

# The Art of Modelling: Introduction to Physics

Ryan D. Martin

August 4, 2017



# Contents

---

<b>1</b>	<b>The Scientific Method and Physics</b>	<b>3</b>
1.1	Science and the Scientific Method . . . . .	3
1.2	Theories and models . . . . .	5
1.3	Fighting intuition . . . . .	6
1.4	The scope of Physics . . . . .	6
1.4.1	Classical Physics . . . . .	6
	Mechanics . . . . .	6
	Electromagnetism . . . . .	7
1.4.2	Modern Physics . . . . .	7
	Quantum mechanics and particle physics . . . . .	7
	The Special and General Theories of Relativity . . . . .	8
	Cosmology and astrophysics . . . . .	8
	Particle astrophysics . . . . .	8
1.5	Thinking like a physicist . . . . .	8
1.6	Summary . . . . .	10
<b>2</b>	<b>Comparing Model and Experiment</b>	<b>11</b>
2.1	Orders of magnitude . . . . .	11
2.2	Units and dimensions . . . . .	12
2.2.1	Base dimensions and their SI units . . . . .	13
2.2.2	Dimensional analysis . . . . .	14
2.3	Making measurements . . . . .	16
2.3.1	Measurement uncertainties . . . . .	18
	Determining the central value and uncertainty . . . . .	18
	Random and systematic sources of error/uncertainty . . . . .	20
	Propagating uncertainties . . . . .	20
2.3.2	Reporting measured values . . . . .	24
2.3.3	Comparing model and measurement - discussing a result . . . . .	24
2.4	Using computers to facilitate data analysis . . . . .	25
2.4.1	Python: a quick intro to programming . . . . .	25
2.4.2	Using QExpy for propagating uncertainties . . . . .	26
2.4.3	Using QExpy for plotting . . . . .	27
2.5	Summary . . . . .	30

<b>3</b>	<b>Describing motion</b>	<b>31</b>
3.1	Motion in one dimension . . . . .	31
3.1.1	Motion with constant speed . . . . .	32
3.1.2	Motion with constant acceleration . . . . .	36
	Visualizing motion with constant acceleration . . . . .	39
3.1.3	Speed versus velocity . . . . .	40
3.1.4	Generalized motion and instantaneous velocity . . . . .	41
	Using calculus to obtain acceleration from position . . . . .	42
	Using calculus to obtain position from acceleration . . . . .	44
3.2	Summary . . . . .	49



# 1

## The Scientific Method and Physics

---

### Learning Objectives:

- Understand the Scientific Method
- Define the scope of Physics
- Understand the difference between theory and model
- Have a sense of how a physicist thinks

### 1.1 Science and the Scientific Method

Science is an attempt to *describe* the world around us. It is important to note that describing the world around us is not the same as *explaining* the world around us<sup>1</sup>. Science aims to answer the question “How?” and not the question “Why?”. As we develop our description of the physical world, you should remember this important distinction and resist the urge to ask “Why?”.

The Scientific Method is a prescription for coming up with a description of the physical world that anyone can challenge and improve through performing experiments. We can get some insight into the Scientific Method through a simple example.

Imagine that we wish to describe how long it takes for a tennis ball to reach the ground after being released from a certain height. One way to proceed is to describe how long it takes for a tennis ball to drop 1 m, and then to describe how long it takes for a tennis ball to drop 2 m, etc. One could generate a giant table showing how long it takes a tennis ball to drop for any given height. Someone would then be able to perform an experiment to measure how long a tennis ball takes to drop 1 m or 2 m and see if their measurement is consistent with the tabulated values. If we collected the descriptions for all possible heights, then we would effectively have a valid and testable “scientific theory” that describes how long it takes tennis balls to drop from any height.

Suppose that a budding scientist, let’s call her Chloë, then came along and noticed that

---

<sup>1</sup>Some might say this is the role of religion

there is a pattern in the theory that can be described much more succinctly and generally than by using a giant table. In particular, suppose that she notices that, mathematically, the time,  $t$ , that it takes for a tennis ball to drop a height,  $h$ , is proportional to the square root of the height:

$$t \propto \sqrt{h}$$

**Example 1-1:** Use Chloë's Theory ( $t \propto \sqrt{h}$ ) to determine how much longer it will take for an object to drop by 2 m than it would to drop by 1 m:

*When we have a proportionality law (with a  $\propto$ ) sign, we can always change this to an equal sign by introducing a constant, which we will call  $k$ :*

$$\begin{aligned} t &\propto \sqrt{h} \\ \rightarrow t &= k\sqrt{h} \end{aligned}$$

*Let  $t_1$  be the time to fall a distance  $h_1 = 1$  m, and  $t_2$  be the time to fall a distance  $h_2 = 2$  m. In terms of our unknown constant,  $k$ , we have:*

$$\begin{aligned} t_1 &= k\sqrt{h_1} = k\sqrt{(1 \text{ m})} \\ t_2 &= k\sqrt{h_2} = k\sqrt{(2 \text{ m})} \end{aligned}$$

*By taking the ratio,  $\frac{t_1}{t_2}$ , our unknown constant  $k$  will cancel:*

$$\begin{aligned} \frac{t_1}{t_2} &= \frac{\sqrt{(1 \text{ m})}}{\sqrt{(2 \text{ m})}} = \frac{1}{\sqrt{2}} \\ \therefore t_2 &= \sqrt{2}t_1 \end{aligned}$$

*and we find that it will take  $\sqrt{2} \sim 1.41$  times longer to drop by 2 m than it will by 1 m.*

Chloë's "Theory of Tennis Ball Drop Times" is appealing because it is succinct, and it also allows us to make **verifiable predictions**. That is, using this theory, we can predict that it will take a tennis ball  $\sqrt{2}$  times longer to drop from 2 m than it will from 1 m, and then perform an experiment to verify that prediction. If the experiment agrees with the prediction, then we conclude that Chloë's theory adequately describes the result of that particular experiment. If the experiment does not agree with the prediction, then we conclude that the theory is not an adequate description of that experiment, and we try to find a new theory.

Chloë's theory is also appealing because it can describe not only tennis balls, but the time it takes for other objects to fall as well. Scientists can then set out to continue testing her theory with a wide range of objects and drop heights to see if it describes those experiments

as well. Inevitably, they will discover situations where Chloë's theory fails to adequately describe the time that it takes for objects to fall (can you think of an example?).

We would then develop a new “Theory of Falling Objects” that would include Chloë's theory that describes most objects falling, and additionally, a set of descriptions for the fall times for cases that are not described by Chloë's theory. Ideally, we would seek a new theory that would also describe the new phenomena not described by Chloë's theory in a succinct manner. There is of course no guarantee, ever, that such a theory would exist; it is just an optimistic hope of scientists to find the most general and succinct description of the physical world.

This example highlights that applying the Scientific Method is an iterative process. Loosely, the prescription for applying the Scientific Method is:

1. Identify and describe a process that is not currently described by a theory.
2. Look at similar processes to see if they can be described in a similar way.
3. Improve the description to arrive at a “Theory” that can be generalized to make predictions.
4. Test predictions of the theory on new processes until a prediction fails.
5. Improve the theory.

## 1.2 Theories and models

For the purpose of this textbook, we wish to introduce a distinction in what we mean by “theory” and by “model”. We will consider a “theory” to be a set of statements that gives us a broad description, applicable to several phenomena and that allow us to make verifiable prediction. We will consider a “model” to be a situation-specific description of a phenomenon *based on a theory*. Using the example from the previous section, our theory would be that the fall time of an object is proportional to the square root of the drop height, and a model would be applying that theory to describe a tennis ball falling by 4.2 m.

This textbook will introduce the theories from Classical Physics, which were mostly established and tested in the seventeenth and nineteenth centuries. We will take it as given that readers of this textbook are not likely to perform experiments that challenge those well-established theories. The main challenge will be, given a theory, to define a model that describes a particular situation, and then to test that model. This introductory physics course is thus focused on thinking of “doing physics” as the task of correctly modelling a situation.

**Checkpoint 1-1:** Models cannot be scientifically tested, only theories can be tested.

- A) True
- B) False



## 1.3 Fighting intuition

It is important to remember to fight one's intuition when applying the scientific method. Certain theories, such as Quantum Mechanics, are very counter-intuitive. For example, in Quantum Mechanics, it is possible for an object to be in two locations at the same time. In the Theory of Special Relativity, it is possible for two people to disagree whether two clocks are synchronized.

There is no requirement in science that a theory be pretty or intuitive. The only requirement is that a theory describes experimental data. One should then take care in not forcing one's preconceived notions into interpreting a theory. For example, Quantum Mechanics does not actually predict that objects can be in two locations at once, only that objects behave *as if* they were in two locations at once. A famous example is Shrödinger's cat, which can be modelled as being both alive and dead at the same time. However, just because we model it that way does not mean that it really is alive and dead at the same time.

## 1.4 The scope of Physics

Physics describes a wide range of phenomena within the physical science, ranging from the microscopic particles that make up matter to the evolution of the entire Universe. We often speak of classical and modern physics depending on when the theories were developed, and we can further subdivide these area of physics depending on the scale or the type of the phenomena that are described.

### 1.4.1 Classical Physics

This textbook is focused on classical physics, which corresponds to the theories that were developed before 1905.

#### **Mechanics**

Mechanics describes most of our everyday experiences, such as how objects move, including how planets move under the influence of gravity. Isaac Newton was the first to formally develop a theory of mechanics, using his "Three Laws" to describe the behaviour of objects in our everyday experience. His famous work published in 1687, "Philosophiae Naturalis Principia Mathematica" ("The Principia") also included a theory of gravity that describes the motion of celestial objects.

Following the 1781 discovery of the planet Uranus by William Herschel, astronomers noticed that the orbit of the planet was not well described by Newton's theory. This led Urbain Le Verrier (in Paris) and John Couch Adams (in Cambridge) to predict the location of a new planet that was disturbing the orbit of Uranus rather than to claim that Newton's theory was incorrect. The planet Neptune was subsequently discovered by Le Verrier in 1846, one year after the prediction, and seen as a resounding confirmation of Newton's theory.

In 1859, Urbain Le Verrier also noted that Mercury's orbit around the Sun is different than that predicted by Newton's theory. Again, a new planet was proposed, "Vulcan", but that planet was never discovered and the deviation of Mercury's orbit from Newton's prediction remained unexplained until 1915, when Albert Einstein introduced a new, more complete,

theory of gravity, called “General Relativity”. This is a good example of the scientific method; although the discovery of Neptune was consistent with Newton’s theory, it did not prove that the theory is correct, only that it correctly described the motion of Uranus. The discrepancy that arose when looking at Mercury ultimately showed that Newton’s theory of gravity fails to provide a proper description in the proximity of very massive objects (Mercury is the closest planet to the Sun).

### **Electromagnetism**

Electromagnetism describes electric charges and magnetism. At first, it was not realized that electricity and magnetism were connected, and it was Charles Augustin de Coulomb that published in 1784 the first description of how electric charges attract and repel each other. Magnetism was discovered in the ancient world, when people noticed that lodestone (rocks made from magnetized magnetite mineral) could attract iron tools. In 1819, Oersted discovered that moving electric charges could influence a compass needle, and several subsequent experiments were carried out to discover how magnets and moving electric charges interact.

In 1865, James Clerk Maxwell published “A Dynamical Theory of the Electromagnetic Field”, wherein he first proposed a theory that unified electricity and magnetism as two facets of the same phenomenon. One important concept from Maxwell’s theory is that light is an electromagnetic wave with a well-defined speed. This uncovered some potential issues with the theory as it required an absolute frame of reference in which to describe the propagation of light. Experiments in the late 1800s failed to detect the existence of this frame of reference.

### **1.4.2 Modern Physics**

In 1905, Albert Einstein published three major papers that set the foundation for what we now call “Modern Physics”. These papers covered the following areas that were not well-described by classical physics:

- A description of Brownian motion that implied that all matter is made of atoms
- A description of the photoelectric effect that implied that light is made of particles
- A description of the motion of very fast objects that implied that mass is equivalent to energy, and that time and distance are relative concepts

In order to accommodate Einstein’s descriptions, physicists had to dramatically re-formulate new theories.

### **Quantum mechanics and particle physics**

Quantum mechanics is a theory that was developed in the 1920s to incorporate Einstein’s conclusion that light is made of particles (or rather, quantized lumps of energy called quanta) and describe Nature at the smallest scales. This could only be done at the expense of determinism, leading to a theory that could not predict how particular situations evolve in time, but only the probabilities that certain outcomes will be realized. Quantum mechanics was further refined during the twentieth century into Quantum Field Theory, which led to the Standard Model of particle physics that describes our current understanding of matter.

### The Special and General Theories of Relativity

In 1905, Einstein published his “Special Theory of Relativity”, which described how light propagates without the need for an absolute frame of reference, thus solving the problem introduced by Maxwell. This required physicists to consider space and time on an equal footing (“Space-time”), rather than two independent aspects of the natural world, and led to a flurry of odd, but verified, experimental predictions. One such prediction is that time flows slower for objects moving fast, which has been experimentally verified by flying precise atomic clocks on air planes and satellites. In 1915, Einstein further refined his theory into General Relativity, which is our best current description of gravity and includes a description of Mercury’s orbit which was not described by Newton’s theory.

### Cosmology and astrophysics

Cosmology describes processes at the largest scales and is mostly based on applying General Relativity to the scale of the Universe. For example, cosmology describes how our Universe started from the Big Bang and how large scale structures, such as galaxies and clusters of galaxies, have formed and evolved into our present day Universe.

Astrophysics is focused on describing the formation and the evolution of stars, galaxies, and other “astrophysical objects” such as neutron stars and black holes.

### Particle astrophysics

Particle astrophysics is a relatively new field that makes use of subatomic particles produced by astrophysical objects to learn both about the objects *and* about the particles. For example, the 2015 Nobel Prize in Physics was awarded to Art McDonald (a Canadian physicist from Queen’s University) for using neutrinos<sup>2</sup> produced by the Sun to both learn about the nature of neutrinos and about how the Sun works.



*Figure 1.1: A galaxy in the Coma cluster of galaxies (credit:NASA).*

## 1.5 Thinking like a physicist

In a sense, physics can be thought of as the most fundamental of the sciences, as it describes the interactions of the smallest constituents of matter. In principle, if one can precisely describe how protons, neutrons, and electrons interact, then one can completely describe how a human brain thinks. In practice, the theories of particle physics lead to equations that are too difficult to solve for system that include as many particles as a human brain. In fact, they are too difficult to solve exactly for even rather small systems of particles such as atoms bigger than helium (containing several protons, neutrons and electrons).

We have a number of other fields of science to cover complex systems of particles interacting. Chemistry can be used to describe what happens to systems consisting of many atoms and

---

<sup>2</sup>Neutrinos are the lightest subatomic particles that we know of

molecules. In a living being, it is too difficult to keep track of systems of atoms and molecules, so we use Biology to describe living systems.

One of the key qualities required to be an effective physicist is an ability to understand how to apply a theory and develop a model to describe a phenomenon. Just like any other skill, it takes practice to become good at developing models. Students that graduate with a physics degree are thus often sought for jobs that require critical thinking and the ability to develop quantitative models, which covers many fields from outside of physics such as finance or Big Data. This textbook thus tries to emphasize practice with developing models, while also providing a strong background in the theories of classical physics.

## 1.6 Summary

- Science attempts to *describe* the physical world (answers the question “How?”, not “Why?”).
- The Scientific Method provides a prescription for arriving at theories that describe the physical world and that can be experimentally verified.
- The Scientific Method is necessarily an iterative process where theories are continuously updated as new experimental data are acquired.
- An experiment can only disprove a theory, not confirm it in any general sense.
- Theories are typically valid only in well-defined situations.
- Physics covers a wide scale of phenomena ranging from the Universe down to subatomic particles.
- Classical physics encompasses the theories developed before 1905, when Einstein introduced the need for Quantum Mechanics and the Theorie(s) of Relativity.

# 2

## Comparing Model and Experiment

---

In this chapter, we will learn about the process of doing science and lay the foundations for developing skills that will be of use throughout your scientific careers. In particular, we will start to learn how to test a model with an experiment, as well as learn to estimate whether a given result or model makes sense.

### Learning Objectives:

- Be able to estimate orders of magnitude
- Understand units
- Understand the process of building a model and performing an experiment
- Understand uncertainties in experiments
- Be able to use a computer for simple data analysis

### 2.1 Orders of magnitude

Although one should try to fight intuition when building a model to describe a particular phenomenon, one should not abandon critical thinking and should always ask if a model (or a prediction of the model) makes sense. One of the most straightforward ways to verify if a model makes sense is to ask whether it predicts the correct order of magnitude for a quantity. Usually, the order of magnitude for a quantity can be determined by making a very simple model, ideally one that you can work through in your head. When we say that a prediction gives the right “order of magnitude”, we usually mean that the prediction is within a factor of “a few” (up to a factor of 10) of the correct answer.

**Example 2-1:** How many ping pong balls can you fit into a school bus? Is it of order 10,000, or 100,000, or more?

*Our strategy is to estimate the volumes of a school bus and of a ping pong ball, and then calculate how many times the volume of the ping pong ball fits into the volume of*

the school bus.

We can model a school bus as a box, say  $20\text{ m} \times 2\text{ m} \times 2\text{ m}$ , with a volume of  $80\text{ m}^3 \sim 100\text{ m}^3$ . We can model a ping pong ball as a sphere with a diameter of  $0.03\text{ m}$  ( $3\text{ cm}$ ). When stacking the ping pong balls, we can model them as little cubes with a side given by their diameter, so the volume of a ping pong ball, for stacking, is  $\sim 0.000\,03\text{ m}^3 = 3 \times 10^{-5}\text{ m}^3$ . If we divide  $100\text{ m}^3$  by  $3 \times 10^{-5}\text{ m}^3$ , using scientific notation:

$$\frac{100\text{ m}^3}{3 \times 10^{-5}\text{ m}^3} = \frac{1 \times 10^2}{3 \times 10^{-5}} = \frac{1}{3} \times 10^7 \sim 3 \times 10^6$$

Thus, we expect to be able to fit about three million ping pong balls in a school bus.

**Checkpoint 2-1:** Fill in the following table giving the order of magnitude in meters of the size of different physical objects. Feel free to Google these!

Object	Order of magnitude
Proton	
Nucleus of atom	
Hydrogen atom	
Virus	
Human skin cell	
Width of human hair	
Human	1 m
Height of Mt. Everest	
Radius of Earth	
Radius of the Sun	
Distance to the Moon	
Radius of the Milky Way	

## 2.2 Units and dimensions

In 1999, the NASA Mars Climate Orbiter disintegrated in the Martian atmosphere because of a mixup in the units used to calculate the thrust needed to slow the probe and place it in orbit about Mars. A computer program provided by a private manufacturer used units of pounds seconds to calculate the change in momentum of the probe instead of the Newton seconds expected by NASA. As a result, the probe was slowed down too much and

disintegrated in the Martian atmosphere. This example illustrates the need for us to **use and specify units** when we talk about the properties of a physical quantity, and it also demonstrates the difference between a dimension and a unit.

“Dimensions” can be thought of as types of measurements. For example, length and time are both dimensions. A unit is the standard that we choose to quantify a dimension. For example, meters and feet are both units for the dimension of length, whereas seconds and jiffies<sup>1</sup> are units for the dimension of time.

When we compare two numbers, for example a prediction from a model and a measurement, it is important that both quantities have the same dimension *and* be expressed in the same unit.

**Checkpoint 2-2:** The speed limit on a highway:

- A) has dimension of length over time and can be expressed in units of kilometers per hour
- B) has dimension of length can be expressed in units of kilometers
- C) has dimension of time over length and can be expressed in units of meters per second
- D) has dimension of time and can be expressed in units of meters

### 2.2.1 Base dimensions and their SI units

In order to facilitate communication of scientific information, the International System of units (SI for the french version *Système International d’unités*) was developed. This allows us to use a well-defined convention for which units to use when expressing quantities. For example, the SI unit for the dimension of length is the meter and the SI unit for the dimension of time is the second.

In order to simplify the SI unit system, a fundamental (base) set of dimensions was chosen and the SI units were defined for those dimensions. Any other dimension can always be re-expressed in terms of the base dimensions shown in table 2.1 and thus in terms of the corresponding base SI units.

From the base dimensions, one can obtain “derived” dimensions such as “speed” which is a measure of how fast an object is moving. The dimension of speed is  $\frac{L}{T}$  (length over time) and the corresponding SI unit is  $\text{m/s}^2$  (meters per second) and corresponds to a measure of how much distance an object can cover per unit time (the faster the object, the larger the distance covered per unit time). Table 2.2 shows a few derived dimensions and their corresponding SI units.

By convention, we can indicate the dimension of a quantity,  $X$ , by writing it in square

<sup>1</sup>A jiffy is a unit used in electronics and generally corresponds to either 1/50 or 1/60 seconds.

<sup>2</sup>Note that we can also write meters per second as  $\text{m}\cdot\text{s}^{-1}$ , but we often use a divide by sign if the power of the unit in the denominator is 1.



Dimension	SI unit
Length [L]	meter [m]
Time [T]	seconds[s]
Mass [M]	kilogram [kg]
Temperature [ $\Theta$ ]	kelvin [K]
Electric current [I]	ampère [A]
Amount of substance [N]	mole [mol]
Luminous intensity [J]	candela [cd]
Dimensionless [0]	unitless []

Table 2.1: Base dimensions and their SI units with abbreviations.

Dimension	SI unit	SI base units
Speed [L/T]	meter per second [m/s]	[m/s]
Frequency [1/T]	hertz [Hz]	[1/s]
Force [M·L·T <sup>-2</sup> ]	newton [N]	[kg·m·s <sup>-2</sup> ]
Energy [M·L <sup>2</sup> ·T <sup>-2</sup> ]	joule [J]	[N·m=kg·m <sup>2</sup> ·s <sup>-2</sup> ]
Power [M·L <sup>2</sup> ·T <sup>-3</sup> ]	watt [W]	[J/s=kg·m <sup>2</sup> ·s <sup>-3</sup> ]
Electric Charge [I· T]	coulomb [C]	[A· s]
Voltage [M·L <sup>2</sup> ·T <sup>-3</sup> ·I <sup>-1</sup> ]	volt [V]	[J/C=kg·m <sup>2</sup> ·s <sup>-3</sup> ·A <sup>-1</sup> ]

Table 2.2: Example of derived dimensions and their SI units with abbreviations.

brackets,  $[X]$ . For example,  $[X] = I$ , would mean that the quantity  $X$  has dimensions of electric current. Similarly, we can indicate the SI units of  $X$  with  $SI[X]$ ; since  $X$  has dimensions of current,  $SI[X] = A$ .

### 2.2.2 Dimensional analysis

We call “dimensional analysis” the process of working out the dimensions of a quantity in terms of the base dimensions. A few simple rules allow us to easily work out the dimensions of a derived quantity. Suppose that we have two quantities,  $X$  and  $Y$ , both with dimensions. We then have the following rules to find the dimension of a quantity that depends on  $X$  and  $Y$ :

1. You can only add or subtract two quantities if they have the same dimension:  $[X+Y] = [X] = [Y]$

2. The dimension of the product is the product of the dimensions:  $[XY] = [X] \cdot [Y]$
3. The dimension of the ratio is the ratio of the dimensions:  $[X/Y] = [X]/[Y]$

The next two examples show how to apply dimensional analysis to obtain the unit or dimension of a derived quantity.

**Example 2-2:** Given that acceleration has SI units of  $\text{ms}^{-2}$  and that force has dimensions of mass multiplied by acceleration, what are the dimensions and SI units of force, expressed in terms of the base dimensions and units?

*We can start by expressing the dimension of acceleration, since we know from its SI units that it must have dimension of length over time squared.*

$$[\text{acceleration}] = \frac{L}{T^2}$$

*Since force has dimension of mass times acceleration, we have:*

$$[\text{force}] = \frac{M \cdot L}{T^2}$$

*and the SI units of force are thus:*

$$SI[\text{force}] = \frac{\text{kg} \cdot \text{m}}{\text{s}^2}$$

*Force is such a common dimension that it, like many other derived dimensions, has its own derived SI unit, the Newton  $[N]$ .*

**Example 2-3:** Use Table 2.2 to show that voltage has the same dimension as force multiplied by speed and divided by electric current.

*According to Table 2.2, voltage has dimensions:*

$$[\text{voltage}] = M \cdot L^2 \cdot T^{-3} \cdot I^{-1}$$

*while force, speed and current have dimensions:*

$$[\text{force}] = M \cdot L \cdot T^{-2}$$

$$[\text{speed}] = L \cdot T^{-1}$$

$$[\text{current}] = I$$

*The dimension of force multiplied by speed divided by electric charge*

$$\begin{aligned} \left[ \frac{\text{force} \cdot \text{speed}}{\text{current}} \right] &= \frac{[\text{force}] \cdot [\text{speed}]}{[\text{current}]} = \frac{M \cdot L \cdot T^{-2} \cdot L \cdot T^{-1}}{I} \\ &= M \cdot L^2 \cdot T^{-3} \cdot I^{-1} \end{aligned}$$

*where, in the last line, we combined the powers of the same dimensions. By inspection, this is the same dimension as voltage.*

When you derive a model to describe a situation, your model will typically provide a value for a quantity that you are interested in modelling. You should always use dimensional analysis to ensure that the dimension of the quantity your model predicts has the correct dimension. For example, suppose that you model the speed,  $v$ , that an object has after falling from a height of 100 m on the surface of the planet Mars. Presumably,  $v$  will depend on the mass and radius of the planet. You can be guaranteed that your model for  $v$  is incorrect if the dimension of  $v$  is not speed. Dimensional analysis should always be used to check that your model is not incorrect (note that getting the correct dimension is not a guarantee of the model being correct, only that it is “not definitely wrong”). Similarly, you should also use order of magnitude estimates to evaluate whether your model gives a reasonable prediction.

**Checkpoint 2-3:** In Chloë’s theory of falling objects from Chapter , the time,  $t$ , for an object to fall a distance,  $x$ , was given by  $t = k\sqrt{x}$ . What must the SI units of Chloë’s constant,  $k$ , be?

- A)  $\text{T L}^{\frac{1}{2}}$
- B)  $\text{T L}^{-\frac{1}{2}}$
- C)  $\text{s m}^{\frac{1}{2}}$
- D)  $\text{s m}^{-\frac{1}{2}}$

## 2.3 Making measurements

Having introduced some tools for the modelling aspect of physics, we now address the other side of physics, namely performing experiments. Since the goal of developing theories and models is to describe the real world, we need to understand how to make meaningful measurements that test our theories and models.

Suppose that we wish to test Chloë’s theory of falling objects from Chapter 3:

$$t = k\sqrt{x}$$

which states that the time,  $t$ , for any object to fall a distance,  $x$ , from the surface of the Earth is given by the above relation. The theory assumes that Chloë’s constant,  $k$ , is the same for any object falling any distance on the surface of the Earth.

One possible way to test Chloë’s theory of falling objects is to measure  $k$  for different drop heights to see if we always obtain the same value. Results of such an experiment are

presented in Table 2.3, where the time,  $t$ , was measured for a bowling ball to fall distances of  $x$  between 1 m and 5 m. The table also shows the values computed for  $\sqrt{x}$  and the corresponding value of  $k = \frac{t}{\sqrt{x}}$ :

$x$ [m]	$t$ [s]	$\sqrt{x}$ [m $^{\frac{1}{2}}$ ]	$k$ [s m $^{-\frac{1}{2}}$ ]
1.00	0.33	1.00	0.33
2.00	0.74	1.41	0.52
3.00	0.67	1.73	0.39
4.00	1.07	2.00	0.54
5.00	1.10	2.24	0.49

Table 2.3: Measurements of the drop times,  $t$ , for a bowling ball to fall different distances,  $x$ . We have also computed  $\sqrt{x}$  and the corresponding value of  $k$ .

When looking at Table 2.3, it is clear that each drop height gave a different value of  $k$ , so at face value, we would claim that Chloë's theory is incorrect, as there does not seem to be a value of  $k$  that applies to all situations. However, we would be incorrect in doing so unless we understood *the precision of the measurements* that we made. Suppose that we repeated the measurement at a drop height of  $x = 3$  m, and obtained the values in Table 2.4.

$x$ [m]	$t$ [s]	$\sqrt{x}$ [m $^{\frac{1}{2}}$ ]	$k$ [s m $^{-\frac{1}{2}}$ ]
3.00	1.01	1.73	0.58
3.00	0.76	1.73	0.44
3.00	0.64	1.73	0.37
3.00	0.73	1.73	0.42
3.00	0.66	1.73	0.38

Table 2.4: Repeated measurements of the drop time,  $t$ , for a bowling ball to fall a distance  $x = 3$  m. We have also computed  $\sqrt{x}$  and the corresponding value of  $k$ .

This simple example highlights the critical aspect of making any measurement: it is impossible to make a measurement with infinite accuracy. The values in Table 2.4 show that if we repeat the exact same experiment, we are likely to measure different values for a single quantity. In this case, for a fixed drop height,  $x = 3$  m, we obtained a spread in values of the drop time,  $t$ , between roughly 0.6 s and 1.0 s. Does this mean that it is hopeless to do science, since we can never repeat measurements? Thankfully, no! It does however require that we deal with the inherent imprecision of measurements in a formal manner.

### 2.3.1 Measurement uncertainties

The values in Table 2.4 show that for a fixed experimental setup (a drop height of 3 m), we are likely to measure a spread in the values of a quantity (the time to drop). We can quantify this “uncertainty” in the value of the measured time but quoting the measured value of  $t$  by providing a “central value” and an “uncertainty”:

$$t = (0.76 \pm 0.15) \text{ s}$$

where 0.76 s is called the “central value” and 0.15 s the “uncertainty” or the “error”<sup>3</sup>. When we present a number with an uncertainty, we mean that we are “pretty certain” that the true value is in the range that we quote. In this case, the range that we quote is that  $t$  is between 0.61 s and 0.91 s (given by 0.76 s - 0.15 s and 0.76 s + 0.15 s). When we say that we are “pretty sure” that the value is within the quoted range, we usually mean that there is a 68% chance of this being true and allow for the possibility that the true value is actually outside the range that we quoted. The value of 68% comes from statistics and the normal distribution which you can learn about on the internet or in a more advanced course.

#### Determining the central value and uncertainty

The tricky part when performing a measurement is to decide out how to assign a central value and an uncertainty. For example, how did we come up with  $t = (0.76 \pm 0.15) \text{ s}$  from the values in Table 2.4?

Determining the uncertainty and central value on a measurement is greatly simplified when one can repeat the measurement multiple times, as we did in Table 2.4. With repeatable measurements, a reasonable choice for the central value and uncertainty is to use the mean and standard deviation of the measurements, respectively.

If we have  $N$  measurements of some quantity  $t$ ,  $\{t_1, t_2, t_3, \dots, t_N\}$ , then the mean,  $\bar{t}$ , and standard deviation,  $\sigma_t$ , are defined as:

$$\bar{t} = \frac{1}{N} \sum_{i=1}^{i=N} t_i = \frac{t_1 + t_2 + t_3 + \dots + t_N}{N} \quad (2.1)$$

$$\sigma_t^2 = \frac{1}{N-1} \sum_{i=1}^{i=N} (t_i - \bar{t})^2 = \frac{(t_1 - \bar{t})^2 + (t_2 - \bar{t})^2 + (t_3 - \bar{t})^2 + \dots + (t_N - \bar{t})^2}{N-1} \quad (2.2)$$

$$\sigma_t = \sqrt{\sigma_t^2} \quad (2.3)$$

Note that the mean is just the arithmetic average of the values, and that the standard deviation,  $\sigma_t$ , requires one to first calculate the mean, then the variance ( $\sigma_t^2$ , the square of the standard deviation). You should also note that for the variance, we divide by  $N-1$  instead of  $N$ . The standard deviation and variance are quantities that come from statistics and are good measure of how spread out the values of  $t$  are.

---

<sup>3</sup>We use the word error as a synonym for uncertainty, not “mistake”.

**Example 2-4:** Calculate the mean and standard deviation of the values for  $k$  from Table 2.4.

*In order to calculate the standard deviation, we first need to calculate the mean of the  $N = 5$  values of  $k$ :  $\{0.58, 0.44, 0.37, 0.42, 0.38\}$ . The mean is given by:*

$$\bar{k} = \frac{0.58 + 0.44 + 0.37 + 0.42 + 0.38}{5} = 0.44 \text{ s m}^{-\frac{1}{2}}$$

*We can now calculate the variance:*

$$\sigma_k^2 = \frac{1}{4}[(0.58 - 0.44)^2 + (0.44 - 0.44)^2 + (0.37 - 0.44)^2 + (0.42 - 0.44)^2 + (0.38 - 0.44)^2] = 7.3 \times 10^{-3} \text{ s}^2 \text{ m}$$

*and the standard deviation is then given by the square root of the variance:*

$$\sigma_k = \sqrt{0.0073} = 0.09 \text{ s m}^{-\frac{1}{2}}$$

*Using the mean and standard deviation, we would quote our value of  $k$  as  $k = (0.44 \pm 0.09) \text{ s m}^{-\frac{1}{2}}$ .*

Any value that we measure should always have an uncertainty. In the case where we can easily repeat the measurement, we should do so to evaluate how reproducible it is, and the standard deviation of those values is usually a good first estimate of the uncertainty in a value<sup>4</sup>. Sometimes, the measurements cannot easily be reproduced; in that case, it is still important to determine a reasonable uncertainty, but in this case, it usually has to be estimated. Table 2.5 shows a few common types of measurements and how to determine the uncertainties in those measurements.

Type of measurement	How to determine central value and uncertainty
Repeated measurements	Mean and standard deviation
Single measurement with a graduated scale (e.g. ruler, digital scale, analogue meter)	Closest value and half of the smallest division
Counted quantity	Counted value and square root of the value

Table 2.5: Different types of measurements and how to assign central values uncertainties.

For example, we would quote the length of the grey object in Figure 2.1 to be  $L = (2.8 \pm 0.5) \text{ cm}$  based on the rules in Table 2.5, since 2.8 cm is the closest value on the ru-

<sup>4</sup>In practice, the standard deviation is an overly conservative estimate of the error and we would use the error on the mean, which is the standard deviation divided by the square root of the number of measurements.

ler that matches the length of the object and 0.5 mm is half of the smallest division on the ruler. Using half of the smallest division of the ruler means that our uncertainty range covers one full division. Note that it is usually better to reproduce a measurement to evaluate the uncertainty instead of using half of the smallest division.

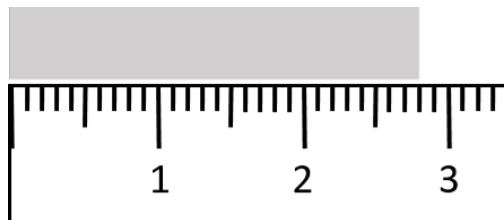


Figure 2.1: The length of the grey rectangle would be quoted as  $L = (2.8 \pm 0.5)$  cm using the rule of “half the smallest division”.

The **relative uncertainty** in a measured value is given by dividing the uncertainty by the central value, and expressing the result in percent. For example, the relative uncertainty in  $t = (0.76 \pm 0.15)$  s is given by  $\frac{0.15}{0.76} = 20\%$ . The relative uncertainty gives an idea of how precisely a value was determined. Typically, a value above 10% means that it was not a very precise measurement, and we would generally consider a value smaller than 1% to correspond to quite a precise measurement.

### Random and systematic sources of error/uncertainty

It is important to note that there are two possible sources of uncertainty in a measurement. The first is called “statistical” or “random” and occurs because it is impossible to exactly reproduce a measurement. For example, every time you lay down a ruler to measure something, you might shift it slightly one way or the other which will affect your measurement. The important property of random sources of uncertainty is that if you reproduced the measurement many times, these will tend to cancel out and the mean can usually be determined to arbitrary accuracy with enough measurements.

The other source of uncertainty is called “systematic”. Systematic uncertainties are much more difficult to detect and to estimate. One example would be trying to measure something with a scale that was not properly tarred (where the 0 weight was not set). You may end up with very small random errors when measuring the weights of object (very repeatable measurements), but you would have a hard time noticing that all of your weights were offset by a certain amount unless you had access to a second scale. Some common examples of systematic uncertainties are: incorrectly calibrated equipment, parallax error when measuring distance, reaction times when measuring time, effects of temperature on materials, etc.

### Propagating uncertainties

Going back to the data in Table 2.4, we found that for a known drop height of  $x = 3$  m, we measured different values of the drop time, which we found to be  $t = (0.76 \pm 0.15)$  s (using the mean and standard deviation). We also calculated a value of  $k$  corresponding to each value of  $t$ , and found  $k = (0.44 \pm 0.09)$  s m $^{-\frac{1}{2}}$ . Suppose that we did not have access to the individual values of  $t$ , but only to the value of  $t = (0.76 \pm 0.15)$  s with uncertainty. How do we calculate a value for  $k$  with uncertainty? In order to answer this question, we need to know how to “propagate” the uncertainties in a measured value to the uncertainty in a valued derived from the measurements. We now look at different methods for propagating uncertainties.

**1. Estimate using relative uncertainties** The relative uncertainty in a measurement gives us an idea of how precisely a value was determined. Any quantity that depends on that measurement should have a precision that is similar; that is we expect the relative uncertainty on  $k$  to be similar to that in  $t$ . For  $t$ , we saw that the relative uncertainty was approximately 20%. If we take the central value of  $k$  to be the central value of  $t$  divided by  $\sqrt{x}$ , we find:

$$k = \frac{(0.76 \text{ s})}{\sqrt{(3 \text{ m})}} = 0.44 \text{ s m}^{-\frac{1}{2}}$$

Since we expect the relative uncertainty in  $k$  to be approximately 20%, then the absolute uncertainty is given by:

$$\sigma_k = 0.2 \cdot k = 0.09 \text{ s m}^{-\frac{1}{2}}$$

which is close to the value obtained by averaging the five values of  $k$  in Table 2.4.

## 2. The Min-Max method

A pedagogical way to determine  $k$  and its uncertainty is to use the “Min-Max method”. Since  $k = \frac{t}{\sqrt{x}}$ ,  $k$  will be the biggest when  $t$  is the biggest, and the smallest when  $t$  is the smallest. We can thus determine “minimum” and “maximum” values of  $k$  corresponding to the minimum value of  $t$ ,  $t^{\min} = 0.61 \text{ s}$  and the maximum value of  $t$ ,  $t^{\max} = 0.91 \text{ s}$ :

$$\begin{aligned} k^{\min} &= \frac{t^{\min}}{\sqrt{x}} = \frac{0.61 \text{ s}}{\sqrt{(3 \text{ m})}} = 0.35 \text{ s m}^{-\frac{1}{2}} \\ k^{\max} &= \frac{t^{\max}}{\sqrt{x}} = \frac{0.91 \text{ s}}{\sqrt{(3 \text{ m})}} = 0.53 \text{ s m}^{-\frac{1}{2}} \end{aligned}$$

This gives us the range of values of  $k$  that correspond to the range of values of  $t$ . We can choose the middle of the range as the central value of  $k$  and half of the range as the uncertainty:

$$\begin{aligned} \bar{k} &= \frac{1}{2}(k^{\min} + k^{\max}) = 0.44 \text{ s m}^{-\frac{1}{2}} \\ \sigma_k &= \frac{1}{2}(k^{\max} - k^{\min}) = 0.09 \text{ s m}^{-\frac{1}{2}} \\ \therefore k &= (0.44 \pm 0.09) \text{ s m}^{-\frac{1}{2}} \end{aligned}$$

which, in this case, gives the same value as that obtained by averaging the individual values of  $k$ . While the Min-Max method is useful for illustrating the concept of propagating uncertainties, we usually do not use it in practice as it tends to overestimate the true uncertainties in a measurement.

**3. The derivative method** In the example above, we assumed that the value of  $x$  was known precisely (and we chose 3 m) which of course is not realistic. Let us suppose that we



have measured  $x$  to within 1 cm so that  $x = (3.00 \pm 0.01)$  m. The task is now to calculate  $k = \frac{t}{\sqrt{x}}$  when both  $x$  and  $t$  have uncertainties.

The derivative method lets us propagate the uncertainty in a general way, so long as the relative uncertainties on all quantities are “small” (less than 10-20%). If we have a function,  $F(x, y)$  that depends on multiple variables with uncertainties (e.g.  $x \pm \sigma_x$ ,  $y \pm \sigma_y$ ), then the central value and uncertainty in  $F(x, y)$  are given by:

$$\begin{aligned}\bar{F} &= F(\bar{x}, \bar{y}) \\ \sigma_F &= \sqrt{\left(\frac{\partial F}{\partial x} \sigma_x\right)^2 + \left(\frac{\partial F}{\partial y} \sigma_y\right)^2}\end{aligned}\tag{2.4}$$

That is, the central value of the function  $F$  is found by evaluating the function at the central values of  $x$  and  $y$ . The uncertainty in  $F$ ,  $\sigma_F$  is found by taking the quadrature sum of the partial derivatives of  $F$  evaluated at the central values of  $x$  and  $y$  multiplied by the uncertainties in the corresponding variables that  $F$  depends on. The uncertainty will contain one term in the sum per variable that  $F$  depends on. At the end of the chapter, we will show you how to calculate this easily with a computer, so do not worry about getting comfortable with partial derivatives (yet!). Note that the partial derivative,  $\frac{\partial F}{\partial x}$  is simply the derivative of  $F(x, y)$  relative to  $x$  evaluated as if  $y$  were a constant. Also, when we say “add in quadrature”, we mean square the quantities, add them, and then take the square root (same as you would calculate the hypotenuse of a right-angle triangle).

**Example 2-5:** Use the derivative method to evaluate  $k = \frac{t}{\sqrt{x}}$  for  $x = (3.00 \pm 0.01)$  m and  $t = (0.76 \pm 0.15)$  s.

Here,  $k = k(x, t)$  is a function of both  $x$  and  $t$ . The central value is easily found:

$$\bar{k} = \frac{t}{\sqrt{x}} = \frac{(0.76 \text{ s})}{\sqrt{(3 \text{ m})}} = 0.44 \text{ s m}^{-\frac{1}{2}}$$

Next, we need to determine and evaluate the partial derivative of  $k$  with respect to  $t$  and  $x$ :

$$\frac{\partial k}{\partial t} = \frac{1}{\sqrt{x}} \frac{d}{dt} t = \frac{1}{\sqrt{x}} = \frac{1}{\sqrt{(3 \text{ m})}} = 0.58 \text{ m}^{-\frac{1}{2}}$$

$$\frac{\partial k}{\partial x} = t \frac{d}{dx} x^{-\frac{1}{2}} = -\frac{1}{2} t x^{-\frac{3}{2}} = -\frac{1}{2} (0.76 \text{ s}) (3.00 \text{ m})^{-\frac{3}{2}} = -0.073 \text{ s m}^{-\frac{3}{2}}$$

And finally, we plug this into the quadrature sum to get the uncertainty in  $k$ :

$$\begin{aligned}\sigma_k &= \sqrt{\left(\frac{\partial k}{\partial x}\sigma_x\right)^2 + \left(\frac{\partial k}{\partial t}\sigma_t\right)^2} = \sqrt{\left((0.073 \text{ s m}^{-\frac{3}{2}})(0.01 \text{ m})\right)^2 + \left((0.58 \text{ m}^{-\frac{1}{2}})(0.15 \text{ s})\right)^2} \\ &= 0.09 \text{ s m}^{-\frac{1}{2}}\end{aligned}$$

So we find that:

$$k = (0.44 \pm 0.09) \text{ s m}^{-\frac{1}{2}}$$

which is consistent with what we found with the other two methods.

We should ask ourselves if the value we found is reasonable, since we also included an uncertainty in  $x$  and would expect a bigger uncertainty than in the previous calculations where we only had an uncertainty in  $t$ . The reason that the uncertainty in  $k$  has remained the same is that the relative uncertainty in  $x$  is very small,  $\frac{0.01}{3.00} \sim 0.3\%$ , so it contributes very little compared to the 20% uncertainty from  $t$ .

The derivative methods leads to a few simple short cuts when propagating the uncertainties for simple operations, as shown in Table 2.6. A few rules to note:

1. Uncertainties should be combined in quadrature
2. For addition and subtraction, add the absolute uncertainties in quadrature
3. For multiplication and division, add the relative uncertainties in quadrature

Operation to get $z$	Uncertainty in $z$
$z = x + y$ (addition)	$\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2}$
$z = x - y$ (subtraction)	$\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2}$
$z = xy$ (multiplication)	$\sigma_z = xy \sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$
$z = \frac{x}{y}$ (division)	$\sigma_z = \frac{x}{y} \sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$
$z = f(x)$ (a function of 1 variable)	$\sigma_z = \left \frac{df}{dx}\sigma_x\right $

Table 2.6: How to propagate uncertainties from measured values  $x \pm \sigma_x$  and  $y \pm \sigma_y$  to a quantity  $z(x, y)$  for common operations.

**Checkpoint 2-4:** We have measured that a llama can cover a distance of  $(20.0 \pm 0.5) \text{ m}$  in  $(4.0 \pm 0.5) \text{ s}$ . What is the speed (with uncertainty) of the llama?

### 2.3.2 Reporting measured values

Now that you know how to attribute an uncertainty to a measured quantity and then propagate that uncertainty to a derived quantity, you are ready to present your measurement to the world. In order to conduct “good science”, your measurements should be reproducible, clearly presented, and precisely described. Here are general rules to follow when quoting a number:

1. Show the units, preferably SI units (use derived SI units, such as newtons, when appropriate)
2. Include a sentence describing how the uncertainty was determined (if it is a direct measurement, how did you choose the uncertainty? If it is a derived quantity, how did you propagate the uncertainty?)
3. Show no more than 2 “significant digits”<sup>5</sup> in the uncertainty and format the central value to the same decimal as the uncertainty.
4. Use scientific notation when appropriate (usually numbers bigger than 1000 or smaller than 0.01).
5. Factor out the power 10 from the central value and uncertainty (e.g.  $(10.123 \pm 310.000) \text{ m}$  would be  $(10.12 \pm 0.31) \times 10^3 \text{ m}$ )

**Checkpoint 2-5:** How many significant digits are in the number 01043.120?

- A) 8
- B) 7
- C) 6
- D) 5

### 2.3.3 Comparing model and measurement - discussing a result

In order to make science advance, we make measurements and compare them to a theory or model prediction. We thus need a precise and consistent way to compare measurements with each other and with predictions. Suppose that we have measured a value for Chloë’s constant  $k = (0.44 \pm 0.09) \text{ s m}^{-\frac{1}{2}}$ . Of course, Chloë’s theory does not predict a value for  $k$ , only that fall time is proportional to the square root of the distance fallen. Isaac Newton’s Universal Theory of Gravity does predict a value for  $k$  of  $0.45 \text{ s m}^{-\frac{1}{2}}$  with negligible uncertainty. In this case, since the model (theoretical) value easily falls within the range given by our uncertainty, we would say that our measurement is consistent (or compatible) with the theoretical prediction.

Suppose that instead, we had measured  $k = (0.55 \pm 0.08) \text{ s m}^{-\frac{1}{2}}$  so that the lowest value compatible with our measurement,  $k = 0.55 \text{ s m}^{-\frac{1}{2}} - 0.08 \text{ s m}^{-\frac{1}{2}} = 0.47 \text{ s m}^{-\frac{1}{2}}$  is not compatible with Newton’s prediction. Would we conclude that our measurement invalidates Newton’s theory? The answer is: it depends... And what “it depends on” should always be discussed any time that you present a measurement (even if it happened that your measurement is compatible with a prediction - maybe that was a fluke). Below, we list a few

<sup>5</sup>Significant digits are those excluding leading and trailing zeroes.

common points that should be addressed when presenting a measurement and that will guide you into deciding whether your measurement is consistent with a prediction:

- How was the uncertainty determined and/or propagated?
- Are there systematic effects that were not taken into account when determining the uncertainty? (e.g. reaction time, parallax, something difficult to reproduce).
- Are the relative uncertainties reasonable?
- What assumptions were made in calculating your value?

In the above, our value of  $k = (0.55 \pm 0.08) \text{ s m}^{-\frac{1}{2}}$  is the result of propagating the uncertainty in  $t$  which was found by using the standard deviation of the values of  $t$ . It is thus conceivable that the true value of  $t$ , and therefore of  $k$ , is outside the range that we quote. Since our value of  $k$  is still quite close to the theoretical value, we would not claim to have invalidated Newton's theory with this measurement. Our uncertainty in  $k$  is  $\sigma_k = 0.08 \text{ s m}^{-\frac{1}{2}}$ , and the difference between our measured and the theoretical value is only  $1.25\sigma_k$ , so very close to the value of the uncertainty.

In a similar way, we would discuss whether two different measurements, each with an uncertainty, are compatible. If the ranges given by uncertainties in two values overlap, then they are clearly consistent and compatible. If on the other side, if the ranges do not overlap, they could be inconsistent, or the discrepancy might instead be the result of how the uncertainties were determined and the measurements could still be considered consistent.

## 2.4 Using computers to facilitate data analysis

In this textbook, we will encourage you to use computers to facilitate making calculations and displaying data. We will make use of a popular programming language called Python, as well as several “modules” from Python that facilitate working with numbers and data. Do not worry if you do not have any programming experience; we assume that you have none and hope that by the end of this book, you will have some capability to decrease your workload by using computers.

### 2.4.1 Python: a quick intro to programming

In Python, as in other programming language, the equal sign is called the assignment operator. Its role is assign the value on right to the variable on the left. The following code does the following:

- assigns the value of 2 to the variable `a`
- assigns the values of `2*a` to the variable `b`
- prints out the value of the variable `b`

*Python Code 2.1: Declaring variables in Python*

```
#This is a comment, and is ignored by Python
a = 2
b = 2*a
print(b)
```

*Output 2.1:*

4

Note that any text that follows a pound sign (#) is intended as a comment and will be ignored by Python. Inserting comments in your code is very important for being able to understand your computer program in the future or if you are sharing your code.

In Python, if you want to have access to “function”, which are more complex series of operations, then typically need to load the module that defines those operations. For example, if you want to be able to take the square root of a number, then you need to load the appropriately named math module, as in the following example:

*Python Code 2.2: Using functions from modules*

```
#First, we load the math module
import math as m
a = 9
b = m.sqrt(a)
print(b)
```

*Output 2.2:*

3

In the above code, we loaded the math module (and renamed it `m`); this then allows us to use the functions that are part of that module, including the square root function (`m.sqrt()`).

### 2.4.2 Using QExpy for propagating uncertainties

QExpy is a Python module that was developed at Queen’s University to handle all aspects of undergraduate physics laboratories. In this section, we look at how to use it to propagate uncertainties. Recall Example 2-5, where we propagated the uncertainties in  $t$  and  $x$  to  $k = \frac{t}{\sqrt{x}}$ . We show below how easily this can be done with QExpy:

*Python Code 2.3: QExpy to propagate uncertainties*

```
#First, we load the QExpy module
import qexpy as q
#Now define our measurements with uncertainties:
t = q.Measurement(0.76, 0.15) #0.76 +/- 0.15
x = q.Measurement(3, 0.1) #3 +/- 0.1
#Now define k, which depends on t and x:
k = t/q.sqrt(x)
#Print the result:
print(k)
```

*Output 2.3:*

0.44 +/- 0.09

which is the result that we obtained when manually applying the derivative method. Note that we used the square root function from the QExpy module, as it “knows” how to take the square root of a value with uncertainty (a “Measurement” in the language of QExpy).

Also recall that in Table 2.4, we have 5 different measurements of time that we used to

calculate the mean and standard deviation of  $t$  to use as central value and uncertainty. We can do this very easily in QExpy, by setting our value of  $t$  to be equal to a list of measurements:

*Python Code 2.4: QExpy to calculate mean and standard deviation*

```
#First, we load the QExpy module
import qexpy as q
#We define $t$ as a list of values (note the square brackets):
t = q.Measurement([1.01, 0.76, 0.64, 0.73, 0.66])
#Choose the number of significant figures to print:
q.set_sigfigs(2)
#Print the result:
print("t = ", t)
```

*Output 2.4:*

```
t = 0.76 +/- 0.15
```

### 2.4.3 Using QExpy for plotting

Recall Table 2.3, where we measured the time for an object to drop from different heights. One of the easiest way to look at the data is to visualize them on a graph. In this case, we measured the time,  $t$ , that it took to drop different heights,  $x$ . Chloë's Theory stated that the time,  $t$ , is proportional to the square root of the distance fallen,  $x$ , and we introduced a constant of proportionality  $k$ :

$$t = k\sqrt{x}$$

This means that if we make a graph of  $t$  versus  $\sqrt{x}$ , we should expect that the points fall on a straight line that goes through zero, with a slope of  $k$ . We can easily use QExpy to make this plot of the data in Table 2.3.

*Python Code 2.5: Using QExPy for plotting*

```
#First, we load the QExpy module:
import qexpy as q

#Then we enter the data:
#start with the values for the square root of height:
sqx = [1. , 1.41, 1.73, 2., 2.24]
#and then, the corresponding times:
t = [ 0.33, 0.74, 0.67, 1.07, 1.1 ]

#Let us attribute an uncertainty of 0.15 to the measured values of t:
terr = 0.15

#We now make the plot, first, we create the plot object
fig = q.MakePlot( xdata = sqx, xname = "sqrt(distance)", xunits = "sqrt(m)",
                  ydata = t, yerr = terr, yname = "time", yunits = "s",
                  data_name = "Data1")

#Ask QExpy to also show the line of best fit
fig.fit("linear")
```

```
#Then, we show it:
fig.show()
```

*Output 2.5:*

```
-----Fit results-----
Fit of Data1 to linear
Fit parameters:
Data1_linear_fit0_fitpars_intercept = -0.24 +/- 0.22,
Data1_linear_fit0_fitpars_slope = 0.61 +/- 0.13

Correlation matrix:
[[ 1.    -0.968]
 [-0.968  1.    ]]

chi2/ndof = 2.04/2
-----End fit results-----
```

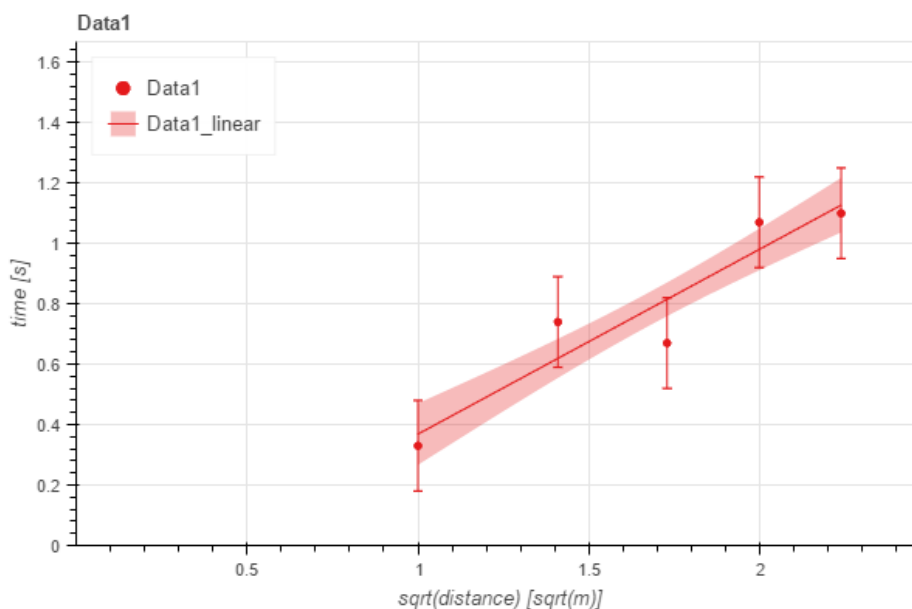


Figure 2.2: QExpy plot of  $t$  versus  $\sqrt{x}$ .

The plot in Figure 2.2 shows that the data points do indeed appear to fall on a straight line. We’ve also asked QExpy to show us the line of best fit to the data, represented by the line with the shadowed area. When we asked for the line of best fit, QExpy not only drew the line, but also gave us the values for the slope and the intercept of the line. The shaded area around the line corresponds to line that one would obtain using different possible values of the slope and offset within their uncertainties. The output also provides a line that tells us that  $\text{chi2}/\text{ndof} = 2.04/2$ ; although you do not need to understand the details, this is a measure of how well the data are described by the line of best fit. Generally, the fit is assumed to be “good” if this ratio is close to 1 (the ratio is called “the chi-squared  $\chi^2$  per degree of freedom”). The “correlation matrix” tells us how the best fit value of the slope is linked to the best fit value of the intercept.

Since we expect the slope of the data to be  $k$ , this provides us a method to determine  $k$  from the data as  $(0.61 \pm 0.13) \text{ s m}^{-\frac{1}{2}}$ . When we used Table 2.4 to determine  $k$  using repeated measurements at a drop height of 3.0 m, we obtained  $k = (0.44 \pm 0.09) \text{ s m}^{-\frac{1}{2}}$ , which is consistent with what we get from the slope of the best fit line. Finally, we expect the intercept to be equal to zero. The best fit line from QExpy has an intercept of  $(-0.24 \pm 0.22) \text{ s}$ , which is slightly below, but consistent, with zero. From these data, we would conclude that our measurements are consistent with Chloë's Theory.



## 2.5 Summary

- Measurable quantities have dimensions and units
- A physical quantity should always be reported with units, preferably SI units
- When you build a model to predict a physical quantity, you should always ask if the prediction makes sense (Does it have a reasonable order of magnitude? Does it have the right dimensions?)
- Any quantity that you measure will have an uncertainty.
- The best way to determine an uncertainty is to repeat the measurement and use the mean and standard deviation of the measurements as the central value and uncertainty.
- Relative uncertainties tell you whether your measurement is precise.
- When propagating uncertainties, use a computer [#becauseits2015](#)
- If you expect two measured quantities to be linearly related (one is proportional to the other), plot them to find out! Use a computer to do so!

# 3

## Describing motion

---

In this chapter, we will introduce the tools required to describe motion. In later chapters, we will use the theories of physics to model the motion of objects, but first, we need to make sure that we have the tools to describe the motion. We generally use the word “kinematics” to label the tools for describing motion (e.g. speed, acceleration, position, etc), whereas we refer to “dynamics” when we use the laws of physics to predict motion (e.g. what motion will occur if a force is applied to an object).

### Learning Objectives:

- Understand how to describe motion in 1D, 2D, and 3D
- Understand the meaning of position, velocity, speed, and acceleration
- Understand how to use vectors
- Have a working understanding of derivatives and integrals

### 3.1 Motion in one dimension

The most simple type of motion to describe is that of a particle that is constrained to move along a straight line; much like a train along a straight piece of track. When we say that we want to describe the motion of the particle (or train), what we mean is that we want to be able to say where it is at what time. Formally, we want to know the particle’s **position as a function of time**, which we will label as  $x(t)$ . The function will only be meaningful if:

- we specify an axis along the direction of motion
- we specify an origin where  $x = 0$
- we specify a direction along the axis of motion corresponding to increasing values of  $x$
- we specify the units for the quantity,  $x$ .

That is, unless all of these are specified, you would have a hard time describing the motion

of an object to one of your friends over the phone (or by Facebook chat).

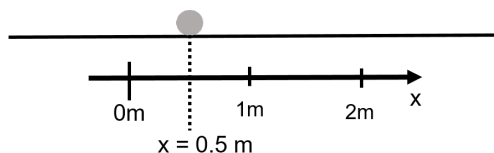


Figure 3.1: In order to describe the motion of the grey ball along a straight line, we introduce the  $x$ -axis, represented by an arrow to indicate the direction of increasing  $x$ , and the location of the origin, where  $x = 0$  m. Given our choice of origin, the ball is currently at a position of  $x = 0.5$  m.

Consider Figure 3.1 where we would like to describe the motion of the grey ball as it moves along a straight line. In order to quantify where the ball is, we introduce the “ $x$ -axis”, illustrated by the black arrow. The direction of the arrow corresponds to the direction where  $x$  increases (i.e. becomes more positive). We have also chosen a point where  $x = 0$ , and by convention, we choose to express  $x$  in units of meters (the S.I. unit for the dimension of length).

Note that we are completely free to choose both the direction of the  $x$ -axis and the location of the origin. The  $x$ -axis is a mathematical construct that we introduce in order to describe the physical world; we could have just as easily chosen for it to point in the opposite direction with the origin corresponding to the current position of the ball. Since we are completely free to choose where we define the  $x$ -axis, we should always try to choose the option that is most convenient to us.

### 3.1.1 Motion with constant speed

Now suppose that the ball in Figure 3.1 is rolling, and that we recorded its  $x$  position every second in a table and obtained the values in Table 3.1 (we will ignore measurement uncertainties and pretend that the values are exact). The easiest way to visualize the values

Time [s]	X position [m]
0.0 s	0.5 m
1.0 s	1.0 m
2.0 s	1.5 m
3.0 s	2.0 m
4.0 s	2.5 m
5.0 s	3.0 m
6.0 s	3.5 m
7.0 s	4.0 m
8.0 s	4.5 m
9.0 s	5.0 m

Table 3.1: Position of a ball along the  $x$ -axis recorded every second.

in the table is to plot them on a graph. Plotting position as a function of time is one of the most common graphs to make in physics, since it is often a complete description of the

motion of an object. We can easily plot these values in Python:

*Python Code 3.1: QExpy to calculate mean and standard deviation*

```
#First, we load the QExpy module
import qexpy as q
#We define t as a list of values (note the square brackets):
t = [0.0, 1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0]
#Similarly, we define the corresponding positions:
position = [0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0]
#Define the plot, and show it:
fig = q.MakePlot(xdata=t, xunits="s", xname="time",
                 ydata=position, yunits="m", yname="position",
                 data_name="position vs time")
fig.show()
```

*Output 3.1:*

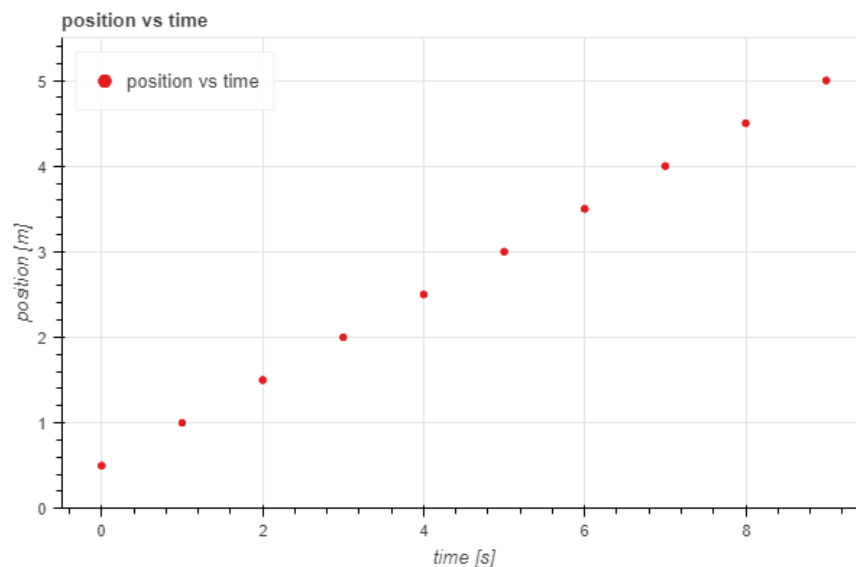


Figure 3.2: Plot of position as a function of time using the values from Table 3.1.

The data plotted in Figure 3.2 show that the  $x$  position of the ball increases linearly with time (i.e. it is a straight line). This means that in equal time increments, the ball will cover equal distances. Note that we also had the liberty to choose when we define  $t = 0$ ; in this case, we chose that time is zero when the ball was at  $x = 0.5$  m.

**Checkpoint 3-1:** Using the data from Table 3.1, at what position along the  $x$ -axis will the ball be when time is  $t = 9.5$  s, if it continues its motion undisturbed?

Since the position as a function of time for the ball plotted in Figure 3.2 is linear, we can summarize our description of the motion using a function,  $x(t)$ , instead of having to tabulate

the values as we did in Table 3.1. The function will have the functional form:

$$x(t) = a + b \times t$$

The constant  $a$  is the “offset” of the function, the value that the function has at  $t = 0$  s. The constant  $b$  is the slope and gives the rate of change of the position as a function of time. We can determine the values for the constants  $a$  and  $b$  by choosing any two rows from Table 3.1 (to determine 2 unknown quantities, you need 2 equations), and obtain 2 equations and 2 unknowns. For example, choosing the points where  $t = 0$  s and  $t = 2.0$  s:

$$\begin{aligned} x(t = 0 \text{ s}) &= 0.5 \text{ m} = a + b \times (0 \text{ s}) \\ x(t = 2.0 \text{ s}) &= 1.5 \text{ m} = a + b \times (2.0 \text{ s}) \end{aligned}$$

The first equation immediately gives  $a = 0.5$  m, which we can substitute into the second equation to get  $b$ :

$$\begin{aligned} 1.5 \text{ m} &= a + b \times (2.0 \text{ s}) = 0.5 \text{ m} + b \times (2.0 \text{ s}) \\ \therefore b &= \frac{(1.5 \text{ m}) - (0.5 \text{ m})}{(2.0 \text{ s})} = 0.5 \text{ m/s} \end{aligned}$$

which gives us the functional form for  $x(t)$ :

$$x(t) = (0.5 \text{ m}) + (0.5 \text{ m/s}) \times t$$

where you should note that  $a$  and  $b$  have different dimensions. Since  $a$  is added to something that must then give dimensions of length (for position,  $x$ ),  $a$  has dimensions of length.  $b$  is multiplied by time, and that product must have dimensions of length as well;  $b$  thus has dimensions of length over time, or “speed” (with S.I. units of m/s).

We can generalize the description of an object whose position increases linearly with time as:

$$\boxed{x(t) = x_0 + v_x t} \tag{3.1}$$

where  $x_0$  is the position of the object at time  $t = 0$  s ( $a$  from above), and  $v_x$  corresponds to the distance that the object covers per unit time ( $b$  from above) along the x-axis. We call  $v_x$  the “speed” of the object. If  $v_x$  is large, then the object covers more distance in a given time, i.e. it moves faster. If  $v_x$  is a negative number, then the object moves in the negative  $x$  direction.

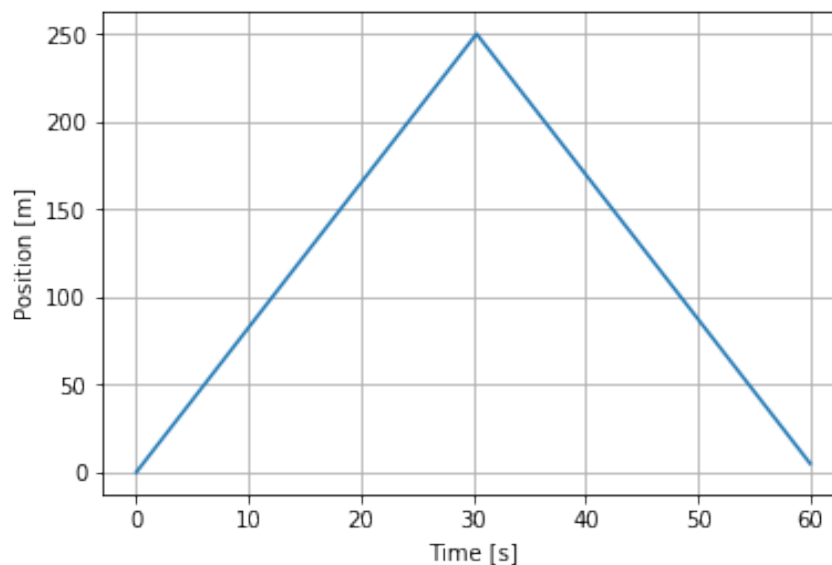


Figure 3.3: Position as a function of time for an object.

**Checkpoint 3-2:** Referring to Figure 3.3, what can you say about the motion of the object?

- A) The object moved faster and faster between  $t = 0$  s and  $t = 30$  s, then slowed down to a stop at  $t = 60$  s.
- B) The object moved in the positive x-direction between  $t = 0$  s and  $t = 30$  s, and then turned around and moved in the negative x-direction between  $t = 30$  s and  $t = 60$  s.
- C) The object moved with a higher speed between  $t = 0$  s and  $t = 30$  s than it did between  $t = 30$  s and  $t = 60$  s.

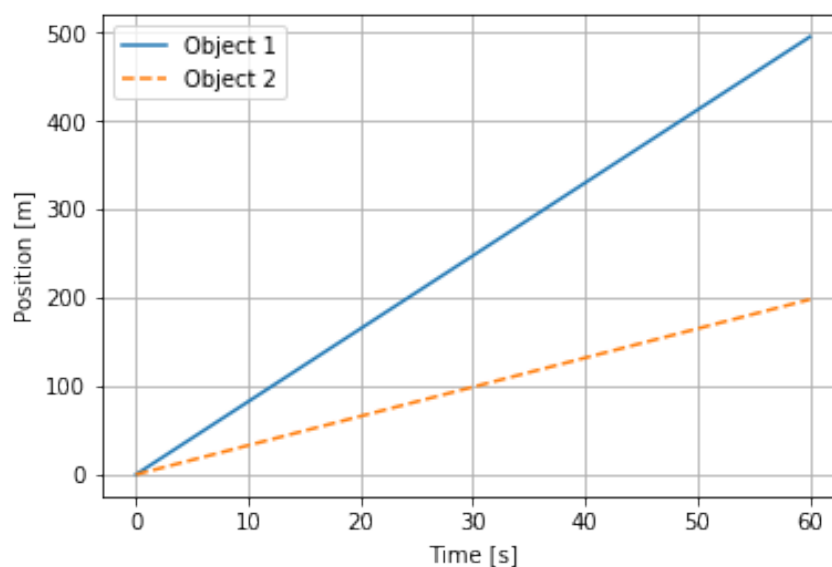


Figure 3.4: Positions as a function of time for two objects.

**Checkpoint 3-3:** Referring to Figure 3.4, what can you say about the motion of the two objects?

- A) Object 1 is slower than Object 2
- B) Object 1 is more than twice as fast as Object 2
- C) Object 1 is less than twice as fast as Object 2

### 3.1.2 Motion with constant acceleration

Until now, we have considered motion where the speed is a constant (i.e. where speed does not change with time). Suppose that we wish to describe the position of a falling object that we released from rest at time  $t = 0$  s. The object will start with a speed of 0 m/s and it will **accelerate** as it falls. We say that an object is “accelerating” if its speed is not constant. As we will see in later chapters, objects that fall near the surface of the Earth experience a constant acceleration (their speed changes at a constant rate).

Formally, we define acceleration as the rate of change of speed. Recall that speed is the rate of change of position, so acceleration is to speed what speed is to position. In particular, we saw that if the speed,  $v_x$ , is constant, then position as a function of time is given by:

$$x(t) = x_0 + v_x t \quad (3.1)$$

In analogy, if the acceleration is constant, then the speed as a function of time is given by:

$$\boxed{v_x(t) = v_{x0} + at} \quad (3.2)$$

where  $a$  is the “acceleration” and  $v_{x0}$  is the speed of the object at  $t = 0$ . We can work out the dimensions of acceleration for this equation to make sense. Since we are adding  $v_{x0}$  and  $at$ , we need the dimensions of  $at$  to be speed:

$$\begin{aligned} [at] &= \frac{L}{T} \\ [a][t] &= \frac{L}{T} \\ [a]T &= \frac{L}{T} \\ [a] &= \frac{L}{T^2} \end{aligned}$$

Acceleration thus has dimensions of length over time squared, with corresponding S.I. units of m/s<sup>2</sup> (meters per second squared or meters per second per second).

Now that we have an understanding of acceleration, how do we describe the position of an object that is accelerating? We cannot use equation 3.1, since it is only correct if the speed is constant.

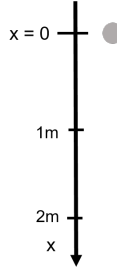


Figure 3.5:  $x$ -axis for an object that starts at rest at  $x = 0$  m when  $t = 0$  s and falls downwards (in the direction of increasing  $x$ ).

Let us work out the corresponding equation for position as a function of time for accelerated motion using the  $x$ -axis depicted in Figure 3.5. We will determine  $x(t)$  for the grey ball that starts at rest ( $v_{x0} = 0$  m/s) at the position  $x = 0$  m at time  $t = 0$  s with a constant positive acceleration  $a = 10$  m/s<sup>2</sup>. We would like to use equation 3.1, but we cannot because it only applies if the speed is constant. To remedy this, we pretend (we “approximate”) that for a very small amount of time, the speed is almost constant. Let us take a very small interval in time, say  $\Delta t = 0.001$  s, and approximate that the speed is constant during that interval. We can then use equation 3.1 for that small interval,  $\Delta t$ .

At  $t = 0$  s, we have  $x = 0$  m,  $v_{x0} = 0$  m/s and  $a = 10$  m/s<sup>2</sup>. At  $t = \Delta t$  (at the end of the interval), the speed will have increased from  $v = 0$  m/s to  $v = a\Delta t$ .

The average speed during the first interval,  $v_{avg}$  is then given by averaging the speeds at the beginning and at the end of the interval:

$$v_{avg}(t = \Delta t) = \frac{1}{2}(0 + a\Delta t) = \frac{1}{2}a\Delta t = \frac{1}{2}(10 \text{ m/s}^2)(0.001 \text{ s}) = 0.005 \text{ m/s}$$

We can use this average speed to find the position at time  $t = \Delta t$ :

$$x(t = \Delta t) = x_0 + v_{avg}(t = \Delta t)\Delta t = \frac{1}{2}a(\Delta t)^2 = \frac{1}{2}(10 \text{ m/s}^2)(0.001 \text{ s})^2 = 0.000005 \text{ m}$$

We can proceed to the next interval in time. At the beginning of the second interval, the speed is  $v(t = \Delta t) = a\Delta t$  and at the end of the second interval, it is  $v(t = 2\Delta t) = 2a\Delta t$ . The average speed during the second interval is thus

$$v_{avg}(t = 2\Delta t) = \frac{1}{2}(a\Delta t + 2a\Delta t) = \frac{3}{2}a\Delta t = \frac{3}{2}(10 \text{ m/s}^2)(0.001 \text{ s}) = 0.015 \text{ m/s}$$

The position at the end of the second interval is thus given by the position at the end of the first time interval plus the distance covered in the second interval:

$$\begin{aligned} x(t = 2\Delta t) &= x(t = \Delta t) + v_{avg}(t = 2\Delta t)\Delta t = \frac{1}{2}a(\Delta t)^2 + \frac{3}{2}a(\Delta t)^2 \\ &= \frac{1}{2}a(2\Delta t)^2 = \frac{1}{2}(10 \text{ m/s}^2)(2 \times 0.001 \text{ s})^2 = 0.00002 \text{ m} \end{aligned}$$



where in the last line, we kept the  $\frac{1}{2}$  factored out and brought a factor of 2 inside the parenthesis with the  $\Delta t$ . You can carry out this exercise to ultimately find the position at any time. However, if you carry it out over a few more intervals, you may notice a pattern. For the  $N$ th interval when  $t = N\Delta t$  at the end of the interval, we have:

$$\begin{aligned} v(t = (N-1)\Delta t) &= a(N-1)\Delta t && \text{(at beginning of interval } N) \\ v(t = N\Delta t) &= aN\Delta t && \text{(at end of interval } N) \\ v_{avg}(t = N\Delta t) &= \frac{1}{2}(a(N-1)\Delta t + aN\Delta t) = \frac{1}{2}a(2N-1)\Delta t \\ x(t = N\Delta t) &= \frac{1}{2}a(N\Delta t)^2 \end{aligned}$$

The last line gives us exactly what we were after, namely the position as a function of time for a fixed acceleration,  $a$ , if the object started at rest at a position of  $x = 0$  m:

$$x(t) = \frac{1}{2}at^2 \quad (3.3)$$

If at  $t = 0$ , the object had an initial position along the  $x$ -axis of  $x_0$ , then the position  $x(t)$  would be shifted by an amount  $x_0$ :

$$x(t) = x_0 + \frac{1}{2}at^2 \quad (3.4)$$

Finally, if the object had an initial speed  $v_{x0}$  at  $t = 0$ , one can easily reproduce the iteration above to find that we need to add an additional term to account for this. We arrive at the general description of the position of an object moving in a straight line with acceleration,  $a$ :

$$\boxed{x(t) = x_0 + v_{x0}t + \frac{1}{2}at^2} \quad (3.5)$$

Note that equation 3.1 is just a special case of the above when  $a = 0$ .

**Example 3-1:** A ball is thrown upwards with a speed of 10 m/s. After which distance will the ball stop before falling back down? Assume that gravity causes a constant downwards acceleration of 9.8 m/s<sup>2</sup>.

*We will solve this problem in the following steps:*

1. *Setup a coordinate system (define the  $x$ -axis).*
2. *Identify the condition that corresponds to the ball stopping its upwards motion and falling back down.*
3. *Determine the distance at which the ball stopped.*

*Since we throw the ball upwards with an initial speed upwards, it makes sense to choose*

an  $x$ -axis that points up and has the origin at the point where we release the ball. With this choice, referring to the variables in equation 3.5, we have:

$$\begin{aligned}x_0 &= 0 \\v_{x0} &= +10 \text{ m/s} \\a &= -9.8 \text{ m/s}^2\end{aligned}$$

where the initial speed is in the positive  $x$ -direction, and the acceleration,  $a$ , is in the negative direction (the speed will be getting smaller and smaller, so its rate of change is negative).

The condition for the ball to stop at the top of the trajectory is that its speed will be zero (that is what it means to stop). We can use equation 3.2 to find what time that corresponds to:

$$\begin{aligned}v(t) &= v_{x0} + at \\0 &= (10 \text{ m/s}) + (-9.8 \text{ m/s}^2)t \\\therefore t &= \frac{(10 \text{ m/s})}{(9.8 \text{ m/s}^2)} = 1.02 \text{ s}\end{aligned}$$

Now that we know that it took 1.02 s to reach the top of the trajectory, we can find how much distance was covered:

$$\begin{aligned}x(t) &= x_0 + v_{x0}t + \frac{1}{2}at^2 \\x &= (0 \text{ m}) + (10 \text{ m/s})(1.02 \text{ s}) + \frac{1}{2}(-10 \text{ m/s}^2)(1.02 \text{ s})^2 = 5.0 \text{ m}\end{aligned}$$

and we find that the ball will rise by 5 m before falling back down.

### Visualizing motion with constant acceleration

When an object has a constant acceleration, its speed and position as a function of time are described by the two equations:

$$\begin{aligned}v(t) &= v_{x0} + at \\x(t) &= x_0 + v_{x0}t + \frac{1}{2}at^2\end{aligned}$$

where the speed changed linearly with time, and the position changes quadratically with time (it goes as  $t^2$ ). Figure 3.6 shows the position and the speed as a function of time for the ball from example 3-1 for the first three seconds of the motion.

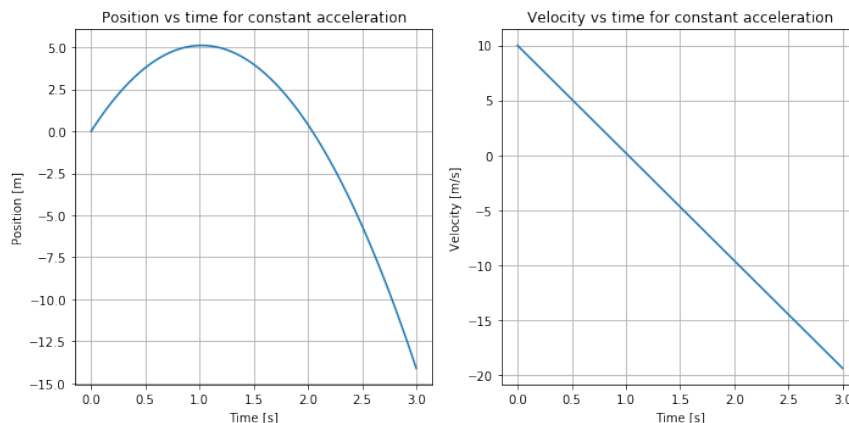


Figure 3.6: Position and speed as a function of time for the ball in example 3-1.

We can divide the motion into three parts:

**1) Between  $t = 0$  s and  $t = 1$  s**

At time  $t = 0$  s, the ball starts at a position of  $x = 0$  m (left) and a speed of  $v_{x0} = 10$  m/s (right). During the first second of motion, the position continues to increase (the ball is moving up), until the position stops increasing at  $t = 1$  s. During that time, the speed decreases linearly from 10 m/s to 0 m/s due to the constant negative acceleration from gravity. At  $t = 1$  s, the speed is 0 m/s and the ball is momentarily at rest (as it reaches the top of the trajectory before falling back down).

**2) Between  $t = 1$  s and  $t = 2$  s**

At  $t = 1$  s, the speed continues to decrease linearly (it becomes more and more negative) as the ball falls back down faster and faster. The position also starts decreasing just after  $t = 1$  s, as the ball returns back down to the point of release. At  $t = 2$  s, the ball returns to the point from which it was thrown, and the ball is going with the same speed (10 m/s) as when it was released, but in the opposite direction (downwards).

**3) After  $t = 2$  s**

If nothing is there to stop the ball, it continues to move downwards with ever increasing speed. The position continues to become more negative and the speed continues to become larger in magnitude and more negative.

### 3.1.3 Speed versus velocity

In the previous example, our language was not quite as precise as it should be when conducting science. Specifically, we need a way to distinguish the situation when the speed is decreasing (becoming more negative), while the object is actually going faster and faster (after  $t = 1$  s in Figure 3.6). We will use the term **speed** to refer to how fast an object is moving (how much distance it covers per unit time), and we will use the term **velocity** to

also indicate the direction of the motion. In other words, the speed is the absolute value of the velocity<sup>1</sup>. The speed is thus always positive, whereas the velocity can also be negative.

With this vocabulary, the speed of the ball in Figure 3.6 decreases between  $t = 0$  s and  $t = 1$  s, and increases thereafter. On the other hand, the velocity continuously decreases (it is always becoming more and more negative). Velocity is thus the more precise term since it tells us both the speed and the direction of the motion.

### 3.1.4 Generalized motion and instantaneous velocity

Most objects do not have a constant velocity or acceleration. We thus need to generalize our description of the position and velocity of an object to a more general case. This can be done in much the same way as we introduced accelerated motion; namely by pretending that during a very small interval in time,  $\Delta t$ , the velocity and acceleration are constant, and then considering the generalized motion as the sum over many small intervals. In the limit that  $\Delta t$  tends to zero, this will be an accurate description.

During a small interval in time,  $\Delta t$ , the object will cover a distance,  $\Delta x$ . The velocity in that interval is given by:

$$v = \lim_{\Delta t \rightarrow 0} \frac{\Delta x}{\Delta t}$$

which is exact in the limit  $\lim_{\Delta t \rightarrow 0}$ . We call this the **instantaneous velocity**, as it is the velocity only in that small instant in time where we choose  $\Delta x$  and  $\Delta t$ .

Another way to read this equation is that the velocity,  $v$ , is the slope of the graph of  $x(t)$ . Recall that the slope is the “rise over run”, in other words the change in  $x$  divided by the corresponding change in  $t$ . If we go back to Figure 3.6, we can indeed see that the graph of the velocity versus time ( $v(t)$ ) corresponds to the instantaneous slope of the graph of position versus time ( $x(t)$ ). For  $t < 1$  s, the slope of the  $x(t)$  graph is positive but decreasing (as is  $v(t)$ ). At  $t = 1$  s, the slope of  $x(t)$  is instantaneously 0 m/s (as is the velocity). Finally, for  $t > 1$  s, the slope of  $x(t)$  is negative and increasing in magnitude, as is  $v(t)$ .

Leibniz and Newton were the first to develop mathematical tools to deal with calculations that involve quantities that tend to zero, as we have here for our time interval  $\Delta t$ . Nowadays, we call that field of mathematics “calculus”, and we will make use of it here. Using the vocabulary of calculus, rather than saying that “instantaneous velocity is the slope of the graph of position versus time at some point in time”, we say that “instantaneous velocity is the time derivative of position as a function of time”. We also use a slightly different notation so that we do not have to write the limit  $\lim_{\Delta t \rightarrow 0}$ :

$$\boxed{v(t) = \frac{dx}{dt} = \frac{d}{dt}x(t)} \quad (3.6)$$

---

<sup>1</sup>This is true for one-dimensional motion, whereas in two or more dimensions, velocity is a vector and speed is the magnitude of that vector.

where we can really think of  $dt$  as  $\lim_{\Delta t \rightarrow 0} \Delta t$ , and  $dx$  as the corresponding change in position over an *infinitesimally* small time interval  $dt$ .

Similarly, we introduce the **instantaneous acceleration**, as the time derivative of  $v(t)$ :

$$a(t) = \frac{dv}{dt} = \frac{d}{dt}v(t) \quad (3.7)$$

### Using calculus to obtain acceleration from position

Suppose that we know the function for position as a function of time, and that it is given by our previous result:

$$x(t) = x_0 + v_{x0}t + \frac{1}{2}at^2$$

With our calculus-based definitions above, we should be able to recover that:

$$\begin{aligned} v(t) &= v_{x0}t + at \\ a(t) &= a \end{aligned}$$

Let us start by determining  $v(t)$ :

$$v(t) = \frac{dx}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\Delta x}{\Delta t}$$

Knowing our function  $x(t)$ , we can re-write this as:

$$\frac{dx}{dt} = \lim_{\Delta t \rightarrow 0} \frac{x(t + \Delta t) - x(t)}{\Delta t} \quad (3.8)$$

We will proceed as follows, as illustrated in Figure 3.7:

1. Use  $x(t)$  to determine  $\Delta x$  for a small interval  $\Delta t$
2. Divide  $\Delta x$  by  $\Delta t$
3. Take the limit,  $\lim_{\Delta t \rightarrow 0}$

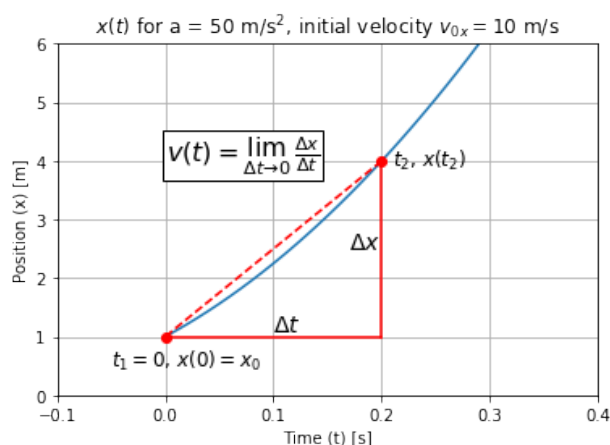


Figure 3.7: Obtaining the instantaneous velocity from the graph of position as a function of time. As  $\Delta t \rightarrow 0$ ,  $\frac{\Delta x}{\Delta t}$  approaches the instantaneous velocity  $v(t)$ , which is the slope of the curve  $x(t)$  at time  $t$ .

To obtain  $\Delta x$ , we introduce a start time,  $t_1$ , and an end time  $t_2$  for our time interval, such that  $t_2 - t_1 = \Delta t$ , centred about a time  $t$ . The change in position is then given by:

$$\begin{aligned}
 \Delta x &= x(t_2) - x(t_1) \\
 &= \left( x_0 + v_{x0}t_2 + \frac{1}{2}at_2^2 \right) - \left( x_0 + v_{x0}t_1 + \frac{1}{2}at_1^2 \right) \\
 &= v_{x0}(t_2 - t_1) + \frac{1}{2}a(t_2^2 - t_1^2) \\
 &= v_{x0}\Delta t + \frac{1}{2}a(t_2 - t_1)(t_2 + t_1) \\
 &= v_{x0}\Delta t + \frac{1}{2}a\Delta t(t_2 + t_1)
 \end{aligned}$$

which we divide by  $\Delta t$  to get  $v(t)$ :

$$\begin{aligned}
 v(t) &= \frac{\Delta x}{\Delta t} \\
 &= \frac{v_{x0}\Delta t + \frac{1}{2}a\Delta t(t_2 + t_1)}{\Delta t} \\
 &= v_{x0} + \frac{1}{2}a(t_2 + t_1)
 \end{aligned}$$

We now take the limit where  $\Delta t \rightarrow 0$ , namely when  $t_2 - t_1$  is very small. As we make the interval small,  $t_2$  and  $t_1$  will both approach the same value of time, say  $t$ , corresponding to the time at the centre of the interval. In particular, the average of  $t_1$  and  $t_2$ , given by  $\frac{1}{2}(t_1 + t_2)$ , will approach the time at the centre of the interval,  $t$ . We thus recover the equation for instantaneous velocity:

$$v(t) = v_{x0} + \frac{1}{2}at$$

Of course, once you become more familiar with calculus, you will be able to directly use the formulas for derivatives to recover the answer:

$$\begin{aligned}
 v(t) &= \frac{d}{dt} \left( x_0 + v_{x0}t + \frac{1}{2}at^2 \right) \\
 &= v_{x0} + at
 \end{aligned}$$

Similarly, we can now confirm that the acceleration is a constant, independent of time:

$$\begin{aligned}
 a(t) &= \frac{dv}{dt} = \frac{d}{dt} (v_{x0} + at) \\
 &= a
 \end{aligned}$$

You can easily verify that you obtain this result by first calculating  $v(t)$  at two different times,  $t_1$  and  $t_2$ , taking the difference,  $\Delta v = v(t_2) - v(t_1)$ , and then taking the limit of  $\lim_{\Delta t \rightarrow 0} \frac{\Delta v}{\Delta t}$  to get  $a(t)$ .

### Using calculus to obtain position from acceleration

Now that we saw that we can use derivatives to determine acceleration from position, we will see how to do the reverse and use acceleration to determine position. Let us suppose that we have a constant acceleration,  $a(t) = a$ , and that we know that at time  $t = 0$  s, the object had a speed of  $v_{x0}$  and was located at a position  $x_0$ .

Since we only know the acceleration as a function of time, we first need to find the velocity as a function of time, before we can find the function for the position. We start with:

$$a(t) = a = \frac{d}{dt}v(t)$$

which tells us that we know the slope (derivative) of the function  $v(t)$ , but not the actual function. In this case, we must do the opposite of taking the derivative, which in calculus is called taking the “anti-derivative” with respect to  $t$  and has the symbol  $\int dt$ . In other words, if:

$$\frac{d}{dt}v(t) = a(t)$$

then:

$$v(t) = \int a(t)dt + C$$

where as we will see later, the constant  $C$  is required in order for the function  $v(t)$  to go through the point  $v(t = 0) = v_{x0}$ . All we need now is to determine how to calculate the anti-derivative,  $\int a(t)dt$ . Since in this case,  $a(t)$  is a constant,  $a$ , we can determine the anti-derivative quite easily.

For a small interval in time,  $\Delta t$ , the velocity will change by a small amount,  $\Delta v$ , such that:

$$\begin{aligned} a &= \frac{\Delta v}{\Delta t} \\ \Delta v &= a\Delta t \end{aligned}$$

If we label the start time of the interval as  $t_1$  and the end of the interval as  $t_2$ , we have:

$$\begin{aligned} \Delta t &= t_2 - t_1 \\ \Delta v &= v(t_2) - v(t_1) = a\Delta t \\ \therefore v(t_2) &= v(t_1) + a\Delta t \end{aligned}$$

If we set  $t_1 = 0$  s to correspond to the point where  $v(t) = v_{x0}$ , then we can write the velocity at  $t_2$  as:

$$v(t_2) = v(t = \Delta t) = v_{x0} + a\Delta t$$

and we see that the velocity changed by an amount  $a\Delta t$  over a period of time  $\Delta t$ . Since  $a$  is the same at all times, this is always true, and after a period of time  $t = N\Delta t$ , the velocity

will have changed by  $Na\Delta t$ , and we recover the original equation for velocity as a function of time when acceleration is constant:

$$\begin{aligned} v(t = N\Delta t) &= v_{x0} + Na\Delta t \\ \therefore v(t) &= v_{x0} + at \end{aligned}$$

We can identify the anti-derivative for the case where  $a(t)$  is constant:

$$\begin{aligned} \frac{d}{dt}v(t) &= a \\ \int a dt &= at + C \end{aligned}$$

where the constant,  $C$ , is given by  $v_{x0}$ .

To now obtain position as a function of time, we proceed in the same manner, namely:

1. Define a small interval in time  $\Delta t$
2. Calculate the corresponding change in position  $\Delta x$
3. Add  $\Delta x$  to our original position, and repeat.

Again, we have the time-derivative of the position equal to a function of time:

$$\frac{d}{dt}x(t) = v(t) = v_{x0} + at$$

and we need to find the anti-derivative:

$$x(t) = \int (v_{x0} + at) dt$$

given that, at time  $t = 0$ , the position was  $x = x_0$ . After a small interval in time,  $\Delta t$ , the position will have changed by amount  $\Delta x$ :

$$\Delta x = v(t)\Delta t$$

so that the position at time  $t = \Delta t$  will be given by:

$$\begin{aligned} x(t = \Delta t) &= x_0 + \Delta x \\ &= x_0 + v(t)\Delta t \end{aligned}$$

The problem here is to evaluate  $v(t)$  since the velocity changes throughout the interval. One possible choice is to evaluate the velocity at  $t = \frac{1}{2}\Delta t$ , midway in the interval, as we did

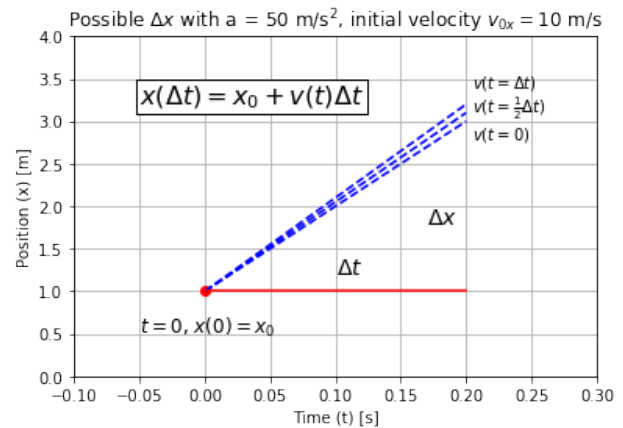


Figure 3.8: Determining  $\Delta x$ , given  $v(t)$  and  $\Delta t$ . Three different choices of  $v(t)$  are shown, depending on whether  $v(t)$  is evaluated at the start of the interval, in the middle, or at the end. As  $\Delta t \rightarrow 0$ , these all become equal.



before. We could also choose to use the velocity at the beginning or at the end of the time interval, as all three choices will converge to the same value when  $\Delta t \rightarrow 0$ , as illustrated in Figure 3.8. For now, we will leave the choice open and simply call the velocity that we use  $v_1$  to indicate that it is the velocity in the first interval. We thus write the position,  $x(t = \Delta t)$ , after a time interval  $\Delta t$  as:

$$x(t = \Delta t) = x_0 + v_1 \Delta t$$

The position,  $x(t = 2\Delta t)$ , after another interval in time  $\Delta t$  will then be given by:

$$\begin{aligned} x(t = 2\Delta t) &= x(t = \Delta t) + v_2 \Delta t \\ &= x_0 + v_1 \Delta t + v_2 \Delta t \end{aligned}$$

where  $v_2$  is the velocity over the second interval in time (different than  $v_1$ , since velocity changes with time). For the  $N$ th interval, we label the position  $x_N = x(t = N\Delta t)$ :

$$\begin{aligned} x_N &= x(t = N\Delta t) = x_0 + v_1 \Delta t + v_2 \Delta t + \cdots + v_N \Delta t \\ &= x_0 + \sum_{i=1}^N v_i \Delta t \end{aligned}$$

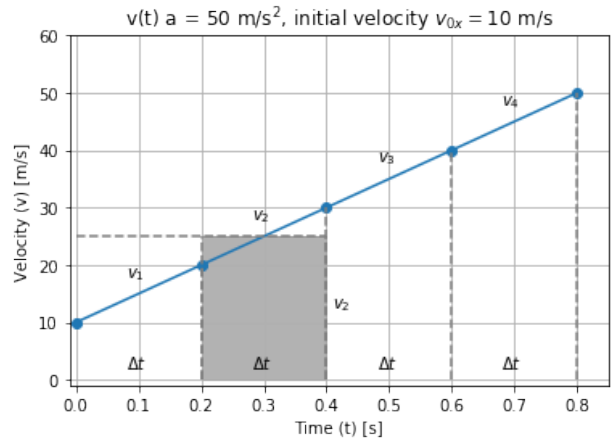
where we made use of the summation notation ( $\sum$ ) to avoid writing out every term. The above equation is only correct in the limit of  $\Delta t \rightarrow 0$ , in which case it must be the anti-derivative of  $v(t)$ :

$$\begin{aligned} x(t) &= \int v(t) dt \\ &= x_0 + \lim_{\Delta t \rightarrow 0} \sum_{i=1}^N v_i \Delta t \end{aligned}$$

so that we can identify:

$$\boxed{\int v(t) dt = \lim_{\Delta t \rightarrow 0} \sum_{i=1}^N v_i \Delta t + C}$$

(3.9) *Figure 3.9: Illustration of the anti-derivative  $\int v(t) dt$  as a sum.*



where it is understood that  $v_i$  is the “average” velocity in the  $i$ th interval, and the constant  $C = x_0$  ensures that  $x(t = 0) = x_0$ . This is now a general definition for the anti-derivative, as we have made no specific assumption about the function  $v(t)$ . Equation 3.9 tells us that the anti-derivative of a function (in this case,  $v(t)$ ) can be obtained from a sum.

The sum is illustrated in Figure 3.9, which shows the function  $v(t)$  and several intervals  $\Delta t$ . Over each interval,  $i$ , we labeled the average velocity,  $v_i$ . As the intervals shrink,  $\Delta t \rightarrow 0$ , the average velocity  $v_i$  approaches the instantaneous velocity,  $v(t)$ , at the centre for the interval. Since the function  $v(t)$  is linear, the speed at the middle of an interval is exactly

equal to the average speed in the interval. Taking  $v_i$  as the speed in the middle of the interval, we then see that each term in the sum,  $v_i\Delta t$ , is equal to the area of between the  $v(t)$  and the  $t$ -axis. This is illustrated for the second term in the sum,  $v_2\Delta t$ , with the grey rectangle in Figure 3.9.

The anti-derivative of a function is thus related to the area between the function and the horizontal axis. If we specify limits on the horizontal axis between which we calculate the area, then the anti-derivative is called an **integral**. For example, if we wish to calculate the sum of  $v_i\Delta t$  for values between  $t_a$  and  $t_b$ , we would write the integral as:

$$\int_{t_a}^{t_b} v(t)dt = \lim_{\Delta t \rightarrow 0} \sum_{i=1}^N v_i\Delta t$$

where the sum is such that for  $i = 1$ ,  $v_i$  is close to  $v(t = t_a)$  and for  $i = N$ ,  $v_N$  is close to  $v(t = t_b)$ . An illustration of taking the integral of  $v(t)$  is shown in Figure 3.10 where the sum is shown for two different values of  $\Delta t$ . It is clear that as  $\Delta t$ , the sum becomes equal to the area between the curve and the horizontal axis.

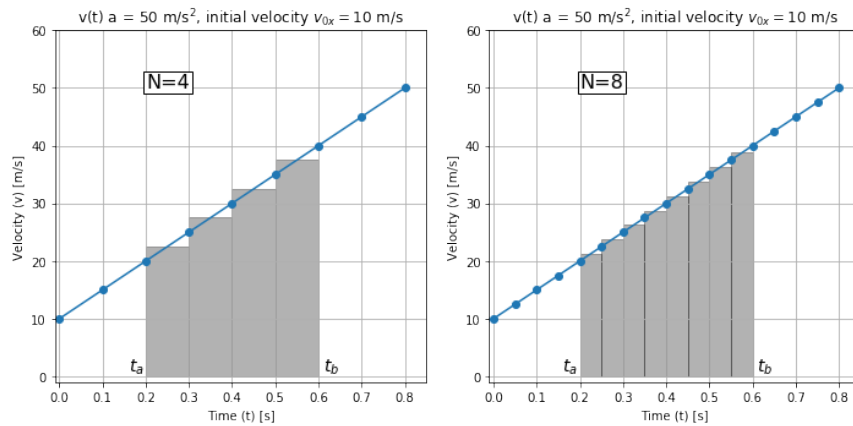


Figure 3.10: Integral of  $v(t)$  between  $t_a = 0.2$  s and  $t_b = 0.6$  s illustrated as a sum with of 4 terms when  $\Delta t = 0.1$  s or of 8 terms when  $\Delta t = 0.05$  s.

Since  $v(t)$  is a linear function when acceleration is constant, we can easily calculate the area between the curve and the horizontal axis. In the case of a linear function,  $v(t) = v_{x0} + at$ , the area is a trapezoid, and we have:

$$\begin{aligned} \int_{t_a}^{t_b} v(t)dt &= \text{base} \times \text{average height} \\ &= (t_b - t_a) \times \frac{1}{2} (v(t_a) + v(t_b)) \\ &= (t_b - t_a) \frac{1}{2} (v_{x0} + at_a + v_{x0} + at_b) \\ &= \left( \frac{1}{2} (v_{x0}t_b + at_at_b + v_{x0}t_b + at_b^2) \right) - \left( \frac{1}{2} (v_{x0}t_a + at_a^2 + v_{x0}t_a + at_bt_a) \right) \\ &= \left( v_{x0}t_b + \frac{1}{2}at_b^2 \right) - \left( v_{x0}t_a + \frac{1}{2}at_a^2 \right) \end{aligned}$$

If we define a new function,  $V(t) = v_{x0}t + \frac{1}{2}at^2$ , then we have:

$$\int_{t_a}^{t_b} v(t)dt = V(t_b) - V(t_a)$$

In other words, given a function,  $v(t)$ , the integral of that function between two values  $t_a$  and  $t_b$  can be found by evaluating a different function,  $V(t)$ , at the end points  $t_a$  and  $t_b$ <sup>2</sup>. As you will see in your calculus course, the function  $V(t)$  is precisely what we call the anti-derivative:

$$\int v(t)dt = V(t) + C$$

which has derivative:

$$\frac{dV}{dt} = v(t)$$

Note that when taking the integral, the constant  $C$  always cancels.

---

<sup>2</sup>Note that we only explicitly showed that this true if  $v(t)$  is linear, but the result is in fact general.

## 3.2 Summary

- Something interesting