

Contents

Intro	i
1 Conceptual Overview	1
1.1 Classical Addition of Velocities	2
1.2 Inertial Reference Frame	3
1.3 Classical Inertial Reference Frames	4
1.4 The Aether	7
1.5 Speed of Light	9
1.6 Position and Time	10
1.7 Doppler Effect	18
1.8 Aberration	19
1.9 Relativistic Beaming	21
1.10 A Source's Retarded Vector Field of Light	22
1.11 Velocity	24
1.12 Summary	25
1.13 Definitions	27
2 Position, Time and Velocity	30
2.1 Classical Relativity	31
2.2 Postulates of Special Relativity	34
2.3 Time Dilation	35
2.4 Length Contraction	38
2.5 Position and Time Transformation	41
2.6 Velocity	48
2.7 Acceleration	49
2.8 Summary	51
2.9 Variables	52
3 Observers Delayed Worldview	53
3.1 Relativistic Observers Delayed view	54
3.2 Transformed Retarded Coordinates	56
3.3 Perceived Velocity	58
3.4 Summary	60
3.5 Variables	60

4 Spherical light pulses	61
4.1 Spherical Polar Coordinates	61
4.2 Surface Element on a Sphere	63
4.3 Light Velocity Transform	64
4.4 Doppler Effect	65
4.5 Aberration in Special Relativity	66
4.6 Relativistic Beaming	68
4.7 A Moving Source's Field of light	70
4.8 Retarded Light Propagation	71
4.9 Flux of light from Point Source	74
4.10 Variables	76
5 Invariant Quantities	77
5.1 Space-Time Interval	77
5.2 Energy-Momentum Relation	78
5.3 Energy-Momentum Relation For Photons	81
5.4 Variables	81
Outro	83
Resources	84

Intro

Classical physics accurately predicts the movement of objects in our everyday world. However, there are effects that it does not account for, and as objects approach the speed of light, these effects that are negligible at low speeds become increasingly significant. This leads to two observers moving at high speeds relative to each other becoming unable to agree on the lengths of objects and the times in which events occur. Crucially though, one thing remains constant: the speed of light relative to each observer is the same, regardless of their movement. We must redefine our classical understanding of time and space to find out how this can be. This is where special relativity comes into play.

Special relativity is vital in particle physics, nuclear physics, and astrophysics, used for calculating nuclear reaction output power and measuring galaxy speeds. It is also needed for precision timekeeping on fast-moving GPS satellites to calculate positions accurately.

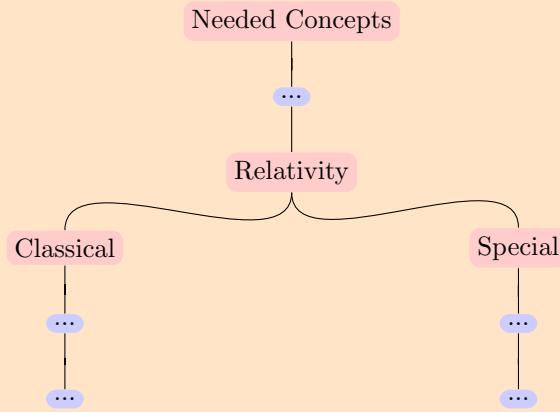
Despite its importance, special relativity continues to confuse even those well-versed in physics and mathematics. It can be as confusing to visualize and understand as quantum mechanics is. The theory requires letting go of some deeply held intuitions about the absolute nature of space and time that differ from our usual everyday experience of the physical world.

This handbook is meant to offer an intuitive, visual, and comprehensive overview of special relativity. The aim is to maintain simplicity and avoid unnecessary abstraction in explanations. For example, time in diagrams is shown either using animations or multiple diagrams given one after another in a timeline. This

approach avoids the use of unnecessary space-time diagrams, which represent time on one of the axes and make special relativity more abstract and confusing to visualize.

The first chapter is readable even for those without a math or physics background. It provides a non-mathematical, visual overview of the concepts. This chapter builds an intuitive understanding of the core ideas step-by-step, providing a solid foundation for developing the essential mathematical framework in later sections. The mathematical portion will begin with deriving how space and time coordinates function in special relativity.

This will be followed by velocity addition, the Doppler effect, and the energy-momentum relationship. In addition to these topics, we will also explore some interesting subjects not typically covered in standard resources. These include how a source's emitted light is concentrated in the direction of its movement, aberration, and how a delay in light signals to an observer means they are currently viewing sources of light in their past positions. These topics will probably not mean much to you yet, but they will give interesting insights.



Now, let us begin!

Chapter 1

Conceptual Overview

To make sense of special relativity, we will first look at the classical understanding of the world. Then, we will see how we must change these classical laws of physics to account for light's constant speed relative to all observers. Which in turn will lead to many interesting consequences. The chapter begins with an overview of the core concepts to help us understand why we need special relativity and what it means, leaving the mathematical details and deeper insights for the following chapters.

This chapter's concepts will establish the mathematical framework for later chapters, using physical objects, such as moving trucks on roads, to represent the different perspectives of various observers. Whereas, in the mathematical chapters, we will use coordinate axes to represent the different perspectives (frames of reference).

Later, you will see the parallels between this conceptual overview and the mathematical chapters. This chapter will provide an overview of all the required concepts for special relativity, building up the needed knowledge bit by bit, starting with our classical understanding of how perspectives (frames of reference) work and showing how our understanding must change if we are to allow all observers to view light as moving at the same constant speed.

1.1 Classical Addition of Velocities

Imagine a truck, carrying a cannon, is moving at 20 meters per second, and the cannon fires a ball at 100 meters per second relative to the truck. Classically, we add the truck and the ball's speed, to find the ball moves at 120 meters per second relative to the road. So, in classical physics, velocities are directly added. The ball relative to the road would move at the speed of the ball relative to the truck plus the truck's speed relative to the road, as shown in Figure (1.1).

However, this classical understanding is just a very good approximation for objects in our everyday life that we are used to seeing, which are moving much slower than the speed of light (roughly 300 000 000 meters per second), but we will get to this later. For now, in the following sections, we will explain what an inertial reference frame is and how this works in the classical context so that you can see more clearly the differences between classical and special relativity.

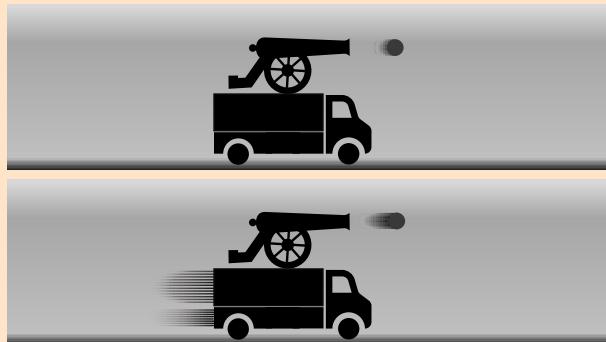


Figure 1.1: **Addition of cannonball and truck's velocities.** A diagram, showing the speed of a cannonball from two different perspectives, (top) is from the truck's perspective with the cannon at rest and the road moving backwards, the cannonball is shot and moves forward. The Figure (bottom) is from the road's perspective, with the truck moving relative to it. When the cannonball is shot, it looks like it is moving forward faster, with the previous cannonball's speed combined with the speed of the cannon that it was shot from. *** add labels of ball, truck and road speed or There could be a background object like a stop sign/person that is blurred in the top and not blurred in the bottom.

1.2 Inertial Reference Frame

A reference frame is a grid of coordinates that we can overlay space with, so that we can state with respect to it where and when an event occurs. If this grid is stationary or moving at a constant velocity, it is called an inertial reference frame. In the previous truck example, we could associate a reference frame with the truck, by introducing a grid of coordinates that is stationary relative to the truck and, hence, moving along with it. This perspective, fixed to the truck, is the truck's frame of reference. A second grid of coordinates, stationary relative to the road, can also be introduced, representing the road's reference frame. These grids of coordinates are shown from the road's reference frame in Figure (1.2).

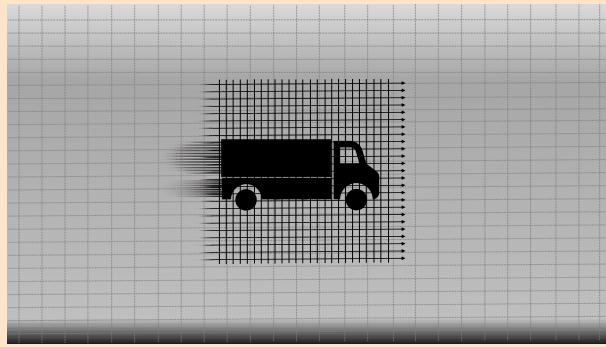


Figure 1.2: **Reference frames of a truck and road.** A diagram, showing the reference frame of a truck as a grid of coordinates that is moving with the truck relative to the road's reference frame shown as a stationary dashed gray grid of coordinates.

An inertial frame of reference is a reference frame that is not being accelerated. You can tell if you are being accelerated as you will feel a force. For example, in an accelerating car, you will feel the seat being accelerated into you, with your body slightly lagging behind in the acceleration. This principle is used in accelerometers to measure acceleration, as shown in Figure (1.3).

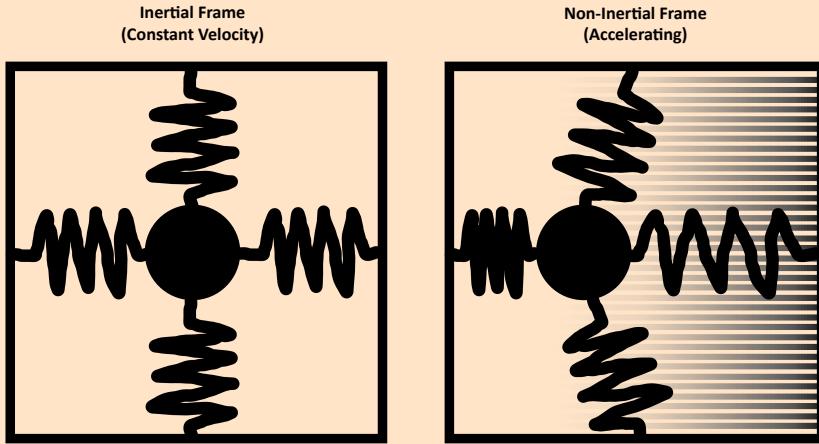


Figure 1.3: **Inertial vs non-inertial frames.** A diagram, showing a ball attached to the walls of a box by springs, with the ball centered in the box in the inertial frame, with no acceleration (left), and in a non-inertial frame (right) where the box is now accelerating to the right, The ball lags behind as the box accelerates.

We will now look at how movement changes in classical relativity when we change between these inertial frames.

1.3 Classical Inertial Reference Frames

To see how to swap between reference frames in special relativity, we will first have to introduce the classical approach.

To illustrate this, we will look at a setup of two rats on a treadmill with a platform as shown in Figure (1.4), with three rocks hanging above the platform, which are at rest relative to the room. The two frames of reference will be the treadmill's platform and the room in which the treadmill is in. Both rats start under the same rock. One rat runs to a rock positioned in the forward direction of the treadmill. The other rat runs to an equally distanced rock to the side. After this, they return to the starting rock.

Here, the platform can be seen as the medium in which the rats move. Both rats travel at the same constant speed relative to the platform, with the platform

either at rest or at a slower speed than the rat's motion to allow them to get to the rocks.

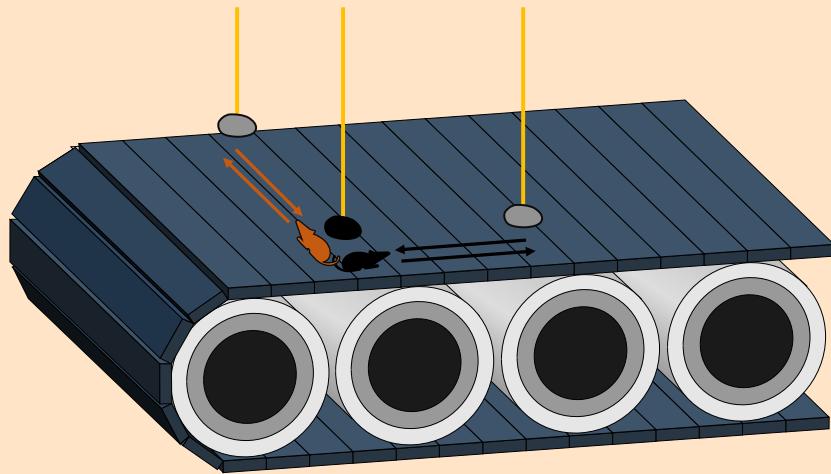


Figure 1.4: **Rats on a treadmill.** A diagram, showing a 3d view of a treadmill platform, with two rats and three hanging rocks, with arrows representing the rats moving from a starting rock to the other two equally distanced rocks, with the paths perpendicular to each other.

If the platform is at rest, they will return to the starting rock simultaneously. But if the treadmill is turned on and the platform is moving, the rats will also have to work against the movement of the platform to get to the rocks. This will lead to different distances the rats have to travel relative to the platform, and as a result, the rats get back to the starting rock at different times. Figure (1.5) shows what the direction of movement of the rats and rocks look like in each of the reference frames when the platform is moving.

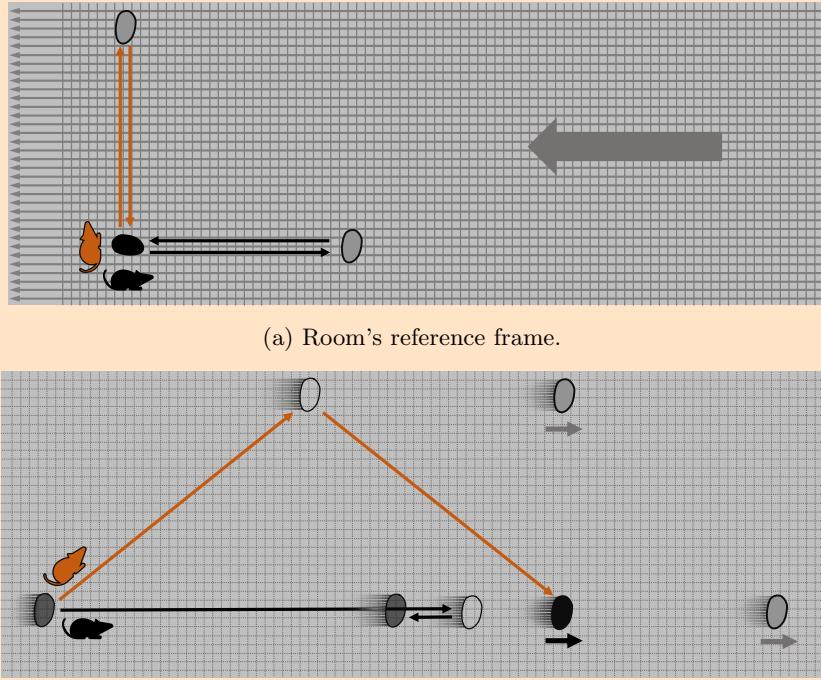


Figure 1.5: Different reference frames of rats on treadmill. Rats running on a moving platform of a treadmill from two different reference frames, with the paths of the rats represented as arrows.

From the room's perspective, the platform is moving backwards. The rat that is moving along the direction of the treadmill's motion will have impedance to its movement from the backwards pull of the platform. But after it turns around, it will have a boost from the platform to return to the starting rock. Meanwhile, the sideways-moving rat will have to move sideways and balance out the platform's backwards pull to keep its movement in the direction of the hanging rocks. This gives different lengths of paths depending on the reference frame, as shown in Figure (1.5). The rats in the platform's reference frame, move with the same speed for all the paths, represented by the arrows. Whereas, in the room's reference frame, we have a slower constant speed of the sideways-moving rat and two different speeds for the forward and backward paths of the other rat. This leads to different times for the rats to return to the starting

rock.

The rats here use the medium of the treadmill's platform to move, and similarly, it was thought that light needed a medium, called the aether, to travel through space. If there was this medium, we could test how fast it moved relative to Earth by finding the difference between the return times of light emitted in two perpendicular directions, just as in the treadmill example. So, next, we will look at the notion of a universal medium referred to as the aether and whether we can have the same frame swapping when it comes to light traveling in two different frames of reference, as we did in this section.

1.4 The Aether

In the theory electromagnetism from the late 1800s, it was believed that light was a wave and, therefore, would need a medium that filled the vacuum of space for it to travel through, called "the luminiferous aether", and that light traveled at a constant speed relative to this aether, like how the rats in the previous section moved at a constant speed relative to the medium of the treadmill's platform. It was from the theory of electromagnetism and the experiments performed from it, that showed that light should have a constant speed, atleast relative to this aether, but it was yet to be known that light would actually have a constant speed relative to all things.

If the aether existed, the Earth would be moving through it, illustrated in Figure (1.6).

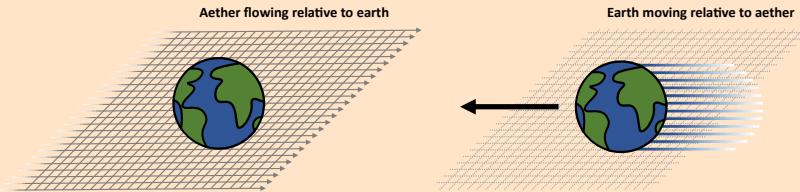


Figure 1.6: **The earth and the aether.** A diagram showing the proposed aether's and Earth's movement relative to each other.

So, an experimental setup by Michelson and Morley, shown in Figure (1.7),

called an interferometer, was devised to measure Earth's movement through the aether [1], by measuring how it affected the return times of light emitted in different directions, as observed in Earth's reference frame. It did this by splitting a single light beam into two perpendicular paths that are then reflected back to be recombined and sent toward a light detector. By rotating a platform that the setup is on, the two light paths could be aligned, one parallel and the other perpendicular to the Earth's motion through the presumed aether. They reasoned that if the speed of light was constant with respect to the proposed aether, just like in the rat experiment from the previous section, the split light beams would recombine at different times.

From the previous section's system of rats on a treadmill, we can see that the treadmill's platform is analogous to the aether, and the rats are analogous to the light. Both have the room where these experiments are done as the other inertial reference frame.

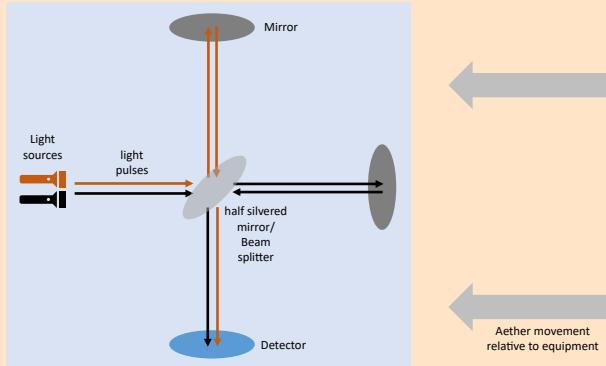


Figure 1.7: Experiment to detect the aether. A diagram showing the Michelson-Morley experiment, we can take the part of the paths between the beam splitter and two mirrors to be analogous to the case of the paths in the previous section of the rat and treadmill system.

However, when Michelson and Morley performed the experiment, they found no difference in travel time to the detector for both paths, indicating that there was no difference in the speed of light in any direction in Earth's reference frame. Hence, light's speed is not dependent on the supposed aether. This null result

seriously discredited the aether theories. It ultimately led to the proposal by Einstein in 1905 that the speed of light (in a vacuum) is a universal constant and is independent of the motion of the observer or source.

To allow for us to have this universal speed of light (in a vacuum), it will require us to change our ideas of how time and positions are perceived by different observers.

1.5 Speed of Light

The experiments within a vacuum, showed light does not just travel at a constant velocity relative to the proposed medium (the aether), and instead travels at a constant velocity regardless of the motion of the reference frame. The observed speed of light is independent of how fast the source and receiver of that light are moving in the frame, as shown with a moving truck's headlights in Figure (1.8). Light does move slower in objects such as glass, but this is due to the interactions between the light and the material, impeding the movement of light. Light itself moves extremely fast compared to any other everyday speeds we are used to (roughly 300 000 000 meters per second; it can travel the world's diameter in the blink of an eye). This is why we do not notice any delay in the light arriving from the things we see in everyday life, and we assume everything to be located where we are currently seeing it, though we should remember that there is this delay. For example, the light we currently see from the sun was emitted eight minutes ago for it to reach us now, and the further an object is located from us, the further back in time we are currently seeing it because of this delay.

When we look at the same truck setup as in Figure (1.1), but now have it in a vacuum, where the cannon firing a cannonball is swapped for headlights emitting light, we will find the same speed of light when measured relative to the truck or the road, but how can this be true?

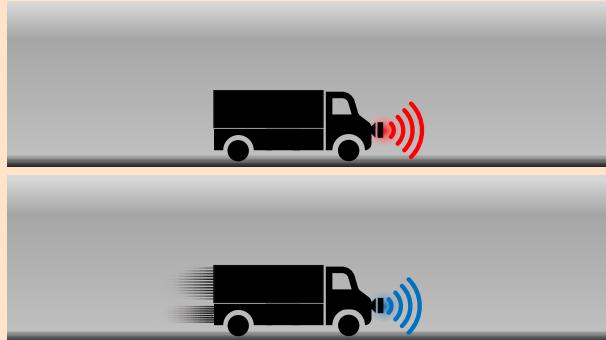


Figure 1.8: **Emitted light in different frames.** A diagram, showing light emitted from a truck in two different reference frames, with the emitted light having the same speed in each frame, though with different frequencies and energies due to the Doppler effect, which will be explained later.

For this to be true, we need a new way of thinking about velocity addition. This is because the velocities of objects must be added in a way that is consistent with the requirement that the speed of light is constant but also gives the approximate classical addition at speeds of objects at much less than the speed of light, like how we observed in the situation with the cannon and truck, shown in Figure (1.1). Since the speed of the light depends only on the units of time and positions, the only way to correct for this is to have the measured positions and times of objects transform differently when swapping between reference frames than the classical way. This is what we will talk about next. For the curious, it was the experiment by Ole Rømer that showed that light traveled at a finite speed rather than being instantaneously emitted and received.

1.6 Position and Time

1.6.1 Time Dilation

Consider a simple clock, as shown in Figure (1.9), made of a light pulse moving back and forth between two mirrors on a moving truck's floor and roof. Each tick of the clock is defined as the time it takes for the light to travel from one mirror to the other and back. For an observer in the truck, they will see the

light go straight up and down between the mirrors, but an observer stationary relative to the road will see not just the light traveling up and down but also with the direction of the truck.

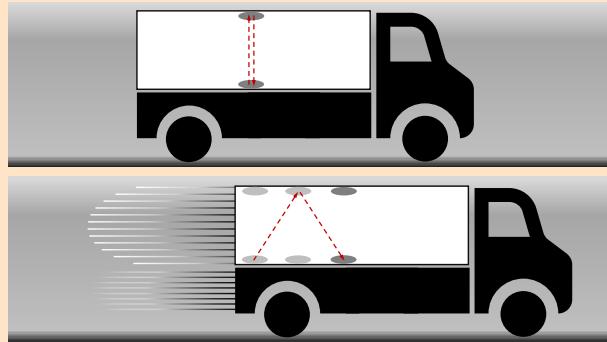


Figure 1.9: **Light clock on a truck.** A diagram, showing a clock consisting of light being reflected between the floor and roof of a truck to keep time, in the trucks frame (top), and in the road’s frame with the truck moving (bottom).

Given that the distance traveled by the light in the frame with the truck moving is longer, yet experimentally we know that the speed of light is the same for both reference frames, as shown in previous sections, to keep the speed of light invariant: the clock itself must tick slower in moving objects,

There was also the possibility that the height of the moving truck could be contracted, to give the same path length for light so that the ticks happen at the same rate, but this leads to paradoxes as shown in Figure (1.10).

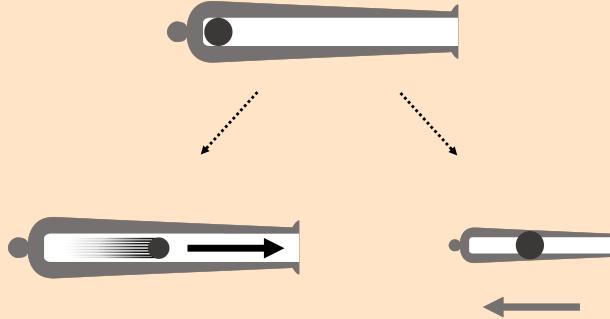


Figure 1.10: Paradox of height and width contraction. A diagram showing why there must not be a contraction in the perpendicular direction to the frame's relative motion. The top Figure shows a ball and cannon at rest. The bottom figures show the cannon ball being fired in the frame of the cannon (left), where the moving ball has a contracted width, and (right) in the frame of the ball, where the cannon is now moving and has a contract width, with the ball at rest being the same size. Both frames would contradict each other if there was a change in the width of moving objects, as the walls of the cannon and the surface of the ball would overlap in one frame and have a gap between them in the other. So, we require no size change perpendicular to the object's movement.

From this paradox, we are only left with the possibility that the perceived travel time of the light must be longer in the road's reference frame, Meaning the light clock ticks slower to the observer on the road watching it as it moves.

How much slower the time is passing in the moving truck relative to the road observer can be solved using the ratio of the lengths of the paths in each frame as shown in Figure (1.9), as this is the same ratio as the time between the ticks. Importantly, this time dilation phenomenon isn't limited to light clocks; it applies universally to all types of clocks. If you were to play a movie on the truck, it would take a longer time for it to play through from start to finish for an observer watching it from the road's reference frame. An observer on the road and in the truck experience time at different rates. This is not just an illusion, it is physical reality.

The closer the truck goes to the speed of light, the more significant the difference in how fast the road frame's time flows relative to the truck frame's time. This can be seen as a longer horizontal stretch in the path taken for the light from

the perspective of someone standing still on the road, with the time between each tick approaching infinity as the speed of the truck approaches the speed of light.

If now, we instead had two trucks moving at the same speed relative to the road but in opposite directions. Each truck would see its light moving vertically up and down with the other truck's light moving diagonally. Hence, each truck perceives the other truck's clock ticking slower. An observer in the road's reference frame would see both trucks move at the same speed and, hence, the passage of time between ticks as being the same for both trucks. The slowing of time of moving objects is the same regardless of its relative direction of motion, only the relative speed changes the path distance of the ticking light. This concept is confusing and difficult to visualize, but it will get easier with practice.

For now, though, we will continue building up the concepts of special relativity.

1.6.2 Simultaneity

Imagine a truck in its rest frame, with a light bulb in the middle and mirrors on the front and back walls. If the light bulb gives off a pulse of light, the light will travel from the center of the truck to the mirrors simultaneously and back to reach the light bulb also simultaneously. But an observer on the road watching the truck drive past will see the light bulb simultaneously emit light in both directions and simultaneously return to the bulb again. However, the light reflects off each wall's mirrors at different times in the road's frame to do this, as shown in Figure (1.11). This is because the speed of light is the same for both directions, but the truck is moving, meaning the back of the truck is moving towards where the bulb was when the light pulse was emitted, making the distance traveled shorter. The light travels a longer distance to the front wall due to the wall moving away from where it was emitted. Because of this, the observer on the road will see the light hit the mirrors at different times.

Consequently, times of events, such as when the light reaches either mirror, are different for observers in different reference frames. There is no one true

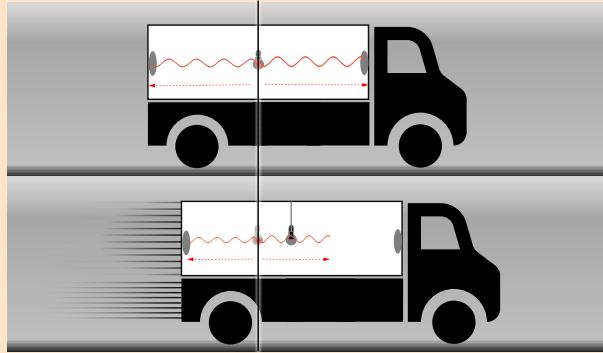


Figure 1.11: **Simultaneity of events in different frames.** A diagram showing how two events of light reaching the two walls of the truck are simultaneous in one frame but happening at different times in another, due to truck's movement in the second frame and the speed of light remaining the same.

order of events. For example, an observer in a faster-moving truck moving in an opposite direction would see the light reach the front wall first. However, for all observers, the light will return to the central light bulb simultaneously. If two events happen in the same position simultaneously, then this happens simultaneously in all frames of reference. This means the light returns to the center of the light bulb simultaneously in each frame.

A note here is that we are talking about when the events **actually** happen in each frame and not when the light from these events reaches the observer, which is later due to the time it takes the light to travel to the observer. Also, I have left out any mention of length contraction of the truck, which will be introduced in the next section, and why it needs to be equal in both directions from the lightbulb for the light to simultaneously return to the light bulb.

1.6.3 Length Contraction

Consider a truck, in its rest frame carrying a square container. If a light source at the container's center emits light in all four directions it will bounce off the mirrored sides of the truck and return to the center simultaneously. We require that in the moving frame, they also all return to the center simultaneously,

as multiple events that happen at a single point simultaneously in one frame, must happen simultaneously in all other frames. The light reflects off the walls at different locations, so the order of reflections in time will depend upon the relative observer, as described in the previous section. The time between the light being emitted and absorbed is the dilated time, that was described in section (1.6.1).

To achieve this simultaneity in the return of the light to the bulb, the total length of the path of light in each of the four directions must be the same. As its speed is the same in all directions. We can work out the length of the upward path from the time dilation section (1.6.1), and this is the length the path needs to be in the horizontal directions as well,

the paths can only have this length if the truck's length is contracted when moving,

*** explain why is length contraction is none contradictory/paradoxical while we have previously seen perpendicular contraction to be so

*** could say that an option is to look at the possibility of length contraction,
*** or just state that when it is worked out where the light (which is moving at a constant speed) must be reflected by the walls to reach the center simultaneous the point at which the horizontal light is reflected must be such that the lorry is observed to be contracted *** and mention that this may seem like a non possible option as the width contract leads to contradictions, but it turns out this length contract does not suffer from the same paradox

the exact amount of contraction can be worked geometrically, and it turns out that the ratio of the increase in the amount of time that passes before the light is reabsorbed is inversely proportional to the amount the truck is contracted in the direction of its movement.

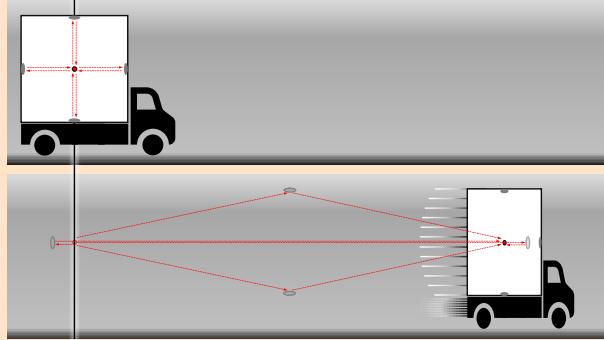


Figure 1.12: **Length contraction of moving truck.** A diagram, showing a truck with a square container in its rest frame (top), emitting light from a central bulb in the four directions, with all light being reflected by the mirrored sides back to the center. In the second frame with the truck now moving (bottom), we see the truck's length contracted and what the light paths would be in this frame. ... *** The bottom image could be confusing to some. As the vertical and horizontal path lengths should be equivalent for simultaneity, though the diagram appears to show that the vertical path is longer... make diagram clearer

So now we know that moving objects must shrink in the direction they move to allow for a consistent speed of light in each frame. But this is not noticeable when objects move much slower than light speed. If you are unsatisfied with this, the next section will reason the need for length contraction in another way.

Another Illustration of Length Contraction

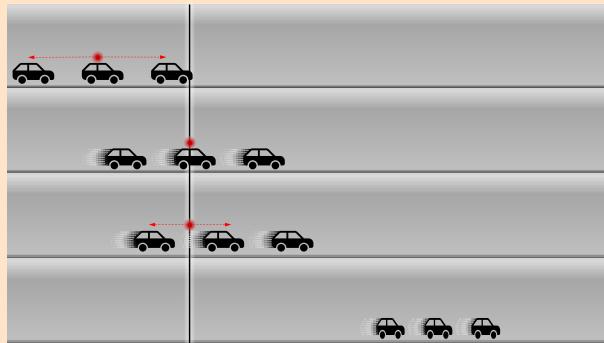


Figure 1.13: **Length contraction between three accelerating cars.** A diagram, illustrating an experiment with three cars initially at rest and equally spaced on a road. The middle car emits light that reaches the front and back cars simultaneously, triggering all cars to accelerate for a predetermined time. Now another light is emitted from the middle car after the acceleration has finished, the road observer sees this second light pulse reach the back car first, as it is moving towards where the light had been emitted, this causes it to begin accelerating before the front car. This results in the cars being closer together after acceleration, demonstrating Lorentz length contraction. However, to the car observer, the distance between the cars remains unchanged from the initial distances.

Suppose we have three equally distanced cars at rest on the road, as shown in Figure (1.13). If the middle car sends out a pulse of light, it will reach the front and back car at the same time for the road observer and the observers in the car. When it reaches them both, all the cars accelerate for a predetermined fixed amount of time to the same speed. Following this acceleration period, the cars have maintained equal spacing relative to both the car observers and the observer on the road. As for all observers, the light reached the front and back cars at the same time.

Now, suppose we do this a second time, with the middle car releasing another pulse. In this case, we have in the frame of the observers in the cars: again, with the front and back drivers receiving the pulses simultaneously, meaning that they will start accelerating simultaneously. After the acceleration has finished, the distances between the cars will be the same for the car observers. But for the

road observer, they see the back car receive the signal first, as that car is moving towards the point where the light pulse was emitted, and the front car is moving away from it. This would mean the back car would start accelerating first to get to the final constant velocity and get closer to the front car before the front car begins to accelerate to that same final speed that the back car had already reached. This results in the cars being closer together after the acceleration. The observers in the car and the road observer disagree with what the distances between the cars are. This contraction of length between cars, in the road's frame in which the cars are moving, is called Lorentz length contraction, which means that objects that are moving faster become shorter, and the distances between the objects also become shorter.

1.7 Doppler Effect

If we have a source at rest, emitting circular pulses of light with equal times between each pulse, we will have concentric circular pulses in this frame. But, if we move to a frame where the source is now moving to the right, each circular pulse is emitted from a different position. Due to the source moving, each pulse will be emitted closer to the right-hand side of the previous pulses, creating a bunching up of the pulses (increase in frequency) in the direction of movement and a spreading out (decrease in frequency) of the pulses in the opposite direction. This also happens in the classical version of the Doppler effect. For example, you will notice that an ambulance or police car sounds different when driving towards or away from you due to the bunching up and spreading out of the sound waves in the direction and opposite direction of the moving vehicle. However, in special relativity, we also must take the time dilation of the pulses into account, as there will be a longer time between each subsequent pulse. This is due to objects moving relative to an observer and therefore having their perceived time moving more slowly. This is called the second order Doppler effect most notably seen in optical clocks. This has a decreasing effect on the frequency in all directions. However, directly in the direction of movement of the

source, this is outweighed by the frequency increase from the previous bunching-up effect. Since the energy of the light is proportional to the frequency, it is also increased in the direction of motion of the source due to the Doppler effects and decreased in the opposite direction.

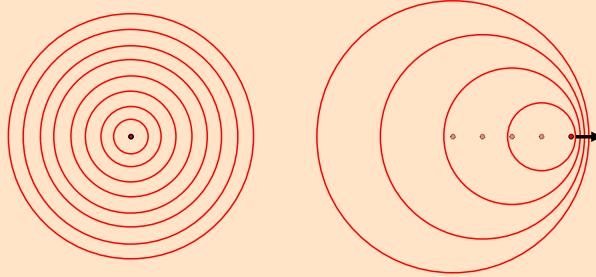


Figure 1.14: **Spherical waves from a moving source.** A diagram showing (left) a central source at rest emitting several circular pulses of light with equal time between each pulse, (right) the same source in a frame where it is now moving and emitting circular pulses of light, but each subsequent pulse is emitted from a different position as the source is moving, marked by a faded dot. *** could be done in latex instead

One thing not mentioned yet in this picture so far is how the light is also affected by what is called the aberration, which is the change in the angular distribution of the light at each part of the spherical pulses, which will be explained in the next section.

1.8 Aberration

Here, we will talk about what the Doppler effect in the previous section has yet to show, which is the effect on the angular distribution of the light in each part of the spherical pulses. We will show that there is a higher concentration of light in the direction of the source's movement, as shown in Figure (1.15).

With the help of the length contraction section (1.6.3), Consider a truck with a spherical mirrored container, with a central bulb emitting light in all directions. In the rest frame all light reaches the spherical wall simultaneously and returns to the center bulb simultaneously as well. Then in the moving frame, we have

the spherical container length contracted and the light moving at the same speed but reflecting off the same parts of the walls but at different times, and then returning to the center bulb simultaneously. For this to be true, the directions of the light must be aberrated (have their direction of propagation changed) in the way shown in the diagram to allow for this simultaneous return to the bulb.

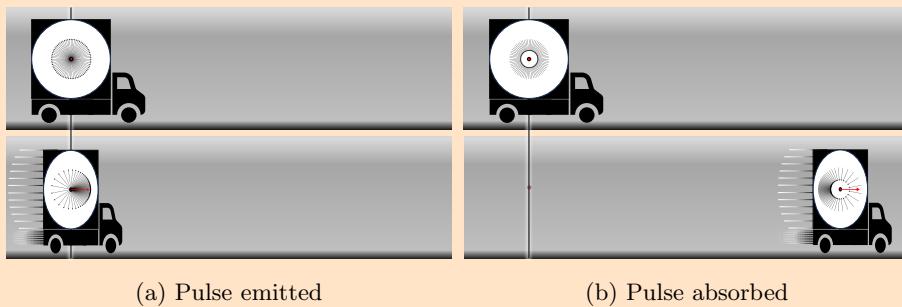


Figure 1.15: **Aberration of light on a truck.** Two diagrams of a truck with a spherical container in two different frames, a light pulse is emitted from the center of the container (left) and reflected back by the mirrored wall to be absorbed in the center again (right), the truck in its rest frame (top of both figures) shows an evenly distributed outward light pulse, but in the frame, with the moving truck (bottom of both figures) the emitted light is now more concentrated in the direction of the truck's movement, which is required to allow the light to be reflected by the same part of the container and returned to the center of the truck simultaneously.... *** could add the 45 degree angle (rest frame) light rays evolving in this scene as shown in previous figure, as the current picture is difficult to follow... Maybe stack a and b vertically...

From the diagrams, you can see how light's direction of propagation is changed when emitted or absorbed by a moving source relative to its rest frame, this is called relativistic aberration. The faster the source is moving, the more the direction of each part of the light pulse is changed/aberrated.

If the speed of a source was to approach the speed of light, all emitted light's propagation direction would approach the direction of movement of the source. If the source could theoretically reach the speed of light, all light would be emitted in the direction of the source but also move at the same speed. That is, the source would move with the emitted light, and it would not leave the vicinity of the source. The rate at which it emits the light would also tend to zero. If a

photon theoretically had mass and its influence of the gravitational force moved at the speed of light, then we would not be able to feel any gravitational effects from this theoretical photon with mass outside its vicinity. This is due to all its gravitational field being propagated in the direction of its movement and at the same speed as the photon.

Another note is that in astrophysics, it must be taken into account that the Earth's view of the universe is distorted by aberration when looking through telescopes, with the aberrational effect depending on where it is in its rotation around the sun.

We will look at the combination of the Doppler effect and aberration next.

1.9 Relativistic Beaming

When the Doppler effect and relativistic aberration are taken together, it is known as relativistic beaming. If we have a source in its rest frame emitting a spherical pulse of light with evenly distributed angles. Then in a second frame with the source now moving, the doppler effect has the spherical pulse's wavelength bunched up in the direction of movement (giving a change in the color of the light), with the aberrational effect also giving the angular spread of the light at a higher concentration in the direction of the movement of the source, as shown in figures (1.16).

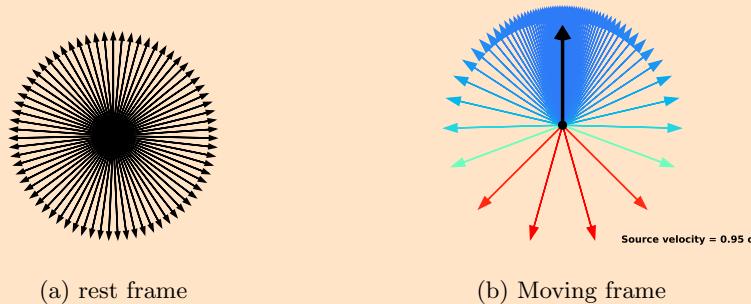


Figure 1.16: **Relativistic beaming of emitted light.** Outward light pulse from source in a rest and moving frame.

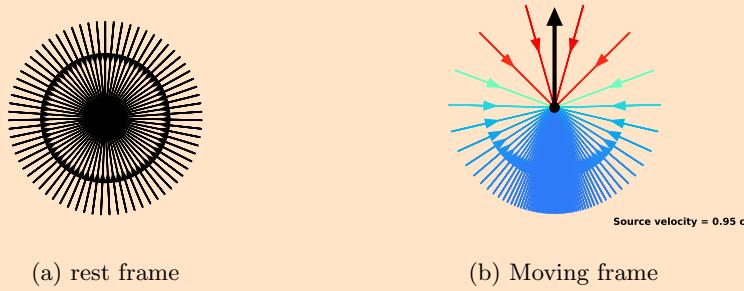


Figure 1.17: **Relativistic beaming received light.** Inward light pulse from source in its rest and moving frame.

For astrophysics, we need to consider this beaming when looking at distant stars that are moving relative to us so that we can accurately calculate their brightness and the spectrum of light that they are emitting, along with some other things. We also need to consider that the position we are currently seeing the stars in is their past position due to the delay in light traveling to us on Earth, which is referred to as the retarded position. We will talk about this next.

1.10 A Source's Retarded Vector Field of Light

We can make a vector field plot, which is a plot of arrows that shows the position and direction of light in space. So if we have a source in its rest frame that continually emits pulses of light with a constant time between each pulse, we get what is shown in Figure (??). If we wanted to know what this light vector field looked like in a frame with the source moving, we would need to take into account that the position the source emitted each pulse from is its past position (retarded position). We also have the dilated time between pulses in this moving frame, so using these and the beaming effect from the previous section we get the light's vector field as shown in Figure (??). From this, we can work out the distribution and, hence, the concentration of the light for any moving source relative to when it is at rest. If you look closely at the diagrams,

you will see the aberrated rings of all the pulses, all with different points from which they were emitted, with them being more compact in the direction of the source's movement.

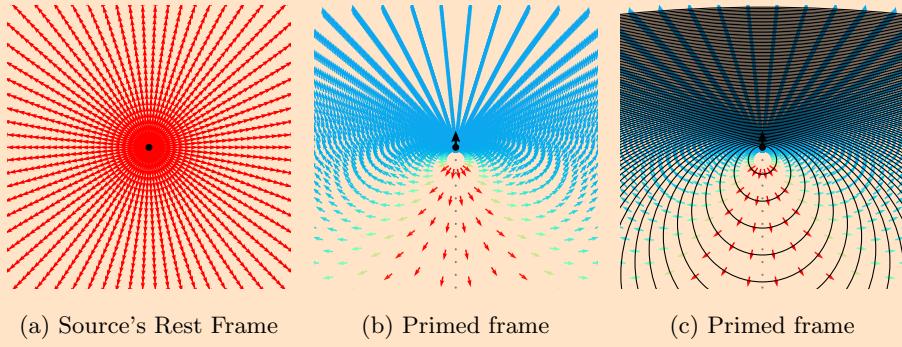


Figure 1.18: **Vector field of light pulses from a source.** A diagram, showing multiple spherically symmetrical pulses of light from a source at rest, fig (a), and in a corresponding inertial frame where the source is now moving, fig (b), the colors show the Doppler effect on the light, and the grey dots show the past positions of source when it had emitted each pulse, there is also a diagram showing the circles of each pulse, fig (c), This shows the spherical waves being closer together in the direction of movement.

In particle physics, the photon is the particle that composes light. It is the carrier of the electromagnetic force between charges. So how the light's propagation transforms when swapping between frames is necessary to understand electromagnetism, and how the magnetic field comes about for non stationary sources. The diagram of the moving source shows you where the source and light currently are. But, if you were an observer at a particular position in the vector field, you would see the source at its retarded position and not where it currently is, as the light has to travel to you before you see it. This will be used in the next section to show that you can perceive things moving faster than light, even if they are not.

1.11 Velocity

We looked at how velocities at the speed of light change between different frames of reference, that is, how their directions of propagation are rotated (aberrated). However, for objects moving at velocities slower than that of light, its speed will change in addition to the direction being transformed when swapping between frames. This is due to us needing the relativistic transform to tend towards the classical transform as we look at smaller and smaller speeds of objects. To understand the general velocity transformation, some mathematical detail is needed, so we will need to brush over this for now, and instead look at how the perceived speeds of objects can be more than that of light, even if they are moving slower.

1.11.1 Perceived vs Actual Speed

Imagine that a ball one light year away (the distance it takes light to travel in a year) is fired directly towards you, as shown in Figure (1.19). It will take one year for the first light from when the ball starts moving to reach you. If the ball moves at three-quarters of the speed of light, it will hit you in four-thirds of a year (a year and four months) after it is fired. The last light from the ball will reach you just as the ball hits you. As you perceive it, the time between the first and last light from the moving ball is four months. During those four months, you will see the ball start at its initial position and travel a distance of one light-year. So, to you, the ball appears to have been moving three times faster than light. This is just how it appears to the observer due to the delay in the light signal, giving the latency in how the system is observed.

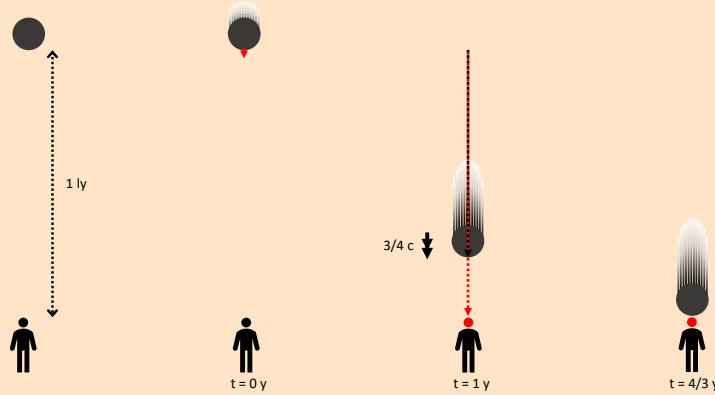


Figure 1.19: **Perceived faster than light movement.** A diagram demonstrating the perceived speed of a ball vs its actual speed ... the red arrow represents the light being emitted from when the ball starts moving from its position one light year away...

This shows how important it is to consider the delay in the light from objects when observing relativistic systems. This view is called the delayed/retarded view, meaning what we see now are objects in their past positions, and the further things are from us, the further into the past we are currently seeing them.

1.12 Summary

A key early concept was that light traveled through a medium called the luminiferous aether. The Michelson-Morley experiment attempted to measure Earth's motion relative to this aether. Surprisingly, no difference was found in the speed of light regardless of direction through this proposed aether. This conflicted with the intuitive addition of velocities, discredited the aether theory, and led Einstein to propose light's constant speed in all inertial frames of reference as a fundamental principle.

Special relativity emerged from this insight that light's speed in a vacuum is constant for all observers, regardless of the light source's speed relative to each observer. This required rethinking the concepts of time intervals and distances

between points to accommodate light's fixed speed. This led to the requirement that a clock moving relative to an observer ticks slower from the observer's perspective while also being contracted in the direction of its motion. This was shown in Figure (1.9) and Figure (1.12). The faster an object moves relative to an observer, the more its length contracts and its time dilates to the observer. It also led to the simultaneity of events not being absolute. For example, light emitted from the middle of a moving truck reaches the front and back walls simultaneously for an observer at rest in the truck. But an observer standing still on the road sees the light hit both walls at different times, as shown in Figure (1.11). There is no universal "now" at a distance-observers relate times of events differently. It is only events that happen simultaneously at the same location that happen simultaneously in all frames of reference.

Classical physics is verified by our day to day experience, though this does not reflect the full reality, as relativistic effects only become readily noticeable at speeds close to that of light. With the speed of light being 300 million meters per second, astronomically faster than what occurs in our daily experience.

When it comes to velocities, they combine differently than in classical physics. However, we must ensure that the transforms reduce to the intuitive addition of velocities of classical physics at low speeds while accommodating light's constant rate in all frames of reference. This leads to initially bizarre outcomes such as **apparent** faster-than-light motion emerging from how relativistic optical effects play out across distances, which is due to the delay in light signals from objects at a distance. Nevertheless, special relativity gives a truer, more accurate picture of reality.

Due to space-time effects on wave propagation, we also have seen that the emitted light's frequencies shift (doppler effect) for a moving source, and that light is concentrated in the direction of the source's motion (aberrational effect). Together, both intensely focus light in the direction of rapidly moving objects, giving relativistic beaming.

These effects will all be explained more fully with math in the coming chapters, along with other effects that require the math to fully understand, as well as

new energy, mass, and momentum conservation laws. We will derive everything step by step, starting with the transformation of positions and times between reference frames.

1.13 Definitions

1.13.1 Mathematical Framework

Reference frame: An abstract coordinate system that is taken to be at rest consisting of three spatial and one time axis. The origin, orientation, and scale of its axis are specified by a set of points in space. The purpose of it is to provide a standardized means of defining the position and orientation of objects at any instant in time.

Inertial reference frame: A reference frame that is not undergoing any acceleration. An observer's reference frame is inertial if they do not feel a force acting on them.

Proper/Rest frame: This is the reference frame of the observer or object itself.

Primed frame: A reference frame that is moving relative to the current frame.

frame velocity: The velocity of a reference frame relative to the current reference frame.

Frame transform: The changing of coordinates and other quantities in a reference frame into their corresponding values in another reference frame.

Galilean/Classical frame transform: Transformation of coordinates and physical quantities according to Newtonian physics between two frames of reference.

Relativistic frame transform: The changing of the spatial coordinates, the time coordinate, and other physical quantities into their corresponding values in a different inertial reference frame using equations from special relativity.

Observer: An entity/thing with an associated frame of reference/system of coordinates, with respect to which you can measure objects' position, orientation,

and other properties.

Event: An occurrence that can be specified by a unique combination of the three spatial and one time coordinates. These four coordinates, together, form a point in spacetime.

Simultaneity: Within a reference frame, two or more events are simultaneous if they occur at the same time within it. Classically, if any two events are simultaneous in one reference frame, they are also simultaneous in another. This is not true in special relativity, and the order of events can depend on the frame of reference.

Retarded position: The previous position of a field's source, that an observer sees it at, due to the time delay between when a moving source emits a field and when the observer detects that field.

1.13.2 The Mechanics

Aether: A proposed medium that filled the vacuum of space that light propagated through.

Vacuum: Empty space with no matter and negligible amount of effects from fields.

Lorentz length contraction: The shortening of an object along the direction of its motion relative to an observer (relative to the object at rest).

Time dilation: The slowing of the passing of time for objects in motion relative to an observer, compared to the passing of time for the observer.

Doppler effect: The change of a wave's frequency due to the motion of its source relative to an observer. This is due to the bunching up of the wave in the direction of movement of the source and the stretching of the wave in the opposite direction, as well as the dilation of time between two parts of the emitted wave.

Aberration of light: The changing of the direction of the propagation of light when moving into another reference frame.

Relativistic beaming: The accumulation of the Doppler effect and aberration

of light, and how it affects the amount of light and its frequency emitted at different angles from a moving source. This effect can be extended to any wave or field emitted at the speed of light from a moving source.

Retarded field: A moving source's field, where at each point has been propagated at the speed of light from the retarded position of the source...

Field: ...

Light year: ...

1.13.3 Definitions for Later

Transform Invariant: ...

Postulates:

*** I could add a section about Skewing and Terrell Rotation in Special Relativity

References

- [1] Morley Williams Michelson Abraham. "On the Relative Motion of the Earth and the Luminiferous Ether". In: *American Journal of Science* 34 (1887).

Chapter 2

Position, Time and Velocity

Now that we have learned about the concepts of special relativity, we can use them and let the mathematics begin. Starting with classical relativity, which will give us an understanding of how to transform coordinates between reference frames. This will give us something to compare special relativity to later, and help with our intuition.

What we learn here will be directly used in all of the following chapters, to derive more of the foundational concepts of the theory. As everything else is all built upon this chapter, which is about how the positions and times are transformed between two inertial frames to allow the speed of light to stay constant.

In this chapter, we will start with how swapping between two inertial frames works classically. Then using the two assumed facts, the postulates, that are used as the basis of reasoning for special relativity, we will find out how to swap between frames in special relativity, by following the train of thought that we did in the previous conceptual chapter. By first deriving the time dilation, and length contraction equations in order to then find the full transform of the space and time coordinates between two inertial frames. We will come across a very useful quantity which we will denote with a γ , this quantity will be used throughout this book to simplify equations. After this, we will look at the transform for infinitesimal changes in distances and use this to derive how velocities

and accelerations change when swapping between two inertial frames. By the end of this chapter, you will understand the coordinate transforms conceptually and mathematically when swapping between two frames, and how to derive the velocity and acceleration transforms.

2.1 Classical Relativity

Classical (Galilean) relativity was the earlier theory of how to transform coordinates in different inertial frames.

It was based on two assumptions, which are that **in each frame, you would observe the following:**

- The distance between objects is the same.
- Time flows at the same rate, giving the same time for all events.

Before special relativity, it was assumed that distances between two points were the same for all observers. That is, the first observer would see a stick and measure it to be the same length as another observer who is moving relative to the first. The other assumption is that the flow of time is the same for everyone. That is, clocks move at the same rate for each observer and if, say, a particle was to emit light at a particular time, this light would be emitted at the same time for all observers.

For classical relativity, if we have the positions and times of a particle's motion relative to one observer in their frame, and we want to find the motion of the particle relative to a second observer who is moving at constant velocity relative to the first. We need to know how to transform these into this second observer's frame. That is, if we have how one observer would describe the coordinates of a particle over time relative to them and we want to now know the coordinates of the particle with respect to the other observer, we need to know how the transformation of coordinates works.

So let us see how the mathematics works for the classical coordinate system transform between two inertial observers frames, by looking at the coordinates

of an event in an initial reference frame with a second inertial frame's coordinate axis moving relative to the initial frame's, illustrated in Figure (2.1):

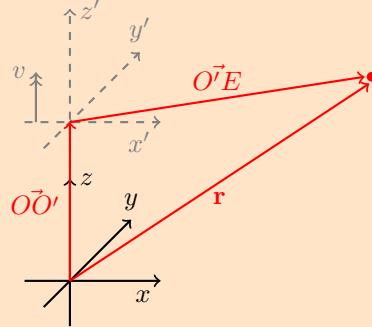


Figure 2.1: An event in initial frame with a primed frame's axis moving relative to its axis. This shows an initial frame with its associated x, y, z -coordinate system at rest in this frame, with origin O . There is also an x', y', z' -coordinate system with origin O' associated with the "primed" frame, at position $O\vec{O}' = (0, 0, vt)$ in the initial frame, where v is the speed of the primed frame in the z -direction relative to the initial frame, and t is the time in the initial frames. An event \vec{E} in this initial frame is described by a set of spatial and time coordinates $\mathbf{r} = (x, y, z)$ and t . We choose the time to be zero in both frames when both axes overlapped, so that t is also the time from when they overlapped. We also show $O'\vec{E}$ which represents the vector from the primed axis origin to the event in the initial frame.

As shown in the diagram, we are in the initial frame with origin O defined by coordinates $(0, 0, 0)$ and the primed frame is moving in the z -direction at velocity v such that the position of the primed origin relative to the initial frame's axis is

$$O\vec{O}' = (0, 0, vt) \quad (2.1)$$

at time t from when the origins overlapped. If there is an event at position

$$\mathbf{r} = (x, y, z) \quad (2.2)$$

the event is described by

$$\vec{E} = (x, y, z, t) \quad (2.3)$$

then in the primed frame, we will have the event described by the primed coordinates that we want to transform into, as

$$\vec{E}' = (x', y', z', t') \quad (2.4)$$

To swap into the primed frame's coordinates, we need to recall from the previous assumptions that distances and times are conserved, and therefore the length and directions of all the vectors between O , O' and E are the same in each frame, and time flows at the same rate ($t' = t$), from this knowledge and Figure (2.1) we have the event's coordinates described in the primed frame by its coordinate axis as

$$\mathbf{r}' = (x', y', z') = \mathbf{r} - O\vec{O}' = (x, y, z - vt) \quad (2.5)$$

This gives us the full description of the position and time transform for an event as

$x' = x$ $y' = y$ $z' = z - vt$ at time $t' = t$	
--	--

$$(2.6)$$

These are the equations that define classical relativity, we can compare this to the equations of the coordinate transforms in special relativity later. They define the event's position relative to whichever frame that we are in's axis. The special relativistic equations later will look similar to these, but with an extra factor in the z-component and with the time will no longer be consistent in both

frames.

We can also write the initial frame's event coordinates in terms of the primed coordinates, by rearranging the transformation equations giving the inverse transform as

$$\begin{aligned}x &= x' \\y &= y' \\z &= z' + vt'\end{aligned}\tag{2.7}$$

at time

$$t = t'$$

So the inverse transformation, going from primed frame to original is just a matter of swapping all primed and initial coordinates and changing the sign on the frame velocity. This can be seen at time $t = 0$ when we have the frame's origins overlapping and there is a symmetry from here, as we have the freedom to choose either frame to be the original, as no frame is particularly special and depending which one you pick the relative velocity is either v or $-v$. We will have the same rules in special relativity later on when we want to get the inverse transform.

You might have noticed that the direction of the second frame is in the upward z-direction instead of being to the right in the case of the truck diagrams from the previous chapter, this is to help us when we use spherical polar coordinates later and to help see the symmetries. But that is a couple of chapters away, for now we will look at the assumptions of special relativity and their consequences, which we will then apply to the same notion of an event that we have in the Figure (2.1).

2.2 Postulates of Special Relativity

For classical relativity, we started with two assumptions. It turns out that these assumptions need to be thrown away, as they contradict the experimental

evidence that the speed of light is constant for every observer and frame. We replace them with two new assumptions, called the **postulates of special relativity**:

1. The speed of light in a vacuum is the same in all inertial/non-accelerating frames of reference.
2. The laws of physics and outcome of experiments are the same in all inertial frames of reference. If an observer in an inertial reference frame carries out an experiment, they would find the same result if they had carried it out within a different inertial frame instead.

The first postulate here will be the most important to us in deriving the transform equations of special relativity and will be used to show how time is dilated/slowed down for moving objects. The second postulate has a consequence that if events happen simultaneously at an overlapping position in one frame, then they will also happen simultaneously in an overlapping position in any other inertial frame. We will now use these assumptions with the concept of a light clock to show the effect of length contraction.

2.3 Time Dilation

Time passes more slowly for things moving relative to you, this is called time dilation. To explain why it happens we will use the simple concept of a light clock. It consists of light bouncing between two mirrors. It keeps time by measuring when light returns to the first mirror, counting it as one tick of its clock, as shown in Figure (2.2).

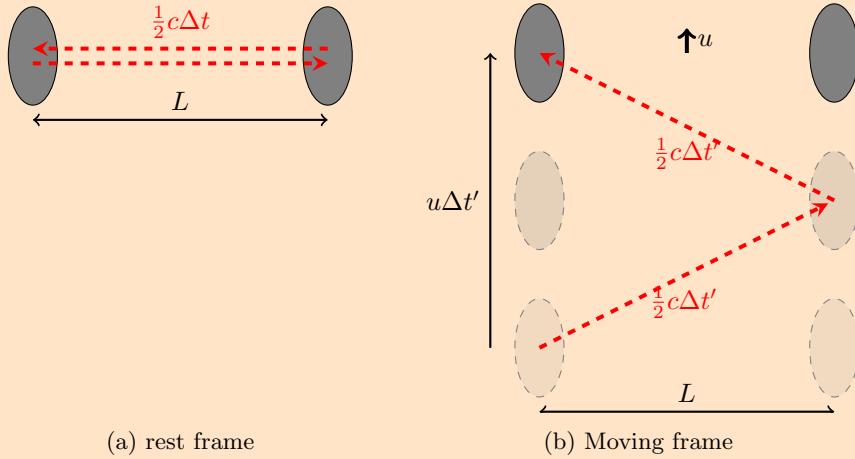


Figure 2.2: **Light clock showing time dilation.** The light clock keeps time by counting each time the light makes a return journey when bouncing between two mirrors with length $L = \frac{1}{2}c\Delta\tau$ between them, were $\Delta\tau$ is the time for the return journey in the rest frame, the diagram shows the mirror's rest frame (left) and a frame with them moving at speed $u = -v$ (right). In this second frame, there is a distance $u\Delta t'$ between where the light was initially reflected by the left mirror and where it returned to it. For this, the light had to travel a total diagonal distance of $c\Delta t'$ to return, were $\Delta\tau$ is the time for the return journey in the primed frame.

The light in the clock, takes a time $\Delta\tau$ in the clock's rest frame to make the return journey. We can see from Figure (2.2a) that in the rest frame, the distance the light travels to the mirror and back in this time is

$$2L = c\Delta\tau. \quad (2.8)$$

The distance the mirrors move during this return journey in the moving frame is

$$u\Delta t' = -v\Delta t', \quad (2.9)$$

where $\Delta t'$ is the time it took light to return in the moving frame and u is the speed of the light clock, which is in the opposite direction of the movement of the primed frame relative to the rest frame, that is $u = -v$.

The distance light travels in this time is

$$c\Delta t' \quad (2.10)$$

This is the hypotenuse of a right angled triangle formed by $u\Delta t' = -v\Delta t'$ and $2L$, so we have the square of the hypotenuse equal to the square of the other two lengths:

$$c^2\Delta t'^2 = v^2\Delta t'^2 + c^2\Delta\tau^2 \quad (2.11)$$

rearranging to find the time it took in the primed frame relative to the rest frame, we find

$$\Delta t' = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}\Delta\tau = \gamma\Delta\tau \quad (2.12)$$

with

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (2.13)$$

This factor, γ , is used throughout special relativity to simplify equations. We can see from the diagram and the equations, that the time for the return journey in the frame with the clock moving is longer relative to its rest frame,

$$\Delta t' \geq \Delta\tau \quad (2.14)$$

We take this outcome as the time taken between two events at the same position in the moving frame takes longer than in the rest frame of the clock, due to the longer distance light must travel at the same speed. Hence systems that are moving seem to have their time flowing slower. One example of this, is that unstable particles that are moving faster have longer livespans before they decay.

2.4 Length Contraction

An object's size gets smaller in the direction of movement compared to when it is at rest, this is called length contraction. To explain why it happens we will add a second perpendicular light path to previous section's experimental setup. So there is now light moving perpendicular as well as parallel the the primed frame's movement, as shown in Figure (2.3)

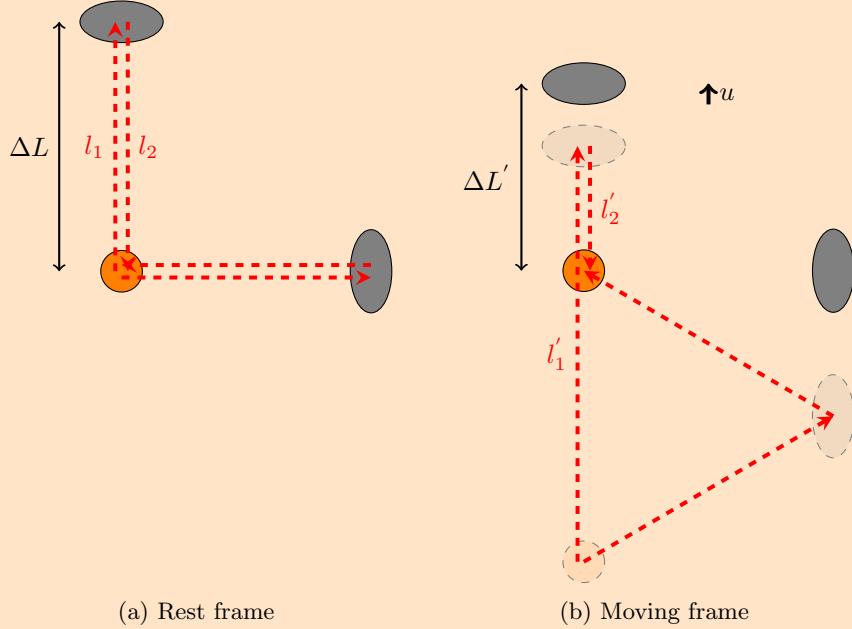


Figure 2.3: **Light paths showing the need for length contraction.** A central orange bulb emits light in two directions, with all light being reflected by the mirrors back to reach the central bulb simultaneously. This is shown at rest, fig (a), and in another inertial frame with the system now moving at speed u , fig (b), with its length now having to be contracted in the direction of movement to allow for the simultaneous return of light to the bulb. The length between the bulb and front mirror is shown as ΔL and $\Delta L'$ in each frame. The paths the light took to the front mirror l_1 and back l_2 are shown in the rest frame with the corresponding paths l'_1 and l'_2 in the moving frame.

If we have a system as described in Figure (2.3). Then since the light returns simultaneously in the rest frame, we need the light to also return simultaneously in the second frame. Multiple events that happen at a single point simultane-

ously in one frame, must happen simultaneously at a single point in all other frames. This is to allow for the second postulate to hold true. The time between the light being emitted and absorbed will be the dilated time that we worked out in the previous section. This can be seen as the sideways light path in this and the light clock's setup from previous section is the same path. In the moving frame, to achieve this simultaneity in the return of the light to the bulb, the length of the full path of light from the central bulb to the front mirror and back ($2\Delta L' = l'_1 + l'_2$) must be the same as the perpendicular path, due to the speed of light being the same for both paths. We can work out the length of the sideways paths from the time dilation section, $c\Delta t'$, and this is the length the total path needs to be in the forward directions as well. We will find that the paths can only have this length, if the distance between the bulb and the front mirror is contracted in the direction of movement. Note that there can not be a contraction in the perpendicular direction to the movement due to the paradox explained in Figure (1.10).

In the moving frame, the pulse sent forward has a path length equal to the speed of light times the time taken for the path. Which is also equal to the primed frame's distance between the bulb and mirror plus the distance the mirror has moved away from the point it was emitted in this time, that is

$$l'_1 = c\Delta t'_1 = \Delta L' + u\Delta t'_1 \quad (2.15)$$

where u is the speed that the front mirror is moving away from the point it was emitted, $\Delta t'_1$ is the time taken to get to the mirror, and $\Delta L'$ is the distance between the bulb and the mirror in front in the primed frame. After reaching the mirror, the light then must travel back to the bulb a distance

$$l'_2 = c\Delta t'_2 = \Delta L' - u\Delta t'_2. \quad (2.16)$$

Where $\Delta t'_2$ is the time taken to return to the bulb from the front mirror, with the source moving twowards the reflected light.

Both of these give the two times as

$$\Delta t'_1 = \frac{\Delta L'}{c - u} \quad (2.17)$$

and

$$\Delta t'_2 = \frac{\Delta L'}{c + u}. \quad (2.18)$$

Adding both these times together gives the total dilated time of return as

$$\Delta t' = \Delta L' \left(\frac{1}{c + u} + \frac{1}{c - u} \right) = \frac{\Delta L'}{c} \frac{2}{1 - u^2/c^2} = 2\gamma^2 \frac{\Delta L'}{c} \quad (2.19)$$

where we used $u = -v$ and then the γ factor to simplify equation. Now from this and using the time dilation equation, we have

$$c\Delta t' = c\gamma\Delta t = 2\gamma\Delta L = 2\gamma^2\Delta L' \quad (2.20)$$

Finally, this leads to the length from the bulb to the front mirror in the primed frame to be

$$\Delta L' = \frac{1}{\gamma} \Delta L. \quad (2.21)$$

This means that for a system moving relative to an observer, the distances for the system in the direction of movement are contracted compared to the rest frame of the object, that is

$$\Delta L' \leq \Delta L. \quad (2.22)$$

Objects and the distance between them are therefore shorter in a frame with them moving compared to in the rest frame of the objects.

With this knowledge and the last section, we can now derive the coordinate transform equations.

2.5 Position and Time Transformation

In the last two sections we have looked at the dilation between two different times at the same location and the contraction between two different positions at the same time, when we swap between frames. Now we will use both these concepts to help us see how the combination of the time and the position coordinate, known as an event, transforms when we swap frames. For the derivation of the coordinate transformation of an event between two inertial frames, we will be starting with the same diagram as we had in the classical relativity in Section (2.1) and apply what we have learned from the time dilation and length contraction sections so that we can be consistent with the postulates of special relativity. This will allow us to derive the coordinate transformation equations, known as the Lorentz transformation.

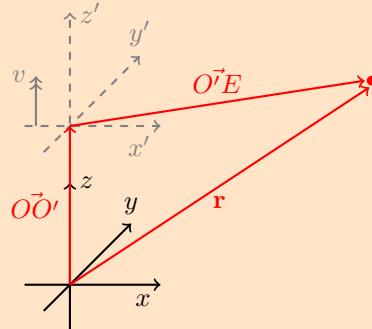


Figure 2.4: An event in initial frame with a primed frame's axis moving relative to it. This shows an initial frame with its associated x, y, z -coordinate system at rest in this frame, with origin O . There is also an x', y', z' -coordinate system with origin O' associated with the "primed" frame, at position $\vec{OO'} = (0, 0, vt)$ in the initial frame, where v is the speed of the primed frame in the z -direction relative to the initial frame, and t is the time in the initial frames. An event \vec{E} in this initial frame is described by a set of spatial and time coordinates $\mathbf{r} = (x, y, z)$ and t . We choose the time to be zero in both frames when both axes overlapped, so that t is also the time from when they overlapped. We also show $\vec{O'E}$ which represents the vector from the primed origin to the event in the initial frame.

As shown in the diagram, we are in the initial frame with origin O defined by coordinates $(0, 0, 0)$ and the primed frame is moving in the z -direction at

velocity v such that the position of the primed origin in the initial frame is

$$\vec{OO'} = (0, 0, vt) \quad (2.23)$$

at time t from when the origins overlapped. If there is an event described by

$$\vec{E} = (x, y, z, t) \quad (2.24)$$

Then the displacement between the initial frame's origin and the event's position in the initial frame is

$$\mathbf{r} = (x, y, z) \quad (2.25)$$

and displacement between the primed frame's origin and the event's coordinates in the initial frame is

$$\vec{O'E} = (x, y, z - vt) \quad (2.26)$$

In the primed frame, we will have the event described by the primed coordinates that we want to transform into, as

$$\vec{E}' = (x', y', z', t') \quad (2.27)$$

with the displacement between the primed frame's origin and the event's position in the primed frame given as

$$\mathbf{r}' = (x', y', z') \quad (2.28)$$

These are just the position coordinates of the event in the primed frame.

The primed frame's axis and coordinates are moving relative to the initial frame, so from the previous length contraction section we would have the displacement between the primed axis origin and the event in the initial frame $\vec{O'E}$, length contracted, in the z -direction, compared to the primed position coordinates in the primed frame \mathbf{r}' , that is

$$O'E = (x, y, z - vt) = (x', y', \frac{z'}{\gamma}) \quad (2.29)$$

The x and y -components are unaffected, with the paradox of it they were effected, explained in Figure (1.10). Rearranging for the z -component, we have

$$(x', y', z') = \mathbf{r}' = (x, y, \gamma(z - vt)) \quad (2.30)$$

There is a symmetry between the frames because if we were originally in the primed frame we would have the initial frame moving instead at $-v$ in the z -direction relative to the primed frame and we could follow the above steps starting with this and the primed coordinate instead, and find the inverse transformation is the same apart from it replaces each frame's coordinates with the other and has the negative of the previous frame velocity instead. That is

$$(x, y, z) = \mathbf{r} = (x', y', \gamma(z' + vt')) \quad (2.31)$$

We can substitute both z -coordinate transforms into each other to find the time transform

$$z = \gamma(\gamma(z - vt) + vt') \quad (2.32)$$

and rearranging gives

$$t' = \gamma \left(\left(\frac{1}{\gamma^2} - 1 \right) \frac{z}{v} + t \right) \quad (2.33)$$

with $1/\gamma^2 = 1 - v^2/c^2$ leading to the primed frames time for the event given in initial frames coordinates as

$$t' = \gamma \left(t - \frac{vz}{c^2} \right) \quad (2.34)$$

again if we want the inverse transform from primed to initial frame we can swap the primed and initial frame coordinates and put an extra minus in front of the frame velocity.

$$t = \gamma \left(t' + \frac{vz'}{c^2} \right) \quad (2.35)$$

We now have the full description of the transformation of the position and time of an event as

$$\begin{aligned} x' &= x \\ y' &= y \\ z' &= \gamma(z - vt) \end{aligned}$$

at time

$$t' = \gamma \left(t - \frac{vz}{c^2} \right) \quad (2.36)$$

from the overlap of coordinate axis.

with

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

These are the equations that define special relativity. They allow us to find an event's position relative to whichever frame that we want to swap into's axis.

In the time transform equation the $\frac{vz}{c^2}$ term is the cause of the simultaneity of events at two different positions in one inertial frame no longer being simultaneous in another inertial frame, and the γ factor, is the term that causes the overall change in how fast time flows at a point relative to the initial frame. In the spatial coordinate transform equations the vt factor is due to both axes' origins moving away from each other and the γ factor is the cause of the change of lengths in the direction of the frame's movement.

We can see how points representing the coordinates in an initial frame look like in a primed frame of reference, by looking at Figure (2.5).

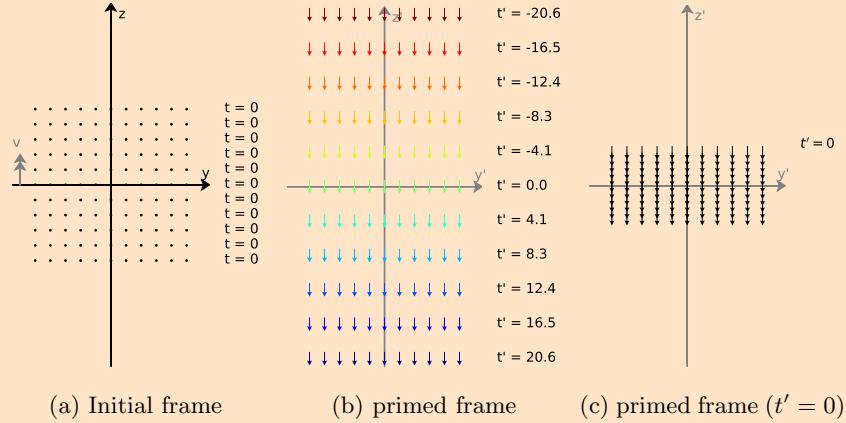


Figure 2.5: Transformation of coordinates into a primed frame. A diagram, showing a grid of points, representing the initial frame's coordinates, at time $t = 0$ that are at rest in their frame (left), giving the direct transformation, using the transform equations, into the moving primed frame (middle) where the points are now moving in the negative z' -direction with $z' = \gamma z$, the colors and labels show the desynchronization of the primed times $t' = -\frac{v}{c^2} \gamma z$, the next figure (right) shows where these points would be when they are synchronized with the primed origin's time $t' = 0$, which can be worked out to be ($z' = \frac{z}{\gamma}$).

This diagram shows stationary points representing coordinates in the initial frame. In the primed frame, these points are transformed to the position $z' = \gamma z$ and are now moving at speed $-v$, they are transformed using the transformation equations to points that are more spread out in the primed frame, but the times they are located at each of these points are different, they have desynchronized times. Therefore to find these points, which represent the coordinates in the initial frame, now in the primed frame at the synchronized time of $t' = 0$, we must propagate the points to that time, which ends up being $z' = \frac{z}{\gamma}$, as shown in figure (2.5c).

This figure is to give you a little bit more intuition on how the coordinates and swapping between frames work. We will later see a more natural interpretation of the coordinate transforms when we introduce the notion of retarded time in Chapter (3), which will give us another reason for why the time transformation is the way it is.

2.5.1 Transforms with Small Primed Frame Velocity

If we have the primed frame velocity much smaller than the speed of light, ($v \ll c$), then we have $v/c \approx 0$ and therefore $\gamma \approx 1$. Hence we get back the previous classical transform equations

$$t' \approx t \quad (2.37)$$

$$z' \approx z - vt \quad (2.38)$$

which is the expected result, as it agrees with the everyday observations we are used to seeing at these speeds. That is, no noticeable difference in times on clocks or lengths of moving objects. We can see when the γ -factor's influence starts to become significant from Figure (2.6).

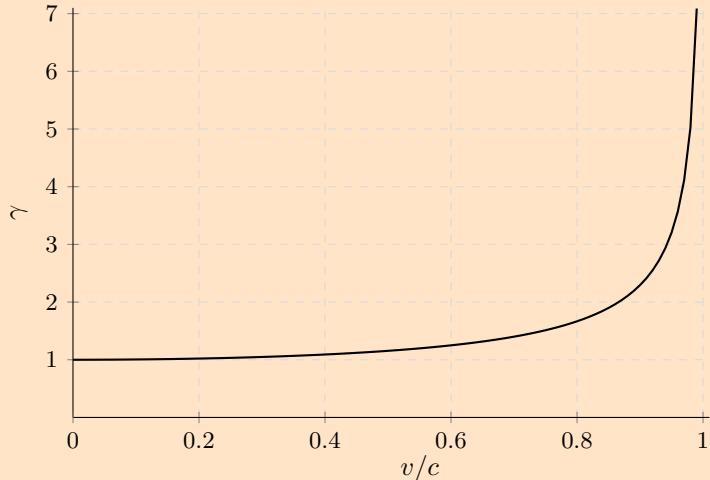


Figure 2.6: **γ -factor.** A diagram, showing the magnitude of γ with increasing ratio of primed frame's speed to the speed of light, with γ tending to infinity as $v/c \rightarrow 1$.

As can be seen, it is not until around when the primed frame velocity is a third of the speed of light, that it significantly differs from being equal to 1. This is a speed unattainable to us, the only thing we can bring up close to this speed

is particles, and nothing humanmade is even a 1000th the speed of light.

2.5.2 Space and Time Intervals

To get the velocity transform equations we will first need the equations to transform the displacement between two positions and the interval between times, then take them in the limit as the changes go to zero. That is, when we have infinitesimal changes of coordinates. If we want to know how the displacement between two coordinates $\vec{R}_1 = (x_1, y_1, z_1)$ and $\vec{R}_2 = (x_2, y_2, z_2)$ transforms, we can take the difference of the transforms of both the coordinates as follows

$$\begin{aligned} x'_2 - x'_1 &= x_2 - x_1 \\ y'_2 - y'_1 &= y_2 - y_1 \\ z'_2 - z'_1 &= \gamma(z_2 - vt_2) - \gamma(z_1 - vt_1) \\ t'_2 - t'_1 &= \gamma\left(t_2 - \frac{v}{c^2}z_2\right) - \gamma\left(t_1 - \frac{v}{c^2}z_1\right) \end{aligned} \tag{2.39}$$

rearranging and replacing the difference of the first and second of each coordinate with $\Delta\vec{R} = (\Delta x, \Delta y, \Delta z) = (x_2 - x_1, y_2 - y_1, z_2 - z_1)$, giving the displacement transform as

$$\begin{aligned} \Delta x' &= \Delta x \\ \Delta y' &= \Delta y \\ \Delta z' &= \gamma(\Delta z - v\Delta t) \\ \Delta t' &= \gamma\left(\Delta t - \frac{v}{c^2}\Delta z\right) \end{aligned} \tag{2.40}$$

This is the same form as the original coordinate transform equations.

From these equations, we can also get the differential quantities when we take infinitesimal changes, which give the same form again, as

$$\begin{aligned}
dx' &= dx \\
dy' &= dy \\
dz' &= \gamma(dz - vdt) \\
dt' &= \gamma \left(dt - \frac{v}{c^2} dz \right)
\end{aligned} \tag{2.41}$$

So when we make an infinitesimally small change in position and/or time within a frame, we know what the corresponding change is in a primed frame, we will use these to get the velocity transform equations. As velocity at any instance is the infinitesimal change in position per infinitesimal change in time. These will also be used later to find, invariant quantities, quantities that remain the same in each frame. For example, the space-time interval and the energy-momentum relationship in Chapter (5).

2.6 Velocity

The velocity at any instant is the infinitesimal change in position over the infinitesimal change in time $\frac{d\vec{R}}{dt}$. So using the infinitesimals from Equation (2.41), we get the velocity in each direction as

$$\begin{aligned}
u_x' &= \frac{dx'}{dt'} = \frac{dx}{\gamma \left(dt - \frac{v}{c^2} dz \right)} \\
u_y' &= \frac{dy'}{dt'} = \frac{dy}{\gamma \left(dt - \frac{v}{c^2} dz \right)} \\
u_z' &= \frac{dz'}{dt'} = \frac{\gamma(dz - vdt)}{\gamma \left(dt - \frac{v}{c^2} dz \right)}
\end{aligned} \tag{2.42}$$

now dividing the top and bottom of the fraction on the right hand side by dt we can give the equations in their vector form as

$$\mathbf{u}' = \frac{1}{\gamma \left(1 - \frac{v}{c^2} u_z\right)} \begin{pmatrix} u_x \\ u_y \\ \gamma(u_z - v) \end{pmatrix} \quad (2.43)$$

at position \mathbf{r}' and time t'

This is used to find the velocity of anything in another frame at the corresponding primed position and time. We will put this equation to work in a later section when we look at how the light emitted from a source particle has its velocity direction changed/aberrated and the effect this has on its intensity in all directions. Next, we will use this to get the acceleration transform.

Side note:

For those interested; The generalized velocity transform (which we will not derive here) for a frame velocity in any direction and not just the z-direction is

$$\mathbf{u}' = \frac{1}{\gamma} \frac{\mathbf{u} + \left[\frac{\gamma - 1}{\|\mathbf{v}\|^2} (\mathbf{u} \cdot \mathbf{v}) - \gamma \right] \mathbf{v}}{1 - \frac{\mathbf{u} \cdot \mathbf{v}}{c^2}} \quad (2.44)$$

with \mathbf{v} being the frame's velocity in any generalized direction.

2.7 Acceleration

To find how the acceleration (known as the 3-acceleration) of a particle transforms for observers in two different frames, we take the differential of the velocity transform in equation (2.43) and use the differentiation rules for two generic functions f and g : $d(gf) = f dg + g df$, and $d[f(g(x))] = dg(x) * df(g(x))$

$$\begin{aligned}
d\mathbf{u}' &= \frac{1}{\gamma \left(1 - \frac{v}{c^2} u_z\right)} \begin{pmatrix} du_x \\ du_y \\ \gamma du_z \end{pmatrix} + \frac{\gamma \frac{v}{c^2} du_z}{\gamma^2 \left(1 - \frac{v}{c^2} u_z\right)^2} \begin{pmatrix} u_x \\ u_y \\ \gamma(u_z - v) \end{pmatrix} \\
&= \dots \\
&= \frac{1}{\gamma \left(1 - \frac{v}{c^2} u_z\right)^2} \begin{pmatrix} du_x + \frac{v}{c^2}(u_x du_z - u_z du_x) \\ du_y + \frac{v}{c^2}(u_y du_z - u_z du_y) \\ \frac{1}{\gamma} du_z \end{pmatrix}
\end{aligned} \tag{2.45}$$

Now, divide by dt' as we did for the velocity in the previous section, and we get the acceleration transform as

$$\mathbf{a}' = \frac{1}{\gamma^2 \left(1 - \frac{v}{c^2} u_z\right)^3} \begin{pmatrix} a_x + \frac{v}{c^2}(u_x a_z - u_z a_x) \\ a_y + \frac{v}{c^2}(u_y a_z - u_z a_y) \\ \frac{1}{\gamma} a_z \end{pmatrix} \tag{2.46}$$

at position \mathbf{r}' and time t'

So if we know how a particle accelerates in one frame we will know its corresponding acceleration in another frame at the corresponding primed time and position.

Side note:

Again for those interested; The generalized acceleration transform (which we will not derive here) for a frame velocity in any direction and not just the z-direction is

$$\mathbf{a}' = \frac{1}{\gamma^3 \left(1 - \frac{(\mathbf{u} \cdot \mathbf{v})}{c^2}\right)^3} \left[\gamma \left(1 - \frac{(\mathbf{u} \cdot \mathbf{v})}{c^2}\right) \mathbf{a} + \frac{\gamma(\mathbf{a} \cdot \mathbf{v})}{c^2} \mathbf{u} + (1 - \gamma)(\mathbf{a} \cdot \hat{\mathbf{v}}) \cdot \hat{\mathbf{v}} \right] \tag{2.47}$$

at position \mathbf{r}' and time t'

This is the 3-acceleration, not to be confused with the 4-acceleration that will

be mentioned in a later chapter, which is a more abstract quantity.

2.8 Summary

We derived the mathematics for the transformation of the spatial and time coordinates for classical relativity to help us understand the derivation of special relativity. We then introduced the assumptions, known as the postulates of special relativity, which we use to derive how time slows down due to time dilation and lengths of moving objects get smaller due to length contraction. This allowed us to derive the transformation equations, which in turn allowed us to calculate the velocity and acceleration transforms. We were also introduced to γ , which is a factor used throughout relativity to simplify equations.

From this chapter, we will now be able to transform from one frame into another, in special relativity. If you have a system of particles in an initial frame and you want to transform their positions to a new frame, you need to first define where this initial axis' origin is. You are free to choose where it is. The primed frame's axis will overlap with this initial frame's axis at your chosen starting time, that is when you define the time as $t = 0$ (the time of the overlap is also $t' = 0$ in the primed frame). You can now transform the positions into the primed frame. These particles have these primed frame positions at the calculated primed frame times, these are defined in Equations (2.36). To transform the velocities and acceleration to these primed frame positions at the corresponding primed frame times, you can use equations (2.46) and (2.43).

An easy mistake to make is to forget that t and t' are the times from the overlap of the frame origins, in their respective frames.

This chapter shows the foundations of special relativity, but there is still more to come. Before we go further though, we will take a look at the delayed view of the world to an observer due to the delay in light from their surrounding. This will help us understand why we need the time transformation to be the way it is.

2.9 Variables

$'$: This apostrophy on a quantity represents the corresponding quantity in a second primed reference frame.

O, O' : Position of proper and primed axis' origin, given by the set of spatial and time coordinates.

E : An event given by the set of spatial and time coordinates x, y, z, t .

$\mathbf{r} = (x, y, z)$: The position coordinates of an event in the current inertial reference frame

$\mathbf{r}' = (x', y', z')$: The corresponding position coordinates of an event in another inertial reference frame, referred to as the primed frame, and denoted with an apostrophe $'$.

$t, t' : \dots$

v : The speed of the second reference frame, primed frame, relative to the first reference frame.

$\mathbf{v} = v \cdot \hat{\mathbf{v}}$: The velocity of the second reference frame, primed frame, relative to the first reference frame, v is the speed and $\hat{\mathbf{v}}$ is the directional unit vector.

c : The speed that light travels in empty space when it is not impeded by matter and fields.

γ : A relativistic factor used to simplify equations, it is given by $1/\sqrt{1 - v^2/c^2}$ which determines how much of an effect the transform of the coordinates deviates from the classical transform. It is only significant at relative frame speeds close to that of light.

$\Delta\tau$: The time between events in the rest frame of object.

$(\Delta x, \Delta y, \Delta z, \Delta t), (\Delta x', \Delta y', \Delta z', \Delta t') : \dots$

$d\mathbf{r} = (dx, dy, dz), dt : \dots$

$\mathbf{u} = (u_x, u_y, u_z), \mathbf{u}' = (u'_x, u'_y, u'_z) : \dots$

$d\mathbf{u}, d\mathbf{u}' : \dots$

$\mathbf{a} = (a_x, a_y, a_z), \mathbf{a}' = (a'_x, a'_y, a'_z) : \dots$

Chapter 3

Observers Delayed Worldview

Light travels extremely fast compared to other everyday speeds we are used to, roughly at 300,000,000 m/s. In the blink of an eye, light could travel around the world twice. We do not need to normally worry about the delay in the light signal, as the time to propagate the distances in our everyday world is extremely short, so the signal can be approximated as being instantaneous, ignoring the more complex modeling of the delayed signal. But when it comes to particle physics or astronomy, we might be dealing with speeds close to that of light or huge distances. For example, light has taken 152,000 years to travel from the Andromeda galaxy, so we are currently seeing this galaxy where it was and how it looked 152,000 years ago. These past positions that we currently see are called the retarded/delayed view positions. The topics in this chapter are not normally included in Special relativity texts. But they are needed to get you a more full story and understanding of relativity. They come with extra insights into the reason for why the coordinate transform equations have to be the way that they are. As we will see when transforming retarded coordinates.

3.1 Relativistic Observers Delayed view

When looking at moving objects we see them where they were in their past positions, due to the delay in the light signal from them. So we will look at what the transform would be if we wanted to transform this delayed view position that is called the retarded coordinate, that consists of the retarded position and retarded time of this position. Imagine that all coordinates act like a source of emitting light. Let us describe a coordinate emitting a photon to an observer at the origin, showing the retarded position the observer sees the coordinate source at.

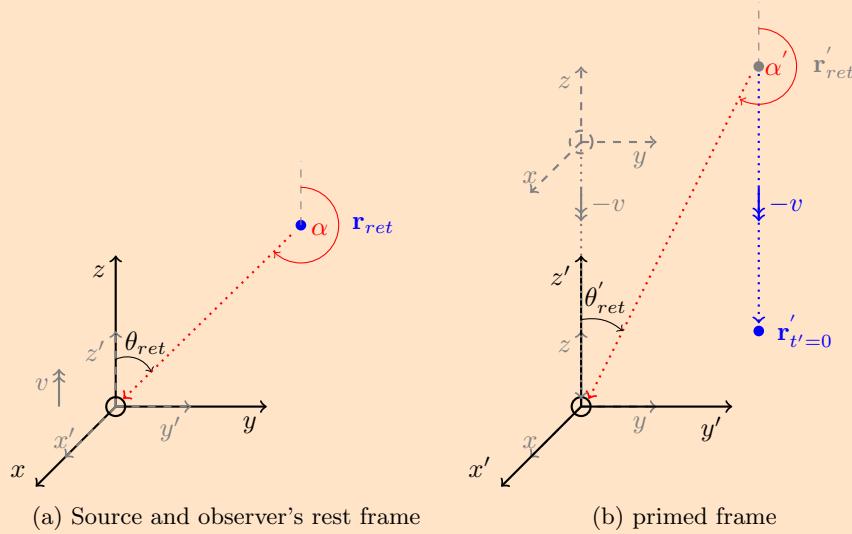


Figure 3.1: Retarded coordinates. Diagram of a stationary source of light that represents the coordinate $\mathbf{r} = \mathbf{r}_{ret}$ (blue) emitting light (red) at an angle α and α' towards the origin, which is received by an observer positioned at the origin at time $t = t' = 0$ with the observer and source at rest in the first inertial frame (left) and then shown moving in a primed frame moving at speed v in the z -direction (right). The past, retarded position of the rest frames coordinate \mathbf{r}'_{ret} at its retarded time is shown in gray, with its current position $\mathbf{r}'_{t'=0}$ shown in blue. The path light takes is shown in red. The corresponding axis in each frame is also shown.

As shown in the diagram, if we take the rest frame of an observer and a source that emits a pulse of light from its position at \mathbf{r} so that it is received by the

observer at the origin, at time $t = 0$. The retarded time the pulse of light was emitted is $t_{ret} = -\frac{\|\mathbf{r}\|}{c}$, which is the negative of the time it took for light to propagate along \mathbf{r} to the origin. The velocity of the light that propagated to the origin is

$$\mathbf{c} = c \begin{pmatrix} 0 \\ \sin \alpha \\ \cos \alpha \end{pmatrix} \quad (3.1)$$

Where α is the direction of propagation. This gives the retarded position the light was emitted from as

$$\mathbf{r}_{ret} = \mathbf{r} = ct_{ret} \begin{pmatrix} 0 \\ \sin \alpha \\ \cos \alpha \end{pmatrix} \quad (3.2)$$

The source's position and retarded position are the same in the rest frame as it is not moving. Also, the position vector and light's velocity vector are in opposite directions, The emitted time being negative flips the direction of the velocity vector here. Now inputting the retarded position and times in the first frame into the coordinate transform we get the retarded position and time in the primed frame as

$$\mathbf{r}'_{ret} = ct_{ret} \begin{pmatrix} 0 \\ \sin \alpha \\ \gamma (\cos \alpha - \frac{v}{c}) \end{pmatrix} = \begin{pmatrix} 0 \\ y \\ \gamma (z - vt_{ret}) \end{pmatrix} \quad (3.3)$$

at retarded time

$$t'_{ret} = \gamma \left(1 - \frac{v}{c} \cos \alpha \right) t_{ret}$$

If instead we use the angle of the retarded coordinate from the z-axis, $\theta_{ret} = \alpha - \pi$, we have that $\cos \alpha = \cos \theta_{ret} + \pi = -\cos \theta_{ret}$ and therefore the retarded time is given as

$$t'_{ret} = \gamma \left(1 + \frac{v}{c} \cos \theta_{ret} \right) t_{ret} \quad (3.4)$$

With the light being received at the origin of the primed frame axis at time $t' = 0$.

The coordinate transform here for retarded coordinates is the same as normal coordinate transform, except the time transform simplifies and allows us to take the retarded time out of brackets with just the factor $\bar{\Lambda} = \gamma (1 - \frac{v}{c} \cos \alpha) = \gamma (1 + \frac{v}{c} \cos \theta_{ret})$ now in front of it. This factor will also come up in the doppler effect, in aberration and when finding the flux of a field. We will use it to simplify equations like we did with γ . We will call $\bar{\Lambda}$ the aberrational factor, but it is also referred to as the relativistic doppler factor.

The retarded view of the world is due to the delay in the light from the world. So at the moment all the light that is being received by the origin is what it sees. When we transform the retarded coordinates when swapping to another frame, the angle the light is being received now to the primed frames origin, which is at the same location at $t = t' = 0$, is changed, this is called the relativistic aberration of light. This aberration will be described fully in the next chapter.

Side Note:

This retarded position is at the retarded time, if you wanted to find the position of where the source currently is, we would need to propagate this to the current time of zero of the observer, which after a bit of working out gives

*** double check this formula:

$$\mathbf{r}'_{t'=0} = \mathbf{r}'_{ret} - \mathbf{v}t'_{ret} = ct_{ret} \begin{pmatrix} 0 \\ \sin \alpha \\ \frac{1}{\gamma} \cos \alpha \end{pmatrix} \quad (3.5)$$

3.2 Transformed Retarded Coordinates

This section here is just a visualisation of how spatial coordinates of an initial frame transform into their retarded positions in a primed frame, showing how

they are spread, as shown in figure (3.2).

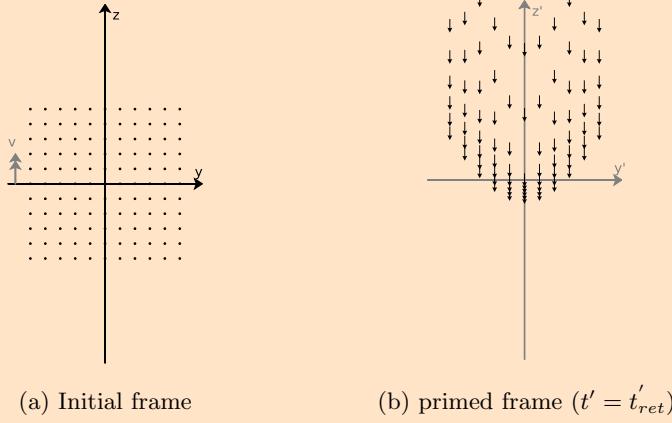


Figure 3.2: Retarded coordinate transform. A diagram, showing an initial small grid of points, representing some position coordinates that are at rest in their proper frame, with an observer at the origin (left), using the retarded time $t_{ret} = \frac{\|\mathbf{r}\|}{c}$ for these coordinates, we can transform these to their retarded positions in a primed frame using the coordinate transform, with the original frames coordinates now moving at velocity $-v = -0.9$, the initial frames coordinates are transformed to where they would be seen to be at by an observer at the origin with the observers time $t' = 0$, these positions are due to the delay in the light from these points taking time to get to the origin, they are also known as the retarded positions.

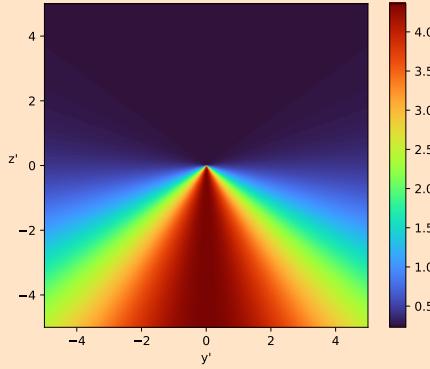


Figure 3.3: Density of initial frame's retarded coordinates in primed frame. This shows the density of the initial frames coordinates when they are in their retarded positions in the primed frame, with $v = 0.9$, and the scale shown on the right hand side is the density.

if the initial frame is where we see everything at when we are still, then if there was another observer overlapping with us moving a v they would see all the points distributed as they are shown in their primed frame.

3.3 Perceived Velocity

From Section (1.11.1) we seen that objects can be perceived to be going faster than light even if they are not. Here we will show how to find the perceived speed of objects.

So if we have a moving particle positioned at \mathbf{r} , in a infinitesimal time dt it will move

$$d\mathbf{r} = \mathbf{u} dt \quad (3.6)$$

as shown in the diagram

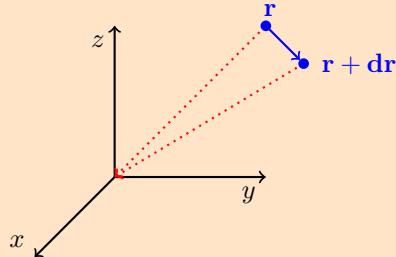


Figure 3.4: **Perceived velocity.** Diagram showing a particle (blue) moving from \mathbf{r} by displacement of $d\mathbf{r} = \mathbf{u} dt$ and the paths that light (red) has to travel to get to the origin from both these positions. leading to a different retarded time between the positions than the actual time difference. This gives a different perceived velocity.

The difference between the retarded times of light from each position reaching the origin is

$$dt_{ret} = \left(dt + \frac{\|\mathbf{r} + d\mathbf{r}\|}{c} \right) - \frac{\|\mathbf{r}\|}{c} \quad (3.7)$$

we can simplfy this expression by using

$$\begin{aligned}
\|\mathbf{r} + d\mathbf{r}\| &= \sqrt{(\mathbf{r} + d\mathbf{r}) \cdot (\mathbf{r} + d\mathbf{r})} \\
&= \sqrt{\mathbf{r} \cdot \mathbf{r} + 2\mathbf{r} \cdot d\mathbf{r} + d\mathbf{r} \cdot d\mathbf{r}} \\
&\approx \sqrt{\mathbf{r} \cdot \mathbf{r} + 2\mathbf{r} \cdot d\mathbf{r}} \\
&= \|\mathbf{r}\| \sqrt{1 + \frac{2\mathbf{r} \cdot d\mathbf{r}}{\mathbf{r} \cdot \mathbf{r}}} \tag{3.8}
\end{aligned}$$

using the Taylor expansion $\sqrt{1+x} \approx 1 + \frac{x}{2}$ for $x \ll 1$:

$$\approx \|\mathbf{r}\| \left(1 + \frac{\mathbf{r} \cdot d\mathbf{r}}{\|\mathbf{r}\|}\right) = \|\mathbf{r}\| + \mathbf{r} \cdot d\mathbf{r}$$

giving the precieved time between these events at the origin as

$$\begin{aligned}
dt_{ret} &= dt + \frac{r d\mathbf{r}}{\|\mathbf{r}\| c} \\
&= dt + \frac{\hat{\mathbf{r}} \cdot \mathbf{u} dt}{c} \\
&= \left(1 + \frac{\mathbf{u} \cdot \hat{\mathbf{r}}}{c}\right) dt \tag{3.9}
\end{aligned}$$

The distances between the points is the same to the observer as the true distance between them, its just the precieved times that are different. Therefore we can get the precieved/retarded velocity to be

$$\mathbf{u}_{ret} = \frac{d\mathbf{r}}{dt_{ret}} = \frac{d\mathbf{r}}{(1 + \frac{\mathbf{u} \cdot \hat{\mathbf{r}}}{c}) dt} = \frac{\mathbf{u}}{1 + \frac{\mathbf{u} \cdot \hat{\mathbf{r}}}{c}} \tag{3.10}$$

For speeds of particles between zero and the speeds of light, the denominator can range between zero and two meaning that we can have perceived speeds between zero and infinity. For something moving at the speed of light, the slowest we could precieve it, which is when its directly moving away, is $0.5c$. This section is to just show an interesting concept, that should be taken into account when making some experimental observations.

3.4 Summary

Here we introduced the retarded positions that are due to the delay in light signal to an observer, and there associated retarded times and derived the transform of these retarded coordinates to the prime frames coordinates. The retarded coordinate transform and also the density of the transformed intial frames retarded coordinates in the primed frame were shown in figure (3.2) and figure (3.3). We then showed what the precieved velocity of an observer at the origin would be. What we have learned here will be used when transforming a field of light pulses later in the next chapter.

3.5 Variables

$$\mathbf{r}_{ret}, \mathbf{r}'_{ret} : \dots$$

$$t_{ret}, t'_{ret} : \dots$$

$$\theta_{ret}, \theta'_{ret} : \dots$$

$$\mathbf{r}'_{t'=0} : \dots$$

$$\mathbf{c} = c \cdot \hat{\mathbf{c}} : \dots$$

$$\alpha : \dots$$

$$u_{ret} : \dots$$

Chapter 4

Spherical light pulses

If a point source at rest emits light in all directions, how will this be effected when swaping to a frame where the source is now moving. This will be the topic of this chapter. We will look at the velocity transform specifically for the speed of light, which will be used throughout this chapter, starting with the doppler effect on the frequency of light that a source emits. Next we will get to relativistic aberration, which is the how the angle the emitted light is propagated at, is effected. Both of these effects together gives relativistic beaming, allowing us to transform a field of light pulses from a source. We will then find equations that describe this field of light, like the flux of the light and the intensity of the light at any point. But we will start with looking into the mathematics of spherical polar coordinates that we will allow us to look at all of this. The topics of relativistic aberration and beaming are not topics that are touched upon in most special relativity sources, but they are interesting and hold some important insights, that you will hopefully appreciate.

4.1 Spherical Polar Coordinates

Before we go any further with special relativity, we will need to take some time making sure we know the math needed for what comes next. This section will

be about spherical polar coordinates and its differential elements, which will be used to get what is called a surface element, and in turn find out what the solid angle is and how to derive it. This will then lead into finding the flux of light through this solid angle element. which will be needed when looking at aberration.

So lets start with how to describe a vector in terms of the Spherical Polar Coordinates, shown in the following diagram

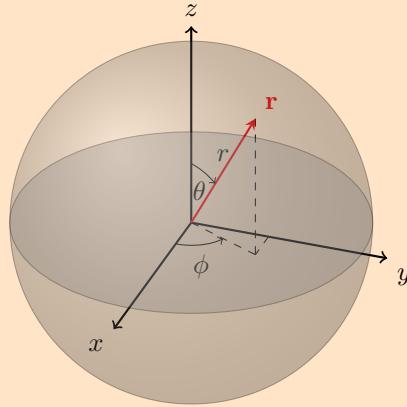


Figure 4.1: **Spherical polar coordinates.** Diagram of an x, y, z -axis showing a position vector \mathbf{r} with its associated spherical polar coordinates (r, θ, ϕ) , where θ is the angle from the z -axis, ϕ is the angle rotated around z -axis starting from the x -axis, and r is the magnitude .

As the figure shows, for a given position \mathbf{r} we have the Cartesian coordinates (x, y, z) , but we can also write this position in terms of spherical polar coordinates (r, θ, ϕ) . Where the angles from the Z-Axis and X-axis are θ (between 0 and π) and ϕ (between 0 and 2π) respectively, and r is the displacement to the position from the origin.

The position can be written using either coordinate system as follows

$$\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = r \begin{pmatrix} \cos \phi \sin \theta \\ \sin \phi \sin \theta \\ \cos \theta \end{pmatrix} \quad (4.1)$$

with

*** check if these are even needed

$$\begin{aligned} r &= \|\mathbf{r}\| = \sqrt{x^2 + y^2 + z^2} \\ \cos \theta &= \frac{z}{r} \\ \cos \phi &= \frac{x}{\sqrt{x^2 + y^2}} \end{aligned} \tag{4.2}$$

4.2 Surface Element on a Sphere

If we define a sphere with radius r and want to find an infinitesimal portion of the surface of the sphere at angles θ and ϕ as shown on this figure

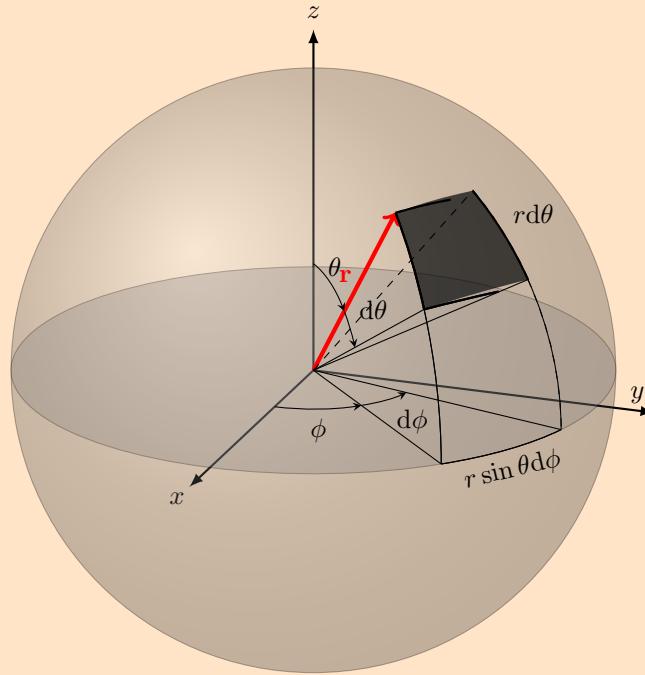


Figure 4.2: Differential surface element on a sphere This shows a surface element on a sphere of radius r that spans from θ to $\theta + d\theta$ and ϕ to $\phi + d\phi$ in spherical polar coordinates, the surface element has sides $rd\theta$ and $r \sin \theta d\phi$, giving the infinitesimal surface area of $r^2 \sin \theta d\theta d\phi$.

The surface element at a constant radius r , spanning from θ to $\theta + d\theta$ and ϕ to

$\phi + d\phi$ on a spherical surface has sides rdr and $r \sin \theta d\phi$ giving the surface area as

$$dS = r^2 \sin \theta d\theta d\phi \quad (4.3)$$

We will also introduce the idea of a solid angle Ω which is a measure of the amount of the field of view from some particular point that a given object covers. For the differential surface element from the figure we have the differential solid angle

$$d\Omega = \frac{dS}{r^2} = \sin \theta d\theta d\phi \quad (4.4)$$

If we were to integrate this over the whole sphere $\int_0^{2\pi} \int_0^\pi \sin \phi d\theta d\phi$, we would get this equal to 4π , which is the full covering of the view. The fraction of the field of view covered from a point by a solid angle is given by $\Omega/4\pi$.

4.3 Light Velocity Transform

The speed of light remains the same after transforming between frames, but the velocity does not, as the direction changes. The velocity transformation in the case where the speed is that of light, $\|u\| = c$, with the direction in the initial frame at an angle α from the z-axis and φ from the x-axis, so that the velocity of light in spherical polar coordinates is

$$\mathbf{c} = c \cdot \hat{\mathbf{e}} = c \begin{pmatrix} \cos \varphi \sin \alpha \\ \sin \varphi \sin \alpha \\ \cos \alpha \end{pmatrix}, \quad (4.5)$$

then from the velocity transform from equation (2.43) we get the primed light velocity as

$$\mathbf{c}' = c \begin{pmatrix} \cos \varphi \sin \alpha' \\ \sin \varphi \sin \alpha' \\ \cos \alpha' \end{pmatrix} = \frac{c}{\gamma(1 - \frac{v}{c} \cos \alpha)} \begin{pmatrix} \cos \varphi \sin \alpha \\ \sin \varphi \sin \alpha \\ \gamma(\cos \alpha - \frac{v}{c}) \end{pmatrix} \quad (4.6)$$

at position \mathbf{r}' and time t'

This still has the same speed of light, which can be shown by taking the magnitude of the velocity vector (that is the square root of the squares of all its elements). However, the angle of propagation has been changed.

4.4 Doppler Effect

When we swap between frames there is also a change in light's wavelength. It is called the doppler effect, which we will describe using the help of the following simple figure of light being emitted from a source in its rest frame and a frame that it is moving.



Figure 4.3: **The Doppler effect.** Diagrams showing light being emitted by a source at velocity \mathbf{c} in the source rest frame (left) and a frame where the source is now moving at velocity $\mathbf{u}' = -\mathbf{v}$ with the light now being emitted at an angle α' giving the light's velocity now as \mathbf{c}' . The different color of light in each frame is showing the effect on the light's frequency and hence color from the doppler effect.

In the rest frame of the source, if it emits a wave at the speed of light in any direction, we have the length the wave has traveled from the source in the infinitesimal time $d\tau$ as being

$$d\lambda = cd\tau \quad (4.7)$$

If we change to a frame where the source is now moving at velocity \mathbf{u}' in the z-direction and we have the wave now emitted at an angle α' from the z-axis, traveling a distance cdt' in a time dt' , with the source moving a distance $u' \cos \alpha' dt'$ in the direction of this emitted wave, this leads to a bunching up of the wave in the direction of the movement of the source, as the same amount of the wave has been emitted in this primed time, but the distance it has traveled from source is

$$d\lambda' = cdt' - u' \cos \alpha' dt' = (1 - \frac{u'}{c} \cos \alpha')cdt' \quad (4.8)$$

we have $dt' = \gamma d\tau$, so the ratio of space the wave is taking up in that direction compared to in the rest frame is

$$Ratio = \frac{d\lambda'}{d\lambda} = \gamma \left(1 - \frac{u'}{c} \cos \alpha' \right) \quad (4.9)$$

since wavelength is inversely proportional to the frequency, and this bunching up at the front will proportionally decrease the wavelength, the frequency ratio ν'/ν is then

$$\Phi_r = \frac{\nu'}{\nu} = \frac{d\lambda}{d\lambda'} = \frac{1}{\gamma \left(1 - \frac{u'}{c} \cos \alpha' \right)} \quad (4.10)$$

This gives us a formula to find out how the frequency of a moving source's emitted light differs from when it is at rest. This is used in astronomy to determine how fast stars and galaxies are moving relative to us on earth.

4.5 Aberration in Special Relativity

A moving light source would emit more of its radiation in the direction of its movement, which we will show by how much here. If a source in its rest frame emits light evenly in all directions, then each emitted light's angle of propagation is transformed using equation (4.6). So that the angles of the lights propagation

from the z-axis can be found to be

$$\cos \alpha' = \frac{\mathbf{c}'_z}{\|\mathbf{c}'\|} = \frac{\cos \alpha + \frac{u'}{c}}{1 + \frac{u'}{c} \cos \alpha}. \quad (4.11)$$

This is the relativistic aberration formula [2], it shows how the light's propagation direction transforms as shown in Figure (4.4).

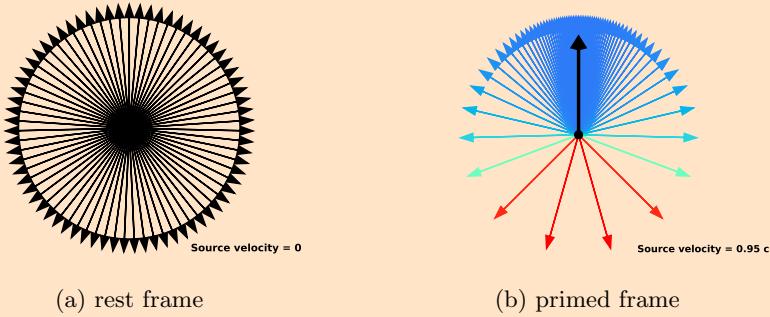


Figure 4.4: **Transforming emitted light.** Diagram showing the directions of propagation of an evenly distributed light pulse in the rest frame of a source (left) and the corresponding distribution in a primed frame in which the source is now moving (right), with the colors showing the Doppler effect on the light's frequency from previous section.

If we want to see the opposite, light being absorbed, we can look at Figure (4.5) which shows what would happen in the case of incoming light from all directions.

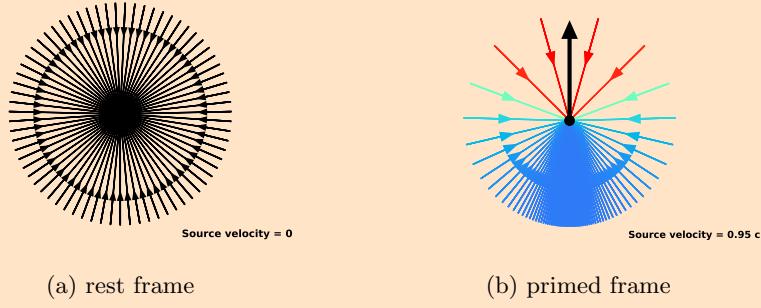


Figure 4.5: **Transforming absorbed light.** Diagram showing the directions of evenly distributed light being absorbed by a receiver in its rest frame (left) and the corresponding distribution in a primed frame in which the receiver is now moving (right), with the colors showing the Doppler effect on the light's frequency.

It can be seen from equation (4.11) that as a source's speed tends to the speed of light, u'/c tends to 1 and as a result, $\cos \alpha'$ tends to 1. Hence all emitted or absorbed light in the source's rest frame tends to the direction of the source's movement. But we also have to remember that the flow of time of the moving source relative to the primed frame tends to point where it runs infinitely slowly. The act of emittance of the light stalls. For both these reasons the light would remain in the vicinity of the source.

4.6 Relativistic Beaming

Relativistic beaming is the combination of both of the two previous sections. If we have a single pulse of light that is evenly spherically distributed, then after a certain time it will look like what is shown in the following figure, due to its aberrated emittance and Doppler effected frequency.

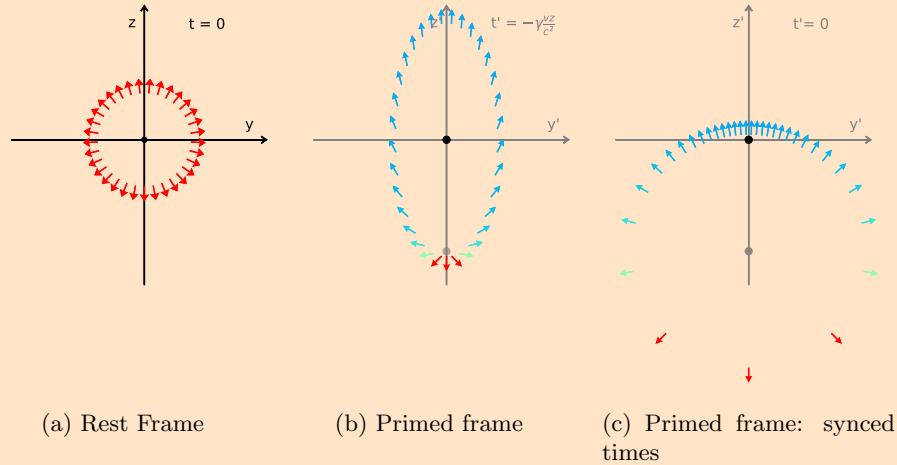


Figure 4.6: **Relativistic beaming.** Diagram showing, the position of an evenly spherically distributed light pulse after propagating for a certain amount of time, with its current time $t = 0$, in the rest frame (left), it is also shown in the primed frame using the direct coordinate transform (middle) giving an ellipsoid with unsynchronised times, and also shown in the primed frame with each part of light pulse propagated forward or backwards to the origins time $t' = 0$ (right). The black dot shows the current location of the source, with grey dot showing its position when the pulse was emitted. all the propagation of the light is in the direction away from this point.

The pulse positions in the primed frame with their time synced with the origin are shown in figure (4.6c) and can be found from the transformed primed frame coordinates from figure (4.6b) by propagating the light from these coordinates forward or backwards in time to be synced with the origins time of $t' = 0$, that is

$$\mathbf{r}'_{t'=0} = \mathbf{r}' - \mathbf{c}' t' \quad (4.12)$$

This can be extended to multiple pulses as we will do in the next section.

4.7 A Moving Source's Field of light

Let us look now at a source continuously emitting light in all directions in its rest frame. To transform each part of the emitted light, we use the position and time $t = 0$ in the coordinate transformation equation to get the light in the primed frame at its desynchronised times and propagate it forward or backwards in time to find its location at the primed time t' . This transform is shown in the following figure.

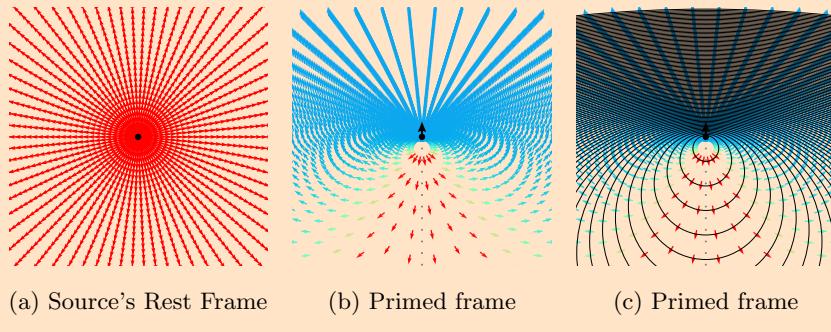


Figure 4.7: **Source's field of emitted photons.** A diagram showing, in the y - z plane, multiple spherically symmetrical pulses of light from a source at rest (left), and in a corresponding primed inertial frame where the source is moving (middle), the colors show the Doppler effect on lights light and the grey dots show the past positions of source when it had emitted each pulse, there is also a diagram showing the circles of each pulse (right), This shows the spherical waves being closer together in the direction of movement.

4.8 Retarded Light Propagation

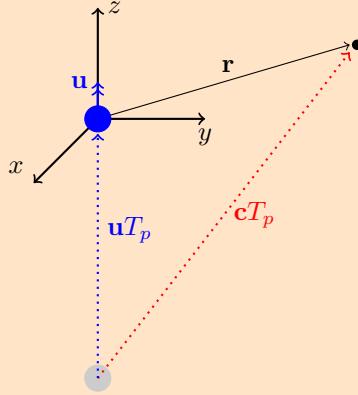


Figure 4.8: The diagram shows a source particle moving at velocity $\mathbf{u} = (0, 0, u)$ in the z -direction in this frame, with the particle currently at the origin with time $t = 0$, the particle and attached proper axis were at the retarded position \mathbf{s}_{ret} shown in grey, the emitted light was then propagated at an angle α from the z -axis, along $\mathbf{c}T_p$ to reach \mathbf{L} at $t = 0$. It has not yet reached the corresponding Lorentz transformed coordinate \mathbf{r} .

$$\mathbf{c}T_p = \mathbf{u}T_p + \mathbf{r} \quad (4.13)$$

$$\hat{\mathbf{c}} = \frac{1}{cT_p} \begin{pmatrix} x \\ y \\ z + uT_p \end{pmatrix} \quad (4.14)$$

...

The magnitude of the retarded field displacement is equal to the distance the field, propagating at the speed of light, travels in the corresponding retarded time. This gives

$$\begin{aligned} (-cT_p)^2 &= \|\vec{R}_{ret}\|^2 \\ &= (x^2 + y^2 + z^2) + u^2 T_p^2 + 2uT_p z, \end{aligned} \quad (4.15)$$

rearranging this, we get the quadratic

$$T_p^2 - \left(2\gamma^2 \frac{u}{c^2} z\right) T_p - \frac{\gamma^2}{c^2} (x^2 + y^2 + z^2) = 0, \quad (4.16)$$

where $\gamma = (1 - u^2/c^2)^{-1/2}$. Now taking the solution for the past time, and making use of the identity

$$\gamma^2 = 1 + \gamma^2 \frac{u^2}{c^2}, \quad (4.17)$$

we get the result

$$\begin{aligned} T_p &= \gamma^2 \frac{u}{c^2} z + \sqrt{\left(-\gamma^2 \frac{u}{c^2} z\right)^2 + \frac{\gamma^2}{c^2} (x^2 + y^2 + z^2)} \\ &= \gamma^2 \frac{u}{c^2} z + \frac{\gamma}{c} \sqrt{x^2 + y^2 + \left(1 + \gamma^2 \frac{u^2}{c^2}\right) z^2} \\ &= \gamma^2 \frac{u}{c^2} z + \frac{\gamma}{c} \sqrt{x^2 + y^2 + \gamma^2 z^2}. \end{aligned} \quad (4.18)$$

...

$$T_p = \gamma_u^2 \frac{u}{c^2} z + \frac{\gamma_u}{c} \sqrt{x^2 + y^2 + \gamma_u^2 z^2} \quad (4.19)$$

4.8.1 Rate of Change of Direction of Light from Moving Source

Now to see the rate of change of the direction that the light propagates, by taking the partial derivative of $\hat{\mathbf{c}}$ with respect to each coordinate

$$\begin{aligned}
\frac{\partial \hat{\mathbf{c}}}{\partial x} &= \frac{1}{\gamma_u M (\frac{u}{c} \frac{\gamma_u z}{M} + 1)^2} \begin{pmatrix} 1 - \frac{x^2}{M^2} + \frac{u}{c} \frac{\gamma_u z}{M} \\ -\frac{x}{M} \frac{y}{M} \\ -\frac{x}{M} \frac{z}{M} \end{pmatrix} \\
\frac{\partial \hat{\mathbf{c}}}{\partial y} &= \frac{1}{\gamma_u M (\frac{u}{c} \frac{\gamma_u z}{M} + 1)^2} \begin{pmatrix} -\frac{y}{M} \frac{x}{M} \\ 1 - \frac{y^2}{M^2} + \frac{u}{c} \frac{\gamma_u z}{M} \\ -\frac{y}{M} \frac{z}{M} \end{pmatrix} \\
\frac{\partial \hat{\mathbf{c}}}{\partial z} &= \frac{1}{\gamma_u M (\frac{u}{c} \frac{\gamma_u z}{M} + 1)^2} \begin{pmatrix} \gamma_u (-\frac{u}{c} - \frac{\gamma_u z}{M}) \frac{x}{M} \\ \gamma_u (-\frac{u}{c} - \frac{\gamma_u z}{M}) \frac{y}{M} \\ \frac{x^2 + y^2}{M^2} \end{pmatrix} \\
\frac{\partial \hat{\mathbf{c}}}{\partial t} &= \frac{-u}{\gamma_u M (\frac{u}{c} \frac{\gamma_u z}{M} + 1)^2} \begin{pmatrix} \gamma_u (-\frac{u}{c} - \frac{\gamma_u z}{M}) \frac{x}{M} \\ \gamma_u (-\frac{u}{c} - \frac{\gamma_u z}{M}) \frac{y}{M} \\ \frac{x^2 + y^2}{M^2} \end{pmatrix}
\end{aligned} \tag{4.20}$$

$$mag = \sqrt{(1 - \frac{x^2}{M^2} + \frac{u}{c} \frac{\gamma_u z}{M})^2 + 2 \frac{x^2 y^2}{M^4} + \frac{x^2 z^2}{M^4} + (1 - \frac{y^2}{M^2} + \frac{u}{c} \frac{\gamma_u z}{M})^2 + \frac{y^2 z^2}{M^4} + (\gamma_u (\frac{u}{c} + \frac{\gamma_u z}{M}) \frac{x}{M})^2 + (\gamma_u (\frac{u}{c} + \frac{\gamma_u z}{M}) \frac{y}{M})^2} \tag{4.21}$$

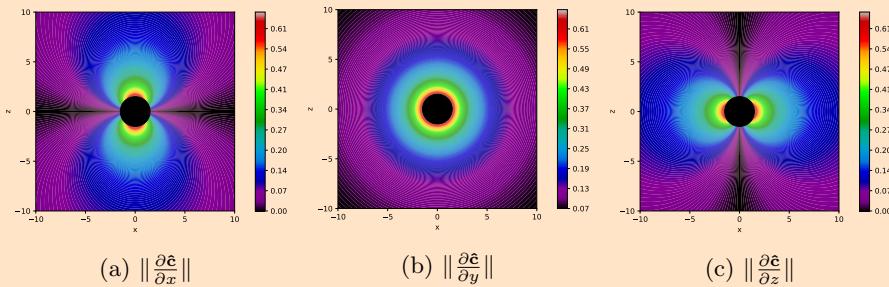


Figure 4.9: The rate of change of the velocity vector field of light from a source at rest. main caption.

these diagrams show the rate of rotation of the lights propagation direction with respect to each coordinate, but the do not show if this rate of change is in the direction of the positive or negative direction of the corresponding axis.

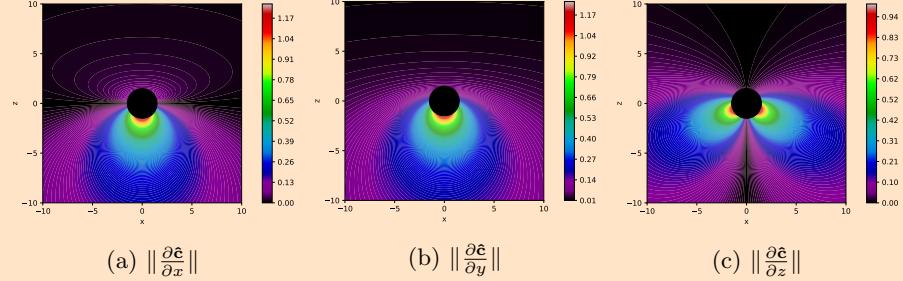


Figure 4.10: **The rate of change of the velocity vector field of light from a moving source.** main caption.

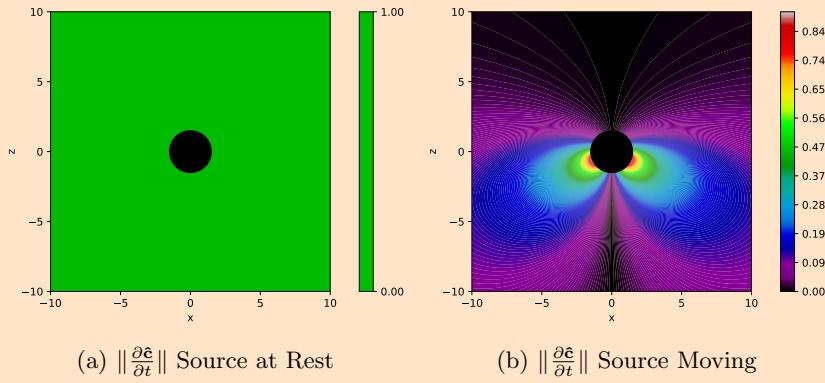


Figure 4.11: **The rate of change of the velocity vector field of light from source.** main caption.

4.9 Flux of light from Point Source

*** What is the density of the field at a certain point, using $k \frac{1}{R^2}$ in a proper frame, and then taking into account the distance from coordinate to retarded position of the source and then the aberration and Doppler effect

*** - Add more meaning to Flux section

-flux describes the rate at which a quantity flows across a surface (flowrate per unit area)

4.9.1 Flux of light Through Surface Element

... radial flux through surface element (search radial flux)

4.9.2 Flux of Light

A proper frame of a particle's differential solid angle element (definition given in previous chapter (maybe move that subsection to here))

Using the equation for the solid angle (what is solid angle?) with $\theta = \alpha$ and $\phi = \varphi$ we have

$$d\Omega = \sin \alpha d\alpha d\varphi, \quad (4.22)$$

encompasses a certain amount of the wavefront, this is the same amount of the wavefront that is encompassed by the coinciding aberrated differential solid angle

$$d\Omega' = \sin \alpha' d\alpha' d\varphi'. \quad (4.23)$$

We can calculate this element by differentiating both sides of eq. (4.11) with respect to α [3], which after a bit of work gives

$$\sin \alpha' d\alpha' = \frac{1}{\gamma^2 \left(1 + \frac{u}{c} \cos \alpha\right)^2} \sin \alpha d\alpha \quad (4.24)$$

this took into account that $v = -u$, now using this and $d\varphi' = d\varphi$ (as the angle φ is always perpendicular to the motion of the particle and hence unaffected by transformation) we have the solid angle in the primed frame given as

$$d\Omega' = \frac{1}{\gamma^2 \left(1 + \frac{u}{c} \cos \alpha\right)^2} \sin \alpha d\alpha d\varphi. \quad (4.25)$$

The relative primed wavefront strength at a given angle is taken as being proportional to the amount of the wavefront per solid angle in the primed frame

relative to that in the proper frame, referred to here as the aberrational wave-front strength weighting, given as

$$\Phi_\Omega = \frac{d\Omega}{d\Omega'} = \hat{A}^2. \quad (4.26)$$

(define what I mean by strength weighting)

*** This flux does not take into account that the source keeps changing position, so at a change $d\mathbf{s}$ from position field is coming from a different retarded position of the source, so there is a rotation of direction and a different flux, there is gradients and rotations

*** This flux transform is the change in flux in the emitted direction from the source only not what another particle would feel

4.10 Variables

$\mathbf{r} = (r, \theta, \phi) : \dots$

$d\theta, d\phi : \dots$

$dS : \dots$ surface element might be confusing to have as an S due to next chapter

$\Omega : \dots$

$d\Omega : \dots$

$d\lambda, d\lambda' : \dots$

$d\tau : \dots$

$\Phi_r, \Phi_\Omega : \dots$

$\nu, \nu' : \dots$

$\hat{A} : \dots$

References

- [2] Albert Einstein et al. “On the electrodynamics of moving bodies”. In: *Annalen der physik* 17.10 (1905), pp. 891–921.
- [3] David W Hogg. “Special relativity”. In: *Lecture Notes* (1997), pp. 39–40.

Chapter 5

Invariant Quantities

Quantities that are frame-independent, i.e. do not change when we transform between inertial reference frames, are called Lorentz invariant.

In this chapter, we will see the invariant quantity related to space-time coordinates and use this to find invariant properties associated with velocity, which will lead to the more abstract conservation relation of momentum energy and mass. We will find a new quantity with the units of energy that has a term which approximates to the kinetic energy of a particle at low velocities and another term that suggests particles have a new type of energy, that was not previously in classical physics, that depends on the mass of a particle (mass-energy equivalence), this will all lead to a new conservation law of energy. The reason we care about the concept of energy is that it has a conservation relation that can be used as a tool to help us make calculations and model more complex systems.

5.1 Space-Time Interval

We define a quantity called the space-time interval ΔS between two events in a generic inertial reference frame as:

$$\Delta S^2 = (c\Delta t)^2 - \Delta x^2 - \Delta y^2 - \Delta z^2$$

(5.1)

All inertial frames of reference agree on this interval as it is invariant between frames, as $\Delta S = \Delta S'$, which can be shown when we sub the values from equation (2.40) into the primed version of this last equation

$$\begin{aligned}
 \Delta S'^2 &= (c\Delta t')^2 - \Delta x'^2 - \Delta y'^2 - \Delta z'^2 \\
 &= c^2 \gamma \left(\Delta t - \frac{v}{c^2} \Delta z \right)^2 - \Delta x^2 - \Delta y^2 - \Delta (\gamma(\Delta z - v\Delta t))^2 \\
 &= (c\Delta t)^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 \\
 &= \Delta S^2.
 \end{aligned} \tag{5.2}$$

5.2 Energy-Momentum Relation

From here forward we will be making some jumps in logic and assertions, leading to claims of what we will call the total energy and rest mass energy, but I will try to give as much reasoning as I can for these choices in assertions.

We can see from the previous section that if a particle is moving in a particular inertial reference frame, we can define the space-time interval of its coordinates in this frame as

$$\Delta S_p^2 = (c\Delta t_p)^2 - \Delta x_p^2 - \Delta y_p^2 - \Delta z_p^2 \tag{5.3}$$

and if we divide by the square of the proper time that passes for the particle in this interval $\Delta\tau$, so that now taking the limit as the change in time in the interval goes to zero, we get an abstract invariant quantity that we will call the space-time velocity

$$\begin{aligned}
\left(\frac{dS_p}{d\tau}\right)^2 &= \left(c \frac{dt_p}{d\tau}\right)^2 - \left(c \frac{dx_p}{d\tau}\right)^2 - \left(\frac{dy_p}{d\tau}\right)^2 - \left(\frac{dz_p}{d\tau}\right)^2 \\
&= c^2 \gamma_u^2 \left(1 - \frac{u_x^2 + u_y^2 + u_z^2}{c^2}\right) = c^2 \gamma_u^2 \left(1 - \frac{u^2}{c^2}\right) \\
&= c^2
\end{aligned} \tag{5.4}$$

It can be seen that this is still an invariant quantity as the speed of light is constant in all frames. Now by multiplying both sides by c^2 and the square of the rest mass m of the particle we have the square of an invariant quantity with the units of energy, that is equal to mc^2 and give the equation to be

$$\begin{aligned}
m^2 c^2 \left(\frac{dS_p}{d\tau}\right)^2 &= \gamma_u^2 m^2 c^4 - \gamma_u^2 m^2 (u_x^2 + u_y^2 + u_z^2) c^2 \\
(mc^2)^2 &= \gamma_u^2 m^2 c^4 - \gamma_u^2 m^2 u^2 c^2
\end{aligned} \tag{5.5}$$

rearranging, we get

$$(\gamma_u mc^2)^2 = (mc^2)^2 + (\gamma_u muc)^2 \tag{5.6}$$

using the Taylor expansion of γ_u on the quantity $\gamma_u mc^2$, we have for the squared term on the lefthand side

*** this needs to be taylor expansion of squared term on lhs, not the none squared term

$$\gamma_u = 1 + \frac{1}{2} \frac{u^2}{c^2} + \frac{3}{8} \frac{u^4}{c^4} + \dots \tag{5.7}$$

$$\begin{aligned}
\gamma_u mc^2 &= mc^2 \left(1 + \frac{1}{2} \frac{u^2}{c^2} + \frac{3}{8} \frac{u^4}{c^4} + \dots\right) \\
&= mc^2 + \frac{1}{2} mu^2 + \frac{3}{8} m \frac{u^4}{c^2} + \dots
\end{aligned} \tag{5.8}$$

This equation's first quantity is an energy that does not depend on the speed of the particle, and only on its rest mass, we shall call this the rest mass energy, and this can be taken as a new type of energy all masses have even when at rest. The second quantity is the classical kinetic energy formula, the rest may be seen as higher-order kinetic energy terms that are very small for $u \ll c$, so

this seems to be a formula for the total energy of a particle in the absence of a potential field (i.e. no change in potential energy) with added terms which become non-negligible at larger particle speeds, we will denote this total energy E . Taking this assertion we have

$$E^2 = (mc^2)^2 + (\gamma_u mu)^2 c^2 \quad (5.9)$$

so this total energy is equal to its rest energy squared plus a second energy squared that depends on its speed, this second quantity along with the Taylor expansion of γ_u gives

$$\begin{aligned} \gamma_u muc &= muc \left(1 + \frac{1}{2} \frac{u^2}{c^2} + \frac{3}{8} \frac{u^4}{c^4} + \dots \right) \\ &= c \left(mu + \frac{1}{2} m \frac{u^3}{c^2} + \frac{3}{8} m \frac{u^5}{c^4} + \dots \right) \end{aligned} \quad (5.10)$$

which at small particle velocities is approximately muc with mu being the classical momentum, we could assert that this whole energy is momentum in the relativistic context, i.e.

$$p = \gamma_u mu \quad (5.11)$$

as it has the same units and is approximately the classical momentum at small particle speeds.

$$E_0 = mc^2 \quad (5.12)$$

This is the energy it has from its rest mass, if a particle at rest was to decay into photons, the photons emitted from this would be equal to this energy. The final energy-momentum equation is then given as

$$E^2 = E_0^2 + (pc)^2 \quad (5.13)$$

*** Outside the scope of SR, should we state that the mass is a measure of its

energy content instead?

5.3 Energy-Momentum Relation For Photons

*** First associate the energy of photon $E = hf$ then find its equivalence in mass, then sub this in, and then say momentum for the photon is then hf/c . For a Photon we have the mass equal to zero and the momentum as $p_{light} = \frac{hf}{c}$ which is taken from quantum mechanics and the energy-momentum relation as

$$E = cp_{light} = h\nu \quad (5.14)$$

where h is Planks constant and ν is the frequency of the light.
nothing is derived here, just states what to do for a photon

5.3.1 Energy-Mass Equivalence

*** Imagine we have a particle at rest emitting light in all directions so that afterward it is still at rest, but in another frame we have this particle moving and the emitted light aberrated, but we still need the particle to remain at a constant velocity after emission. (maybe include a rotation in the impulse/effect of light in each direction on the particle of has)

***-if we have a ball in its rest frame that emits a flash of light evenly in all directions, we will have the ball's energy decrease by the amount of energy of the emitted light, and it will remain at rest after due to the light being emitted evenly in all directions -its kinetic energy is zero in this frame

5.4 Variables

$\Delta S : \dots$

$\Delta S_p : \dots (\delta x_p, \delta y_p, \delta z_p, \delta t_p)$

$m : \dots$

$\gamma_u : \dots$

$\delta\tau$: ...

E : ...

E_0 : ...

p : ... momentum

...

Outro

meaningful conclusion:

- [] summarise the key points that were discussed in the book - [] discuss implications and applications of these points - [] propose areas of expansion for this topic/ future research - [] what the reader can do next to expand knowledge (i.e. internal and external resources)

*** this was meant as an introduction to physical concepts and mathematical tools *** What was left out of this book *** Things still to be added to this book *** Questions you may have *** General confusions ***

Resources

Bibliography

- [1] Morley Williams Michelson Abraham. “On the Relative Motion of the Earth and the Luminiferous Ether”. In: *American Journal of Science* 34 (1887).
- [2] Albert Einstein et al. “On the electrodynamics of moving bodies”. In: *Annalen der physik* 17.10 (1905), pp. 891–921.
- [3] David W Hogg. “Special relativity”. In: *Lecture Notes* (1997), pp. 39–40.