

5 | Statistical Mechanics

Uptil now we have only looked at systems with a small number of particles – either one or two. In principle, Newton’s laws allow us to describe systems containing arbitrary number of particles. In practice however, it becomes increasingly difficult to describe large systems like gas contained within a room or the molecules in a glass of water. Statistical mechanics deals with such situations. It is an expansive and integral field of study which has applications in almost all the other fields of physics.

5.1 Gas in a box

In class, we estimated the number of particles in the lecture room to be around 10^{28} and although the lecture room is huge, in a typical room there would be around 10^{27} molecules of air. We want to make general statements about the behaviour of these particles rather than trying to describe the motion of each particle in great detail. What do we know about the gas from experience?

- If you’ve noticed the particles of dust dancing about in a ray of sunlight entering a room, you know that their motion is completely random. Although we cannot see the gas molecules, it is the collision with these molecules that make a dust particle move about randomly. This is called Brownian motion.
- We also know from experience that if initially the gas is confined to a small area in the room, eventually all of the room will be filled homogenously with this gas. More importantly, we never see the opposite happen – a gas uniformly occupying the volume of a room never returns back to a small area in the corner of the room.

In statistical mechanics we would like to study simple, isolated systems. So instead of looking at a room – which contains chairs and doors and windows and dust and your overdue assignments – we look at a gas in a box. The gas is made up of only one type of particle and the box is isolated from the environment, i.e., there is no flow of particles or energy or volume between the system and the environment (figure 5.1). We additionally assume that the particles are like tiny spheres and only interact via collisions and not in some complex way – there are no forces on the system like the electromagnetic force or the van der Waals force *etc.*



Figure 5.1: The system is isolated. The walls don't move and are insulated.

To describe this system at a macro level, we need some relevant parameters. For this system these are: the total energy of the particles (E), the volume of the box (V) and the total number of identical particles (N). Now let's imagine many copies of this system having the same values for E, V, N . And let's say that they start in more or less the same initial configuration where all the molecules are confined to one corner (figure 5.2).

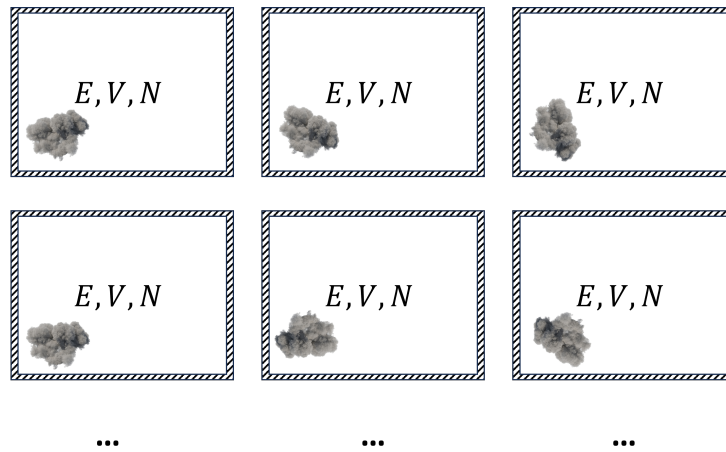


Figure 5.2: An ensemble of systems with similar initial states.

The question is: what do we expect after some time has passed and the system is left to evolve? In what respects do we expect the systems to be different to each other? In what respects do we expect the systems to be similar to each other? Experience tells us that the **macroscopic state** of the systems will be identical (they'll have a similar temperature, pressure, density *etc.*) but the **microscopic state** of each system will be different (not all copies of the same particle will have the same location or the same momentum). The differences in the systems must be microscopic. Quantities like temperature, pressure and density are macroscopic quantities which make no sense if the number of particles is small. In some sense, the idea of temperature (or pressure or density) doesn't exist when talking about a few particles. These are "emergent" quantities.

Even if we start this ensemble of systems in different configurations, they still end up in a similar state called the state of *thermodynamic equilibrium*. At thermodynamic equilibrium, the average properties (T, P, ρ) become the same everywhere but depending on the number of

particles fluctuations still exist. Another way of setting initial configurations is to initially have two halves of a system at different temperatures and then removing the boundary between them (or putting a thermally conducting wall between the halves).

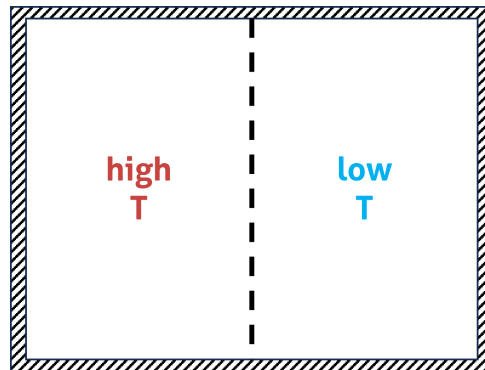


Figure 5.3: The dashed line represents a thermally conducting wall which allows for the transfer of energy but not particles.

There will be an exchange of energy between the two halves. At equilibrium however this macroscopic flow will die away. Again if we imagine having an ensemble of such systems and allow them to reach equilibrium, in all copies P , T and ρ will be the same. They have the same **macrostate**. Nevertheless they'll be in different microscopic states (**microstates**). In other words: a single macrostate corresponds to many microstates.

5.2 The Phase Space

How do we define these microstates? Classically we know that to define the trajectory of a particle completely we need both its position and momentum. We can represent a microstate of a system as a point in a space called the “phase space”. For instance if we have one particle which can only move along one axis, we can represent its microstate as follows:

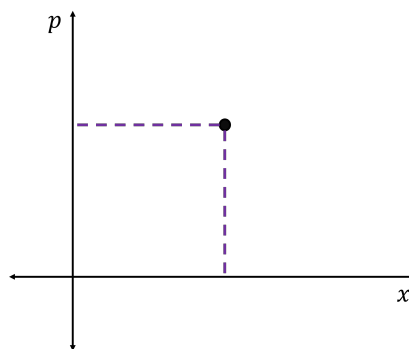


Figure 5.4: When we draw this point in phase space, we have completely defined the microstate.

The coordinates of this point are (x, p) . Note that this point in phase space is for one particle at one instant of time. We can plot a point for multiple particles and multiple instants of time in the same space. For a particle in two dimensions we'll have four coordinates (x, y, p_x, p_y) and hence a 4D phase space. For *two* particles in three dimensions we would require the

following 12 coordinates:

$$(x_1, y_1, z_1, p_{x1}, p_{y1}, p_{z1}, x_2, y_2, z_2, p_{x2}, p_{y2}, p_{z2})$$

In general for N particles we need a point in $6N$ dimensional phase space to define a microstate.

Consider N particles in a cubical box of dimension L . In the $6N$ -dimensional phase space, all coordinates that define position are confined between 0 and L . Is there any such restriction for the momentum? Well, no. Each particle could have any value of momentum in principle. But only such that the following is true:

$$\sum_{i=1}^N \frac{|p_i|^2}{2m} = E \quad (5.1)$$

The total kinetic energy of the particles shouldn't exceed the limit we imposed on the system to begin with. It is impossible to imagine what equation 5.1 represents geometrically. But if you realize that the equation of a circle is $x^2 + y^2 = r^2$ and that the equation of a sphere is $x^2 + y^2 + z^2 = r^2$, you can at least mathematically call equation 5.1 the equation of some-sort-of-higher-dimensional-sphere with radius $r = \sqrt{2mE}$. The formal term for this is a "hypersphere". So, equation 5.1 represents a $3N$ -dimensional hypersphere.

Once again consider an ensemble of such systems that have fixed E, V, N . This is called a *microcanonical ensemble*. We already know that the initial conditions of these systems don't matter – the systems "forget" the initial conditions in that they arrive at the same macrostate. Once all the systems are at equilibrium if we take a snapshot of each system and look at the microstates corresponding to each system we'll find that they uniformly populate the available phase space. This means that all accessible microstates are equally likely to be accessed in this system. Now consider we only look at *one* system in thermodynamic equilibrium. But now instead of looking at the system for one moment in time, we let it evolve and move around in this equilibrium state and note down all the microstates it goes through. Note that even though the system is in thermodynamic equilibrium its microstate will keep changing at impressive speeds. And given enough time, the system will also access *all* the accessible microstates. This is known as the **ergodic hypothesis**. Ergodicity is where the time average of a process is the same as its ensemble average.

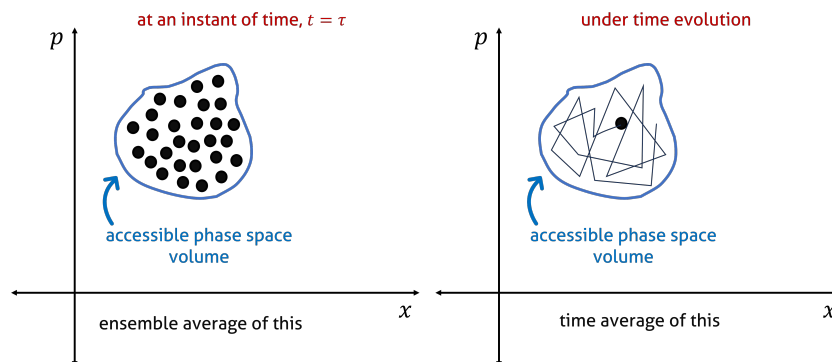


Figure 5.5: Ergodicity of phase space.

5.3 Boltzmann's Entropy

In 1887, Ludwig Boltzmann provided the current definition of entropy, $S = k_B \ln \Omega$, where Ω is the number of microstates accessible by a system. The concept of entropy existed even before Boltzmann gave his equation. Entropy was something that was known to increase when a system moves towards equilibrium. Here is a quote from the Wikipedia article on entropy:

The thermodynamic definition of entropy was developed in the early 1850s by Rudolf Clausius and essentially describes how to measure the entropy of an isolated system in thermodynamic equilibrium with its parts. Clausius created the term entropy as an extensive thermodynamic variable that was shown to be useful in characterizing the Carnot cycle. Heat transfer in the isotherm steps (isothermal expansion and isothermal compression) of the Carnot cycle was found to be proportional to the temperature of a system (known as its absolute temperature). This relationship was expressed in an increment of entropy that is equal to incremental heat transfer divided by temperature.

Boltzmann was the one who connected the idea of entropy to the number of accessible microstates of a system. To see this, consider the following situation (figure 5.6).

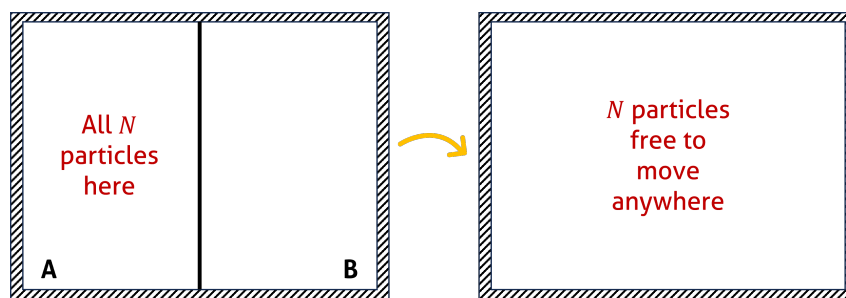


Figure 5.6: N particles are initially restricted to half of the box's volume. Then we remove the wall separating the two halves.

Initially all the particles are in half-A. The phase space in this case would be a hypersphere in the $3N$ momentum dimensions and a hypercube of side length $L/2$ in the $3N$ position dimensions. When we remove the wall, the hypersphere will remain the same but the accessible phase space volume will double – each position coordinate of any particle is not restricted between $(0, L/2)$ anymore but between $(0, L)$. Hence the action of removing the wall has increased the number of accessible microstates and therefore the entropy of the system. In the process of evolution – which is a non-equilibrium state where all particles go from filling one half to filling the whole volume – the new entropy at each point in time will be greater than the original entropy. But in what way does the number of microstates increase? Let's think of a discrete version of this situation to make our life easier. Consider that initially your system is just one die with six possible microstates (the numbers 1 to 6). Now if we “double the system” by considering two dice, the number of microstates accessible by the system will be $6 \times 6 = 36$. The microstates have not doubled but have been squared. Figure 5.7 shows the general case.

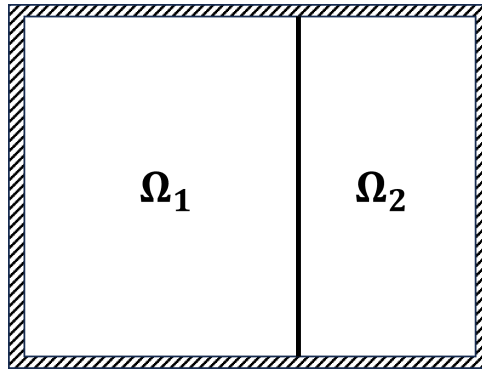


Figure 5.7: The number of microstates accessible in the left half is Ω_1 and the number of microstates accessible in the right half is Ω_2

If the number of microstates accessible in the left half of the system Ω_1 and the number of microstates accessible in the right half of the system is Ω_2 , then after removing the wall the total number of microstates accessible by the combined system will be $\Omega_1 \times \Omega_2$. Going back to the case in figure 5.6 we can see that the entropy doubles if we double the volume. Since the number of microstates accessible by the new system is just $\Omega \times \Omega$,

$$S_{\text{full}} = k_B \ln \Omega^2 = 2 \cdot k_B \ln \Omega = 2 \cdot S_{\text{half}}$$

Such quantities which double when the system is doubled in size¹ are known as *extensive quantities*. Other extensive quantities include volume, mass, internal energy *etc.* On the other hand, quantities like temperature and pressure which do not depend on the size of the system are known as *intensive quantities*.

Exercise 5.1: Never go back

Once the wall in figure 5.6 is removed and equilibrium has been attained, the probability that you'll find *all* the N particles in one half of the box will be,

$$P(\text{all } N \text{ particles in one half}) = \left(\frac{1}{2}\right)^N$$

This is because the probability of each particle being in one half of the box will be $1/2$ (proportional to the required volume). Hence, to find the probability of all N particles to be in one half we just multiply $1/2$, N times. Calculate the following probabilities:

1. finding all particles in a third of the box.
2. finding all particles in 99.9999% volume of the box.

We saw that for a typical real-life room we might have about $N = 10^{28}$ molecules of air in it. Input this value in part 2 above and convince yourself that even if we increase the available physical volume by 0.00001%, the probability that you'll find all particles in the old volume will be vanishingly small.

¹or are halved when the system is halved in size or tripled when the system is tripled in size and so on. In other words, extensive quantities scale linearly with system size.

5.4 What is the temperature?

Us humans obviously have an experiential understanding of what temperature is; one doesn't have to know what temperature *means* to reflexively remove their hand from a hot stove. Experience also tells us that something called "heat energy" will flow from a hot body to a nearby cold body until they have the same *temperature*. We'll see presently how the concept of temperature arises naturally if we allow for energy exchange between two subsystems. Consider figure 5.8.

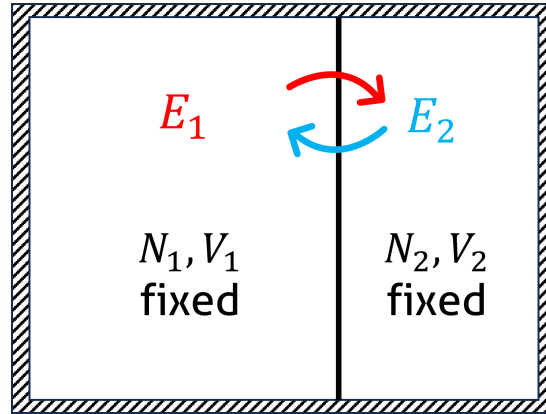


Figure 5.8: There is a thermally conducting wall separating the two parts of the box. The box in its entirety is still a microcanonical ensemble because the total energy $E = E_1 + E_2$ is fixed.

There are two systems each of which have a fixed volume and a fixed number of particles. These two systems are separated by a partition which allows for exchange of energy (but not particles). We assume that the total energy of the composite system is conserved, $E = E_1 + E_2$. If we allow the transfer of energy, we know that the total entropy will keep on increasing until it reaches a maximum. When this happens and the entropy doesn't change anymore, thermodynamic equilibrium has been reached.

What is the relevant parameter with respect to which the entropy will change? The only things changing here are the energies of the individual systems E_1 and E_2 (but since the total energy is conserved, the energy of only one system is an independent variable). Hence, we'll look at how entropy changes with E_1 (or E_2 just as well). We know,

$$S = S_1 + S_2 \quad (5.2)$$

In terms of energy dependence,

$$S(E_1) = S_1(E_1) + S_2(E_2) \quad (5.3)$$

where $E_2 = E - E_1$ is a function of E_1 . To see how one quantity changes with another we use derivatives. Now,

$$\left. \frac{dS}{dE_1} \right|_V = \left. \frac{dS_1}{dE_1} \right|_{V_1} + \left. \frac{dS_2}{dE_1} \right|_{V_2} \quad (5.4)$$

but, $dE_1 = -dE_2$. Hence,

$$\left. \frac{dS}{dE_1} \right|_V = \left. \frac{dS_1}{dE_1} \right|_{V_1} - \left. \frac{dS_2}{dE_2} \right|_{V_2} \quad (5.5)$$

At equilibrium, $dS = 0$. Which means that at equilibrium $dS/dE_1 = 0$. From equation 5.5 we get,

$$\left. \frac{dS_1}{dE_1} \right|_{V_1} = \left. \frac{dS_2}{dE_2} \right|_{V_2} \quad (5.6)$$

This tells us that if two systems are allowed to exchange energy, they reach equilibrium when a certain function – the derivative of entropy with respect to energy – takes the same numerical value for both systems. Clearly, this function must be related to the temperature of either system, since we know that systems in thermodynamic equilibrium possess the same temperature. It turns out,

$$\frac{1}{T} = \left. \frac{dS}{dE} \right|_{N,V} \quad (5.7)$$

And so, temperature is “that thing that becomes numerically equal at equilibrium for two systems when they are allowed to exchange energy”.

Does the form of temperature as seen in equation 5.7 agree with our experience? Equation 5.5 tells us that,

$$\frac{dS}{dE_1} = \frac{1}{T_1} - \frac{1}{T_2} \quad (5.8)$$

Imagine a situation where we slightly perturb the eventual equilibrium configuration of the situation in figure 5.8. Now the system will move towards equilibrium and the entropy will increase, $dS > 0$. We want to verify that if energy flows from system 1 to system 2 we better find that $T_1 > T_2$. Let's see if this works. If energy flows from system 1 to system 2 then $dE_1 < 0$. Hence,

$$\frac{dS}{dE_1} < 0; \quad \text{because } dS > 0$$

which means,

$$\frac{1}{T_1} - \frac{1}{T_2} < 0 \Rightarrow T_1 > T_2 \quad (5.9)$$

So yes, this way of defining temperature makes experiential sense.

5.5 The Canonical Ensemble

...to be added.

6 | Mathematical Physics

This chapter contains the mathematical tools and methods which are needed for understanding the physics being taught.

6.1 Matrices

Consider the following coordinate systems which share an origin but are mutually rotated with respect to each other by angle θ

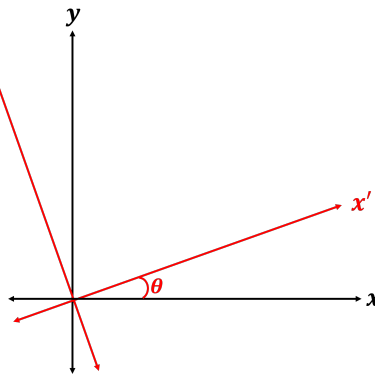


Figure 6.1

We can relate the two coordinates using the following *linear transformation*,

$$\begin{aligned}x' &= ax + by \\y' &= cx + dy\end{aligned}\tag{6.1}$$

We can write the equations 6.1 in a compact form. Conventionally, we write:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}\tag{6.2}$$

This is true only if equation 6.1 applies.

Exercise 6.1: The Rotation Matrix

Find the coefficient a, b, c, d from equation 6.1. The following figure will help you in figuring^a it out.

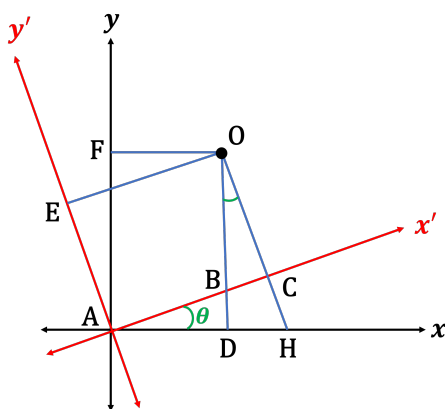


Figure 6.2

Hints:

- On the four axes, x, y, x' and y' , identify what the *lengths* x, y, x' and y' are. For example, $AE = y'$.
- Consider triangles ODH , ABD and ACH .

Write a, b, c, d as a matrix (equation 6.2). Let $\theta \rightarrow -\theta$. This matrix is the classic rotation matrix.

^apun unintended

Now, imagine another coordinate system (x'', y'') which is rotated with respect to (x', y') . The situation is as follows:

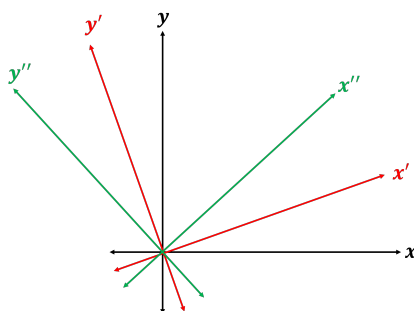


Figure 6.3: Three mutually-rotated coordinate systems

Instead of presuming the rule of matrix multiplication, using this schematic, we can see it naturally emerge from simple algebraic manipulations. We see that, (x'', y'') and (x', y') are related using the same form as in equation 6.1,

$$\begin{aligned} x'' &= a_2 x' + b_2 y' \\ y'' &= c_2 x' + d_2 y' \end{aligned} \quad (6.3)$$

The matrix form of this equation is,

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \quad (6.4)$$

We use a slightly different notation for the relation 6.1 by adding the subscript of 1 to each coefficient. Now,

$$\begin{aligned} x' &= a_1 x + b_1 y \\ y' &= c_1 x + d_1 y \end{aligned} \quad (6.5)$$

The matrix form of this equation is,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (6.6)$$

Let's try to find how (x'', y'') is related to (x, y) . Simply substitute equation 6.5 into equation 6.3 to get (**verify**),

$$\begin{aligned} x'' &= (a_2 a_1 + b_2 c_1)x + (a_2 b_1 + b_2 d_1)y \\ y'' &= (c_2 a_1 + d_2 c_1)x + (c_2 b_1 + d_2 d_1)y \end{aligned}$$

Remember that we can write this equation in matrix form as well,

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} a_2 a_1 + b_2 c_1 & a_2 b_1 + b_2 d_1 \\ c_2 a_1 + d_2 c_1 & c_2 b_1 + d_2 d_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (6.7)$$

Take a look at equation 6.4. Notice that we can use equation 6.6 to write the RHS as,

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Now, compare this equation with equation 6.7. We discover that,

$$\begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} = \begin{pmatrix} a_2 a_1 + b_2 c_1 & a_2 b_1 + b_2 d_1 \\ c_2 a_1 + d_2 c_1 & c_2 b_1 + d_2 d_1 \end{pmatrix} \quad (6.8)$$

This is the rule for matrix multiplication!

Exercise 6.2: Commutativity of matrix multiplication

Now that you know the rule for multiplying matrices, let matrix $A = \begin{pmatrix} 2 & 3 \\ -4 & 16 \end{pmatrix}$ and matrix $B = \begin{pmatrix} 5 & 1 \\ 0 & 5 \end{pmatrix}$. Using equation 6.8, calculate $A \times B$ (also denoted simply as AB). Now, calculate BA . What do you notice?

We can also invert the relation between (x', y') and (x, y) , by writing the unprimed coor-

ordinates on the LHS. Consider once again equation 6.1. Show that,

$$\begin{aligned} x &= \left[\frac{d}{ad-bc} \right] x' + \left[\frac{-b}{ad-bc} \right] y' \\ y &= \left[\frac{-c}{ad-bc} \right] x' + \left[\frac{a}{ad-bc} \right] y' \end{aligned} \quad (6.9)$$

is the inverted form of equation 6.1. In matrix form,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{D} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \quad (6.10)$$

where the common factor, $D = ad - bc$ is known as the *determinant* of the matrix. If M is a matrix such that,

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (6.11)$$

then the inverse of the matrix, M^{-1} , is determined as,

$$M^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \quad (6.12)$$

To calculate the inverse of a 2×2 matrix then, follow these steps:

- Interchange the top-left and bottom-right element.
- Flip the sign of the top-right and bottom-left elements.
- Divide by the determinant, D .

Exercise 6.3: $x \rightarrow M$

i. Calculate the inverse of the following matrices:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 3 & 6 \\ 2 & 4 \end{pmatrix} \quad \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

ii. The inverse of the inverse of a matrix should give you the original matrix, i.e. $(M^{-1})^{-1} = M$. Let, $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, and verify this fact.

6.2 Graphs, derivatives, integrals

To study classical mechanics, we need to describe the motion of bodies in space. Suppose you are doing an experiment wherein you track the x -position of an object with time. So, you have a stopwatch and a ruler using which you take some observations. How can you organize this data? You can start by listing all the values in a tabular form like below,

time, t [s]	position, x [m]
0	1
1	2
3	4.6
4	5.4
5	3.8

Notice that this table is “ordered” in time, but nothing should stop you from ordering it with respect to positions. In principle, you can know everything about the motion – including the rate at which it is moving – if someone gives you this table. But in practice, we have a better way of visualizing motion: using graphs. The table above is represented using a graph as below.

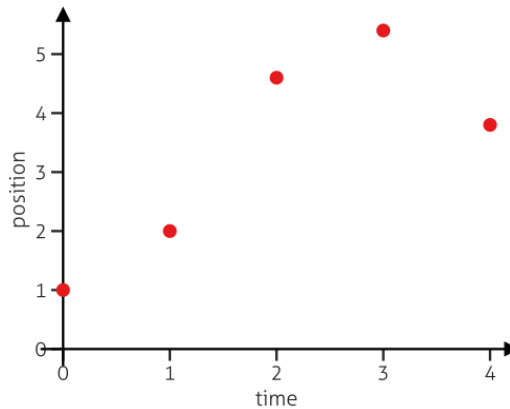


Figure 6.4: A graph depicting an object moving in one dimension

To find the rate of change of position with time, we can naively join the points with a straight line and take the slope between these points (figure 6.5). This slope – which is the ratio of position interval to time interval – is the velocity of the object. But, what time point should we assign to this velocity, since velocity is defined using *two* time points? For the first slope,

$$v(0 \text{ or } 1?) = \frac{\Delta x}{\Delta t} = \frac{x(1) - x(0)}{t(1) - t(0)} = \frac{2 - 1}{1 - 0} = 1 \text{ ms}^{-1}$$

By convention, we assign this to the time in between 0 and 1, which is 0.5 s.

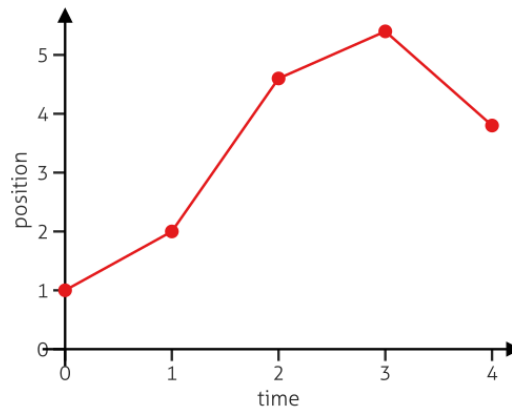


Figure 6.5: The slope of the line between the points is the velocity

Note that the velocity we calculate here is an *average* velocity. We do not know exactly what is happening in between the given times. So, how do we get velocities at a given time point? The simplest way is to reduce the interval between successive time points – instead of measuring the position after every second, measure it after every half a second. We can keep doing this by taking smaller and smaller time intervals, until we get to the limit where the time intervals approach zero. Hence, the velocity now become the limit:

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

It is convention to define Δx in terms of time intervals as $\Delta x = x(t + \Delta t) - x(t)$. The velocity now becomes,

$$v(t) = \lim_{\Delta t \rightarrow 0} \frac{x(t + \Delta t) - x(t)}{\Delta t} \quad (6.13)$$

This relation in a general form is what is called a *derivative*. So, in the limit of billions of entries in the table, i.e. billions of time and position measurements, we get a smooth line in our graph.

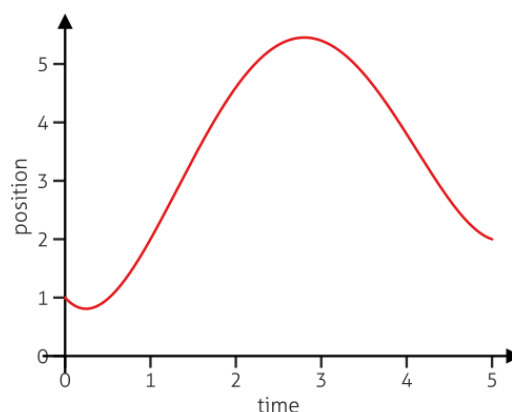


Figure 6.6: Position-time curve in the limit of billions of measurements

With this picture, the meaning of the velocity at a given time point is as follows – if you take a very small portion of the curve (if you zoom in enough), it looks like a straight line. The

slope of this straight line is the velocity at that point. This is called the *instantaneous velocity*. You can also think of the velocity at a time point as the tangent to the curve at that point.

Exercise 6.4: Curve surfing

Given the position-time graph, roughly sketch the velocity-time curve. (**Hint:** See how the slope of the tangents to the curve change as you go from the left of the curve to the right of the curve)

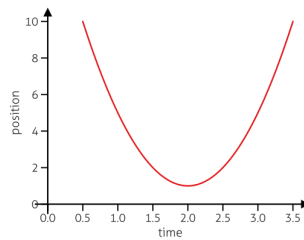


Figure 6.7

6.3 Circular Motion

Consider an object moving along a circle with a constant angular velocity, i.e., the object covers equal arc lengths in equal intervals of time. Mathematically: the angle subtended at the center of the circle by the object changes linearly with time, $\theta = \omega t$.

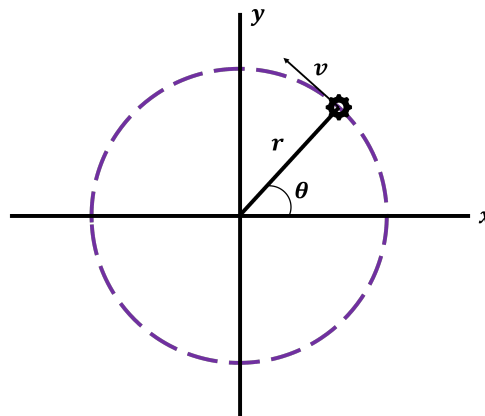


Figure 6.8

What is its acceleration under such motion? First we find what the constant speed, v is. To see this, consider a small change in angle, $\Delta\theta$.

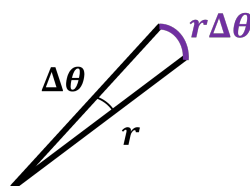


Figure 6.9

The arc length is thus, $\Delta l = r\Delta\theta$. For a tiny instant Δt , the speed is given by:

$$v = \frac{\Delta l}{\Delta t} = \frac{r\Delta\theta}{\Delta t} = r \left(\frac{\Delta\theta}{\Delta t} \right) = r\omega$$

To find the acceleration, we look at the x and y positions of the object in Cartesian coordinates. Here,

$$\begin{aligned} x &= r \cos \theta = r \cos \omega t \\ y &= r \sin \theta = r \sin \omega t \end{aligned}$$

Differentiating with respect to t ,

$$\begin{aligned} \dot{x} &= -r\omega \sin \omega t \\ \dot{y} &= r\omega \cos \omega t \end{aligned}$$

To get acceleration, differentiate once again,

$$\begin{aligned} \ddot{x} &= -r\omega^2 \cos \omega t = -\omega^2 x \\ \ddot{y} &= -r\omega^2 \sin \omega t = -\omega^2 y \end{aligned}$$

Finally, we go back to polar coordinates:

$$a = -\omega^2 r = -\frac{v^2}{r^2} r = -\frac{v^2}{r} \quad (6.14)$$

In terms of direction, we see from the equation above that \vec{a} points in the $(-\hat{r})$ direction, opposite to the radius vector, i.e., towards the center at all times. This is an important point about circular motion: no matter what causes it (gravity, a spring, a hamster etc.), the object is always accelerating with $a = -v^2/r$ pointing inward.

6.4 Vectors

A vector is a quantity with a magnitude and direction. It is basically a collection of multiple numbers which makes sense when given alongside an origin. We can describe a point in space using a vector known as the *displacement vector*, \vec{r} .

$$\vec{r} = (x, y, z) = x\hat{x} + y\hat{y} + z\hat{z}$$

Here, $\hat{x}, \hat{y}, \hat{z}$ are *unit vectors*: vectors of length 1 which point in the x, y, z directions respectively. All unit vectors are represented by writing a *hat* on top, \hat{u} . We can write any vector in terms of these unit vectors by providing the relevant scaling factors. In polar coordinates,

$$\vec{r} = r\hat{r}$$

Dot product ($\vec{r}_1 \cdot \vec{r}_2$): We define the dot or scalar product of two vectors as,

$$\vec{A} \cdot \vec{B} = A_x B_x + A_y B_y + A_z B_z$$

As the name suggests, the result of the dot product operation is a scalar. **Dot product of two radial vectors:** Consider the following two vectors,

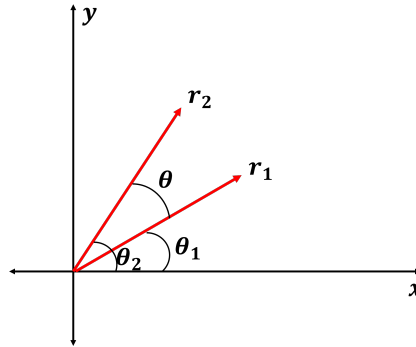


Figure 6.10

$$\vec{r}_1 \cdot \vec{r}_2 = r_{1x}r_{2x} + r_{1y}r_{2y} + r_{1z}r_{2z} = r_{1x}r_{2x} + r_{1y}r_{2y} + 0$$

In terms of their components,

$$r_{1x} = r_1 \cos \theta_1$$

$$r_{1y} = r_1 \sin \theta_1$$

Hence,

$$\begin{aligned} \vec{r}_1 \cdot \vec{r}_2 &= r_1 r_2 \cos \theta_1 \cos \theta_2 + r_1 r_2 \sin \theta_1 \sin \theta_2 \\ &= r_1 r_2 \cos(\theta_2 - \theta_1) \\ &= r_1 r_2 \cos \theta \end{aligned}$$

Orthogonal vectors are those which are perpendicular to each other. The above form tells us that the dot product of two orthogonal vectors is 0. In case of motion of an object around a circle, the radial vector, \vec{r} points radially outwards and the velocity vector, \vec{v} points in the $\hat{\theta}$ direction, i.e., always tangent to the circle. It follows that,

$$\vec{r} = r\hat{r}$$

$$\vec{v} = v\hat{\theta}$$

$$\vec{r} \cdot \vec{v} = 0$$

Cross product ($\vec{r}_1 \times \vec{r}_2$): We define the cross or vector product of two vectors as,

$$\vec{A} \times \vec{B} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} = \hat{x}(A_y B_z - A_z B_y) + \hat{y}(A_z B_x - A_x B_z) + \hat{z}(A_x B_y - A_y B_x)$$

As easy way to remember this rule is to notice the cyclic nature of the subscripts. In the special case where $A_z = B_z = 0$,

$$\vec{A} \times \vec{B} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ A_x & A_y & 0 \\ B_x & B_y & 0 \end{vmatrix} (A_x B_y - A_y B_x) \hat{z}$$

The result of a cross product is a vector which points in the mutually perpendicular direction of the two participating vectors. In polar coordinates,

$$A_x = A \cos \theta_1, A_y = A \sin \theta_1$$

$$B_x = B \cos \theta_2, B_y = B \sin \theta_2$$

Hence,

$$\begin{aligned}\vec{A} \times \vec{B} &= AB \cos \theta_1 \sin \theta_2 - AB \sin \theta_1 \cos \theta_2 \\ &= AB(\cos \theta_1 \sin \theta_2 - \sin \theta_1 \cos \theta_2) \\ &= AB \sin(\theta_2 - \theta_1) \hat{z}\end{aligned}$$

6.5 Relativistically moving infinite rods

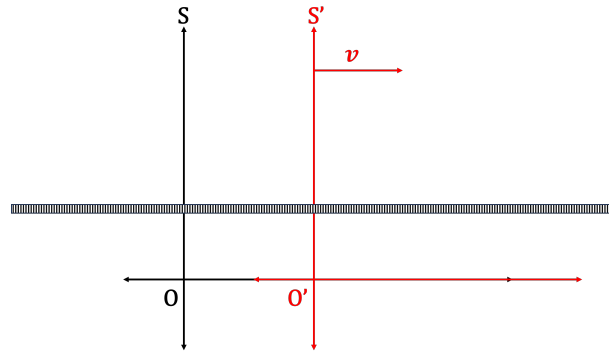


Figure 6.11

Let the linear density of notches in S be n . For observers in S' , the density of notches will be n' . According to the formula for length contraction,

$$n' = \gamma n = \frac{n}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Now, consider the following situation:

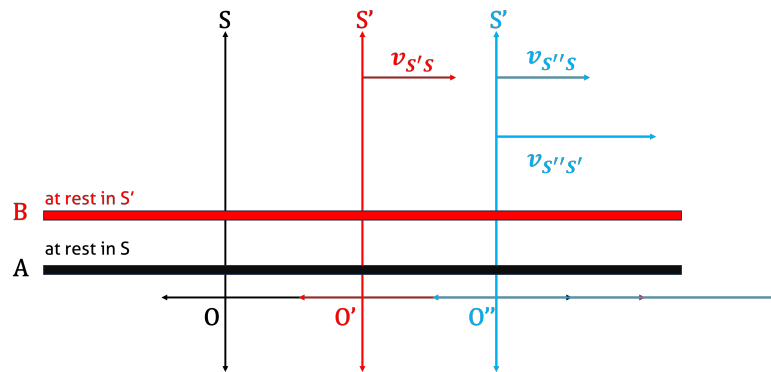


Figure 6.12

Here, n_{B0} is the number of notches per length on rod B as seen from frame S' and n_{A0} is the number of notches per length on rod A as seen from frame S. For observers in S' ,

$$\begin{aligned} n'_B &= n_{B0} \\ n'_A &= \gamma_{S'S} n_{A0} \end{aligned}$$

For observers in S'' ,

$$\begin{aligned} n''_A &= \gamma_{S''S} n_{A0} \\ n''_B &= \gamma_{S''S'} n_{B0} \end{aligned}$$

But remember that the rules of application of the Lorentz transformations when moving between frames in neither of which the object is at rest. Hence,

$$\gamma_{S''S} = \gamma_{S''S'} \gamma_{S'S} \left(1 + \frac{v_{S''S'} v_{SS'}}{c^2} \right)$$

Then,

$$\begin{aligned} n''_A &= \gamma_{S''S'} \gamma_{S'S} \left(1 + \frac{v_{S''S'} v_{SS'}}{c^2} \right) n_{A0} \\ &= \gamma_{S''S'} \left(1 + \frac{v_{S''S'} v_{SS'}}{c^2} \right) n'_A \end{aligned}$$