

2.9: Three Spatial Dimensions (Part 2)

Lorentz Boosts Causing Rotations

As a quantitative example, consider the following thought experiment. Put a gyroscope in a box, and send the box around the square path shown in figure d at constant speed. The gyroscope defines a local coordinate system, which according to classical physics would maintain its orientation. At each corner of the square, the box has its velocity vector changed abruptly, as represented by the hammer. We assume that the hits with the hammer are transmitted to the gyroscope at its center of mass, so that they do not result in any torque. Nonrelativistically, if the set of gyroscopes travels once around the square, it should end up at the same place and in the same orientation, so that the coordinate system it defines is identical with the original one.

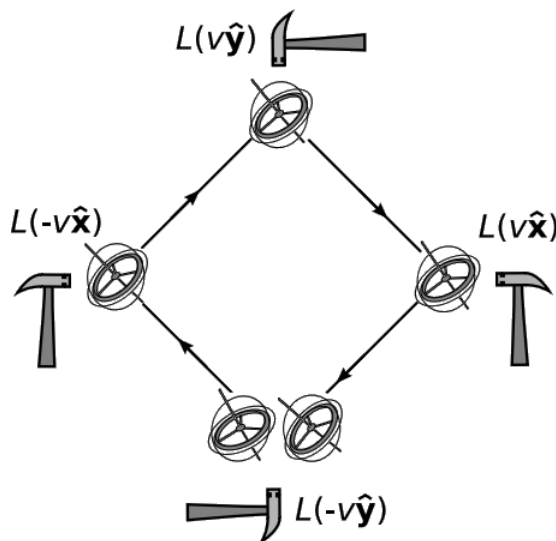


Figure 2.9.4: Nonrelativistically, the gyroscope should not rotate as long as the forces from the hammer are all transmitted to it at its center of mass.

For notation, let $L(v\hat{x})$ indicate the boost along the x axis described by the transformation earlier. This is a transformation that changes to a frame of reference moving in the *negative* x direction compared to the original frame. A particle considered to be at rest in the original frame is described in the new frame as moving in the *positive* x direction. Applying such an L to a vector \mathbf{p} , we calculate $L\mathbf{p}$, which gives the coordinates of the event as measured in the new frame. An expression like $ML\mathbf{p}$ is equivalent by associativity to $M(L\mathbf{p})$, i.e., ML represents applying L first, and then M .

In this notation, the hammer strikes can be represented by a series of four Lorentz boosts,

$$T = L(v\hat{x})L(v\hat{y})L(-v\hat{x})L(-v\hat{y}), \quad (2.9.1)$$

where we assume that the square has negligible size, so that all four Lorentz boosts act in a way that preserves the origin of the coordinate systems. (We have no convenient way in our notation $L(\cdot)$ to describe a transformation that does not preserve the origin.) The first transformation, $L(-v\hat{y})$, changes coordinates measured by the original gyroscope-defined frame to new coordinates measured by the new gyroscope-defined frame, after the box has been accelerated in the positive y direction.

calculation of T matrices (Maxima Software)

The calculation of T is messy, and to be honest, I made a series of mistakes when I tried to crank it out by hand. Calculations in relativity have a reputation for being like this. Figure 2.9.5 shows a page from one of Einstein's notebooks, written in fountain pen around 1913. At the bottom of the page, he wrote "zu umstaendlich," meaning "too involved."

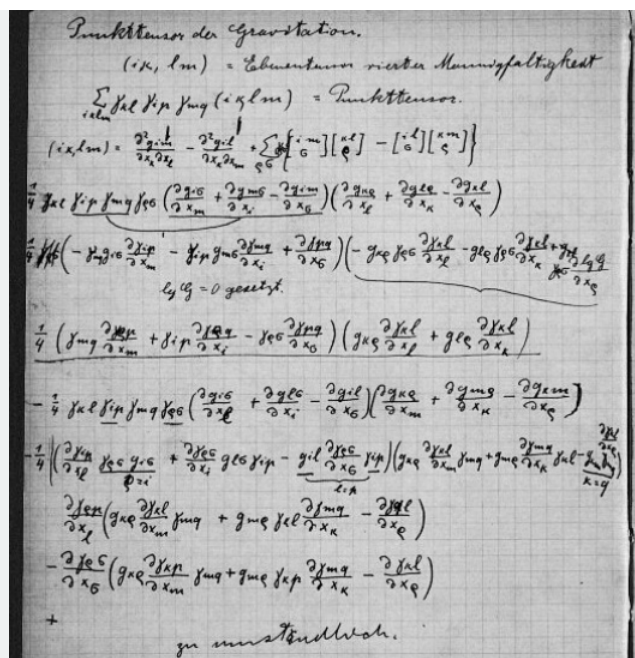


Figure 2.9.5: A page from one of Einstein's notebooks.

Luckily we live in an era in which this sort of thing can be handled by computers. Starting at this point in the book, I will take appropriate opportunities to demonstrate how to use the free and open-source computer algebra system Maxima to keep complicated calculations manageable. The following Maxima program calculates a particular element of the matrix T .

```

1 /* For convenience, define gamma in terms of v: */
2 gamma:1/sqrt(1-v*v);
3 /* Define Lx as L(x-hat), Lmx as L(-x-hat), etc.: */
4 Lx:matrix([gamma, gamma*v, 0],
5 [gamma*v, gamma, 0],
6 [0, 0, 1]);
7 Ly:matrix([gamma, 0, gamma*v],
8 [0, 1, 0],
9 [gamma*v, 0, gamma]);
10 Lmx:matrix([gamma, -gamma*v, 0],
11 [-gamma*v, gamma, 0],
12 [0, 0, 1]);
13 Lmy:matrix([gamma, 0, -gamma*v],
14 [0, 1, 0],
15 [-gamma*v, 0, gamma]);
16 /* Calculate the product of the four matrices: */
17 T:Lx.Ly.Lmx.Lmy;
18 /* Define a column vector along the x axis: */
19 P:matrix([0],[1],[0]);
20 /* Find the result of T acting on this vector,
21 expressed as a Taylor series to second order in v: */
22 taylor(T.P,v,0,2);
  
```

Statements are terminated by semicolons, and comments are written like `/* ... */`. On line 2, we see a symbolic definition of the symbol `gamma` in terms of the symbol `v`. The colon means “is defined as.” Line 2 does not mean, as it would in most programming languages, to take a stored numerical value of `v` and use it to calculate a numerical value of γ . In fact, `v` does not have a numerical value defined at this point, nor will it ever have a numerical value defined for it throughout this program. Line 2 simply means that whenever Maxima encounters the symbol `gamma`, it should take it as an abbreviation for the symbol $1/\sqrt{1-v^2}$. Lines 5-16 define some 3×3 matrices that represent the L transformations. The basis is $\hat{t}, \hat{x}, \hat{y}$. Line 18 calculates the product of the four matrices; the dots represent matrix multiplication. Line 23 defines a vector along the x axis, expressed as a column matrix (three rows of one column each) so that Maxima will know how to operate on it using matrix multiplication by T .

Finally line 26 outputs the result of T acting on P :

```

19
20
21 (%o9)/T/
22
23
24

```

$$\begin{bmatrix} 0 + \dots \\ 1 + \dots \\ 2 \\ -v + \dots \end{bmatrix}$$

I've omitted some output generated automatically from the earlier steps in the computation. The (%o9) indicates that this is Maxima's output from the ninth and final step.

In other words,

$$T \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -v^2 \end{pmatrix} + \dots, \quad (2.9.2)$$

where \dots represents higher-order terms in v . Suppose that we use the initial frame of reference, before T is applied, to determine that a particular reference point, such as a distant star, is along the x axis. Applying T , we get a new vector TP , which we find has a nonvanishing y component approximately equal to $-v^2$. This result is entirely unexpected classically. It tells us that the gyroscope, rather than maintaining its original orientation as it would have done classically, has rotated slightly. It has precessed in the counterclockwise direction in the x - y plane, so that the direction to the star, as measured in the coordinate system defined by the gyroscope, appears to have rotated clockwise. As the box moved clockwise around the square, the gyroscope has apparently rotated by a counterclockwise angle $\chi \approx v^2$ about the z axis. We can see that this is a purely relativistic effect, since for $v \ll 1$ the effect is small. For historical reasons discussed [later](#), this phenomenon is referred to as the Thomas precession.

The particular features of this square geometry are not necessary. I chose them so that (1) the boosts would be along the Cartesian axes, so that we would be able to write them down easily; (2) it is clear that the effect doesn't arise from any asymmetric treatment of the spatial axes; and (3) the change in the orientation of the gyroscope can be measured at the same point in space, e.g., by comparing it with a twin gyroscope that stays at home. In general:

A gyroscope transported around a closed loop in flat spacetime changes its orientation compared with one that is not accelerated.

This is a purely relativistic effect, since a Newtonian gyroscope does not change its axis of rotation unless subjected to a torque; if the boosts are accomplished by forces that act at the gyroscope's center of mass, then there is no nonrelativistic explanation for the effect.

The effect can occur in the absence of any gravitational fields. That is, this is a phenomenon of special relativity.

The composition of two or more Lorentz boosts along different axes is not equivalent to a single boost; it is equivalent to a boost plus a spatial rotation.

Lorentz boosts do not commute, i.e., it makes a difference what order we perform them in. Even if there is almost no time lag between the first boost and the second, the order of the boosts matters. If we had applied the boosts in the opposite order, the handedness of the effect would have been reversed.

Exercise 2.9.2

Self-check: If Lorentz boosts did commute, what would be the consequences for the expression $L(v\hat{x}) L(v\hat{y}) L(-v\hat{x}) L(-v\hat{y})$?

The Velocity Disk

Figure 2.5.6 shows a useful way of visualizing the combined effects of boosts and rotations in 2+1 dimensions. The disk depicts all possible states of motion relative to some arbitrarily chosen frame of reference. Lack of motion is represented by the point at the center. A point at distance v from the center represents motion at velocity v in a particular direction in the x - y plane. By drawing little axes at a particular point, we can represent a particular frame of reference: the frame is in motion at some velocity, with its own x and y axes are oriented in a particular way.

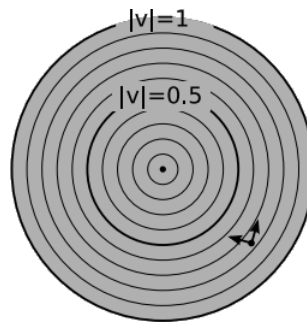


Figure 2.9.6 - The velocity disk.

It turns out to be easier to understand the qualitative behavior of our mysterious rotations if we switch from the low-velocity limit to the contrary limit of ultrarelativistic velocities. Suppose we have a rocket-ship with an inertial navigation system consisting of two gyroscopes at right angles to one another. We first accelerate the ship in the y direction, and the acceleration is steady in the sense that it feels constant to observers aboard the ship. Since it is rapidities, not velocities, that add linearly, this means that as an observer aboard the ship reads clock times τ_1, τ_2, \dots , all separated by equal intervals $\Delta\tau$, the ship's rapidity changes at a constant rate, η_1, η_2, \dots . This results in a series of frames of reference that appear closer and closer together on the diagram as the ship approaches the speed of light, at the edge of the disk. We can start over from the center again and repeat the whole process along the x axis, resulting in a similar succession of frames. In both cases, the boosts are being applied along a single line, so that there is no rotation of the x and y axes.

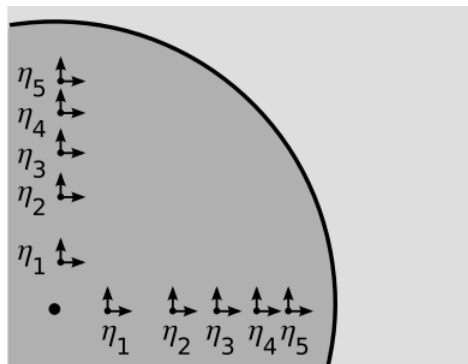


Figure 2.9.7 - Two excursions in a rocketship: one along the y axis and one along x .

Now suppose that the ship were to accelerate along a route like the one shown in Figure 2.5.8. It first accelerates along the y axis at a constant rate (again, as judged by its own sensors), until its velocity is very close to the speed of light, A. It then accelerates, again at a self-perceived constant rate and with thrust in a fixed direction as judged by its own gyroscopes, until it is moving at the same ultrarelativistic speed in the x direction, B. Finally, it decelerates in the x direction until it is again at rest, O. This motion traces out a clockwise loop on the velocity disk. The motion in space is also clockwise.

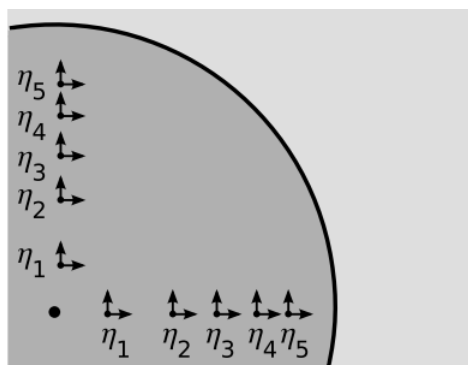


Figure 2.9.8 - A round-trip involving ultrarelativistic velocities. All three legs are at constant acceleration.

We might naively think that the middle leg of the trip, from A to B, would be a straight line on the velocity disk, but this can't be the case. First, we know that non-collinear boosts cause rotations. Traveling around a clockwise path causes counterclockwise rotation, and vice-versa. Therefore an observer in the rest frame O sees the ship (and its gyroscopes) as rotating as it moves from A

to B. The ship's trajectory through space is clockwise, so according to O the ship rotates counterclockwise as it goes A to B. The ship is always firing its engines in a fixed direction as judged by its gyroscopes, but according to O the ship is rotating counterclockwise, its thrust is progressively rotating counterclockwise, and therefore its trajectory turns counterclockwise. We conclude that leg AB on the velocity disk is concave, rather than being a straight-line hypotenuse of a triangle OAB.

We can also determine, by the following argument, that leg AB is perpendicular to the edge of the disk where it touches the edge of the disk. In the transformation from frame A to frame O, y coordinates are dilated by a factor of γ , which approaches infinity in the limit we're presently considering. Observers aboard the rocket-ship, occupying frame A, believe that their task is to fire the rocket's engines at an angle of 45 degrees with respect to the y axis, so as to eliminate their velocity with respect to the origin, and simultaneously add an equal amount of velocity in the x direction. This 45-degree angle in frame A, however, is not a 45-degree angle in frame O. From the stern of the ship to its bow we have displacements Δx and Δy , and in the transformation from A to O, Δy is magnified almost infinitely. As perceived in frame O, the ship's orientation is almost exactly antiparallel to the y axis.²¹

Note

Although we will not need any more than this for the purposes of our present analysis, a longer and more detailed discussion by Rhodes and Semon, www.bates.edu/~msemon/RhodesSemonFinal.pdf, Am. J. Phys. 72(7)2004, shows that this type of inertially guided, constant-thrust motion is always represented on the velocity disk by an arc of a circle that is perpendicular to the disk at its edge. (We consider a diameter of the disk to be the limiting case of a circle with infinite radius.)

As the ship travels from A to B, its orientation (as judged in frame O) changes from $-\hat{y}$ to \hat{x} . This establishes, in a much more direct fashion, the direction of the Thomas precession: its handedness is contrary to the handedness of the direction of motion. We can now also see something new about the fundamental reason for the effect. It has to do with the fact that observers in different states of motion disagree on spatial angles. Similarly, imagine that you are a two-dimensional being who was told about the existence of a new, third, spatial dimension. You have always believed that the cosine of the angle between two unit vectors \mathbf{u} and \mathbf{v} is given by the vector dot product $u_x v_x + u_y v_y$. If you were allowed to explore a two-dimensional projection of a three-dimensional scene, e.g., on the flat screen of a television, it would seem to you as if all the angles had been distorted. You would have no way to interpret the visual conventions of perspective. But once you had learned about the existence of a z axis, you would realize that these angular distortions were happening because of rotations out of the x-y plane. Such rotations really conserve the quantity $u_x v_x + u_y v_y + u_z v_z$; only because you were ignoring the $u_z v_z$ term did it seem that angles were not being preserved. Similarly, the generalization from three Euclidean spatial dimensions to 3+1-dimensional spacetime means that three-dimensional dot products are no longer conserved.

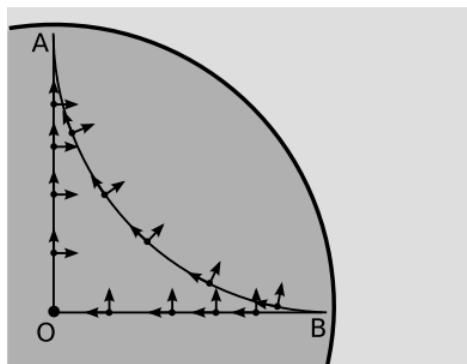


Figure 2.9.9 - In the limit where A and B are ultrarelativistic velocities, leg AB is perpendicular to the edge of the velocity disk. The result is that the x - y frame determined by the ship's gyroscopes has rotated by 90 degrees by the time it gets home.

The General Low-v Limit

Let's find the low-v limit of the Thomas precession in general, not just in the highly artificial special case of $\chi \approx v^2$ for the example involving the four hammer hits. To generalize to the case of smooth acceleration, we first note that the rate of precession $\frac{d\chi}{dt}$ must have the following properties.

- It is odd under a reversal of the direction of motion, $\mathbf{v} \rightarrow -\mathbf{v}$. (This corresponds to sending the gyroscope around the square in the opposite direction.)
- It is odd under a reversal of the acceleration due to the second boost, $\mathbf{a} \rightarrow -\mathbf{a}$.

- It is a rotation about the spatial axis perpendicular to the plane of the \mathbf{v} and \mathbf{a} vectors, in the opposite direction compared to the handedness of the curving trajectory.
- It is approximately linear in \mathbf{v} and \mathbf{a} , for small \mathbf{v} and \mathbf{a} .

The only rotationally invariant mathematical operation that has these symmetry properties is the vector cross product, so the rate of precession must be $k\mathbf{a} \times \mathbf{v}$, where $k > 0$ is nearly independent of \mathbf{v} and \mathbf{a} for small \mathbf{v} and \mathbf{a} .

To pin down the value of k , we need to find a connection between our two results: $\chi \approx v^2$ for the four hammer hits, and $\frac{d\chi}{dt} \approx k\mathbf{a} \times \mathbf{v}$ for smooth acceleration. We can do this by considering the physical significance of areas on the velocity disk. As shown in Figure 2.5.10, the rotation χ due to carrying the velocity around the boundary of a region is additive when adjacent regions are joined together. We can therefore find χ for any region by breaking the region down into elements of area dA and integrating their contributions $d\chi$. What is the relationship between dA and $d\chi$? The velocity disk's structure is nonuniform, in the sense that near the edge of the disk, it takes a larger boost to move a small distance. But we're investigating the low-velocity limit, and in the low-velocity region near the center of the disk, the disk's structure is approximately uniform. We therefore expect that there is an approximately constant proportionality factor relating dA and $d\chi$ at low velocities. The example of the hammer corresponds geometrically to a square with area v^2 , so we find that this proportionality factor is unity, $dA \approx d\chi$.

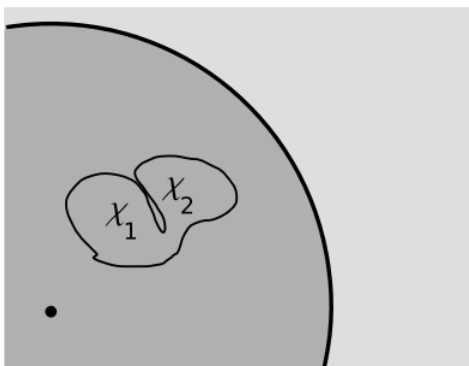


Figure 2.9.10 - If the crack between the two areas is squashed flat, the two pieces of the path on the interior coincide, and their contributions to the precession cancel out ($\mathbf{v} \rightarrow -\mathbf{v}$, but $\mathbf{a} \rightarrow +\mathbf{a}$, so $\mathbf{a} \times \mathbf{v} \rightarrow -\mathbf{a} \times \mathbf{v}$). Therefore the precession χ obtained by going around the outside is equal to the sum $\chi_1 + \chi_2$ of the precessions that would have been obtained by going around the two parts.

To relate this to smooth acceleration, consider a particle performing circular motion with period T , which has $|\mathbf{a} \times \mathbf{v}| = \frac{2\pi v^2}{T}$. Over one full period of the motion, we have $\chi = \int k|\mathbf{a} \times \mathbf{v}| dt = 2\pi k v^2$, and the particle's velocity vector traces a circle of area $A = \pi v^2$ on the velocity disk. Equating A and χ , we find $k = \frac{1}{2}$. The result is that in the limit of low velocities, the rate of rotation is

$$\boldsymbol{\Omega} \approx \frac{1}{2} \mathbf{a} \times \mathbf{v}, \quad (2.9.3)$$

where $\boldsymbol{\Omega}$ is the angular velocity vector of the rotation. In the special case of circular motion, this can be written as $\boldsymbol{\Omega} = (\frac{1}{2})v^2\omega$, where $\omega = \frac{2\pi}{T}$ is the angular frequency of the motion.

An Experimental Test: Thomas Precession in Hydrogen

If we want to see this precession effect in real life, we should look for a system in which both \mathbf{v} and \mathbf{a} are large. An atom is such a system.

The Bohr model, introduced in 1913, marked the first quantitatively successful, if conceptually muddled, description of the atomic energy levels of hydrogen. Continuing to take $c = 1$, the over-all scale of the energies was calculated to be proportional to $m\alpha^2$, where m is the mass of the electron, and $\alpha = \frac{ke^2}{\hbar} \approx \frac{1}{137}$, known as the fine structure constant, is essentially just a unitless way of expressing the coupling constant for electrical forces. At higher resolution, each excited energy level is found to be split into several sub-levels. The transitions among these close-lying states are in the millimeter region of the microwave spectrum. The energy scale of this fine structure is $\sim m\alpha^4$. This is down by a factor of α^2 compared to the visible-light transitions, hence the name of the constant. Uhlenbeck and Goudsmit showed in 1926 that a splitting on this order of magnitude was to be expected due to the magnetic interaction between the proton and the electron's magnetic moment, oriented along its spin. The effect they calculated, however, was too big by a factor of two.

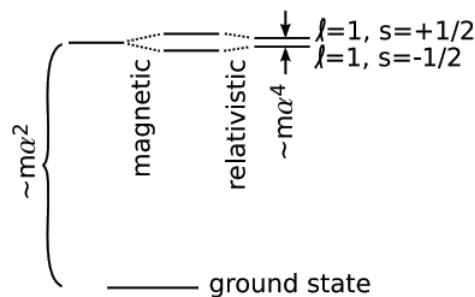


Figure 2.9.11: States in hydrogen are labeled with their ℓ and s quantum numbers, representing their orbital and spin angular momenta in units of \hbar . The state with $s = +\frac{1}{2}$ has its spin angular momentum aligned with its orbital angular momentum, while the $s = -\frac{1}{2}$ state has the two angular momenta in opposite directions. The direction and order of magnitude of the splitting between the two $\ell = 1$ states is successfully explained by magnetic interactions with the proton, but the calculated effect is too big by a factor of 2. The relativistic Thomas precession cancels out half of the effect.

The explanation of the mysterious factor of two had in fact been implicit in a 1916 calculation by Willem de Sitter, one of the first applications of general relativity. De Sitter treated the earth-moon system as a gyroscope, and found the precession of its axis of rotation, which was partly due to the curvature of spacetime and partly due to the type of rotation described earlier in this section. The effect on the motion of the moon was noncumulative, and was only about one meter, which was much too small to be measured at the time. In 1927, however, Llewellyn Thomas applied similar reasoning to the hydrogen atom, with the electron's spin vector playing the role of gyroscope. Since gravity is negligible here, the effect has nothing to do with curvature of spacetime, and Thomas's effect corresponds purely to the special-relativistic part of de Sitter's result. It is simply the rotation described above, with $\Omega = (\frac{1}{2})v^2\omega$. Although Thomas was not the first to calculate it, the effect is known as Thomas precession. Since the electron's spin is $\frac{\hbar}{2}$, the energy splitting is $\pm(\frac{\hbar}{2})\Omega$, depending on whether the electron's spin is in the same direction as its orbital motion, or in the opposite direction. This is less than the atom's gross energy scale ω by a factor of $\frac{v^2}{2}$, which is $\sim \alpha^2$. The Thomas precession cancels out half of the magnetic effect, bringing theory in agreement with experiment.

Uhlenbeck later recalled: "...when I first heard about [the Thomas precession], it seemed unbelievable that a relativistic effect could give a factor of 2 instead of something of order $\frac{v}{c}$... Even the cognoscenti of relativity theory (Einstein included!) were quite surprised."

This page titled [2.9: Three Spatial Dimensions \(Part 2\)](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Benjamin Crowell](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.