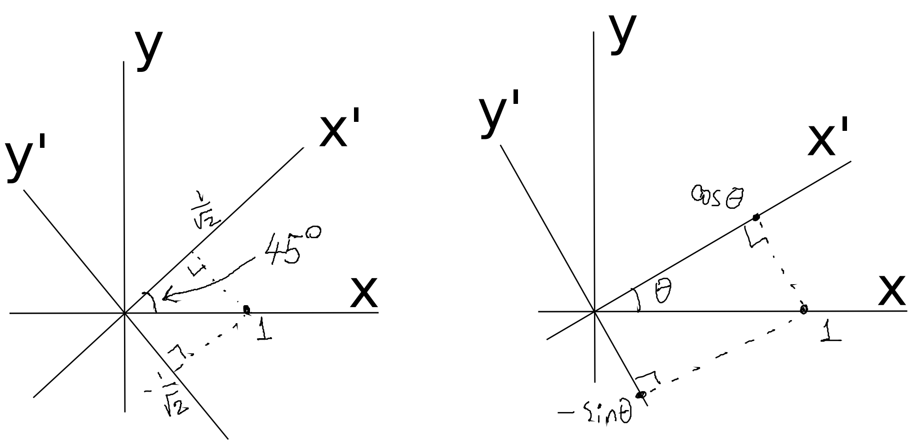
Scalar product has other names: **dot product** and **inner product**. Be careful not to change the order of a matrix product, since that is generally not allowed! In the current case, definitely $\vec{A}^t \vec{B} \neq \vec{B} \vec{A}^t$. Actually, $\vec{B} \vec{A}^t$ results in a square matrix of dimensions $D \times D$ (i.e., D rows and D columns)! This is the so-called the **outer product** of two vectors. Finally, a third kind of vector product is possible: this is the **vector product** or the **cross product**: $\vec{A} \times \vec{B}$. We will discuss the vector product later, when we need to. Different from other products, the vector product is defined only in 3D or 7D.

Often, a description of a linear coordinate transformation is given, and you need to construct the transformation matrix fast. Here is how you do it.

[3] It should be noted that the left hand side has a center dot, while the right hand side (being a matrix multiplication) does not. This difference in the notation must be noted with care.

How to construct a transformation matrix fast?

- Figure out how old unit vectors, $\hat{x}$ first, then $\hat{y}$ and so on, are represented in the new coordinate system.
- Write your answers as column vectors from left to right.
- Voilà, you have the transformation matrix.



For example, consider rotating a 2D coordinate system by 45° (the left figure) with the position vectors themselves not changing. This sort of transformation is sometimes called a *passive transformation.*[4] In any case, let us suppose that a vector is represented as $\vec{r} = \binom{x}{y}$ in the xy coordinate system, and as $\vec{r}\,' = \binom{x'}{y'}$ in the $x'y'$ coordinate system. What is the matrix $\overleftrightarrow{O}$ that will take $\vec{r}$ and then map it to $\vec{r}\,'$ for a general vector: $\vec{r}\,' = \overleftrightarrow{O}\vec{r}$? Due to the linear nature of the rotation[5] all we need for figuring out the matrix is figuing out how two basis vectors[6] transform. In this case, $\binom{1}{0} \to \frac{1}{\sqrt{2}}\binom{1}{-1}$, and $\binom{0}{1} \to \frac{1}{\sqrt{2}}\binom{1}{1}$. Therefore, the transformation matrix is given

[4] The same situation could be described by an *active transformation* in which the coordinate system remains fixed and position vectors themselves are rotated by -45° degrees. In general, if vectors themselves are changed while the coordinate system remains unchanged, we have an active transformation, while we have a passive transformation if vectors themselves are viewed as unchanging, while the coordinate system changes.

[5] For your satisfaction, prove that rotation is indeed a linear operation, in view of the definition Eq. 1.3

[6] Or, any two linearly independent vectors, generally speaking.

by $\frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 & 1 \\ -1 & 1 \end{smallmatrix}\right)$. In general, rotating a 2D coordinate system by θ is described by the following transformation matrix:

$$\overleftrightarrow{O} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{1.5}$$

In terms of $\vec{o}_1 = \left(\begin{smallmatrix} \cos\theta \\ -\sin\theta \end{smallmatrix}\right) = \overleftrightarrow{O}\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ and $\vec{o}_2 = \left(\begin{smallmatrix} \sin\theta \\ \cos\theta \end{smallmatrix}\right) = \overleftrightarrow{O}\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$, we can re-write this in a more general form with a short-hand notation[7]

$$\overleftrightarrow{O} = \begin{pmatrix} \vec{o}_1 & \vec{o}_2 \end{pmatrix} \tag{1.6}$$

Using this form, we can discuss the general properties of an orthogonal matrix. Now, the fundamental definition of an orthogonal transformation is that the column vectors are orthonormal to each other, $\vec{o}_i^{\,t}\vec{o}_j = \delta_{i,j}$, where $\delta_{i,j}$ is the **Kronecker-delta symbol** (1 if $i = j$ and 0 otherwise). If this is the case, then

$$\overleftrightarrow{O}^t\overleftrightarrow{O} = \begin{pmatrix} \vec{o}_1^{\,t} \\ \vec{o}_2^{\,t} \end{pmatrix}\begin{pmatrix} \vec{o}_1 & \vec{o}_2 \end{pmatrix} = \begin{pmatrix} \vec{o}_1^{\,t}\vec{o}_1 & \vec{o}_1^{\,t}\vec{o}_2 \\ \vec{o}_2^{\,t}\vec{o}_1 & \vec{o}_2^{\,t}\vec{o}_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

This proves that the inverse of an orthogonal matrix is its transpose ($\overleftrightarrow{O}^{-1} = \overleftrightarrow{O}^t$), from which all the rest of the properties of an orthogonal matrix, as listed a few pages back, follow.

What we just showed generalizes to an orthogonal matrix in any dimensions.

Now, let us consider two vectors $\vec{A}$ and $\vec{B}$. Under an orthogonal transformation, $\vec{A}' = \overleftrightarrow{O}\vec{A}$ and $\vec{B}' = \overleftrightarrow{O}\vec{B}$. The scalar product

$$\begin{aligned} \vec{A}' \cdot \vec{B}' &= (\overleftrightarrow{O}\vec{A})^t\overleftrightarrow{O}\vec{B} && (\because \vec{A}\cdot\vec{B} = \vec{A}^t\vec{B}) \\ &= \vec{A}^t\overleftrightarrow{O}^t\overleftrightarrow{O}\vec{B} && (\text{matrix transpose rule } (MN)^t = N^tM^t) \\ &= \vec{A}^t\overleftrightarrow{1}\vec{B} && (\because \overleftrightarrow{O} \text{ is an orthogonal matrix.}) \\ &= \vec{A}^t\vec{B} \\ &= \vec{A}\cdot\vec{B} \end{aligned}$$

[7]Note that in this new notation we are repeating the brackets () for the column vectors and then for the matrix. This is no problem. Those brackets are just visual aids, and not an integral part of the definition of a matrix. A matrix is simply a rectangular array of numbers.

This shows that the scalar product is invariant under any orthogonal transformation. So is the **magnitude of a vector**, since $A \stackrel{def}{=} |\vec{A}| \stackrel{def}{=} \sqrt{\vec{A}\cdot\vec{A}}$, and the **angle between two vectors**, $\angle(\vec{A},\vec{B}) \stackrel{def}{=} \cos^{-1}\left[\frac{\vec{A}\cdot\vec{B}}{|\vec{A}||\vec{B}|}\right]$.

What kind of transformations are linear, but not orthogonal? Stretching or skewing (shearing), with the origin fixed.