Contents

1 Introduction

Most of what we've talked about so far has been physics that was developed from the 17th to 19th centuries, like gravitation, EM, etc. At the end of the 19th century, some thought that there was nothing new to be discovered in physics, and all that was left was to obtain more precise measurements for values. However, some (like Lord Kelvin) believed that there were still unresolved problems, in this case the Michelson-Morley results and the black-body spectrum. The Michelson-Morley experiment tried to find the existence of the ether, and eventually led to Einstein reformulating physics with his postulates (inertial reference frames, constant speed of light). These have then led to the modern understanding of energy and relativity.

Black-body radiation diverged from classical postulates, and led to the thought of energy quantization, which then led to quantum mechanics. Quantum mechanics has now driven much of the technology in the modern age, such as MRIs, electron microscopy, and modern microcircuitry.

These discoveries led to a very fast expansion of physics in the 20th century, with the development of QM, QED, the Higgs mechanism, and QCD, as well as the discovery of many different particles. This culminated in the 1970's with the creation of the Standard Model of particle physics. This also provides us with precise estimates of values, which have been validated experimentally. These have now led to another dry spell, where we don't know what else is left. However, we still have several clouds, namely dark matter, dark energy, unifying GR and the SM, and the Hierarchy problem. These problems have remained at the forefront since the 1970's.

This course covers special relativity, and a small amount of general relativity, key concepts in thermodynamics, and an introduction to quantum mechanics, ending with the Schrodinger equation.

2 Index Manipulation

If we have a vector v, we have indices to represent each element:

$$v = \begin{bmatrix} 1 \\ 5 \\ -2 \end{bmatrix} \rightarrow v_1 = 1, v_2 = 5, v_3 = -2$$

If we have a 2D array (matrix), we use two indices:

$$A = \begin{bmatrix} 2 & 0 & -3 \\ 4 & -2 & 1 \\ 0 & 0 & -1 \end{bmatrix} \rightarrow A_{11} = 2, A_{23} = 1, A_{i3} = \begin{bmatrix} -3 \\ 1 \\ -1 \end{bmatrix}$$

For higher dimensions, we can use more indices, but we don't have a good way of representing them on paper, so we use slices of them instead.

Summations over indices can be very powerful, as they let us represent matrix multiplications and other operations compactly:

$$v_i = \sum_{j=1}^{3} A_{ij} w_j$$

This is just a matrix multiplied by a column vector. If we expand v_1:

$$v_i = A_{11} w_1 + A_{12} w_2 + A_{13} w_3$$

This is just the same as taking the first row of A and multiplying it by the vector $\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$. We can also use summations to write a row vector times a matrix:

$$v_i = \sum_{j=1}^{3} A_{ji} w_j$$

Similarly, we can do a matrix times a matrix:

$$C_{ij} = \sum_{k=1}^{3} A_{ik} B_{kj} = \sum_{k=1}^{3} B_{kj} A_{ik}$$

Note that the matrix with the second index summed over goes first. To avoid mistakes, remember that summed-over indices disappear on the left of the equality, and non-summed-over indices should match on both sides of the inequality. If we have summations and we want to multiply them:

$$a_i = \sum_{j=1}^{2} A_{ij} v_j \qquad b_i = \sum_{j=1}^{2} B_{ij} w_j$$

If we have that $B = A^T$, we know that the transpose just swaps the order of the indices, i.e. $B_{ij} = A^T_{ij} = A_{ji}$. Also note that the summed-over index can be any variable, because we know it disappears after we write out the summation, we can replace it with anything else we want, and the summations will be equivalent. When we multiply a_i and b_i, we need to be careful about using the same index for the summations, as they are independent, and we need to change one of them:

$$a_i b_i = \sum_{j=1}^{2} A_{ij} v_j \sum_{m=1}^{2} B_{im} w_m = \sum_{j=1}^{2} \sum_{m=1}^{2} A_{ij} v_j B_{im} w_m = A_{i1} v_1 B_{i1} w_1 + A_{i1} v_1 B_{i2} w_2 + A_{i2} v_2 B_{i1} w_1 + A_{i2} v_2 B_{i2} w_2$$

Moving on to vectors and rotation, we can prove that the scalar product is conserved under rotations. We start with the scalar product of the two rotated vectors, v' and w':

$$\vec{v}' \cdot \vec{w}' = \sum_i v'_i w'_i = \sum_{i,j,k} O_{ij} v_j O_{ik} w_k = \sum_{i,j,k} O^T_{ji} O_{ik} v_j w_k = \sum_{j,k} (O^T O)_{jk} v_j w_k = \sum_{j,k} \delta_{jk} b_j w_k = \sum_j v_j w_j$$

We have utilized the fact that $O^T = O^{-1}$, and also used the Kronecker delta, which is 0 except when $j = k$ (Note that its just the same as the identity matrix). We have seen that the rotated vectors have the same scalar product as the non-rotated vectors.

3 Galilean Transformations

Back in Galileo's time, it was not obvious whether the laws of nature that we saw on Earth would apply in other places, like the Moon. It was also unknown whether the fact that the Earth was moving ver fast would make a difference on the effects of the laws. Galileo stated the Galilean Relativity Statement.

Theorem 3.1. *"The laws of Motion have the same form in all inertial frames of reference."*

In this case, inertial frames of reference are frames that move at constant relative velocities.

The Galilean transformations rely on two axioms:

1. There exists an absolute space, shared by all frames.
2. All inertial frames share a universal time.

Under these axioms, we can then transform coordinates between two inertial frames (without rotation), S to S':

$$\vec{r}' = \vec{r}_0' + \vec{r} - \vec{u}t$$

Where $\vec{u}$ is the velocity of S' with respect to S. In the case where at $t = 0$, the origins are the same and the moving frame is moving in the direction of $\vec{x}$:

$$t' = t \qquad x' = x - ut \qquad y' = y \qquad z' = z$$

Under a Galilean transformation, the distance between two points is frame invariant, it doesn't change:

$$x_B' - x_A' = (x_B - ut) - (x_A - ut) = x_B - x_A$$

Basically, the length of objects look the same whether you are moving or stationary. Galilean transformations also have frame-dependent velocity:

$$\vec{v}' = \frac{d\vec{r}}{dt'} = \frac{d(\vec{r_0} + \vec{r} - \vec{u}t)}{dt}\frac{dt}{dt'} = \frac{d\vec{r}}{dt} - \vec{u} = \vec{v} - \vec{u}$$

And the acceleration is frame invariant:

$$\vec{a}' = \frac{d^2\vec{r}'}{dt'^2} = \frac{d(\vec{v} - \vec{u})}{dt} = \frac{d\vec{v}}{dt} = \vec{a}$$

This tells us that Newton's Second Law ($F = ma$) is frame invariant, since mass is also invariant.

Let's do an example. If we have a stationary observer and an observer moving at 60 miles per hour, and they observe a train moving with some position function, with some length. We can get the velocity and acceleration of the train via derivatives. To get the position of the train in the moving observer's point of view, we can simply subtract by the change in position of the moving observer. By the frame-invariant length, the moving observer keeps observing the same length of the train, and the velocity and acceleration can be determined via derivatives.

This is the world as we viewed it until the Michelson-Morley experiments.

4 Michelson-Morley Experiments

Maxwell's equations seemed applicable in any inertial reference frame (e.g. if you move a magnet closer to a wire loop vs. moving a wire loop closer to a magnet). However, Maxwell's equations are not invariant under Galilean transformations. For example, EM waves travel at the speed of light no matter what frame of reference you start from:

$$\nabla^2 E - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}E = 0$$

A proposed solution to this was the ether, and that the speed was so fast that moving would barely affect the speed of light. The Michelson interferometer was made to detect these very minor changes in the speed of light, due to the fact that the Earth was moving incredibly fast. The claim was that the Earth was moving about 30 kilometers per second with respect to the ether, and thus if a beam of light moves with the Earth, it will move at a different speed than if the light moved perpendicular to the Earth's motion through the ether.

Michelson invented the Michelson interferometer in 1881 as a way to measure the effect of the ether on the speed of light, and Morley helped improve the experiment in 1887. The interferometer split a beam of light into two directions, perpendicular to each other. These beams are then reflected back, and interfere, which is displayed on the screen. In theory, if the speed of light was affected by the ether, one of the beams would be changed slightly by the movement of the Earth through the ether. This would then cause some form of interference. The light travelling with the Earth would go a different distance than the light going in the direction perpendicular to the Earth. If the experiment was then rotated, they would expect a change in the fringes, but in reality, they observed little to no change in the pattern.

5 Special Relativity

In 1905, Einstein then stated his Relativity principle:

Theorem 5.1. *"All laws of Physics have the same form in all inertial frames of reference."*

Note that the difference between this and the Galilean statement is that the Galilean statement only applied to the laws of motion, whereas Einstein's statement applies to all laws of physics.

This then led to the postulate:

Theorem 5.2. *"The speed of light in vacuum has the same value c in every direction in all inertial frames of reference."*

His 1905 paper was the first complete account of special relativity, but elements had been developed by Poincare and Lorentz.

5.1 Time Dilation

If we have a clock moving with some velocity v, over some time period t, it travels a distance $\frac{vt}{2}$, and the light emitted off of it that bounces off of a plate a distance L away vertically will travel a distance $2\sqrt{L^2 + \frac{v^2t^2}{4}}$. This then gives us that the time measured in the reference frame of the stationary observer:

$$t = \frac{2L}{c}\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

If we then look at the reference frame of the moving clock, the total distance travelled by the light is just $2L$, giving us

$$t' = \frac{2L}{c}$$

We can then define a quantity γ:

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

And then giving a relationship between the two times:

$$t = t'\gamma$$

This tells us that time passes more slowly when moving. Note that we don't see this effect in our lives because the speed of light is incredibly fast (299,792,458 meters per second to be exact, because the meter is defined as the distance travelled by light in a vacuum in 1/299792458 of a second). Let's do an example. If Alice is at rest and Bob is moving at 30 km/h for an entire year, we see that β (normalized speed of Bob, defined by $\beta = \frac{v}{c}$) is roughly 2.7×10^{-8}. This gives us a change in time of roughly 12 nanoseconds. Even if you were on Apollo 11, moving at 40 thousand kilometers per hour, the net change in time is only 22 milliseconds. On the other hand, if we move at the speed of the protons in the LHC (299,792,455 m/s) at $\beta = .9999999$, we have a difference of 364 days, 22 hours, and 46 minutes, almost the full year.

5.2 Simultaneity

One effect of relativity is that simultaneity is frame dependent. If a light goes off in the middle of a moving train (looking at it from the rest frame), the ray of light emitted to the left hits the back of the train before the light emitted to the right hits the front of the train. However, if we look at it from the frame of the train, they both hit at the same time. This shows that whether two events that occurred far apart are simultaneous or not depends on the frame of reference.

5.3 Lorentz Contraction

If we once again look at a moving train, and a light turns on at the back of the train, in the rest frame the time it takes for it to reach to front is:

$$t_1 = \frac{L + vt_1}{c} \to t_1 = \frac{L}{c - v}$$

If it then bounces back:

$$t_2 = \frac{L - vt_2}{c} \to t_2 = \frac{L}{c + v}$$

This gets us a total time of

$$t = t_1 + t_2 = \frac{2L}{c} \frac{1}{1 - \frac{v^2}{c^2}}$$

If we now look at it from the train frame, we have that

$$t' = \frac{2L'}{c}$$

Using time dilation, we know that $t = t'\gamma$, giving us that

$$L = \frac{L'}{\gamma}$$

This gives us that moving objects are contracted along the direction of movement.

5.4 Relativity Paradox

If we have a 2 meter long horse and a 1.5 meter barn, can we fit the horse into the barn? Yes! We can just move the horse at roughly 86% the speed of light, and it now has a length of 1 meter, and can fit inside the barn. However, if we look at it from the point of view of the horse, the barn squeezes down, so how can the horse fit inside the barn? This provides a paradox, as according to relativity, both of these versions are true.

To take into account relativity's facets, we use space-time diagrams and Lorentz transform the points of interest.

5.5 Space-time Diagram

Space-time diagrams show space-time points (events at a specific time and location) and trajectories. Rays of light have slopes of 1, and massive objects have slopes of $\frac{c}{v} > 1$. Note that both axes measure distance, x and ct. The axes are set up in a way so that the fastest objects (light) have the lowest slope, and the slowest objects are vertical lines (slope of ∞). If we look at a 3D space-time diagram, each event has a light cone, one being the past lightcone and the other being the future lightcone. Events outside the lightcones can not affect the event.

Figure 1: Example space-time diagram with the trajectories of light particles.

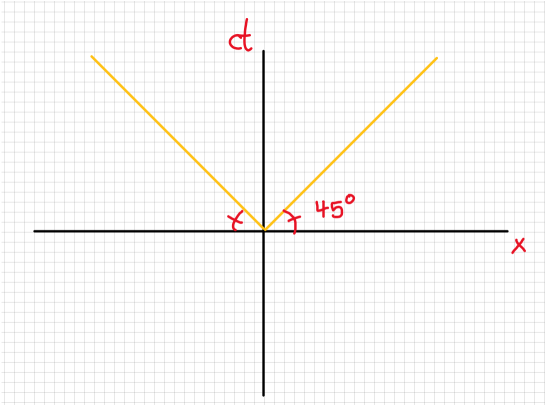
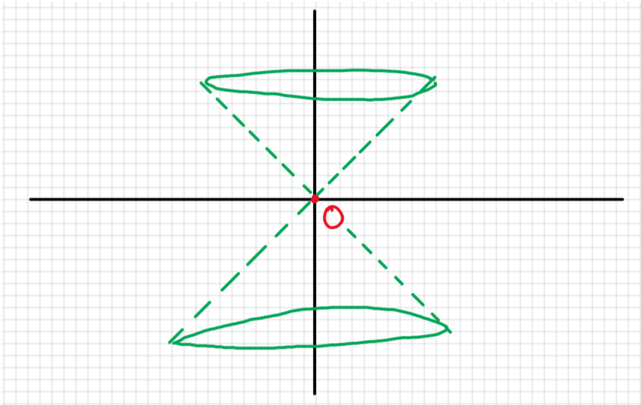


Figure 2: The past and future lightcones of an event at the origin.



We can add a new frame of reference by moving the axes. Remember that the points where $x' = 0$ define the t' axis. If we have two inertial frames in 3-space, both with the same origin but with S being stationary and S' moving with some velocity w in the x direction: To convert this to a space-time diagram, we have a new axis ct' with slope $\frac{c}{w}$. Essentially, we plot the trajectory of the origin. We know that the x' axis is the set of points with $t' = 0$, or points simultaneous with the origin O'. We can do this by shining rays of light in opposite directions, and if they bounce back on 2 events A and B the light rays meet back, A and B occur at the same time. This then gives us x', with slope $\frac{w}{c}$. Note that x' and ct' are not orthogonal axes, they can be oriented with any angle. We can then define a new quantity $\beta = \frac{w}{c}$.

Figure 3: Generating new axes ct' and x'.

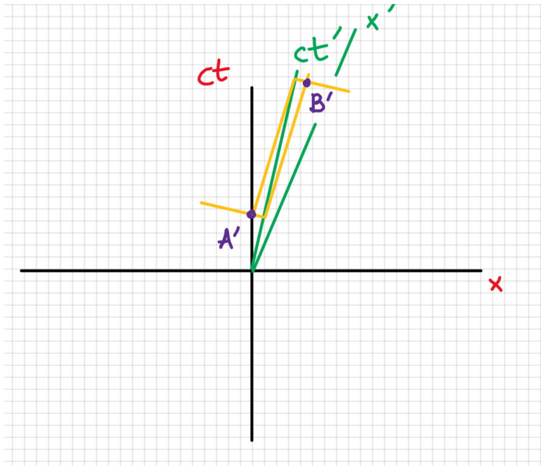
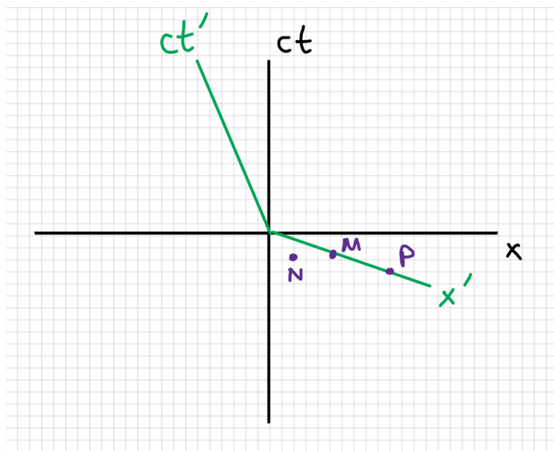


Figure 4: Looking at points from both the viewpoint of the S frame and the S' frame.



Looking at Figure 4, we can see 3 points, N, M, and P. From the perspective of the S frame, events N and M occur simultaneously, and P occurred before N and M. However, according to the S' frame, events M and P occur at the same time and event N occurs before M and P.

How are S and S' coordinates related? We know that the correspondence between them must be linear, as $t' = t\gamma$ and $L' = L\gamma$.

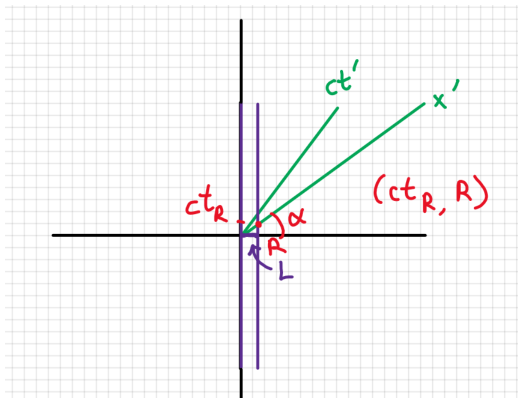
$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \Lambda \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} + \begin{bmatrix} ct'_0 \\ x'_0 \\ y'_0 \\ z'_0 \end{bmatrix}$$

We also know that it must preserve the speed of light, e.g. if an object is travelling at c in S, it must continue travelling at c in S'. We also know that $\Lambda(\beta) \cdot \Lambda(-\beta) = 1$, the inverse of a transformation must be a boost in the opposite direction. The Lorentz boost for two inertial frames with the same origin at $t = 0$ and S' moving at speed v along the x axis is

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix}$$

Let's do an example. If we have a rod of length L at rest in S, how long is it in S' moving at speed v with respect to S?

Figure 5: Determining the length of a rod in a moving frame of reference.



We define length as the distance between two points measured simultaneously. In the S' frame, the length of the rod is measured between the origin and point R. We have to Lorentz transform the points of interest, in this case the point is (ct_R, x_R):

$$\begin{bmatrix} ct'_R \\ x'_R \end{bmatrix} = \Lambda \begin{bmatrix} ct_R \\ x_R \end{bmatrix} = \begin{bmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{bmatrix} \begin{bmatrix} \beta L \\ L \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{L}{\gamma} \end{bmatrix}$$

We have found that $L' = \frac{L}{\gamma}$, the opposite of what we saw earlier, because the rod is now not moving in the S frame, and systematic relativity helps keep track of where the γ goes. Also note that we have a good cross check, ct'_R must be 0 for it to be simultaneous with O'. Note that to find the coordinates, we can just use simple geometry, for an arbitrary point A:

$$ct_A = \frac{1}{\beta} x_A$$

When trying to find distances of objects in the second frame, we must always Lorentz transform the ones that are in S'.

We can use these space-time diagrams to solve problems. To systematically solve problems of relativistic kinematics, we want to first draw the space time diagram, We want to draw the axes of the frames of interest, and draw the trajectories of the objects of interest. When we care about the length of an object, we want to draw the trajectory of both sides of the object. We then want to find the points that we care about, such as the position of the edges of an object at a given time when measuring length. Typically points will be the intersection of straight lines:

$$ct - ct_0 = \rho(x - x_0)$$

We then want to Lorentz transform these points of interest.

Let's look at the Doppler effect. The Doppler effect affects the change in frequency of a wave in relation to an observer who is moving relative to the wave source. This affects light as well.

Let's say we have a spaceship moving away from the Earth at some velocity v, and it sends a light signal to Earth every $\Delta t'$ as measured in the spaceship. How frequently do these signals arrive on Earth as measured in Earth time?

5.6 4-vectors

Euclidean space is an n-dimensional space with the usual scalar product. The norm of the vector is defined as $\vec{v}^2 = \vec{v} \cdot \vec{v}$. We know that rotating or translating vectors in frames have the same length.

However, this only works for space, not for space-time. If we include time, we have 4-vectors, objects with 4 components that Lorentz-transform. We use position 4-vectors x^μ to represent the distance between events (space-time points):

$$x^\mu = \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix}_\mu = \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix}_\mu$$

These are contravariant 4-vectors. We can write the Lorentz transform as

$$x'^\mu = \sum_{\nu=0}^{3} \Lambda^\mu_\nu x^\nu$$

Where Λ^μ_ν is just our regular Lorentz transform. However, because $\Lambda^T \neq \Lambda^{-1}$, we see that the standard length of vectors is not conserved under Lorentz transformations.

We define covariant 4-vectors with subscripts, and lower/raising indices operations with the Minkowski metric $g_{\mu\nu}$:

$$x_\mu = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x^0 \\ -x^1 \\ -x^2 \\ -x^3 \end{bmatrix}$$

Writing this in index expressions:

$$x_\mu = \sum_{\nu=0}^{3} g_{\mu\nu} x^\nu \qquad x^\mu = \sum_{\nu=0}^{3} g^{\mu\nu} x_\nu$$

Where we have

$$g_{\mu\nu} = g^{\mu\nu} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

We can raise/lower any index of any array:

$$A^\mu_\alpha = \sum_{\nu=0}^{3} g^{\mu\nu} A_{\nu\alpha} = [gA_{\nu\alpha}]_{\mu\alpha}$$

We can use this to convert some matrix $A_{\nu\alpha}$ into A^{μ}_{α} (raising one of the indices). If we want to raise the other index now, finding $A^{\mu\beta}$:

$$A^{\mu\beta} = \sum_{\alpha=0}^{3} A^{\mu}_{\alpha} g^{\beta\alpha} = [A^{\mu}_{\alpha} g]_{\mu\beta}$$

We can show that the product of contravariant and covariant 4-vectors is invariant under the Lorentz transformation. We can read contravariant 4-vectors directly off of space-time diagrams, and we have to apply the Minkowski metric to get the covariant vectors (in practice this is just negating the space terms):

$$x_\mu = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x^0 \\ -x^1 \\ -x^2 \\ -x^3 \end{bmatrix}$$

Thus we can basically read contravariant vectors off of a space-time diagram, and then negate the spacial components to get the covariant vector.

We can also prove that the scalar product between a contravariant 4-vector and a covariant 4-vector is invariant under the Lorentz transformation. This is useful to define the norm of a 4-vector, which is also Lorentz invariant:

$$\sum_{\mu} x_\mu x^\mu$$

However, note that this is not technically a norm since it can be negative and 0. If we want to find the length of a vector in a different frame, we can take the norm of the vector instead of having to Lorentz transform it.

The norm Δs^2 of a 4-vector connecting two space-time points is called space-time interval:

$$\Delta s^2 = \sum_{\nu=0}^{3} x'_\nu x'^\nu = \sum_{\nu=0}^{3} x_\nu x^\nu = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$$

5.7 Minkowski Space

Minkowski space is the combination of 3D Euclidean space plus time with the spacetime interval between any two events.

The space-time interval between two events can be different things. If the norm of a vector is greater than 0, the vector is time-like. This will arise when the slope is larger than 1. This indicates a possible causal relationship between the two events.

We can then have light-like vectors, vectors with norms equal to 0. These indicate that the two points are connected by a ray of light.

Finally, vectors with negative norms are space-like, and the ordering of the events depends on the frame of reference.

Let's talk about how simultaneity is not a very meaningful concept, because we can change our frame of reference. The star Betelgeuse dimmed last year. The star is roughly 600 light years away, and thus the star actually began to dim in 1420 AD. However, it could happen at different times based on how fast the observer is moving!

Let's do a couple of definitions. The proper time τ is the time elapsed by the traveler:

$$\tau = \frac{t}{\gamma} = \frac{\sqrt{norm}}{c}$$

Let's do an example of this. A person moves from O to $C(2,3)$. WE could go into the frame of reference of the traveller, and then do a Lorentz transform to find the time coordinate of the traveller. WE would need the point (which we have), and β, which we can see is $\frac{2}{3}$. We could compute γ, and then compute the Lorentz transformation, and we get $\begin{bmatrix}\sqrt{5}\\0\end{bmatrix}$, telling us that he aged $\sqrt{5}$ years. Instead of doing this, we could just take the norm of the vector. For time-like vectors, we just take the square root of the norm, which is just $\sqrt{5}$.

The norm of space-like vectors is a distance squared, with a minus sign. We can use the norm to find the distance between two events in the frame where they are simultaneous. This gets the proper length/distance.

The universe that we live in is a 4-dimensional space with time and Eucliden 3-dimensional space connected by the Minkowski metric. In Euclidean 3-d space, the space of vectors with the same norm trace out spheres, but in Minkowski space, 4-vectors with the same norm lie on hyperbolas. The hyperbola for a time of 5 years is of the form $c^2t^2 - x^2 = 5^2$.

In the case of the Minkowski space, the straight line connecting two events is not the shortest path between them, but rather the longest. If we have a point $C(5,4)$, and compute the norm of the vector from the origin, we see that the traveller ages 3 years. If we instead stayed on Earth for .5 years, and then travel a bit faster, we see that they age half a year on Earth, and then we have the vector $(4, 4.5)$, which has a norm of 2.06 light years. Adding the half year back, we see that the proper time of the nonstraight path ages only 3.56 years. For a general case, it can be proven that

$$\sum_i \frac{\Delta t_i}{\gamma_i} \leq \sum_i \Delta t_i = t_{tot}$$

Essentially, straight lines maximize the time of travel between two points.

5.8 Twin Paradox and Relativistic Dynamics

Since space-like events can be rearranged by changing frames, simultaneity is not a very meaningful concept. However, time-like events are well-ordered, so even if the simultaneity lines can jump around, the time-like points cannot be reordered.

Note that the urban legend that says special relativity is unable to deal with acceleration is false. General relativity is only needed in the presence of gravitation when space-time becomes curved. We can solve problems such as finding how long it takes for a spaceship with constant acceleration g to reach Pluto. We have that $v = gt = 9.8t$, and we have a distance to travel of 7.5 billion kilometers. We don't need relativity to solve this problem:

$$\frac{dx}{dt} = gt \to x = g\frac{t^2}{2} \to t = \sqrt{\frac{2x_P}{g}} = 343.66$$

If we look at how fast the spaceship is when it reaches Pluto, we can again solve this without looking at SR:

$$v_P = gt_P = .04c$$

We see that we should have some relativistic effects. If we want to calculate how much times passes from the perspective of the spaceship? To do this, we need to use special relativity. We care about τ_P, the proper time. We can look at it from the inertial reference frame of the Earth, and have that γ is a function of t:

$$\tau_P = \int_0^{t_P} \frac{dt}{\gamma(t)} = 343.58$$

We can see that they feel 5 minutes of time dilation.

We have talked in depth about relativistic kinematics, but how do we talk about relativistic dynamics?

When we talk about ordinary velocity, we define it as the derivative of the position:

$$\vec{v} = \frac{d\vec{x}}{dt}$$

We have seen the concept of proper time, the time elapsed by the traveller:

$$\tau = \frac{t}{\gamma(v)}$$

The 4-velocity is defined as the space travelled in some frame with respect to the proper time:

$$\eta^\mu = \frac{dx^\mu}{d\tau} = \gamma \frac{dx^\mu}{dt}$$

Since the 4-velocity is a 4-vector, we know that it follows Lorentz transformations. We can take an object moving at velocity v in a frame, and use Lorentz transformations to get the velocity of the object in another moving frame.

$$\eta^\mu = \begin{bmatrix} \gamma(v')c \\ \gamma(v')v'_x \\ \gamma(v')v'_y \\ \gamma(v')v'_z \end{bmatrix} = \begin{bmatrix} \gamma(w) & -\beta(w)\gamma(w) & 0 & 0 \\ -\beta(w)\gamma(w) & \gamma(w) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \gamma(v)c \\ \gamma(v)v_x \\ \gamma(v)v_y \\ \gamma(v)v_z \end{bmatrix}$$

Note that the norm of the velocity 4-vector η^μ is constant:

$$\sum_\nu \eta_\mu \eta^\mu = (\gamma c)^2 - (\gamma v_x)^2 - (\gamma v_y)^2 - (\gamma v_z)^2 = \gamma^2(c^2 - v_x^2 - v_y^2 - v_z^2) = c^2$$

Where we have used the fact that $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$.

We can use these Lorentz transformations to find how velocities transform, such as transformation along the direction of movement:

$$v'_x = \frac{v_x - w}{1 - \frac{wv_x}{c^2}}$$

and transformation perpendicular to the direction of movement:

$$v'_{y,z} = \frac{\frac{v_{y,z}}{\gamma(w)}}{1 - \frac{wv_x}{c^2}}$$

If we take the cases where v,w are both $\ll c$, we recover the Galilean transformations.

Let's do an example. If Superman leaves Earth at $.8c$ at a 45 degree angle, and Supergirl leaves Earth at $.9c$ at a 0 degree angle, what total speed and in which direction does Supergirl fly with respect to Superman?

We can orient the situation so that Superman is moving along the x-axis, giving us that $w = .8c$. We then take the component of Supergirl's velocity in the x direction, which is $.9c\cos 45$. We can then just plug them into the formulae for v'_x and $v'_{y,z}$:

$$v'_x = \frac{.9c\cos(45) - .8c}{1 - \frac{.9c\cos(45)\times .8c}{c^2}}$$

We could also compute v' via the identity $v' = \sqrt{v'^2_x + v'^2_y}$. To get the angle between Superman and Supergirl, we can just use arctan:

$$\theta = \arctan(v'_y, v'_x)$$

Let's say we have two clay balls of equal mass and opposite velocities moving towards each other. We expect them to have a net momentum of $p = mv - mv = 0$. If we look at the left ball's perspective, it is stationary, and the other ball is moving towards it at $\frac{-2v}{1+\frac{v^2}{c^2}}$. What happens to the momentum:

$$p = m \times 0 + m\frac{-2v}{1+\frac{v^2}{c^2}} = \frac{-2mv}{1+\frac{v^2}{c^2}}$$

We see that standard momentum is not conserved in relativity, we would have expected a momentum of $p = 2m(-v) = -2mv$. We introduce 4-momentum:

$$p^\mu = m\eta^\mu = (m\gamma c, m\gamma\vec{v})$$

This is a 4-vector, since m is invariant, and 4-velocity is a 4-vector. At low speeds, the spatial portion looks like standard momentum (Taylor expanding the denominator):

$$m\gamma\vec{v} = \frac{m\vec{v}}{\sqrt{1-\frac{v^2}{c^2}}} = m\vec{v} + m\vec{v}\frac{v^2}{2c^2} + \mathcal{O}\left(\frac{v^4}{c^4}\right)$$

We see that the first order term looks just like momentum.

At low speeds, the time part looks like the kinetic energy plus a constant:

$$m\gamma c = \frac{mc}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{1}{c}\left[mc^2 + m\frac{v^2}{2} + \mathcal{O}\left(\frac{v^4}{c^2}\right)\right]$$

Since it is a 4-vector, if it is conserved in one frame of reference, it is conserved in all frames (which we will not prove). The norm of then momentum 4-vector is always m^2c^2. As a result, if we have $p = |\vec{p}|$:

$$p^\mu = (\sqrt{p^2 + m^2c^2}, \vec{p})_\mu \to E^2 = p^2c^2 + m^2c^4$$

$$p^\nu = \begin{bmatrix} m\gamma c \\ m\gamma\vec{v} \end{bmatrix}_\nu = \begin{bmatrix} \frac{E}{c} \\ \vec{p} \end{bmatrix}_\nu = \begin{bmatrix} \sqrt{p^2 + m^2c^2} \\ \vec{p} \end{bmatrix}_\nu$$

with $\vec{p} = m\gamma\vec{v}$.

We can have multiple different metrics, such as the mostly minuses or West coast metric, $(+,-,-,-)$, which gives us that time-like vectors are positive, giving easier proper time, and positive 4-momentum norms. The other convention is the mostly pluses or East coast metric, $(-,+,+,+)$.

Newton's first law is still valid in relativistic mechanics when using the relativistic momentum:

$$\vec{F} = \frac{d\vec{p}}{dt} = m\frac{d(\gamma\vec{v})}{dt}$$

For example, if we want astronauts in a spaceship to feel a constant force $F = mg$, Newton tells us that

$$\frac{dp}{dt} = F \to p = mv = Ft$$

In the case of relativity:

$$\frac{dp}{dt} = F \to p = \frac{mv}{\sqrt{1-\frac{v^2}{c^2}}} = Ft$$

Let's apply conservation of 4-momentum. Let's look at Compton scattering. A photon γ of energy E_0 bounces off an electron of mass m_e at rest. What would the energy of the outgoing photon be? We know that

$$p^\nu = \begin{bmatrix} \frac{E}{c} \\ \vec{p} \end{bmatrix}$$

We can write the initial 4-momentums for the photon and the electron:

$$\left(\sqrt{p_0^2 + m^2c^2}, \vec{p}_0\right) = (p_0, p_0, 0, 0) = (\frac{E_0}{c}, \frac{E_0}{c}, 0, 0)$$

The electron is at rest initially:

$$(\sqrt{p_0^2 + m_e^2c^2}, \vec{p}_0) = (m_e c, 0, 0, 0)$$

And now looking at the post-collision momentums:

$$\left(\frac{E}{c}, \frac{E}{c}\cos\theta, \frac{E}{c}\sin\theta, 0\right)$$

$$\left(\sqrt{p_e^2 + m_e^2c^2}, p_e\cos\phi, -p_e\sin\phi, 0\right)$$

We now use the fact that the 4-momentum is conserved, and so every component of the 4-momentum is conserved. We can then sum up the components of the photon momentum and the electron momentum on both sides, and set them equal, giving us 4 equations.

Typically in particle physics we measure with electronvolts (eV) to measure energy. This is the energy of an electorn accelerated by a 1 Volt potential. We can then use eV/c for momentum and eV/c^2 for mass. If want to find the energy of an electron with $m_e = 511\ keV/c^2$ that has a momentum of 1 MeV/c?

$$E_e = \sqrt{p_e^2c^2 + m_e^2c^4} = 1.12\,MeV$$

Particle masses can cover a large range, from massless (photon/gluon), up to the GeV range (Higgs is $125\,GeV$).

6 General Relativity

Charge quantifies how strongly an object interacts with standard forces gien a field. For example, if we have an electric field, there is an effect on a particle of charge q. We can extend this to gravity, where if we have a gravitational field $\vec{\nabla}\Phi$ acting on a particle of gravitational charge (mass) m_g, we have some force:

$$\vec{F}_g = m_g \vec{\nabla}\Phi$$

Inertial mass quantifies how much an object is accelerated given a force:

$$\vec{a} = \frac{\vec{F}}{m_i}$$

Essentially, we can think of this as trying to push a feather versus pushing an elephant. Gravitational charge is relative to only the gravitational force, but the inertial mass is with repect to any force, not just gravitation.

If we have an electric field, the force/movement on particles is proportional to their electric charge. If we have a gravitational field, it turns out that the acceleration is independent of the gravitational charge of the object (because we have a $\frac{m_g}{m_i}$ and that is 1). This is not a coincidence, and Einstein put forward his Equivalence principle in 1907.

Theorem 6.1. *It is impossible to detect the existence of a gravitational field by means of local experiments.*

This suggests that gravitation is something intrinsic to the fabric of space-time, and that it affects even objects without mass. We essentially stopped considering gravitation as a force, and instead consider the fabric of space-time is bent, and it affects all objects without taking mass into account. We have gotten rid of the existence of m_g, because we don't consider gravity to be a field.

General relativity uses Riemannian geometry, which is beyond the scope of this course. We use geometry without the fifth Euclidean postulate (parallel postulate).

6.1 Sets and maps

Sets are just collections of objects. Maps associated elements of two sets with each other. $\phi : P \to B$ defines a map ϕ that maps elements from P to elements of B. $\mathbb{R}^n$ is the set of ntuples $x = (x_1, x_2, \ldots, x_n)$. It is easy to define a norm:

$$|\vec{x} - \vec{y}| = \sqrt{\sum_i (x_i - y_i)^2}$$

We like this norm because it lets us define out calculus on it (limits, derivatives, integrals, etc). However, our universe is not $\mathbb{R}^n$. We use charts, which are continuous, invertible maps betweem a set M and $\mathbb{R}^n$. These charts give us ways to transform coordinates from one space to another. An analogy for this is the Mercator projection, which takes a 3-dimensional sphere to a 2 dimensional space. In summary, a manifold is a set M of objects, such as the points on the surface of a sphere, or a torus. Invertible, differentiable charts $\phi : M \to \mathbb{R}^n$ that cover all of M are called an Atlas.

Our contravariant 4-vectors are just vectors. Vectors are actually directional derivatives at a space-point X. They form a linear space known as the Tangent space T_X. This is essentially the space formed by the plane tangent to the manifold. Our covariant 4-vectors are called dual vectors, and they are linear maps (functions) that take vectors from the tangent space T_X and return a number:

$$w : T_X M \to \mathcal{R}$$

6.2 Tensors

A tensor T of rank (k, l) is a multilinear map from a collection of dual vectors and vectors to $\mathbb{R}$:

$$T : T_p^* \times \cdots \times T_p^* \times T_P \times \cdots \times T_p \to \mathbb{R}$$

A dual vector is a tensor of rank $(1, 0)$:

$$w_\mu : T_p M \to \mathbb{R}$$

A vector is a tensor of rank $(0, 1)$, given the scalar product:

The metric is a tensor of rank $(2, 0)$:

$$g(V, W)$$

In general relativity, the energy-momentum tensor $T^{\mu\nu}$ is key. This is the flux of 4-momentum p^μ across a surface of constant x^μ.

6.3 The Metric

The metric is at the heart of GR. The metric allows us to compute the path length and the proper time, with the space-time interval:

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu$$

It lets us determine the "shortest distance" between two points. These are the trajectories of free falling objects, known as geodesics. The metric also gives a notion of the "past" and the "future", due to the $(+, -, -, -)$ signature. The metric also tells us the maximum speed possible, and whether any two points can be in a causal relationship. The metric tensor $g^{\mu\nu}$ replaces the Newtonian gravitational field Φ, and defines the curvature of space-time.

The metric is typically written in terms of the space time interval:

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu$$

In special relativity, in flat space, the metric is

$$g_{\mu\nu}^{SR} = \begin{bmatrix} c^2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

And the metric around stars and black holes is known as the Schwarzchild metric.

We can compute the proper time along time-like trajectories:

$$d\tau = \frac{\sqrt{ds^2}}{c} = \sqrt{\frac{g_{\mu\nu}}{c^2} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda$$

We can use this for SR, such as finding the proper time of a particle moving along the trajectory $t = \frac{3x}{c}$ between $x = 10$ and $x = 100$. We know that the SR metric is $ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$. So we can write out the proper time:

$$\tau = \int_{x_1}^{x_2} \sqrt{\frac{g_{\mu\nu}}{c^2}\frac{dx^\mu}{dx^1}\frac{dx^\nu}{dx^1}}dx^1 = \int_{x_1}^{x_2} \sqrt{c^2(\frac{dt}{dx})^2 - 1(\frac{dx}{dx})^2 - 1(\frac{dy}{dx})^2 - 1(\frac{dz}{dx})^2\frac{dx}{c^2}}$$

We can compute all the derivatives:

$$= \int_{x_1}^{x_2} \sqrt{\frac{8}{c^2}}dx = \sqrt{8}\frac{x_2 - x_1}{c}$$

We can also find the trajectories of light. We know that light always has a space-time interval of 0:

$$ds^2 = 0 = c^2dt^2 - dx^2 \rightarrow \frac{dx}{dt} = \sqrt{c^2} = c$$

The derivative of a vector is coordinate dependent, which is not what we want. Instead, we use covariant derivatives, which as coordinate independent:

$$\nabla_\mu V^\nu = \partial_\mu V^\nu +$$

This leads into Geodesics. These are essentially paths along which the tangent vector is invariant. In flat space, the straight line keeps the tangent vector invariant, hence why they are the fastest path. In curved space however, this is not true. In Newton's world, free falling particles follow straight lines, which are the geodesics of Newton's world. In GR, free falling particles follow geodesics. Even light has to follow geodesics. Geodesics locally maximize the proper time of the object. These geodesics are the reason we see gravitational lensing such as Einstein rings and Einstein crosses.

Curved manifolds means that things don't always work out as nicely as flat space, such as angles of a triangle adding up to 180 degrees, or two points only having 1 geodesic between them.

The Riemann tensor tells us how a vector transforms as you transport it around an infinitesimal loop, and is 0 in flat space. Using this tensor, we can define the Ricci tensor $R_{\mu\nu} = R^\lambda_{\mu\lambda\nu'}$ and the Ricci scalar, $R = R^\mu_\mu = g^{\mu\nu}R_{\mu\nu}$.

7 Thermodynamics

Thermodynamics is not technically a modern branch of physics, but it is a natural segue into quantum mechanics (with the equipartition theorem). We will be looking at the laws of thermodynamics, as well as their applications to heat engines. Statistical mechanics is a m,ore fundamental study than thermodynamics, based on statistical methods and probability theory. Thermodynamics describes the transfer of heat and work based on macroscopic properties of systems, rather than microscopic levels. Thermo is to Stat Mech like Maxwell's E&M is to QFT. QFT is far more fundamental and intricate, but Maxwell's E&M is perfectly valid and useful.

The thermodynamical limit is the limit for systems with a large number of degrees of freedom, in which fluctuations due to the microscopic particles in a medium can be ignored. For example, if we have a system with $N_A = 6.022 \times 10^{23}$ particles, we don't have to worry about slight deviations, and we can just talk about macroscopic variables such as T or p.

A microstate is a different configuration of all the microscopic degrees of freedom. A macrostate is a description of the system based on macroscopic properties. For example, if we have 4 coins, and we shake the box, there are $2^4 = 16$ possible configurations for the microstates. A macrostate could be the total number of heads, or the total number of tails. We see that there is only 1 microstate that corresponds to the macrostate with all 4 being heads, and we can group microstates under which macrostates they represent.

7.1 Zeroth law of Thermodynamics

When cold and warm bodies are placed in thermal contact, energy starts flowing between them. Experiments show that energy flows from the warm body to the cold body. We also know that two bodies are in thermal equilibrium when they are in thermal contact but the energy content and temperature do not change.

Theorem 7.1. *Two systems, each separately in thermal equilibrium with a third, are in equilibrium with each other.*

This is a pretty obvious law, and it is essentially a transitive property.

7.2 Temperature

Imagine two systems with total energy E_1 and E_2, and they can exchange energy between them but not with the outside. The total energy in the system is $E_{tot} = E_1 + E_2$, and is constant. The number of microstates for each one is $\Omega_1(E_1)$ and $\Omega_2(E_2)$. The total number of microstates for the total system will be $\Omega_1(E_1)\Omega_2(E_2)$. This is simply the product rule from combinatorics.

In equilibrium and the thermodynamic limit, the system will choose the macrostate with the largest number of microstates (think about flipping 100 coins, the macrostate with 50 coins will be the most seen):

$$\frac{d}{dE_1}\left(\Omega_1(E_1)\Omega_2(E_2)\right) = 0$$

Using the product and chain rules:

$$\Omega_2(E_2)\frac{d\Omega_1(E_1)}{dE_1} + \Omega_1(E_1)\frac{d\Omega_2(E_2)}{dE_2}\frac{dE_2}{dE_1} = 0$$

We know that the total energy is constant:

$$E = E_1 + E_2 \to dE_1 = -dE_2 \to \frac{dE_2}{dE_1} = -1$$

Inserting this and dividing by $\Omega_1(E_1)\Omega_2(E_2)$, we have that

$$\frac{d\ln\Omega_1}{dE_1} = \frac{d\ln\Omega_2}{dE_2}$$

This is the thermal equilibrium condition, and it allows us to define temperature:

$$\frac{1}{k_B T} = \frac{d\ln\Omega}{dE}$$

7.3 Variables

Some macroscopic variables that we will use for our systems are volume V, particle number N, internal energy U, temperature T, and pressure p. These can be divided into two groups, extensive variables, variables that scale with system size, such as V,U, and N, and intensive variables, which are independent of the system size, such as T and p.

7.4 Ideal Gas

Thermodynamic variables can be constrained by equations of state, such as the most famous one, the ideal gas law:

$$pV = Nk_BT$$

Typically, we write this in terms of the number of moles n and $R = N_Ak_b$:

$$pV = nRT$$

The properties of the ideal gas are that there are no interactions between the gas molecules, and the molecules themselves take up zero volume. In practice, this works for gases under a wide range of conditions, but breaks down at low temperatures and high pressures. The internal energy of an ideal gas is given via

$$U = \frac{3}{2}nRT$$

The ideal gas can only store energy as the kinetic energy of its molecules (which is proportional to T), because the molecules cannot interact. The Van der Waals equation adds two terms that corrects for molecular attraction and the volume of molecules.

7.5 Heat and work

Heat is thermal energy in transit. Like other forms of energy, it is measured in joules. Since it is in transit, you can add heat to an object, but cannot say that the object has heat. Once the transit stops, it becomes a different kind of energy, such as internal energy.

Similarly work is mechanically energy in transit. Work from a force F moving an object a distance Δx is $W = F \cdot \Delta x = p \cdot \Delta V$. Work from n electrical current I across a voltage V_E in a time t is $W = V_E \cdot I \cdot t$.

If we compare warming up your hand by rubbing therm versus setting them on fire, rubbing them is doing work, since we're exerting a force when we rub our hands, and setting them on fire is heat.

Let's now define heat capacity. This is a terrible name, since heat cannot be stored. Heat capacity quantifies how much heat needs to be supplied to an object to raise its temperature by a given amount:

$$C = \frac{dQ}{dT}$$

Heat capacity is measured in joules per kelvin. This is extensive, as it depends on the size of the object (We can tell by looking at the units). Reservoirs are systems that have ∞ heat capacity, so their temperature does not change:

$$dT_{res} = \frac{dQ}{C_{res}} = \frac{dQ}{\infty} = 0$$

We can also define the intensive heat capacity, which is per unit of mass:

$$c = \frac{dQ}{dT}\frac{1}{m} \qquad [c] = JK^{-1}kg^{-1}$$

There are also volumetric and molar heat capacity, but we don't use those that much.

Let's look at water. Liquid water has more heat capacity than ice, because the crystal lattice of ice has fewer degrees of freedom, and thus the heat capacity is lower in ice.

We will have two types of heat capacity, C_V and C_p, which are the heat capacity at a constant volume and the heat capacity at a constant pressure:

$$C_V = \left(\frac{\partial Q}{\partial T}\right)_V \qquad C_p = \left(\frac{\partial Q}{\partial T}\right)_p$$

We always have $C_p > C_V$, because at constant perssure the gase does work against the piston.

7.6 Functions of state

A function of state is any macroscopic observable property that has a definite time-independent value in equilibrium. Examples of these are the internal energy, the pressure, and the temperature. Things that are not functions of state are heat, work, the positions of molecules. Heat and work are not functions of state because they have no meaning in equilibrium, and the position of molecules are not macroscopic observables.

One key fact about functions of state is that the value in equilibrium is independent of the path taken to get to that state:

$$\Delta f = \int_{x_1}^{x_2} df = f(x_2) - f(x_1)$$

This leads us to equations of state, such as the ideal gas equation:

$$f(p, V, T, N) = fV - Nk_BT = 0$$

We can then use differentials:

$$df = pdV + Vdp = 0$$

for constant T and N.

Let's look at an abstract function of state. If we have a system with two properties x and y, and a function of state $df = y\,dx + x\,dy$, does a difference in f between $(0, 0)$ and $(1, 1)$ depend on the path?

$$\Delta f = \int_{(0,0)}^{(1,1)} df = \int_{(0,0)}^{(1,1)} (ydx + xdy) = [xy]_{(0,0)}^{(1,1)} = 1$$

We see that the path does not matter. If we had something like $dg = y\,dx$, we can show that two separate paths returns different values, and thus g is not a function of state.

We have claimed that Wrok and heat are not functions of state:

$$\partial W = -pdV \qquad dQ = TdS$$

We can show that these are inexact differentials, and thus we prove that they are not functions of state.

Internal energy U is the sum of the energy of all the internal degrees of freedom a system has. This is a function of state, and has a definite value in equilibrium. This can have contributions from a bunch of different types of energy, such as kinetic, rotational, etc.

7.7 First Law of Thermodynamics

The first law states that

$$dU = \partial Q + \partial W$$

This is interesting, because we're claiming that a function of state can be written as the sum of two non-functions of state. This is essentially conservation of energy when we only look at thermal and mechanical energy.

If we have an ideal gas, we want to find U and W at $A(1,1)$ and $B(2,2)$, where the x-axis plots V and the y-axis plots p. We have that $pV = nRT$, and $U = \frac{3}{2}nRT$, so for $(1,1)$ we have that $pV = 1 = nRT$, and thus $U = \frac{3}{2}$. At $(2,2)$, we have $U = 6$. We can't find the work, because it is not a function of state, and it is only defined along a certain path, such as the paths between A and B. If we take the 90 degree paths between A and B, what is the work for those paths? We can just use the fact that

$$\partial W = -pdV$$

And if we then were asked to get the heat, we could use the first law, by computing $U_B - U_A$ and setting that equal to $\partial Q + \partial W$, where we have computed ∂W for that path.

In general, the internal energy will be a function of T and V:

$$dU = \left(\frac{\partial U}{\partial T}\right)_V dT + \left(\frac{\partial U}{\partial V}\right)_T dV$$

Using the first law:

$$\partial Q = dU - dW = dU + pdV$$

Dividing by dT and doing stuff out, we find the heat capacity at constant V:

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V = \left(\frac{dU}{dT}\right)_V$$

And the heat capacity at a constant p:

$$C_p = \left(\frac{\partial Q}{\partial T}\right)_p = C_V + \left[\left(\frac{\partial U}{\partial V}\right)_T + p\right]\left(\frac{\partial V}{\partial T}\right)_p$$

For an ideal gas with $pV = nRT$ and $U = \frac{3}{2}nRT$:

$$C_V = \frac{3}{2}nR \qquad C_p = \frac{5}{2}nR$$

7.8 Reversibility

The laws of physics are reversible, which means that if a process is allowed, the time-reversed process is also allowed. However, in real life, many porcesses seem irreversible, such as a falling brick hitting the ground and dissipating heat and sound. These are possible to reverse, but as we saw with macrostates, some are far more probable than others, so reversing these processes is almost impossible, but technically possible.

Generally, when gases are expanded/compressed, energy can be irreversibly converted into heat. It is possible to perform them reversibly, if done slow enough, so that the gas remains in equilibrium during the whole process. Essentially, we semalessly move from one equilibrium point to another, and we can then absorb and emit heat reversibly.

The isothermal expansion is an expansions that is reversible and keeps a constant temperature. This can be done by having the system in contact with a reservoir, and expanding the gas slowly. For an ideal gas, $U = \frac{3}{2}nRT$, and thus $\Delta U = 0$. By the first law:

$$\Delta U = \partial W + \partial Q \to \partial W = -\partial Q$$

We can integrate the changes in heat:

$$\Delta Q = \int \partial Q = -\int \partial W = \int_{V_1}^{V_2} pdV = \int_{V_1}^{V_2} \frac{nRT}{V}\, dV = nRT \ln \frac{V_2}{V_1}$$

The heat is positive, and thus heat is added to the gas.

We also have adiabatic expansions, where they are reversible and adiathermal, which means there is no flow of heat. This is possible with the system being insulated, and since $\partial Q = 0$, we have that $dU = \partial W$. For an adiabatic process, we define the adiabatic index as

$$\gamma = \frac{C_p}{C_V} = 1 + \frac{nR}{C_V}$$

And we have the relationship that pV^γ is a constant.

7.9 Heat Engines and Pumps

Heat engines and heat pumps are devices that operate thermodynamic cycles. Heat engines provide work by extracting heat from a hot reservoir. The efficiency of a heat engine is defined as $\eta = \frac{W}{Q_h}$, where Q_h is the heat provided. Heat pumps are powered by work,m and extract heat from an object. The efficiency is typically called the coefficient of performance.

So far, we have seen the convention that $Q > 0$ means that we are adding heat to the system, and $W > 0$ means that the work is being done on the system (Both of these come from the first law). When we draw diagrams, the sign of a variable will be based on the direction that the arrows are drawn.

7.9.1 Carnot Engine

Heat engines continuously deliver work, and operate in a cyclic process. The Carnot cycle follows two reversible isotherms (AB and CD) and two adiabats (BC and DA).

The heat entering is defined as

$$Q_h = nRT_h \ln \frac{V_B}{V_A}$$

And the heat leaving the system:

$$Q_c = -nRT_c \ln \frac{V_D}{V_C}$$

For the adiabats:

$$\frac{T_h}{T_c} = \left(\frac{V_C}{V_B}\right)^{\gamma-1}$$

$$\frac{T_c}{T_h} = \left(\frac{V_A}{V_D}\right)^{\gamma-1}$$

From these we find that $\frac{V_B}{V_A} = \frac{V_C}{V_D}$. Putting these together, we find that

$$\frac{Q_h}{Q_c} = \frac{T_h}{T_c}$$

The efficiency of the engine is

$$\eta = \frac{W}{Q_h} = \frac{Q_h - Q_c}{Q_h} = 1 - \frac{T_c}{T_h}$$

7.10 Second Law of Thermodynamics

The second law can be expressed in multiple different ways. Rudolf Clausius posited that "No process is possible whose sole result is the transfer of heat from a colder to a hotter body."

Sadi Carnot stated that "Of all the heat engines working between two given temperatures, none is more efficient than a Carnot Engine."

Lord Kelvin stated that "No process is possible whose sole result is the complete conversion of heat into work."

It may not seem clear why these three statements are all equivalent to each other, but they are, as we will show. Lets prove that Carnot is the same as Clausius. We need to prove that if one is violated, the other is also violated. Imagine we have an engine E with $\eta_N > \eta_{Carnot}$ (violating Carnot's statement). If this was the case, we could connect it to a reversed Carnot engine C. This engine will have efficiency η_C:

$$\eta_E > \eta_C \to \frac{W}{Q'_h} > \frac{W}{Q_h} \to Q_h > Q'_h$$

If we look at the total system now, we have from the first law that the net work done is that we move heat from the cold place to the hot place, with no other net work, violating Clausius' statement. This whole argument only hinges on the fact that the Carnot engine is reversible. Doing the violation the other way around is trivial. A corollary of this is that all reversible engines between two temperatures have efficiency η_{Carnot}.

Now if we want to equate Clausius and Kelvin, we start with violating Kelvin's statement, and we can connect it to a reversed Carnot engine, and the net effect is to take Q_c and dump it into T_h, which violates Clausius. Doing the other way around, we can connect it to a Carnot engine and the net effect is that we convert it all to work, which violates Kelvin's claim.

Thus we have proven that all 3 statements are logically equivalent.

If we have two bodies with T-independent heat capacities C_h and C_c. If we use a Carnot engine, what is the maximum work attainable?

We know that heat capacities are defined as $\partial Q_h = -C_h dT_h$, and $\partial Q_c = C_c dT_c$. Note that there is a minus for C_h because from the view of the hot object, heat is leaving the object.

We know that the Carnot efficiency tells us that

$$\frac{\partial Q_h}{T_h} = \frac{\partial Q_c}{T_c}$$

We can insert the definitions of the heat capacities, and we can then integrate:

$$-\int_{T_h}^{T_f} \frac{C_h dT_h}{T_h} = \int_{T_c}^{T_F} \frac{C_c dT_c}{T_c}$$

Doing this out and using $\Delta Q = C\Delta T$:

$$\Delta W = \Delta Q_h - \Delta Q_c = C_h T_h + C_c T_c - (C_h + C_c)T_f$$

We have seen that for a Carnot Engine:

$$\frac{Q_h}{Q_c} = \frac{T_h}{T_c} \to \frac{Q_h}{T_h} + \frac{-Q_c}{T_c} = 0$$

in general, for a reversible cycles:

$$\sum_{cycle} \frac{\Delta Q_{rev}}{T} = 0$$

Imagine a general non-reversible cycle that for each small interval receives heat ∂Q_i from a resrvoir at T_i. Since the internal energy is a function of state, we know that $\Delta U = 0$, and thus $\Delta W = \sum_{cycle} \partial Q_i$. Note that ΔW is positive because we are getting the work leaving the system, not the work being done on the system. If we now add a Carnot engine providing ∂Q_i:

$$\frac{\partial Q_i}{T_i} = \frac{\partial Q_i + \partial W_i}{T} \to \partial W_i = \partial Q_i \left(\frac{T}{T_i} - 1\right)$$

If we now look at the global system, the total amount of work released is $\Delta W + \sum \partial W_i$:

$$0 \geq \Delta W + \sum \partial W_i = \sum \partial Q_i + \sum \partial Q_i \left(\frac{T}{T_i} - 1\right) = T \sum \frac{\partial Q_i}{T_i}$$

Since $T > 0$:

$$\oint \frac{\partial Q}{T} \leq 0$$

This is Clausius' theorem.

Let's do an example with a reversible engine:

$$\oint \frac{\partial Q}{T} = \frac{Q_h}{T_h} - \frac{Q_c}{T_c} = 0$$

For irreversible processes we have that

$$\frac{Q_h}{T_h} - \frac{Q_c}{T_c} < 0$$

This leads to Clausius' statement if we remove the work.

7.11 Entropy

If we replace the two reservoirs in a reversible Carnot cycle with a consecutive succession of reservoirs:

$$\frac{Q_h}{T_h} = \frac{Q_c}{T_c} \to \sum \frac{\Delta Q_{rev}}{T} = 0 \to \oint \frac{\partial Q_{rev}}{T} = 0$$

This lets us define a function of state, as the integral of the quantity is path independent. This function of state

$$dS = \frac{\partial Q_{rev}}{T}$$

is known as entropy.

$$S(B) - S(A) = \int_A^B \frac{\partial Q_{rev}}{T}$$

What is the entropy change after an irreversible process. If we have a cycle with an irreversible process, by Clausius' theorem:

$$\oint \frac{\partial Q}{T} \leq 0$$

Separating the reversible and irreversible paths:

$$\int_A^B \frac{\partial Q}{T} + \int_B^A \frac{\partial Q_{rev}}{T} \leq 0 \to \int_A^B \frac{\partial Q}{T} \leq \int_A^B \frac{\partial Q_{rev}}{T} = S(B) - S(A)$$

For a thermally isolated system, $\partial Q = 0$, which tells us that $dS \geq 0$. This tells us that systems tend to the state that maximizes entropy. If we assume that the universe is a thermally isolate system, and U for the universe is constant, then the entropy of the universe can only increase.

Using the first law, we can find that $dU = TdS - pdV$, for any process, including irreversible ones, as all values in it are functions of state. We can also use entropy to find equilibrium, as the entropy of both objects at equilibrium must be equal in magnitude (sign difference). Systems will maximize entropy until it reaches a point that further exchanges of U and V do not change the entropy:

$$\frac{\partial S_{tot}}{\partial U} = 0 \quad \frac{\partial S_{tot}}{\partial V} = 0$$

If we have an object with temperature T_O and heat capacity C placed in contact with a reservoir with $T = T_R$, how do the entropies of the reservoir, object, and universe change?

We know that $dS = \frac{\partial Q_{rev}}{T} = \frac{\partial Q_{irrev}}{T}$. We also know that $Q_{tot} = C\Delta T$, whch in the differential form is $\partial Q_{tot} = CdT$. We can then set up the integral for the change in entropy:

$$\Delta S_{object} = \int_{T_O}^{T_R} \frac{CdT}{T}$$

For non-gaseous systems that do no work, $dS = \frac{\partial Q}{T}$, so for the reservoir:

$$\Delta S_{res} = \int \frac{\partial Q}{T_R} = \frac{1}{T_R} \int \partial Q = \frac{\Delta Q}{T_R} = \frac{C(T_O - T_R)}{T_R}$$

The entropy of the universe is simply

$$\Delta S_u = \Delta S_{object} + \Delta S_{res}$$

7.12 Joule Expansion

Joule expansion is when there is a violent expansion of an ideal gas in a thermally isolated recipient. If wthe gas is initially confined to the left half, and there is a vacuum on the right half, with a wall in between them, if the wall is quickly removed, the gas will quickly expand.

There can't be any heat transfer, since the system is thermally isolated. There is also no work done, because nothing is being pushed on the right side, since its a vacuum. The system remains at a constant temperature, because $dU \propto dT$, and we know that since there is no heat or work, the change in internal energy is 0. This is irreversible, so how can we compute the change in entropy? We can't use $dS = \frac{\partial Q}{T}$, since its an irreversible process. However, since it is a function of state, we can just find a reversible process that has the same initial conditions and the same final conditions. We can use a reversible isotherm instead of the Joule expansion. We will have a piston that slowly withdraws, while keeping the system in thermal contact with a reservoir at temperature T_0. The change in entropy for both of those two systems will be equivalent.

We can talk a little bit more about entropy. If we take the first law, we see that

$$T = \left(\frac{\partial U}{\partial S}\right)_V \to \frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_V$$

Maximizing the number of microstates Ω we had found a statistical definition of the temperature (using the fact that $U = E$):

$$\frac{1}{k_B T} = \frac{d\ln\Omega}{dE} \to S = k_B \ln\Omega$$

For a Joule expansion, we add 2^N more microstates (each molecule has twice as much volume to be in). The additional entropy is then

$$\Delta S = k_B \ln 2^N = k_B N \ln 2 = k_B n N_A \ln 2 = nR\ln 2$$

7.13 Entropy of Carnot Engines

Find the entropies for ideal power station that operates a Carnot engine between 500C and 20C using 1.61 MW of heat. We are given the values of T_h, T_c, and $\dot{Q}_h$. We can find $\dot{W}$ via the fact that $\dot{W} = \dot{Q}_h \eta_C$. We can then find $\dot{Q}_c$ using the fact that $\dot{Q}_h - \dot{W} = \dot{Q}_c$. To get the entropies of the reservoirs, we can just divide the heat leaving/entering the reservoirs, and divide by the temperatures of the reservoirs. The change in entropy of the engine is 0, because it is a cycle. If we add up all the entropies, we find that the net change in entropy is 0, as all the processes are reversible.

If we have a power station that produces 1 MW of work at the same temperatures, with a 40% efficient engine, we can first find $\dot{Q}_h$ using $\dot{Q}_h = \frac{\dot{W}}{\eta}$. We can then use the same formulae for finding the change in entropies of the reservoirs. Finding the entropy of the engine is still the same, whether or not it is reversible or not. All cyclic engines have $\Delta S = 0$. Irreversible engines waste some hot (high quality) into cold energy, because they cannot convert energy at a high efficiency.

What if we have a Carnot engine power station that is between 500C and 20C, but the hot reservoir is at 600C? The entropy of the Carnot engine is still 0, since its a cycle. However, there will be some change in entropy of the entire system, even though we're using a Carnot engine. Thus even with a reversible engine, we can have an irreversible system. Two objects at different temperatures in thermal contact is always irreversible.

The change in entropy of a given system only depends on its initial and final points. It does not matter whether the process is reversible or irreversible.

Irreversible processes do not produce as much work as they could have. They waste high quality energy into heat.

All cyclic engines have $\Delta S = 0$, whether they are reversible or irreversible. Irreversible engines just waste energy into low quality energy.

7.14 Equipartition Theorem

Imagine a small system with a reservoir that can exchange energy. In theremal equilibrium, the system has energy ϵ, and the reservoir $E - \epsilon$. For each allowed energy of the system, only one microstate contributes. For instance, a system of one molecule has energy $E_{sys} = \frac{1}{2}mv^2$. For each allowed energy E_{res} of the reservoir, a very large number of $\Omega(E_{res})$ of microstates contribute.

$$\ln \Omega(E - \epsilon) = \ln \Omega(E) - \frac{d \ln \Omega(E)}{dE} + \mathcal{O}(\epsilon^2) \approx \ln \Omega(E) - \frac{\epsilon}{k_B T}$$

$$\Omega(E - \epsilon) = \Omega(E) e^{-\epsilon/k_B T}$$

This gives us that

$$P(\epsilon) \propto \Omega(E - \epsilon) \times 1 \propto e^{-\epsilon/k_B T}$$

Where $P(\epsilon)$ is the probability that the object has the energy ϵ. In general, this leads to the equipartition theorem.

Theorem 7.2. *In equilibrium, each quadratic degree of freedom contributes $\frac{1}{2}k_B T$ to the average energy per molecule.*

This can be applied to an ideal gas. Molecules only have 3 degrees of freedom, and from this we actually find out why $U = \frac{3}{2}kT$.

7.15 Black-Body Spectrum

All objects at $T > 0K$ emit EM radiation. This is due to accelerated charges in the movement of molecules in the object. Most of it is infrared, but a high temperatures, we see it in the visible spectrum. Thermal radiation is studied using an ideal object, a black-body, which absorbs all the radiation that falls in its surface.

If we place two black bodies of different shapes so that their radiation is fully absorbed by the other, and wait until they are in thermal equilibrium, they must absorb as much energy as they emit. Black-body radition per unit of time and area does not depend on the size, shape, or composition, but only on the temperature.

We can also place a filter between the two bodies, which only alows certain wavelengths to pass through. They must absorb as much energy as they emit for that certain wavelength. The black-body radiation for each frequency per unit of time and area is the same.

Stars are pretty close to black bodies, but they aren't perfect black-bodies. However, the Cosmic Microwave Background is a better black-body spectrum. In fact, it is the closest to a black-body that we have ever measured. This is leftover radiation from when electrons and protons formed atoms. The initial radiation was around 3000K, and is now redshifted to about 2.73K.

In thermal equilibrium, the EM radiation inside a black-body is in equilibrium. If we have a cubic box with metallic sides, then on the walls, we have that

$$\vec{E} = \vec{B} = 0$$

We can only have standing waves between the walls. There can be two perpendicular polarizations for each standing waves.

The total energy that a black-body radiates is given by

$$R(T) = \sigma T^4$$

And Wien's law gives the peak wavelength:

$$\lambda_{peak} = \frac{2.898 \times 10^6}{T}$$

7.16 Photoelectric Effect

The photoelectric effect occurred when UV light was shone on a charged metal plate. Hallwach noticed that the number of electrons emitted per second depended on the intensity of the incident radiation, which was expected if light was a wave. However, there were other things that didn't match up, like the energy of the electrons depending on the frequency of the radiation, and not the intensity. They also expected a lag between the light and when electrons were emitted, when in reality there was immediate emission. They also expected there to be emissions for all surfaces, but there were no emissions for iron plates.

Einstein had been bothered with the inverse square law for light, claiming that it didn't make sense that the amount of light could be arbitrarily low. In 1905, he extended Planck's quantum hypothesis of light absorption/emission to propagation. He created the photon:

$$E_{photon} = h\nu$$

This theory was in total opposition to the entrenched wave theory. For this, he won the Nobel prize in 1921. It was later shown that Planck's blackbody radiation law can also be obtained using photons.

The electrons bound in the plate require some amount of energy to kick them out of the material. Einstein thought that the photons were providing this energy:

$$E_{kin} = E_{photon} - W = h\nu - W$$

Where W is the work function that defines the amount of energy required to kick the electrons out. Each material has a different work function. For example, Potassium has $W_K \approx 2eV$, Zinc has $W_{Zn} \approx 3.6eV$, and Iron has $W_{Fe} \approx 4.7eV$. This explains why iron was ot able to lose the electrons, as the energy of the photons used was not enough to overcome the work function. We see that if we now think about the system using photons, all the issues that the wave nature had with the photoelectric effect were remedied. In 1916, Millikan used this to measure Planck's constant h, by finding the maximum potential V_0 necessary to stop all the photoelectrons, giving us the first precise measurement of the value of h:

$$h\nu - W = eV_0 \rightarrow V_0 = \frac{h\nu}{e} - \frac{W}{e}$$

Rutherford developed the planetary model via the gold foil experiment, but it had an issue, which was that accelerated charges emite EM radiation, and should spiral into the nucleus. Bohr fixed this by saying that electrons orbit in "stationary states", with no energy loss, and atoms radiate when electrons make transitions between stationary states. The angular momentum of the electrons is quantized in units of $\hbar = \frac{h}{2\pi}$. Typically the energy levels are written in terms of the fine structure constant α:

$$E_n = -\frac{m_e c^2}{2}\alpha^2 \frac{Z^2}{n^2}$$

where $\alpha = \frac{e^2}{4\pi\epsilon_0 \hbar c} = \frac{1}{147.04}$. Also good to know is the mass of the electron is $m_e c^2 = 511\text{keV}$.

However, the Bohr model doesn't explain quantization, why energy levels are stationary, and completely breaks down for more complicated atoms. However, it was very good for describing hydrogen and atoms with single outer-shell electrons, such as Lithium.

7.17 Matter Waves

In 1900, Planck postulates that the only allowed energyies for the light modes in a black-body are $E_N = nh\nu$. Shortly after, in 1905, Einstein develops the notion of a photon, which propagates like a particle with momentum $p = \frac{h\nu}{c}$. This explains the photoelectric effect beautifully. In 1913, Bohr extends the quantization idea to the angular momentum of electrons in the atom. This leads to discrete energy levels of the atom, and the change in energy between the energy levels. This explained the hydrogen emission spectrum. In 1923, Compton proves the notion of a photon, by measuring the Compton effect, the change of momentum after scattering x-rays from electrons.

At tihs point, light had been shown in different ways to act like both a wave, and like a particle. Wave-particle duality was clear in Compton's experiment, as he used light like a photon but measure the wavelength of the outgoing light, treating it like both a particle and a wave. In 1924, Louis de Broglie extended the wave particle duality to matter. He postulated that particles could be described as waves:

$$\nu = \frac{E}{h} \qquad \lambda = \frac{h}{p} = \frac{h}{\gamma m \nu}$$

Typically these wavelengths are tiny, but they matter for things such as electrons. For energies much less than rest mass, we have that $p = \sqrt{2E_{kin}m}$. This actually explained the quantiation of atomic levels, as the electrons must be able to have a "connected" orbit, where there are an integer number of wavelengths in the circumference:

$$2\pi r = n\lambda = n\frac{h}{p}$$

Using $\hbar = \frac{h}{2\pi}$:

$$pr = L = n\hbar$$

Waves can be written in terms of the wave 4-vector k^μ:

$$k^\mu = 2\pi\left(\frac{\nu}{c}, \frac{\vec{n}}{\lambda}\right) = \left(\frac{\omega}{c}, \vec{k}\right) = \frac{p^\mu}{\hbar} = \left(\frac{E}{c\hbar}, \frac{\vec{p}}{\hbar}\right)$$

As a 4-vector, this Lorentz-transforms, and can be used to calculate relativistic Doppler shifts.

In general, the frequency and wavelength are related as

$$E = \sqrt{m^2c^4 + p^2c^2} \to \nu = \sqrt{\frac{m^2c^4}{h^2} + \frac{c^2}{\lambda^2}}$$

7.18 The Schrodinger Equation

Schrodinger developed an equation to describe a wave. For a plane wave moving in the x direction:

$$\Psi(x,y,z,t) = Ae^{-i(\omega t - kx)} = Ae^{-i(2\pi\nu t - \frac{2\pi x}{\lambda})}$$

Using the deBroglie expressions:

$$= Ae^{-i(\frac{Et}{\hbar} - \frac{px}{\hbar})}$$

Taking derivatives with respect to time/position:

$$i\hbar\frac{\partial}{\partial t}\Psi(x,y,z,t) = -i^2\hbar\frac{E}{\hbar}\Psi(x,y,z,t) = E\Psi(x,y,z,t)$$

$$\frac{-\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x,y,z,t) = \frac{p^2}{2m}\Psi(x,y,z,t)$$

Forgetting about the rest mass, for a free particle, $E = \frac{p^2}{2m}$, so the two derivatives are the same. For a particle in a potential V moving in a generic direction, a good guess is

$$E = \frac{p^2}{2m} + V$$

$$i\hbar\frac{\partial}{\partial t}\Psi(x,y,z,t) = \left(-\frac{\hbar^2}{2m}\nabla^2 + V(x,y,z,t)\right)\Psi(x,y,z,t)$$

The solution of this equation gives us Ψ, which tells us everything about the dynamics of a particle. This equation can only be suggested, not derived. $V(x,y,z,t)$ is the potential energy function (or just the potential). This is a partial differential equation (not an ODE like $F = ma$). This is first order in time, so if $\Psi(x,y,z,0)$ is known, then Ψ is known for all times in the future. Also note that Ψ is necessarily a complex valued function.

In 1926, Max Born interprets the wavefunction, and says that $|\Psi(x,t)|^2$ is the probability of finding the particle between x and $x + dx$ at time t. When we measure, we set where the particle is for future measurements. The realist position says that the particle was always there, and QM is just an incomplete theory, and we have hidden variables that we don't know, and they lead to the distribution of probability. The Orthodox position (Copenhagen interpretation) stated that the particle was everywhere, but the act of measurement collapses the wavefunction, and forces the particle to be found in a certain location.

In order for the statistical interpretation to make sense, Ψ must be normalized:

$$\int_{-\infty}^{\infty} |\Psi(x,t)|^2\, dx = 1$$

Essentially, the sum of all probabilities must be equal to 1. We can compute the change of the normalization with time:

$$\frac{d}{dt}\int_{-\infty}^{\infty} |\Psi(x,t)|^2\, dx = \int_{-\infty}^{\infty} \frac{\partial}{\partial t}(\Psi^*\Psi)\, dx = \int_{-\infty}^{\infty} \Psi^*\frac{\partial\Psi}{\partial t} + \frac{\partial\Psi^*}{\partial t}\Psi\, dx$$

Now using the Schrodinger equation and its complex conjugate, we have expressions for both $\frac{\partial\Psi}{\partial t}$ and $\frac{\partial\Psi^*}{\partial t}$, which we can insert into the integral. As a result, we are eventually left with

$$\frac{d}{dt}\int_{-\infty}^{\infty} |\Psi(x,t)|^2\, dx = \frac{i\hbar}{2m}\left(\Psi^*\frac{\partial\Psi}{\partial x} - \frac{\partial\Psi^*}{\partial x}\Psi\right)_{-\infty}^{\infty} = 0$$

This is 0 because Ψ needs to be 0 at infinity in order for it to be normalizable. If $\Psi(x,0)$ is normalized, then $\Psi(x,t)$ stays normalized (actually if any time is normalized, then all times are normalized, not just $t=0$).

For a multiple-particle system, there is only one wavefunction, which has all of the positions as arguments:

$$\Psi(\vec{r}_1, \vec{r}_2, \vec{r}_3, \dots, \vec{r}_n, t)$$

For example, the potential for two electrons in a Coulomb field of a very heavy Helium nucleus and their own is

$$V(\vec{r}_1, \vec{r}_2) = \frac{1}{4\pi\epsilon_0}\left(-\frac{2e^2}{|\vec{r}_1|} - \frac{2e^2}{|\vec{r}_2|} + \frac{e^2}{|\vec{r}_1 - \vec{r}_2|}\right)$$

The normalization of the wavefunction for this remains 1. In general, there is one wavefunction for the whole universe, and we can decouple pieces that don't interact. The probability of finding a particle between a and b is

$$P(a < x < b) = \int_a^b \Psi^*\Psi\, dx$$

The normalization implies that Ψ vanishes faster than $\frac{1}{\sqrt{x}}$ for $x \to \pm\infty$. This makes integration by parts easier:

$$\int_{-\infty}^{\infty} f(x)g'(x)\, dx = f(x)g(x)|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f'(x)g(x)\, dx$$

This first term will typically vanish if $f(x)$ or $g(x)$ are proportional to Ψ.

Since the Schrodinger equation is a 2nd order differential equation, Ψ is continuous. If $V(x)$ is non-infinite, then Ψ is also differentiable. If $V(x) = \infty$ for x in some range, then $\Psi(x) = 0$ in that range. Essentially, if there is infinite potential at a location, the particle cannot be in that region, as it would require infinite energy.

For a particle in state Ψ, its average position is called the expectation value:

$$\langle x \rangle = \int_{-\infty}^{\infty} x|\Psi(x,t)|^2\, dx = \int_{-\infty}^{\infty} \Psi^*[x]\Psi\, dx$$

The expectation value is not where the particle is most likely to be, it is the average value of the particle's position. For the momentum:

$$\langle p \rangle = m\frac{d\langle x \rangle}{dt} = m\frac{d}{dt}\int_{-\infty}^{\infty} x|\Psi|^2\, dx = m\int_{-\infty}^{\infty} \frac{\partial}{\partial x}(x\Psi^*\Psi)\, dx$$

$$= m\int_{-\infty}^{\infty} x\frac{\partial}{\partial x}\left[\frac{i\hbar}{2}\left(\Psi^*\frac{\partial \Psi}{\partial x} - \frac{\partial \Psi^*}{\partial x}\Psi\right)\right]\, dx$$

Integrating by parts:

$$= -\frac{i\hbar}{2}\int_{-\infty}^{\infty}\left(\Psi^*\frac{\partial \Psi}{\partial x} - \frac{\partial \Psi^*}{\partial x}\Psi\right)\, dx = \int \Psi^*\left[-i\hbar\frac{\partial}{\partial x}\right]\Psi\, dx$$

We call $-i\hbar\frac{\partial}{\partial x}$ the momentum operator. The total energy of the system is given by the Hamiltonian operator:

$$H = \frac{p^2}{2m} + V \to \langle H \rangle = \int \Psi^*[]\Psi\, dx$$

7.18.1 Time independent Schrodinger equation

If the potential is independent of time, $V(x)$, we can use separation of variables:

$$\Psi(x,t) = \psi(x)\phi(t)$$

Then the Schrodinger equation becomes

$$i\hbar\psi\frac{d\phi}{dt} = -\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}\phi + V\psi\phi$$

Dividing by $\psi\phi$:

$$i\hbar\frac{1}{\phi}\frac{d\phi}{dt} = -\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}\frac{1}{\psi} + V(x) = E$$

We see that the left side does not depend on position, and the right side does not depend on time. This is a constant E, the energy in the system. Solving the time component:

$$i\hbar\frac{d\phi}{dt} = E\phi \to \phi = e^{-iEt/\hbar}$$

The terms that depend on x give us the time-independent Schrodinger equation (TISE):

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V(x)\psi = E\psi$$

E is the energy of the state. For unbound states, like free particles, the energy spectrum is continuous. For bound states, boundary conditions restrict energy to discrete values. For example, harmonic oscillator solutions have energies $E_n = (n + \frac{1}{2})\hbar\omega$, for $n = 0, 1, 2, \ldots$.

Energy levels for the solutions to the TISE are always real. We can multiple the TISE by ψ_n^* and then integrate:

$$-\frac{\hbar^2}{2m}\int \psi_n^*\nabla^2\psi_n + \int \psi_n^* V\psi_n = E_n\int \psi_n^*\psi_n$$

Integrating by parts twice, we see that the other term is equal to its complex conjugate, and thus it must be real. Therefore E_n must be real.

We can also show that the total energy must be positive:

$$\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} = (V - E)\psi$$

Since ψ is normalizable, it must tend to 0 on both sides. If E is lower than the minimum of V, then $V - E$ is always positive. But this cannot be possible with a normalizable ψ, and thus E cannot be lower than the minimum of $V(x)$.

The stationary state solutions to the TISE are orthonormal:

$$\int \psi_m^*(x)\psi_n(x)\,dx = \delta_{mn}$$

They have to be normal because if $m = n$ we should expect the integral to be 1 (because its just ψ^2). If we multiply the TISE for ψ_n by ψ_j^* and integrate:

$$-\frac{\hbar^2}{2m}\int \psi_j^*\nabla^2\psi_n + \int \psi_j^* V\psi_n = E_N\int \psi_j^*\psi_n$$

If we do the same thing the other way, multiplying the conjugate of the TISE for ψ_j by ψ_n:

$$-\frac{\hbar^2}{2m}\int \psi_n \nabla^2 \psi_j^* + \int \psi_n V \psi_j^* = E_J \int \psi_n \psi_j^*$$

Subtracting the two equations:

$$(E_n - E_j)\int \psi_n \psi_j^* = 0$$

The separable solutions are stationary states, where the probability does not change with time:

$$\Psi(x,t) = \psi(x)e^{-iEt/\hbar} \to |\Psi(x,t)|^2 = \psi^* e^{iEt/\hbar}\psi e^{-iEt/\hbar} = |\psi(x)|^2$$

The general solution is a linear combination of the separable solutions:

$$\Psi = \sum_{n=1}^{\infty} c_n \psi_n(x) e^{-iE_N t/\hbar}$$

If the potential is even, $V(-x) = V(x)$, the probability must be even:

$$|\psi(-x)|^2 = |\psi(x)^2|$$

This leads to a ψ of definite parity, either even or odd:

$$\psi_{odd}(x) = \frac{\psi(x) - \psi(-x)}{\sqrt{2}} \qquad \psi_{even}(x) = \frac{\psi(x) + \psi(-x)}{\sqrt{2}}$$

Operators return the value of an observable:

$$\hat{X}\Psi(x) = x\Psi(x)$$

$$\hat{P}\Psi(x) = -i\hbar\frac{\partial}{\partial x}\Psi(x) = p(x)\Psi(x)$$

$$\hat{H}\Psi(x) = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\Psi(x)$$

When well defined, the function is called an eigenstate and the value is an eigenvalue. The Dirac delta is an eigenstate of the position:

$$\hat{X}\delta(x) = x_0\delta(x)$$

And exponentials are eigenstates of momentum:

$$\hat{P}e^{ip_0x/h} = -i\hbar\frac{\partial}{\partial x}e^{ip_0x/\hbar} = p_0 e^{ip_0x/\hbar}$$

From this, we can already intuit Heisenberg's uncertainty principle.

Writing the TISE in terms of the Hamiltonian:

$$\hat{H}\psi = E\psi$$

The expected value is the energy:

$$\langle H\rangle = \int \psi^* \hat{H}\psi\, dx = \int \psi^* E\psi\, dx = E\int |\psi|^2\, dx = E$$

These are states of definite energy, as the variance is 0:

$$\langle H^2\rangle = \int \psi^* \hat{H}^2\psi\, dx = \int \psi^* E^2\psi\, dx = E^2$$

$$\sigma_H^2 = \langle H^2\rangle - \langle H\rangle^2 = E^2 - E^2 = 0$$

7.19 Infinite Square Well

We have a potential function that is 0 for $[0, a]$, and ∞ otherwise. Outside of the well, we know that $\psi(x) = 0$. Inside the well, we can solve the TISE:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} = E\psi \rightarrow \frac{d^2\psi}{dx^2} = -k^2\psi$$

Where we have set $k = \frac{\sqrt{2mE}}{\hbar}$. We know that $E > 0$, because E has to be larger than V_{min}. Now solving the equation for ψ:

$$\psi = A\sin(kx) + B\cos(kx)$$

This is the general solution of the time-independent equation. can now solve for the constants. Continuity of ψ at $x = 0$ requires that $\psi(0) = 0$, which gets us that $B = 0$. We can impose continuity at $x = a$, and thus $\psi(a) = A\sin(ka) = 0$. This actually gives us k, because A cannot be 0, so we need $k_n = \frac{n\pi}{a}$, which tells us that $E_n = \frac{\hbar^2 k_n^2}{2m}$. To find A, we can use the normalization condition:

$$\int_{-\infty}^{\infty} |\psi|^2\,dx = \int_0^a |A|^2 \sin^2(kx)\,dx = |A|^2\frac{a}{2} = 1 \rightarrow |A|^2 = \frac{2}{a}$$

This tells us that

$$\psi_n(x) = \sqrt{\frac{2}{a}}\sin\left(\frac{n\pi}{a}x\right)$$

Now adding the time portion to get the full solution:

$$\Psi_n(x,t) = \sqrt{\frac{2}{a}}\sin\left(\frac{n\pi}{a}x\right)e^{-iE_nt/\hbar} \quad E_n = \frac{n^2\pi^2\hbar^2}{2ma^2} \quad n = 1, 2, 3, \ldots$$

Note that all real valued wavefunctions $\psi = \psi^*$ have $\langle p \rangle = 0$.

7.20 Finite Square Well

The finite square well can be written with potential $V(x) = 0$ for all $|x| > a$, and $V(x) = -V_0$ for $x \in [-a, a]$. The bound states have $-V_0 < E < 0$.

Solving the equation, we have 3 different cases, because we have 3 different piecewise definitions. For the middle case, in the well:

$$\frac{d^2\psi}{dx^2} = -l^2\psi \quad l = \frac{\sqrt{2m(E + V_0)}}{\hbar}$$

This gives us the solution from the infinite square well:

$$\psi(x) = C\sin(lx) + D\cos(lx)$$

For the left and right wings:

$$\frac{d^2\psi}{dx^2} = k^2\psi \quad k = \frac{\sqrt{-2mE}}{\hbar}$$

Notice that the sign has changed, and this gets us an exponential solution:

$$\psi(x) = Ae^{-kx} + Be^{kx}$$

This gets us 3 solutions, each with two unknown constants and each also relying on the unknown E. To get these constants to be determined, we can first force normalization, which drops a term from each wing function (because otherwise they go to infinity). We can also leverage the fact that $V(x)$ has definite parity, and thus the wavefunction must have definite parity, so we can drop one of the terms from the middle function (either the sine or the cosine), and we also know that the wing terms must have equal coefficients to force symmetry.

This gives us two solutions, even and odd. This brings it down to just 2 constants and E. From here, we can use the fact that ψ must be continuous at $x = a$, and in this case ψ' must be continuous at $x = a$.

These conditions lead to transcendental equations that are solved numerically. Also note that for the finite square well, the solutions go past the walls of the well, with the exponential wings. These wings eventually go to 0, but are different from the infinite square well, which had the endpoints at 0. This extension past the walls leads to tunneling!

7.21 Free Particles

Free particles have $V(x) = 0$ everywhere:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} = E\psi$$

This always has the solution

$$\psi = Ae^{ikx} + Be^{-ikx}$$

with $k = \frac{\sqrt{2mE}}{\hbar}$. k is positive, but we can merge both solutions and let k have any value. There are no boundary conditions constraining k, and so we have a continuous spectrum of possible energy. The solutions of the full equation are plane waves:

$$\Psi(x,t) = \psi(x)e^{-i\frac{Et}{\hbar}} = Ae^{ik(x-\frac{\hbar k}{2m}t)}$$

These correspond to waves travelling at $v = \pm\frac{\hbar|k|}{2m} = \pm\sqrt{\frac{E}{2m}}$. Note that this is not normalizable.

Physical states can be expressed as linear combinations of the plane waves. For instance, a Gaussian wave packet:

$$\Psi(x,t) = \int_{-\infty}^{\infty} Ae^{-\alpha(k-k_0)^2}e^{ik(x-\frac{\hbar k}{2m}t)}\,dk$$

7.22 Tunneling

Tunneling is just the finite square well inverted, where $V(x) = V_0$ in the range $[-a, a]$, and 0 elsewhere. We cna split this into 3 different solutions, with the middle part being complex exponentials, and the wing solutions being sines and cosines. To define the constants, we can use continuity, and continuity of the derivative. Once we look at some solutions, we see that the solution “passes through” the potential “wall”, and when we look at physical solutions using wave packets, we can see that the square of the wavefunction “bounces” off of the wall, but a small amount of probability passes through to the other side of the wall.

7.23 Harmonic Oscillator

The harmonic oscillator potential is one of the most useful potentials in physics:

$$V(x) = \frac{1}{2}m\omega x^2$$

Small oscillations around the minima of general potentials are simple harmonic oscillators, meaning that we can approximate things with harmonic oscillators. This is one of the few quantum mechanical systems for which an exact, analytical solution is known. It can be shown that for high energies, the wavefunction tends to the classical solution.

7.24 Uncertainty Principle

If we generate a wave by holding the end of a rope and shaking it up and down, it is easy to see the wavelength, but hard to tell where the wave is. On the other hand, if we give the rope one single jerk, the wavelength is hard to tell, but the location of the wave is easy to tell.

7.25 Quantum Mechanics Interpretations

The realist position is that particles have defined positions, it's just that QM is incomplete, and there is some hidden variable that is necessary to fully describe the particle.

The Orthodox or Copenhagen interpretation is that the particle is everywhere as described by Ψ. The act of measurement collapse the wavefunction, and forces the particle to take a stand. This is a less satisfying answer, and the wavefunction collapse is still unexplained.

In 1935, Einstein, Podolsky, Rosen published the famous EPR paradox. If we have a particle like π^0, which has no spin, that decays into things that have nonzero spin (e^- and e^+ both have spin $\frac{1}{2}$). The sum of the spins are 0, because we started with a net spin of 0. The Orthodox position tells us that they travel with their spins entangled, and measuring the spin of one child particle collapses the wavefunction and determines the spin of the other no matter how far it is. Einstein thought that this spooky action at a distance was not possible. Was information travelling faster than the speed of light? EPR was made to disprove the Orthodox position, but thought QM was still correct (they just thought Ψ was ot enough). In 1964, John Stewart Bell proved that any local hidden variable theory is incompatible with QM. He predicted the difference between the realist and orthodox expectations for measurement, and it seemed that the orthodox position was closer to what is observed in nature.

In classical mechanics, we just have a state and the equations of motion. In textbook QM (Orthodox position), we have a state (Ψ), and equations of motion, but we also have that measurements return eigenvalues of operators. We also know that the probability of an eigenstate can be given via the magnitude of an inner product. After the measurement, the wavefunction collapses to whatever state is measured. In many-worlds QM, we have the same state function and equations of motions, and nothing else. Some point out that the Orthodox interpretation is less natural, and should be called too-few-worlds QM.