Introduction

This course will be on classical mechanics. Classical mechanics was developed by Isaac Newton in the 17th century and for over two hundred years until near the end of the 19th century it was physics. Over those two hundred years, many great physicists (Lagrange, Hamilton, Poisson and others) developed sophisticated formulations of classical mechanics that allowed complicated problems involving multiple particles subject to multiple constraints to be solved. These were, however, purely mathematical developments. The physics of classical mechanics was completely set forward by Newton.

Despite its great success in describing the physics of our everyday lives, classical mechanics is not a correct theory. It is simply an (excellent) approximation at the macroscopic scale to the deeper theory of quantum mechanics. Since classical mechanics is not a correct theory, it might seem strange that we start our journey with it. Why don’t we instead jump ahead to quantum mechanics? There are many reasons for starting with classical mechanics. In any case, we are in good company. All great physicists, including Albert Einstein and Richard Feynman, first learned classical mechanics. Let’s see why.

Reasons for starting with Classical Mechanics

There are several reasons for beginning our physics journey with classical mechanics. Here are some of the more important ones.

1) Classical mechanics has important historical significance. Although, in general, when studying physics it isn’t very helpful to be concerned with the historical development, nevertheless, it is valuable to know what great physicists of the 18th and 19th centuries were thinking and doing. We can gain valuable insights by studying, or at least being familiar with, their work.

2) Classical mechanics is intuitive. Since the physics of our everyday lives has important survival value, the brains of humans and other animals evolved an intuitive understanding of physics as described by classical mechanics. We intuitively know what will happen when, for example, a ball is thrown in the air. Its behavior doesn’t seem strange to us. Even a dog or a young child are able to catch a thrown ball. They are in a sense solving for the motion of the ball although certainly not in a mathematical form. If our ancestors had not known how a spear would fly through the air or how a tiger would jump, they would not have fared very well. Knowledge of quantum mechanics, on the other hand, had no survival value and so our brains didn’t evolve to have an intuitive sense of it. The quantum mechanical behavior of systems at the microscopic level does not at all follow our intuition or common sense. It is, therefore, best to
start our study of physics with classical mechanics, a theory that "makes sense" to us.

3) There are three important dynamical quantities in physics: energy, momentum and angular momentum. They arise from the basic principles of quantum mechanics but, if we first encountered them there, they would be abstract mathematical quantities. However, at the macroscopic scale approximated by classical mechanics, they become the classical quantities of energy, momentum and angular momentum. From experiences in our everyday lives, we have a concrete understanding of what these are. We know that a moving ball has energy and when it hits us we realize it has momentum. We also intuitively understand the angular momentum of a spinning top or a spinning ice skater. We are then more readily able to grasp the more abstract concepts of energy, momentum and angular momentum at the microscopic scale.

4) Although it isn’t a correct theory, classical mechanics is an extremely good effective theory at the macroscopic scale. Many of the problems that we might want to solve at this scale are best handled by classical mechanics. If we want to describe the flight of a ball, the motion of a roller coaster, the orbit of a planet or the trajectory of a spacecraft, we would best use classical mechanics. Trying to solve problems at this scale using quantum mechanics would be silly.

5) In developing mathematical formulations of the physics first proposed by Newton, physicists of the 18th and 19th century discovered concepts that were far deeper than classical mechanics itself. We will soon see examples of these formulations. They include the principle of stationary action, Lagrangians, Hamiltonians and Poisson brackets. All of these concepts lie deep in physics. The early physicists were getting a glimpse of the deeper level of quantum mechanics on which classical mechanics is based, but they had no realization that was what they were seeing. That would have to wait until the heroic efforts by physicists in the early 20th century. In terms of our study of fundamental physics, this is the most important reason for starting with classical mechanics. Here in the more intuitive realm we will get our first exposure to these deep concepts that will then carry through all of physics.

**Mathematics**

Mathematics is the language of physics. We would not be able to seriously discuss physics without knowing some basic ideas of mathematics. Before we begin our study of classical mechanics, we’ll review most of the mathematics that we will be need. The math subjects that we’ll cover are: differential calculus, integral calculus, complex variables, trigonometry, differential equations, vectors and vector calculus. For some, this will be a review while others might be seeing the topics for the first time. Despite the impression that you might have from high school or college math courses, these are not difficult subjects. There are only a few basics things that you need to know and we will cover them as simply as possible.

**Mathematics** is a collection of topics: calculus, number theory, geometry, etc. Each topic is based on a set of axioms and postulates on which is built a full structure through theorems derived through logical proofs using the axioms and previously de-
rived theorems. Math would seem to have nothing to do with “reality” but rather exists in its own Platonic world based solely on axioms and theorems.

One of the great, perhaps the greatest, mysteries is why mathematics provides a description and explanation of the real world. We don’t know why this relationship between mathematics and physics exists and perhaps we never will. In the meanwhile, it is a continuous source of wonderment that abstract mathematics is somehow fundamentally related to reality and constrains the way nature behaves.

The roles of mathematicians and physicists are very different. Mathematicians build a Platonic world with no concern for reality. They only care that this world is consistent and based on rigorous proofs. Physicist, on the other hand, simply want to use mathematics as a tool to describe nature. They are generally not concerned with the rigor of the mathematics. For them, the correctness of the mathematics comes from its ability to provide a correct description of nature. If the physics that they develop from the mathematics gives a good description of nature, then it is assumed that the math is right and that it could be shown to be rigorous if desired. This is very troubling to mathematicians who don’t have the real world as a test but can rely only on the proven rigor of the mathematics.

Nevertheless, there has been throughout history a close symbiotic relationship between physics and mathematics. Sometimes physicists develop a new field in mathematics to help them solve some problem in physics. A prime example is the creation of calculus by Newton that allowed him to develop a theory of mechanics. Other times physicists will find that some abstract field in mathematics plays a fundamental role in physics. An example of this that we will soon discuss is vectors. Vector spaces are a purely abstract concept developed by mathematicians that was later found to have many important applications in physics. One of the most remarkable examples of what would seem to be purely Platonic mathematics is the fundamental and essential role that complex numbers play in the theory of quantum mechanics.

**Differential Calculus**

**The basic idea of differential calculus**

The first math topic that we will discuss is calculus. Newton invented calculus as part of his effort to develop a theory of mechanics. A working knowledge of calculus is essential for doing physics. There are two types of calculus: differential and integral. We will first discuss differential calculus.

**Differential calculus** is a means of determining the slope of the tangent to the curve of a function. For example, if we have a function, \( f(x) \), the derivative of the \( f(x) \) with respect to \( x \) evaluated at \( x_0 \) gives the slope of the tangent line at \( x_0 \), as shown in the figure below. We express the derivative as:

\[
f'(x) = \frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{\Delta f(x)}{\Delta x} = \lim_{x_1,x_2 \to x_0} \frac{f(x_2) - f(x_1)}{x_2 - x_1}
\]
That’s it. That’s all that differential calculus is about, finding slopes of tangents to curves of functions.

**Rules of differential calculus**

The following are a few rules on finding derivatives that are important to know.

- The derivative is a **linear operation**. That means that:

  \[
  \frac{d}{dx}[af(x)] = a \frac{df(x)}{dx}
  \quad \text{and} \quad
  \frac{d}{dx}[f(x) + g(x)] = \frac{df(x)}{dx} + \frac{dg(x)}{dx}
  \]

- The derivative of \(x^n\) with respect to \(x\) is given by:

  \[
  \frac{dx^n}{dx} = nx^{n-1}
  \]

- The derivative of the product of two functions \(f(x)\) and \(g(x)\) is given by the **Leibniz or product Rule**:

  \[
  \frac{d}{dx}[f(x)g(x)] = \frac{df(x)}{dx}g(x) + f(x)\frac{dg(x)}{dx}
  \]

- The derivative of a function of a function of \(x\), \(g(f(x))\), with respect to \(x\) is given by the so-called **chain rule**:

  \[
  \frac{d}{dx}[g(f(x))] = \frac{dg(f(x))}{df(x)} \frac{df(x)}{dx}
  \]

**Second derivatives**

The **second derivative** of a function is the derivative of the derivative. In other words, it gives the slope of the slope of the function. We use the illogical notation of \(d^2/dx^2\) to denote the second derivative. The second derivative is the derivative operator acting twice.

\[
\frac{d^2f(x)}{dx^2} = \frac{d}{dx}\left(\frac{df(x)}{dx}\right) = \frac{df'(x)}{dx}
\]
Note that the second derivative is also a linear operation:

\[
\frac{d^2 [af(x)]}{dx^2} = a \frac{d^2 f(x)}{dx^2} \quad \text{and} \quad \frac{d^2 [f(x) + g(x)]}{dx^2} = \frac{d^2 f(x)}{dx^2} + \frac{d^2 g(x)}{dx^2}
\]

Keep in mind that the second derivative is simply the derivative operator operating on a function twice. It is not the square of the first derivative.

\[
\frac{d^2 f(x)}{dx^2} \neq \left( \frac{df(x)}{dx} \right)^2
\]

**Partial derivatives**

We will often need the derivative of a multivariable function with respect to one of the variables. This is called the partial derivative. We use the symbol \( \partial / \partial x \) rather than \( d/dx \) to indicate that we are taking the derivative with respect to one variable while the other variables are held fixed. For example:

\[
\frac{\partial f(x, y)}{\partial x}
\]

means to take the derivative of \( f(x, y) \) with respect to \( x \) while holding the value of \( y \) fixed.

**Role of differential calculus in physics**

Differential calculus plays a prominent role in physics. One of the most common is in determining the velocity of a particle by taking the derivative of the position of the particle with respect to time:

\[
v = \frac{dx(t)}{dt}
\]

The velocity of the particle at a given time is the slope of the function \( x(t) \) at that time.

Another important role of differential calculus in physics is in finding the extrema, the local maxima and minima, of a function. As shown in the figure below, at a local maximum or minimum the slope of the function is zero, that is, the derivative of the function is zero. The first derivative alone, however, does not tell us if the extremum is a local maximum or a local minimum. For that, we need to look at the second derivative of the function. If the second derivative is negative then it is a local maximum. If the second derivative is positive then it is a local minimum, as shown in the figure.

**Total time derivative**

If \( L(x, t) \) is a function of both the position of a particle and of time, then the total time derivative of \( L \) is given by:

\[
\frac{dL(x(t), t)}{dt} = \frac{\partial L(x(t), t)}{\partial x} \frac{dx(t)}{dt} + \frac{\partial L(x(t), t)}{\partial t}
\]
The first term, in which the chain rule is used, is the rate of change of $L$ due to the dependence of $x$ on time while the second term is due to the direct dependence of $L$ on time. If $L$ doesn’t depend explicitly on time, the partial derivative of $L$ with respect to $t$ is zero. Nevertheless, if the position of the particle depends on time, the value of $L$ is time dependent and the total derivative of $L$ with respect to time is non-zero.

You now know everything you need to know about differential calculus.

**Integral Calculus**

The basic idea of integral calculus

**Integral calculus** involves integration which is the inverse of differentiation. That is, while differentiation of a function yields the slope of the function; integration of a function gives a function for which the integrated function is the slope. This follows from inverting the formula for differentiation:

$$f'(x) = \frac{df(x)}{dx} \implies df(x) = f'(x)dx$$

where the differential change in a function, $df(x)$, is the slope of the function times the differential change in $x$. 
We then define the integral of a function as:

\[ f(x) = \int df(x) = \int f'(x)dx \]

**The definite integral**

First consider a finite difference in \( x \), \( \Delta x \). Then the change in \( f(x) \) is, \( \Delta f(x) = f(x_2) - f(x_1) \), and is given approximately by:

\[ \Delta f \approx f'(x)\Delta x \]

where \( f'(x) \) is evaluated at the point midway between \( x_1 \) and \( x_2 \). Note that this is also the area of a rectangle of height equal to the midpoint value of \( f'(x) \) and of width equal to \( \Delta x \), as shown in the figure below.

If we want to find the difference in the function \( f(x) \) between any two points \( x_1 \) and \( x_2 \) we can just add up a bunch of rectangles:

\[ f(x_2) - f(x_1) \approx \sum_i f'(x_i)\Delta x \]
In order to get an exact value for \( f(x_2) - f(x_1) \), we simply go the limit of taking the width of the rectangles to zero while summing over an infinite number of rectangles. This is represented by:

\[
f(x_2) - f(x_1) = \lim_{\Delta x \to 0} \sum_i f'(x_i) \Delta x = \int_{x_1}^{x_2} f'(x) \, dx
\]

The symbol \( \int \) indicates that the sum is over an infinite number of terms and \( dx \) is the infinitesimal width as \( \Delta x \to 0 \).

The definite integral of \( f'(x) \) between \( x_1 \) and \( x_2 \) is given by:

\[
\int_{x_1}^{x_2} f'(x) \, dx = f(x) \bigg|_{x_1}^{x_2} = f(x_2) - f(x_1)
\]

It gives both the difference of the function \( f(x) \) between \( x_1 \) and \( x_2 \) and the area under the \( f'(x) \) curve between \( x_1 \) and \( x_2 \). The values \( x_1 \) and \( x_2 \) are called the limits of integration.

The indefinite integral
If no limits of integration are specified, it is called an **indefinite integral** that specifies the function \( f(x) \).

\[
f(x) + C = \int f'(x) \, dx
\]

Here \( C \) is an arbitrary constant called the **constant of integration**. It arises because the derivative of a constant is zero. That is, a constant function has zero slope. Therefore if \( f'(x) \) is the slope of \( f(x) \), it is also the slope of \( f(x) + C \) where \( C \) is any constant.

**Determining integrals**

The derivative of a function depends only upon its local properties, i.e., the slope of the function at the point at which the derivative is being evaluated. For that reason, it is usually straightforward to find the derivative of any well behaved, analytical function. An **analytical function** is one that can be written as an infinite power series:

\[
f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots
\]

where the \( a \)'s are constants that depend on the particular function. We can then use the rule given above for finding the derivative of a power of \( x \) to obtain:

\[
f'(x) = a_1 + 2a_2 x + 3a_3 x^2 + \cdots
\]

The integral, in contrast, depends upon global properties. Namely, we are to find a function whose slope at each point is specified. This is generally very difficult and in most cases an analytical expression for the function cannot be obtained. In practice this is alright since the approximate value of the integral can be obtained numerically. We simply use a computer program that determines the value of the function, \( f'(x) \), that is to be integrated at a set of equal spaced points between the two limits. We then have the computer sum the value of \( f'(x) \) at each of these points times the separation between the points.

\[
f(x) \approx \sum_i f'(x_i) \Delta x
\]

The smaller the separation between the points, the more accurate the approximation will be but of course the more terms there are to be summed and the longer the calculation will take. For our study of fundamental physics, we will not need to calculate many integrals. Applied physics is a different matter.

**Integration by parts**

**Integration by parts** is a technique that we will use often. From the product rule above, we have:

\[
\frac{d[f(x)g(x)]}{dx} = f(x) \frac{dg(x)}{dx} + \frac{df(x)}{dx} g(x)
\]
Integrating both sides, we have:

\[
\int_{x_i}^{x_f} \frac{d[f(x)g(x)]}{dx} \, dx = \int_{x_i}^{x_f} f(x) \frac{dg(x)}{dx} \, dx + \int_{x_i}^{x_f} \frac{df(x)}{dx} \, g(x) \, dx
\]

but we also have that:

\[
\int_{x_i}^{x_f} \frac{d[f(x)g(x)]}{dx} \, dx = f(x)g(x) \bigg|_{x_i}^{x_f}
\]

So

\[
\int_{x_i}^{x_f} f(x) \frac{dg(x)}{dx} \, dx + \int_{x_i}^{x_f} \frac{df(x)}{dx} \, g(x) \, dx = f(x)g(x) \bigg|_{x_i}^{x_f}
\]

A couple of important functions

The exponential function

The most important function in physics is the **exponential function**:

\[
f(x) = e^x
\]

Its importance is due to the fact that its derivative is equal to itself:

\[
f'(x) = \frac{d(e^x)}{dx} = e^x = f(x)
\]

We know this because we define the exponential function such that it is true. Recall that any analytical function can be written as a power series with coefficients \(a_n\) where the coefficients determine the function.

\[
f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots = \sum_{n=0}^{\infty} a_n x^n
\]

The exponential function is given by choosing appropriate coefficients, \(a_n\), so that \(d(e^x)/dx = e^x\).

\[
e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}
\]

Here \(n!\) is \(n\) factorial and defined by:

\[
n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1
\]

Using the rule for the derivative of a power of \(x\), \(dx^n/dx = nx^{n-1}\), it is easy to show that:

\[
\frac{d(e^x)}{dx} = e^x
\]
Using the chain rule, we have for a constant $\alpha$

$$\frac{d(e^{\alpha x})}{dx} = \alpha e^{\alpha x}$$

If $\alpha > 0$, then $f(x) = e^{\alpha x}$ increases exponentially and $f(x) = e^{-\alpha x}$ decrease exponentially as shown in the figures below.

Since the rate of change of the exponential function is proportional to the value of the function, for exponential increase (growth), the greater the value the faster it increases while for exponential decrease (decay) the smaller the value the slower it decreases.

Examples of exponential growth are:
- The growth of the number of bacteria given an unconstrained supply of nutrient
- The salaries of Wall Street executives
- The expansion of the universe

Examples of exponential decrease are:
- The number of radioactive nuclei remaining in a sample
- The worth of a dollar given a constant rate of inflation
- The density of the atmosphere as a function of altitude

The value of the number $e$ is obtained by setting $x = 1$.

$$e = e^1 = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!}$$

It is a real number represented by an infinite, non-repeating sequence of decimals:

$$e = 2.718\ldots$$

Since $e$ is an important number, it is worth remembering it to a few decimal places. This might not be so useful, but you will at least be able to impress your friends.
The logarithmic function

Another important function is the logarithmic function, which is the inverse of the exponential function. That is, if \( y = e^x \) then \( \ln y = x \). The natural logarithm of \( y \) is the power to which \( e \) must be raised to give \( y \). Note that:

\[
e^{\ln x} = x \quad \text{and} \quad \ln(e^x) = x
\]

There are logarithms to other bases such as \( \log_{10} y \) that is the power to which 10 must be raised to give \( y \) but the logarithm to the natural base \( e \), \( \log_e y = \ln y \), is the most important.

We can readily derive the derivative of the logarithmic function by comparing two expressions for the derivative of \( e^{\ln x} \).

First:

\[
\frac{d(e^{\ln x})}{dx} = \frac{d(e^{\ln x})}{d(\ln x)} \frac{d(\ln x)}{dx} = e^{\ln x} \frac{d(\ln x)}{dx} = x \frac{d(\ln x)}{dx}
\]

We also have:

\[
\frac{d(e^{\ln x})}{dx} = \frac{dx}{dx} = 1
\]

Equating these two expressions we have:

\[
x \frac{d(\ln x)}{dx} = 1 \quad \Rightarrow \quad \frac{d(\ln x)}{dx} = \frac{1}{x}
\]

Since the logarithmic function is the inverse of the exponential function, we can plot the logarithmic function by first plotting the exponential function and then interchanging the \( x \) and \( y \) axes. This is equivalent to reflecting the exponential function about the line \( y = x \) as shown in the figure below.

![Graph of logarithmic function](image)