Chapter 4

Events, Worldlines, Intervals

4.1 Introduction

As should have become clear in all the previous discussion, the primary unit in Special Relativity is the event. An event is at a place and a time. It is the problem of coordinatizing to label events. Although we have developed a coordinatizing scheme that all inertial observers can agree on, the labeling of any specific event will vary from one inertial observer to the next. For instance, in Section 3.3.5, when discussing when Harry is at Dorothy. Note that even though there is one event it has different coordinate descriptions depending on the observer. Harry says that when he is at Dorothy they are zero distance apart but that this occurred $\frac{4}{3}$ years he left Sally. Sally says that when Harry is at Dorothy they are one light year away and it was $\frac{5}{3}$ years after she and he were together. This should not be a surprise. It was the case in Newtonian physics; position labels depended on the observer even in Galilean transformations, Equation 3.5. The big difference in the case of the Lorentz transformations is the changes in the time coordinates. It is in this sense that people often talk of the space-time continuum when talking about Special Relativity. There is an intrinsic mixing of space and time labels. It may be worthwhile to make a brief excursion into a discussion of place in the two dimensional plane to remind us of what can be done here.

4.2 Place and Path in the Two Dimensional Plane

The material of this subsection is a trivial diversion from the general development of the kinematic effects of special relativity. It is being set here to provide a background to the development of a more intuitive basis for ideas that are important to understanding the implications of the relativity.

Consider an unmarked plane of places. Our problem is the establishment of a labeling system that is efficient and easy to use. A more general discussion of this problem of labeling was in Section 3.1 and is extended to General Relativity in Section 9.9. We will make a pair of often unstated assumptions about the nature of our plane. All places are the same in the sense that there is nothing that you can do at any place that would differentiate that place from any other place. In addition, we assume that at any place all directions are the same. These assumptions allow us to say that the space of places is homogeneous and isotropic.

In this case, all the places can be labeled simply by choosing two orthogonal directions each with a length scale and a special location called the origin. From our assumptions about the structure of our space, it is clear that the choice of an origin is arbitrary and that this choice cannot play an important part in any analysis of the properties of places. You cannot tell where you are and all that can matter is difference in the labels of the places. Another way to say this is to say that although you can talk about where you are, the coordinate of a place. All important concerns do not involve the coordinates but involve only differences in coordinates, $(\vec{x}_2 - \vec{x}_1)$. This form is unchanged by a translation of the coordinate origin. If you replace the coordinates $\vec{x}_{1.2}$ by the new cooridinates $\vec{x}'_{1,2} = \vec{x}_{1,2} - \vec{a}$,

$$\left(\vec{x}'_2 - \vec{x}'_1\right) = (\vec{x}_2 - \vec{x}_1)$$

; the combination of variables $(\vec{x}_2 - \vec{x}_1)$ is unchanged by the translation. It is called a form invariant for translations. A form invariant is a combination of coordinates that when transformed, although all the coordinate terms change, is itself not changed. In more formal language, the transformation of the coordinates, $\vec{x}' = \vec{x} + \vec{a}$, is a family of transformations called translations. The elements of this family are labeled by the parameters $\vec{a}$. In the invariant form, $(\vec{x}_2 - \vec{x}_1)$, the transformation $\vec{x}' = \vec{x} + \vec{a}$, changes all the elements in the form $(\vec{x}_2 - \vec{x}_1) \rightarrow (\vec{x}'_2 - \vec{x}'_1)$ but, because the label of the transformation, $\vec{a}$, drops out of the form $(\vec{x}_2 - \vec{x}_1) = (\vec{x}'_2 - \vec{x}'_1)$ that particular form is unchanged.

An important issue in the plane is the distance between two points. In the above paragraph, a distance scale has been defined for each coordinate direction. These need not be the same. This may seem to be a bizarre choice but it is does happen. I was born and raised in Philadelphia, a city with row houses. The unit of length was what was called a "block". The trouble with the "block" was that in the two directions the actual block had different lengths. In fact, the shorter direction was about a quarter of the long direction. There was another distance that was used called the "city block" which was the same length as the long direction block. When we talked about how far something was, we used city blocks. In other words, in terms of the coordinate grid,

$$d_{\text{city blocks}} = \sqrt{\Delta\text{coordinate}^2_{\text{long direction}} + 16\Delta\text{coordinate}^2_{\text{short direction}}}$$

where $\Delta\text{coordinate}_{\text{long direction}}$ is shorthand for the difference in the coordinates. You can now measure the distance in "city blocks". Actually, the situation was more bizarre than that because this is the distance in "city blocks" is as the crow would fly. For people, the distance is the total number of intervals over the grid between the places because that is how you have to move in this system.

Let's now follow the more usual construction of using the same distance scale in each of the coordinate directions. In fact, we can go one step further and use the same distance scale for all directions. Then the distance between two points can be found by using the distance scale along the direction set by the two points, as the crow flies. Saying this we now realize that it is assuming the rotation of the distance scale to different orientations does not change it. This is a expression of the underlying isotropy of the space. An alternative approach to finding the distance is to use the coordinate differences. To reproduce the effects of the reorientation of the rod, the measure of distance must reflect the translation and reorientation invariance of the distance measure. This rotational and translation invariance in the definition of distance is expressed as

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} = \sqrt{\Delta x^2 + \Delta y^2}. \tag{4.1}$$

In other words, the transformation of the coordinate system produces changes in the coordinates which for rotations and translations are $(x, y) \rightarrow (x', y')$ with $x' = x\cos(\theta) + y\sin(\theta) + a_x$ and $y' = y\cos(\theta) - x\sin(\theta) + a_y$ where θ is the angle of rotation and the label for the elements of the family of rotations and $\vec{a}$ is the labels for the translations. Equation 4.1, is a form invariant for these transformations.

Distance not only depends on the two points but also the path connecting the points. In our discussion above, we used the word distance for the separation of the points which is the distance over the straight line path between the points, as the crow flys. In the more general case, there can be an arbitrary path connecting the points. Since the number of paths

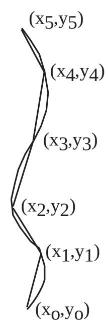

Figure 4.1: **A Path in a Plane** For a curved path, a cumulative distance can be assigned by adding the straight line distance for each segment of a sensibly rectified approximation to the curve, $d\,[traj.] = \sum_{i=0,\mathrm{Path}}^{f-1} \sqrt{(x_{i+1}-x_i)^2 + (y_{i+1}-y_i)^2}$. In the limit of small segments, this cumulative time is the proper time over the trajectory, $d\,[traj.] = \int_{(x_0,t_0),\mathrm{Path}}^{(x_f,t_f)} \sqrt{dx^2+dy^2}$.

between two points is an infinite class that is larger than the class of real numbers, you cannot perform ordinary analysis on the path labels. For this case, the name functional is used instead of function. This extension of the idea of functions to functionals leads to a wealth of new and very powerful mathematics.

For our purposes it is sufficient to consider paths that can be sensibly rectified into a sequence of straight line segments and the total distance is the sum of these intervals. We could measure the length of each segment by taking our length definition and placing it along the straight line segments measuring each length by aligning the length along the segment. This is a place where our assumption of homogeneity and isotropy come into play. The length of the rod is the same no matter how we orient it and where we place it. Alternatively, we can use the coordinate difference method but then we have to be sure the coordinate system reflects these symmetries. Our definition of length, Equation 4.1 accomplishes this if the coordinate directions use the same length. In this case, the path length is the sum of the appropriate straight line separations or

$$d\left[path:(x_0,y_0;x_f,y_f)\right]=\sum_{i=0,\text{Path}}^{f-1}\sqrt{(x_{i+1}-x_i)^2+(y_{i+1}-y_i)^2} \tag{4.2}$$

or in the limit of small intervals

$$d\left[path:(x_0,y_0;x_f,y_f)\right]=\int_{(x_0,y_0)\text{Path}}^{(x_f,y_f)}\sqrt{(dx)^2+(dy)^2}. \tag{4.3}$$

Using the Equation 4.2 and rotational and translational symmetry it is easy to show that the straight line path is the shortest. This takes advantage of another idea that is worth discussing at this point. Our previous discussions of the rotations and translations dealt with changes to the coordinate system. These same ideas can also be applied to the points themselves. You can view the transformation as shifting all the points. When the transformation is on the coordinate axis the transformation is called passive. When it is applied to the points, it is called active. In the case of proving that the straight line is the shortest path between two points, we can use a passive transformation to move the origin to one of the points. Then we can use an active transformation to rotate the the second point so that it is on the y axis. Using Equation 4.2, the path length of any arbitrary path other than a straight line will include δx contributions which can only increase the sum above that of the path with no δx contributions. There for the straight line path is the shortest path.

Although we have used the idea of angle in the above discussion of rotations, we have assumed the usual measure of angle and not discussed the exact nature of the relationship of angle to the rotation transformation. Our only real criteria when viewing the coordinate transformation was to preserve the form of the distance measure,

$$d=\sqrt{(x_2-x_1)^2+(y_2-y_1)^2}=\sqrt{\Delta x^2+\Delta y^2}.$$

As stated above, rotations and translations are the family of transformations that preserve this form,

$$d^2=(x_2-x_1)^2+(y_2-y_1)^2=\left(x_2'-x_1'\right)^2+\left(y_2'-y_1'\right)^2, \tag{4.4}$$

where x' and y' are the new coordinate labels for the point at (x,y). A more general transformation would be

$$\begin{aligned} x' &= ax+by+c \\ y' &= dx+ey+f. \end{aligned} \tag{4.5}$$

The form invariant is preserved by all values of c and f but the other parameters that label the transformation are constrained by the requirement of Equation 4.4. These constraints are

$$\begin{aligned} a^2 + d^2 &= 1 \\ b^2 + e^2 &= 1 \\ 2ab + 2de &= 0. \end{aligned}$$

and the solutions can be written in terms of the single parameter b as $a = \sqrt{1-b^2} = e$ and $d = -b$. In other words, our family of transformations is a three parameter group; two for translations, c and f, and one for rotations, b. Putting this family of transformations into a group adds the requirement that sequential operation of the transformation is a one of the members of the family. There is an added requirement that the family of transformations have an identity transformation. Our family of transformations satisfy these additional requirements. Thus the rotation segment of our transformations can be written as a matrix operating on the doublet (x, y) with the matrix given by

$$\begin{pmatrix} \sqrt{1-b^2} & b \\ -b & \sqrt{1-b^2} \end{pmatrix} \tag{4.6}$$

This is not the only solution to the constraint equations but it is a nice one in that the identity element, do nothing element, is the $b = 0$ element. A complication with this form for describing rotations is the requirement that two rotations are a rotation. This requires that

$$\begin{pmatrix} \sqrt{1-b_2^2} & b_2 \\ -b_2 & \sqrt{1-b_2^2} \end{pmatrix} \begin{pmatrix} \sqrt{1-b_1^2} & b_1 \\ -b_1 & \sqrt{1-b_1^2} \end{pmatrix} = \begin{pmatrix} \sqrt{1-b_3^2} & b_3 \\ -b_3 & \sqrt{1-b_3^2} \end{pmatrix},$$

or

$$b_3 = b_1\sqrt{1-b_2^2} + b_2\sqrt{1-b_1^2}. \tag{4.7}$$

Using the parameter b to label the rotations, you can see that they do not add when rotations are combined. This is an unfortunate statement because it is not really true. The bs add but not simply like numbers. By now, I would guess that many of you smell the rat in this analysis. If we had just used the good old fashioned idea of the angle to label the rotations, the rotation matrix would simply have been

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{4.8}$$

or $b = \sin\theta$ and the condition of Equation 4.7 becomes simply $\theta_3 = \theta_1 + \theta_2$ or usual numeric addition. When you think about it, you realize that the property of simple addition comes from the definition of the angle. The angel θ between two straight lines is

$$\theta \equiv \frac{S}{R} \tag{4.9}$$

where S is the arc length of the circle generated by a distance R as the line of length R sweeps from the one line to the other.

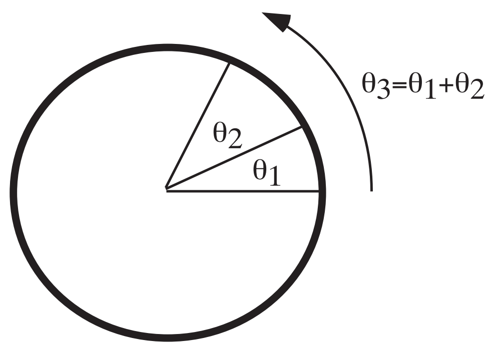

Figure 4.2: **Addition of Angles** The usual definition of angle is $\theta \equiv \frac{S}{R}$ where S is the arc length of a circle. Since this definition of angle uses the length of a segment of the invariant curve for rotations, the circle, and since the arc lengths of curves add simply, this "angle" adds simply.

In our discussion of the label of rotations above a certain curve in the two plane, a circle, played a special role in the definition of angle. The circle acquires its special role because it is the locus of places that is generated by rotations when the transformations are viewed actively. This definition uses a construction that is based on the fact that the active transformation of points generates a set of points on what is called the invariant curve or surface in higher dimensions. In other words, the circle is the invariant curve for our rotations. Using the arc lengths along these curves which are clearly additive in the numeric sense, we obtain an additive measure of rotations.

Our b in the previous analysis was a definition of amount of rotation, "angle," that was $\theta' \equiv \frac{h}{H}$ where h is the height and H is the hypotenuse of a right triangle constructed between the lines generating the "angle." Of course, we recognize this as $b = \sin\theta$ and then the complex addition formula, Equation 4.7 is a reflection of the fact that

$$\arcsin b_1 + \arcsin b_2 = \arcsin\left(b_1\sqrt{1-b_2^2} + b_2\sqrt{1-b_1^2}\right).$$

Let's take this analysis of rotations as active transformations one step further. In old fashioned two space, (x, y), if we consider rotations about the origin, we had a form invariant $x^2 + y^2 = r^2$, i. e. if the coordinate system is rotated, in the new coordinates that same point is now labeled as (x', y') and the combination, $x'^2 + y'^2$ takes on the same value, $x'^2 + y'^2 = r^2$. Now viewed actively, for every point, rotations generate a locus of places with the same distance from the origin satisfy this form invariant and are circles centered on the origin. A rotation will map one point on a circle onto another point on that same circle.

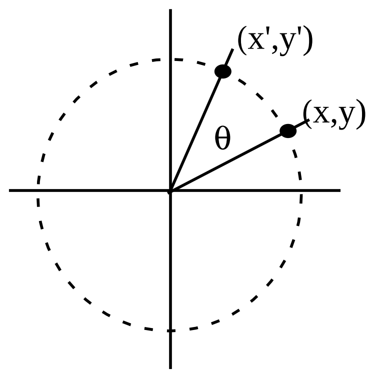

Figure 4.3: Rotations can be treated as a mapping of the points in the plane. Here the rotation θ maps (x, y) onto (x', y')

In the above analysis of the addition of angles, Equation 4.8, the special functions $\cos(\theta)$ and $\sin(\theta)$ did neat things for us. Another related property of these functions is that they satisfy the constraint, $x^2 + y^2 = r^2$, by having $x = r\cos(\theta)$ and $y = r\sin(\theta)$. This constraint is satisfied for any θ since $\cos^2(\theta) + \sin^2(\theta) = 1$. We also can describe the location of a place, (x, y), as a distance and an angle, (r, θ). It is a trivial observation that the rotations connect different places with the same distance. Consider three places, $(r, 0)$, (x_1, y_1) and (x_2, y_2), that are the same distance from the origin, i. e. on the same circle, $r = \sqrt{x_1^2 + y_1^2} = \sqrt{x_2^2 + y_2^2}$. The rotation that maps place $(r, 0)$ on to (x_1, y_1) is labeled by an angle θ_1, $\theta_1 = arctan(\frac{y_1}{x_1})$, and the rotation that maps $(r, 0)$ onto (x_2, y_2) is labeled by an angle θ_2, $\theta_2 = arctan(\frac{y_2}{x_2})$. A rotation with the angle labeled $\theta_2 - \theta_1$ maps (x_1, y_1) onto (x_2, y_2) – and again the angles are additive, see Figure 4.3.

We can use this idea to find the general transformation law for rotations, Equation 4.8. Consider the point $(r_1, 0)$, it is obvious that under the rotation θ this point is mapped to $(r_1\,cos(\theta), r_1, sin(\theta))$. Similarly, a rotation of the same angle θ maps $(0, r_2)$ into the point $(-r_2\, sin(\theta), r_2\, cos(\theta))$.

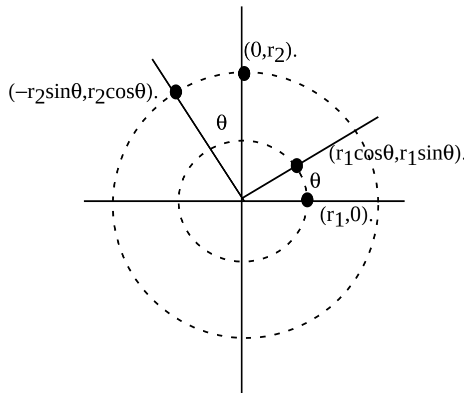

Figure 4.4: Rotating points on the coordinate axis to find the form of the transformations under rotations.

Since the transformations are linear, we can make the general transformation for any point (x, y) by combining these:

$$\begin{aligned} x' &= x\, cos(\theta) - y\, sin(\theta) \\ y' &= x\, sin(\theta) + y\, cos(\theta) \end{aligned} \tag{4.10}$$

You can easily derive the addition formula for the trigonometric functions by having two rotations. Starting with (x, y) and transforming by a rotation, θ_1 to (x', y') and then transforming (x', y') by a rotation θ_2 to (x'', y''), we have the sequence of equations:

$$\begin{aligned} x' &= x\, cos(\theta_1) - y\, sin(\theta_1) \\ y' &= x\, sin(\theta_1) + y\, cos(\theta_1) \end{aligned} \tag{4.11}$$

and

$$\begin{aligned} x'' &= x'\, cos(\theta_2) - y'\, sin(\theta_2) \\ y'' &= x'\, sin(\theta_2) + y'\, cos(\theta_2) \end{aligned} \tag{4.12}$$

and since the angles are additive,

$$\begin{aligned} x'' &= x\, cos(\theta_1 + \theta_2) - y\, sin(\theta_1 + \theta_2) \\ y'' &= x\, sin(\theta_1 + \theta_2) + y\, cos(\theta_1 + \theta_2) \end{aligned} \tag{4.13}$$

Substituting the first equation into the second and reorganizing, we also have

$$\begin{aligned} x'' &= x\,(cos(\theta_1)cos(\theta_2) - sin(\theta_1)cos(\theta_2)) \\ &\quad -y\,(sin(\theta_1)cos(\theta_2) + cos(\theta_1)sin(\theta_2)) \\ y'' &= x\,(sin(\theta_1)cos(\theta_2) + cos(\theta_1)sin(\theta_2)) \\ &\quad +y\,(cos(\theta_1)cos(\theta_2) - sin(\theta_1)cos(\theta_2)) \end{aligned} \tag{4.14}$$

Equating the coefficients of x and y, we have the usual formulas for the addition of angles of the trigonometric functions

$$sin(\theta_1 + \theta_2) = sin(\theta_1)cos(\theta_2) + cos(\theta_1)sin(\theta_2) \tag{4.15}$$

$$cos(\theta_1 + \theta_2) = cos(\theta_1)cos(\theta_2) - sin(\theta_1)sin(\theta_2) \tag{4.16}$$

.

The requirement of additivity is a linear one and thus does not fix the scale of the angles. Since θ is dimensionless, and we require that additivity hold for all r, a natural measure of angle is the radian, $\theta = \frac{arc\ length}{r}$. With this definition, the full circle has angle 2π.

4.3 Minkowski Space-time

In the discussion of Harry, Sally, and Dorothy, Section 3.3.5, we studied the fact that Sally and Harry had different trajectories in space-time. A trajectory is the connected set of events that represent the places and times through which an object moves. Trajectories of material objects and observers are called worldlines. Sally's time axis is her worldline. Sally is also an inertial observer; she experiences no acceleration in the course of her motion. Harry' s worldline, on the other hand, has a bend. He has an acceleration and that is knowable by him, see Section 1.2; he spills his martini on his shirt. Note that any other inertial observers coordinatizing this situation cannot be differentiated from Sally and would have her worldline as straight and his would still have a bend. Thus the idea of an inertial worldline, straight, is the same for all inertial observers and the straightness of the inertial observer worldline is unchanged, the Lorentz transforms, Equation 3.16, map straight lines into straight lines. The space-time that is coordinatized by Sally is an example of a Minkowski space-time. This is a space-time that has a global coordinate system such as Sally's and is also invariant under the Lorents transformations and space and time translations. This large group of transformations is called the Poincare' transformations.

The basic assumption of special relativity is that the events take place in a four dimensional structure that contains a three dimensional Euclidean space and a time like dimension. A $(3,1)$ space that has an invariant measure,

$$\Delta s^2 \equiv \Delta x^2 + \Delta y^2 + \Delta z^2 - c^2 \Delta t^2, \tag{4.17}$$

for the Poincare' transformations. It is easy to show by substitution of the Lorentz transformations, Equation 3.16, that the interval, Equation 4.17 is invariant under Lorentz transformations. Since it is defined by differences in coordinates, it is invariant under translations in space and time.

This $(3,1)$ is different from the Euclidean space plus time of Newtonian physics in that the group of transformations that govern it, the Poincare' transformations, preserve this different measure. The Galilean transformations, Equations 3.4, preserve the usual distance measure of the Pythagorian Theorem, $\Delta l^2 = \Delta x^2 + \Delta y^2 + \Delta z^2$, which is invariant under rotations and spatial translations, see Section 4.2. This is why the three spatial dimensions are a Euclidean space. In the case of the Galilean transformations, the time coordinate is unchanged. That is why the Newtonian world is a space plus time world not a $(3,1)$ world.

The geometry of this four dimensional, $(3,1)$, manifold is important to our understanding of the kinematics of Relativity. Although there are similarities with a four dimensional Euclidean manifold, the differences are important and often at the heart of what seems to be a paradox of Relativity. Since this manifold has translation symmetry, the structure is contained in the relationship between events; since this is a space-time, events are the fundamental element and not points. This is the first important difference since we tend not to think of ourselves as lines, a connected set of events - a sequence of heart beats.

In space-time as in space, we have the concept of a continuous connected set of events. This is called a trajectory. At any event on a trajectory, the slope is the inverse of the velocity relative to some inertial observer; the one that we choose to have a time axis, the $x = 0$ line, straight up and with a perpendicular set of lines of simultaneity, lines of constant t. The trajectories of light rays are straight lines with slope one. The trajectories of inertial observers are straight lines with slopes greater than one.

Space-time around any one event is divided into regions separated by the trajectories of light rays emanating from that event, see Figure 4.5. This separation of events is the same for all Lorentz observers since the light ray trajectories are unchanged by the Lorentz transformations. All the events in the upper light cone are the future of the event in question. This is in the

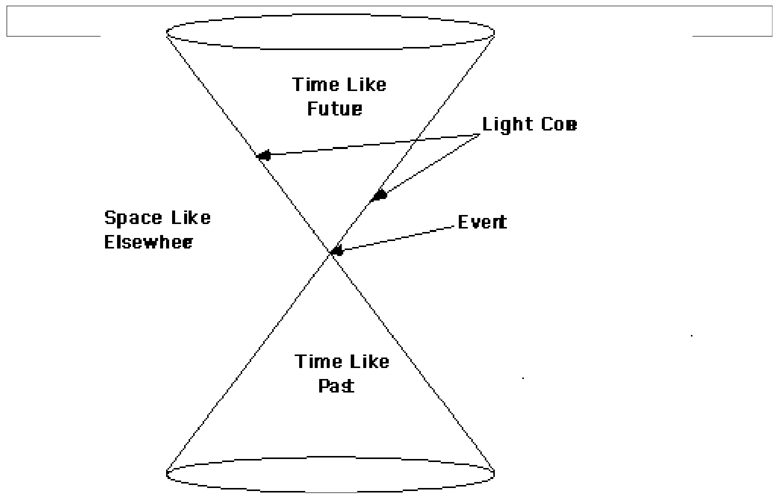

Figure 4.5: Future Past and Elsewhere. For any event, in this case the event at the vertex of the two cones, all the other events in space-time can be categorized into a future, a past, and an elsewhere. Since the trajectories of light rays are unchanged by Lorentz transformation, this classification of the relationship between two events is the same for all inertial observers.

sense that, from the origin event and any event in the future, there exists an inertial observer for whom the interval between these two events is a pure time, $(0, \tau)$, i. e. no spatial separation, and that the time of the other event is after the now of our original event, $\tau > 0$. Any other inertial observer would give the two events labels $(\vec{x}_0, t_0)$ as the original event and $(\vec{x}_1, t_1)$ as the other event and, using the form of the invariant, Equation 4.17, we have

$$c^2\tau^2 = -\Delta s^2 = c^2 (t_1 - t_0)^2 - (x_1 - x_0)^2 - (y_1 - y_0)^2 - (z_1 - z_0)^2 > 0. \tag{4.18}$$

Similarly, events in what is called the backward light cone from our original event are in the past. There exists an inertial observer for whom the second event is a pure time, $(0, \tau)$, but in this case $\tau < 0$. Again, any other inertial observer would give the two events labels $(\vec{x_0}, t_0)$ and $(\vec{x_1}, t_1)$ and, using the form of the invariant, Equation 4.17, we have $c^2\tau^2 = -\Delta s^2 = c^2 (t_1 - t_0)^2 - (x_1 - x_0)^2 - (y_1 - y_0)^2 - (z_1 - z_0)^2 > 0$.

The union of the events in the past and of the events in the future to our origin event is the set of time-like events relative to our origin event. This is all events relative to the original event with intervals in any inertial coordinate system such that the negative of the interval squared, $-\Delta s^2 =$

$c^2(t_1 - t_0)^2 - (x_1 - x_0)^2 - (y_1 - y_0)^2 - (z_1 - z_0)^2 > 0$. If the event under discussion is in the upper light cone or future of our origin event, we chose the positive sign for the square root of the negative of the interval squared and if the event is in the lower or past light cone we chose the negative of the square root of the interval squared. This is called the proper time between the events although we have to be careful because, just as is the case in Euclidean spaces where distance is the corresponding concept and we now realize that distance is path dependent, convention calls this the proper time between the events, see Section 4.2 and discussion later in this section.

There are clearly a large number of events that are not time-like relative to our origin event, see Figure 4.5. These are called elsewhere or space-like events relative to our origin event. Similar to our construction of future and past, for any elsewhere event there exists a Lorentz observer for whom the events are separated by a spatial interval, $(\vec{x}, 0)$. Again, any other inertial observer would give the two events labels $(\vec{x}_0, t_0)$ and $(\vec{x}_1, t_1)$ and, using the form of the invariant, Equation 4.17, we have $d^2 = \vec{x}^2 = \Delta s^2 = (x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2 - c^2(t_1 - t_0)^2 > 0$.

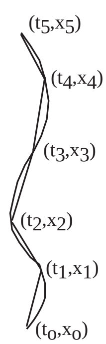

Figure 4.6: **A Time-Like Trajectory in Space-Time** For a curved trajectory to be time-like each segment must be time-like. A cumulative time can be assigned to a time like trajectory by adding the proper time for each segment of a sensibly rectified approximation to the curve, $\tau[traj.] = \sum_{i=0,\text{Traj},}^{f-1} \sqrt{(t_{i+1} - t_i)^2 - \frac{(x_{i+1}-x_i)^2}{c^2} - \frac{(y_{i+1}-y_i)^2}{c^2} - \frac{(z_{i+1}-z_i)^2}{c^2}}$. In the limit of small segments, this cumulative time is the proper time over the trajectory, $\tau[traj.] = \int_{(x_0,t_0),\text{Traj.}}^{(x_f,t_f)} \sqrt{dt^2 - \frac{dx^2}{c^2} - \frac{dy^2}{c^2} - \frac{dz^2}{c^2}}$.

Another important geometric concept deals with trajectories. It makes sense to describe a cumulative interval along a trajectory. Depending on the bending, it is sensible to approximate the cumulative interval by sensibly rectifying the trajectory and adding the intervals of each segment, see Figure 4.6,

$$\tau\,(traj.) = \sum_{i=0}^{f} \sqrt{(t_{i+1}-t_i)^2 - \frac{(x_{i+1}-x_i)^2}{c^2} - \frac{(y_{i+1}-y_i)^2}{c^2} - \frac{(z_{i+1}-z_i)^2}{c^2}}. \tag{4.19}$$

This is the same procedure that is used in the case of paths in an two, three, or even an n dimension Euclidean space, see Section 4.2. The complication here is the fact that intervals squared, Equation 4.17, comes in two varieties, time-like and space-like, negative and positive intervals squared respectively. Although it is possible to have trajectories with some time-like segments and some space-like segments, it does not make any sense to assign a cumulative interval to them. We will see shortly that, in addition, trajectories with space-like segments have problematic causal structure. For these reasons, we require that all trajectories of sensible material objects have time-like trajectories. A time-like trajectory is one in which all the segments are time-like intervals. An equivalent definition is that a trajectory is time-like if for every event on the trajectory all subsequent events are in the future light cone of that event and all previous events are in the past light cone. A special simple case of a time-like trajectory is the trajectory of an inertial observer. The proper time over this trajectory is to within an origin time the coordinate time for that inertial observer; it is the time on his clock. This idea is carried over to the case of any time-like trajectory. The cumulative proper times of the segments is called the proper time over the trajectory and is the time that would be recorded on a clock carried along that trajectory. This is actually what we did in our analysis of the travel time of Harry and Sally in Section 3.3.5. Without comment, we added the times of the segments of Harry's time as recorded on a clock used by an inertial observer that would be comoving with him; it was without comment because it was so eminently plausible. We will look at this question in more detail in a later section, Section 4.7. Note also that the slope of the segment is the inverse of the average velocity in the segment. In this sense the velocity along any time-like segment is always less than the speed of light or any segment that is space-like has an average velocity that is greater than c.

There are also trajectories in space-like directions, i. e, all segments space-like, and a cumulative distance can be assigned. This cumulative distance is called the proper distance of the trajectory. Since there are three spatial directions, there are also space-like surfaces, generated by two non-parallel space-like directions, and space-like volumes, all three sides space-like. These are not possible constructions for time-like situations since there is only one time coordinate.

As stated above the slope at any event is the inverse of the instantaneous velocity of the trajectory at that event. For time-like segments, this velocity can be used to reference a family of inertial observers that have that velocity. These are called comovers. The comover that shares the event is called the local comover. In our example of time differences for two traveler in Section 3.3.5, Sally and Dorothy were comovers. They were not local comovers since they were apart. Harry had two families of comovers, one for the first segment and another one for the second segment. The comover that moved with Harry on the either segment is the local comover for that segment. The anomaly of Harry's travel time is that he uses the clocks of two non-identical comovers. This is also the signature that he is accelerated. We will discuss this situation in more detail in Section 4.7.

As in our discussion of points in space, Section 4.2, transformations of a Minkowski space-time can be viewed both as active or passive. As in that case, the passive view identifies the transformation with a change in the coordinate system, how a relatively moving inertial observer would label the same event, and in the active view the transformation is between related events that have similar properties to an different inertial observer. When viewed actively, the Lorentz transformations are often referred to as Lorentz boosts. For example, consider the event that one observer says is simultaneous with his/her origin event and a distance d away along the positive x axis. Now consider a Lorentz transform with the label v. When viewed as a passive transformation, the event which was labeled $(d, 0)$ is now labeled $\left(\frac{d}{\sqrt{1-\frac{v^2}{c^2}}}, \frac{-\frac{v}{c^2}d}{\sqrt{1-\frac{v^2}{c^2}}}\right)$ Whe this transformation is viewed actively, the second label is an event that is at the same proper distance from the origin as the event $(d, 0)$ and is one that would be labeled as $(d, 0)$ to the inertial observer moving at a speed v in the negative x direction. Similarly, for the time-like separated event from the origin $(0, \tau)$, the Lorentz transformation labeled v produces the label $\left(-\frac{v\tau}{\sqrt{1-\frac{v^2}{c^2}}}, \frac{\tau}{\sqrt{1-\frac{v^2}{c^2}}}\right)$ and is the label for that event for an inertial observer moving with speed v along the positive x axis. When viewed actively, this is an event at the same proper time from the origin event as the original event and one that would be labeled as $(0, \tau)$ for a Lorentz observer who is moving at a speed v in the negative x direction.

4.4 Causality and Trajectories

An obviously important issue is the idea of causality. In Newtonian physics, causality expressed itself as preceding events could influence subsequent events but not visa versa. In relativistic physics, there are more subtle points. Influence is achieved by being able to get an object or message from one event to another. This was actually the basis for our designation of the events in the forward light cone from any event as that events future. These later events are ones that can be connected by a time-like trajectory. In other words, an observer at the origin event could throw a rock or a light ray and it would get to the location of the event before the event happened. This causal relation does not hold for events in the elsewhere of our origin event. There is no material or light signal that can go from our original event to the place of elsewhere events before they occur. Events in each others elsewhere are not causally related.

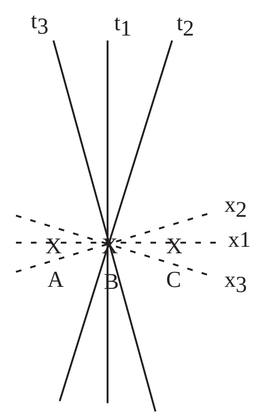

Figure 4.7: **Temporal Order of Space-like Events** Three events labeled A B and C are simultaneous to some inertial observer, Observer 1. Two other inertial observers move relative to 1 one to increased position, observer 2, and one to decreasing position, observer 3, and coincide with 1 at event B. Their lines of simultaneity for event B are shown as x_2 and x_3. To observer 2 event A is after B, $t_{A_2} > 0$, and event C is before B, $t_{C_2} < 0$. On the other hand, observer 3 has event A is before B, $t_{A_3} < 0$, and event C is after B, $t_{C_3} > 0$.

There is a more interesting example of causality breakdown associated with trajectories with space-like intervals. First consider a simple example with three events that are all space-like with respect to each other and simultaneous to some inertial observer. We could take the time of these events to be $t = 0$ for that observer, see Figure 4.7. The line of simultaneity of these events can be taken as $t = 0$ for that observer. There are two other inertial observers moving toward the original observer and all three coincide at the central event and synchronize their clocks to $t = 0$ at that time. For the two later observers, the events are no longer simultaneous. In fact, for the relatively moving observer moving to increasing position the event at positive position occurs before $t = 0$ and the event at negative position is after $t = 0$. For the observer moving to decreasing position, the equivalent situation obtains; the event at negative position occurs before $t = 0$ and the event at positive position occurs after $t = 0$. In other words, for events in the elsewhere of an origin event, the sign of the time of those events will depend on the inertial observer who coordinatizes the events. This is not the case for events that are future or past of the origin event. These are either positively signed for events in the future or negatively signed for events in the past to all inertial observers.

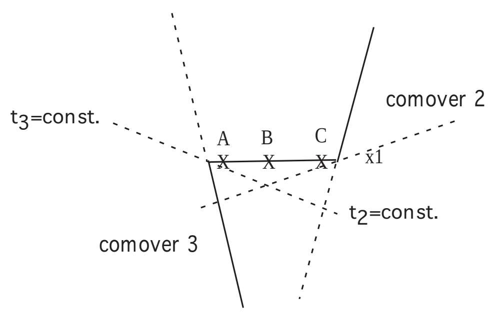

Figure 4.8: **A Trajectory with a Spacelike Interval** The Three events labeled A B and C of Figure 4.7 are part of a continuous trajectory. The inertial observer that is comoving with the first segment of the trajectory would indicate that the trajectory advances smoothly through the space-like interval and it makes sense to assign a direction. To the comover to the second time-like segment, the region has a flow direction that is the reverse of the usual. In fact as seen by the lines of simultaneity for comover 2, $t_2 = const.$, that for times slightly before event C to a time slightly after event A, there are three events that are on the trajectory.

The situation becomes more complex when we connect these space-like related events in a trajectory. Consider a trajectory that has the three events in the previous paragraph and Figure 4.7. The other segments of the trajectory are time-like and comove with the observers 2 and 3. In Figure 4.8, we show this trajectory. These two comoving inertial observers have a very different interpretation of the trajectory. To the comover of the first segment, the trajectory unfolds as a single trajectory with a uniform sense of flow. For the comover to the second time-like segment, the trajectory folds onto itself with an interval of time in which there are three events on the trajectory at any one time. This bizarre behavior makes no sense in classical physics. Suppose the trajectory represented in Figure 4.8 was a message being sent from each of the comovers. Comover 3 says that he/she is sending the message to comover 2 but conversely comover 2 says that he/she is sending the signal. Suppose that at event B the message is destroyed. Which comover did not get the message? In light of this causality problem, we make it an axiom that the trajectories of objects or messages must travel by time-like or for light light-like trajectories. In order to guarantee a coherent idea of causality, there is no signal that travels faster than the speed of light.

4.5 The Hyperbolic Hangle

Can we find a set of functions similar to the sin and cos and an additive measure that satisfy the form invariant for Lorentz transformations, $x^2 - c^2t^2 = d^2$? The answer to this question will be yes. We will do this analysis in a space with only one space and one time dimension for notational simplicity. The extension to higher dimensions is trivial.

Define

$$\cosh(\phi) \equiv \frac{e^{\phi} + e^{-\phi}}{2} \tag{4.20}$$

and

$$\sinh(\phi) \equiv \frac{e^{\phi} - e^{-\phi}}{2} \tag{4.21}$$

Also, define $\tanh(\phi) \equiv \frac{\sinh(\phi)}{\cosh(\phi)} = \frac{e^{\phi}-e^{-\phi}}{e^{\phi}+e^{-\phi}}$. Then $\cosh^2(\phi) - \sinh^2(\phi) = 1$ for all ϕ. These functions are not the only pair of functions that satisfy this relationship but, as we will see, they are the only pair that do that and also satisfy the additivity requirement for Lorentz transformations when they are parameterized by ϕ. Using ϕ and calling it the hyperbolic angle or "hangle," we can develop a set of relations for Lorentz transformations much like that which was accomplished in the previous section for rotations, Section 4.2. Another name for ϕ which we will occasionally us is the "rapidity." Hangle

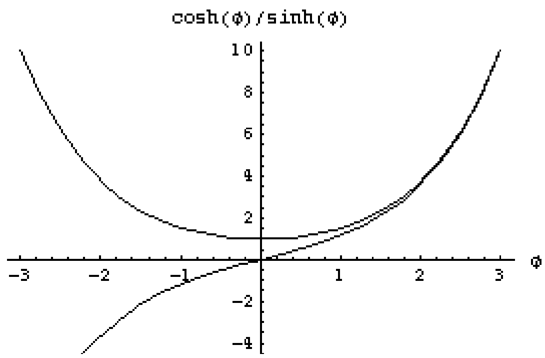

Figure 4.9: The functions $\cosh(\phi)$, topmost, and $\sinh(\phi)$, lower are plotted on the same graph.

reminds us of the relationship to angles and rapidity reminds us of the relationship to velocity. This second relationship will be clarified later.

First let's find the invariant surfaces for Lorentz transformations. This development follows the same pattern as the two space case, Section 4.2. These surfaces must be the values of (x, t) that satisfy the invariant form, $x^2 - c^2t^2$, events that have the same proper time and distance from the origin event. This is better expressed using the four hyperbolas $x = \pm\sqrt{d^2 + c^2t^2}$ for events that are space-like from the origin and $t = \pm\frac{1}{c}\sqrt{d^2 + x^2}$ for events that are time-like, see Figure 4.10.

In particular treating the Lorentz transformations with label $-v$ actively, the event $(d, 0)$ is mapped onto (x_1, t_1) where x_1 and t_1 are:

$$\begin{aligned} x_1 &= d\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \\ t_1 &= \frac{d}{c}\frac{\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}}. \end{aligned} \tag{4.22}$$

Obviously $d^2 = x_1^2 - c^2t_1^2$ and, as v varies from $-c$ to c, (x_1, t_1) moves along the hyperbola, $x = \sqrt{d^2 + c^2t^2}$. Also note that for any (x_1, t_1), the proper distance fo the origin is d. In this sense, the events along the hyperbola, $x = \sqrt{d^2 + c^2t^2}$, are the locus of events that are the same proper distance, d, from the origin. Similarly, the event $(-d, 0)$ generates the hyperbola $x = -\sqrt{d^2 + c^2t^2}$.

A similar argument holds for the event $(0, \frac{d}{c})$ which is on the upper leg of the invariant hyperbola, $t = \frac{1}{c}\sqrt{d^2 + x^2}$, and is mapped onto (x_2, t_2) on

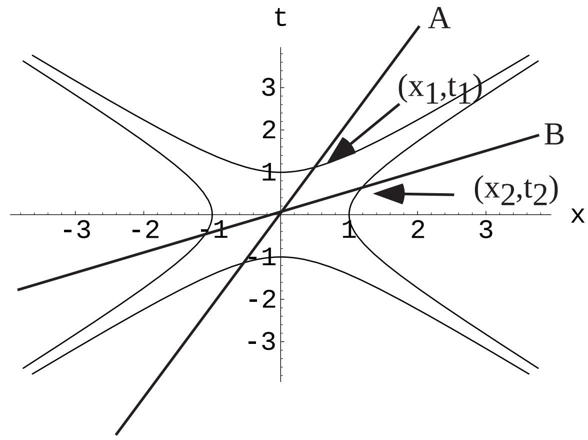

Figure 4.10: **Lorentz Invariant Surface** The four curves considered in a counter clockwise loop starting at the extreme right, $x = \sqrt{d^2 + c^2t^2}$, $t = \frac{1}{c}\sqrt{d^2 + x^2}$, $x = -\sqrt{d^2 + c^2t^2}$, and $t = -\frac{1}{c}\sqrt{d^2 + x^2}$ form the invariant surface for the Lorentz transformations. The events on the two hyperbolas, $x = \sqrt{d^2 + c^2t^2}$ and $x = -\sqrt{d^2 + c^2t^2}$, are in the elsewhere of the origin event. The events on the hyperbola, $t = \frac{1}{c}\sqrt{d^2 + x^2}$, are in the future of the origin event and the hyperbola, $t = -\frac{1}{c}\sqrt{d^2 + x^2}$, are in the past of the origin event. The inertial observer that shares the origin event and (x_1, t_1), Equation 4.22, has as the locus of events that are simultaneous with the origin event the line B passing through the origin and the event (x_2, t_2), Equation 4.23. The use of the term surface for these curves stems from the fact the in a three space one time world these are surfaces. For the figure the value of $d = 1$ was chosen.

that hyperbola as follows:

$$
\begin{aligned}
x_2 &= d\frac{\frac{v}{c}}{\sqrt{1 - \frac{v^2}{c^2}}} \\
t_2 &= \frac{d}{c}\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}
\end{aligned}
\tag{4.23}
$$

It is important to note that, for the inertial observer that passes through the origin and the event (x_1, t_1), the line of events that contains (x_2, t_2) and the origin event is the locus of events that are simultaneous with the origin event to that observer.

We can also identify x_1 and t_1 by

$$\begin{aligned} x_1 &= d\cosh(\phi) \\ t_1 &= \frac{d}{c}\sinh(\phi) \end{aligned} \tag{4.24}$$

and as ϕ develops the point (x_1, t_1) moves up along the hyperbola $x = \sqrt{d^2 + c^2t^2}$. Again, we identify x_2 and t_2 as;

$$\begin{aligned} x_2 &= d\sinh(\phi) \\ t_2 &= \frac{d}{c}\cosh(\phi) \end{aligned} \tag{4.25}$$

and as ϕ increases the point moves outward along the hyperbola. If we identify the line between the events $(0, 0)$ and (x_1, t_1) as the line of simultaneity and the line between the events $(0, 0)$ and (x_1, t_1) as the worldline of the inertial frame moving at speed v relative to the original frame, we have a new labeling for the Lorentz transformations.

In other words, and very similarly to the case of rotations, we can identify the Lorentz transformations of velocity v with a hangle ϕ as:

$$\begin{aligned} x' &= x\cosh(\phi) + ct\sinh(\phi) \\ ct' &= x\sinh(\phi) + ct\cosh(\phi) \end{aligned} \tag{4.26}$$

where $\tanh(\phi) \equiv \frac{v}{c}$.

The great advantage of this labeling of the Lorentz transformations is the additivity of the labeling in ϕ. To show this, we follow the same pattern that was used for the rotations. Consider two subsequent Lorentz transformations:

$$\begin{aligned} x' &= x\cosh(\phi_1) + ct\, sinh(\phi_1) \\ ct' &= x\sinh(\phi_1) + ct\, cosh(\phi_1) \end{aligned} \tag{4.27}$$

and

$$\begin{aligned} x'' &= x'\cosh(\phi_2) + ct'\sinh(\phi_2) \\ ct'' &= x'\sinh(\phi_2) + ct'\cosh(\phi_2) \end{aligned} \tag{4.28}$$

and, if we want the hangles to be additive, the compounding of these transformations should yield:

$$\begin{aligned} x'' &= x\cosh(\phi_1 + \phi_2) + ct\sinh(\phi_1 + \phi_2) \\ ct' &= x\sinh(\phi_1 + \phi_2) + ct\cosh(\phi_1 + \phi_2) \end{aligned} \tag{4.29}$$

Inverting the defining relations, Equation 4.20 and Equation 4.21, we have $e^{\pm\phi} = \cosh(\phi) \pm \sinh(\phi)$. Expanding the definition of $\sinh(\phi_1 + \phi_2)$ and $\cosh(\phi_1 + \phi_2)$ it is easy but tedious to show that these functions satisfy the correct addition relations so that they are equal to two successive Lorentz transformations of magnitude ϕ_1 and ϕ_2. The addition formula is:

$$\sinh(\phi_1 + \phi_2) = \sinh(\phi_1)\cosh(\phi_2) + \cosh(\phi_1)\sinh(\phi_2) \quad (4.30)$$
$$\cosh(\phi_1 + \phi_2) = \cosh(\phi_1)\cosh(\phi_2) + \sinh(\phi_1)\sinh(\phi_2) \quad (4.31)$$

Although the formula for the addition of velocities is rather cumbersome, the addition of hangles is simple.

As with the case of rotations, the hangle is dimensionless and the defining rules do not set a scale. The "natural" scale for ϕ is the ratio of the c times the proper time along the hyperbola $x = \sqrt{d^2 + c^2t^2}$, remember that it is a timelike curve, to the proper distance from the origin to that hyperbola. Calling this unit of hangle the hradian, we have

$$\phi(\text{in hradians}) \equiv \frac{c \times \text{proper time}}{\text{proper distance}} \quad (4.32)$$

Notice that the hangle goes to infinity as the relative velocity goes to c.

From the previous material and realizing that the commover at the event (x, t) has a relative velocity $\frac{v}{c} = \frac{ct}{x} = \tanh(\frac{c\tau}{d})$, we have:

$$x = d\cosh(\frac{c\tau}{d}) \quad (4.33)$$
$$ct = d\sinh(\frac{c\tau}{d}) \quad (4.34)$$

where τ is the proper time from the event $(d, 0)$ to the event (x, t) on the trajectory $x = \sqrt{d^2 + c^2t^2}$.

This rather extended diversion served two purposes. Firstly, it clarified the complex addition formula for collinear velocities. Here the problem was that velocity was not a good label for the family of transformation that is identified as the Lorentz velocity transformations. The additive label is the hangle not the velocity. This is similar to the case in two spatial dimensions discussed in Section 4.2. The second reason will be that the time-like trajectory that is generated by the invariant curve at a distance d from the origin will be seen to be the trajectory of the a uniformly accelerated observer, see Chapter 6. This is a special case of motion but it has great interpretive power and is valuable as an exact analytic solution to a non-trivial state of motion.

4.5.1 The same result directly using calculus

Unfortunately, the standard symbol for the derivative is lower case d; this is also our symbol for the proper distance to the trajectory of the uniformly accelerated object from the origin event, (0,0). In order to avoid confusion, in this section, I will use the symbol D for the proper distance and keep d for the derivative. It should be clear from the context the role the symbol d is playing.

From the definitions of sinh and cosh, Equations 4.21 and 4.20, we can derive

$$\frac{d(\sinh(\phi))}{d\,\phi} = \cosh(\phi) \tag{4.35}$$

$$\frac{d(\cosh(\phi))}{d\,\phi} = \sinh(\phi) \tag{4.36}$$

From the definition of the proper time along the trajectory between the event $(D,0)$ and the event (x,t),

$$c\tau = \int_{traj,(D,0)}^{(x,t)} \sqrt{c^2\,(dt)^2 - (dx)^2} \tag{4.37}$$

and the equation of the trajectory in terms of ϕ, $ct = D\sinh(\phi)$ and $x = D\cosh(\phi)$,

$$\begin{aligned} c\tau &= \int_0^{\phi(x,t)} D\sqrt{\cosh^2(\phi') - \sinh^2(\phi')}d\phi' \\ c\tau &= D\,\phi(x,t) \end{aligned} \tag{4.38}$$

where $\phi(x,t) \equiv \tanh^{-1}(\frac{ct}{x})$ is the hangle from the origin to the event (x,t).

Although this approach seems to be much simpler than the previous derivation, it must be kept in mind that the derivative and integral relations used above depend on the additivity properties that were the central part of the previous discussion. When you think about it you realize that the arc length along the trajectory is additive and therefore must me proportional to ϕ, the measure that is additive along the invariant curve.

4.6 Four Vectors and Invariants

4.6.1 Event Four Vector and Four Velocity

In the previous sections, we developed the idea of a Minkowski space, Section 4.3. In this section, we want to develop an efficient formalism for expressing ideas in Minkowski space. As in Euclidean space, a vector formalism is possible. Given an origin event and inertial observer, a coordinate system can be established. An event is a time and a place, a set of four numbers, $(t, \vec{x})$, that specifies that event in that coordinate system. We can designate the coordinates with an index x^μ with $x^0 \equiv ct$, $x^1 \equiv x$, $x^2 \equiv y$, and $x^3 \equiv z$. In this notation, the Lorentz transformations are expressed as

$$x'^\alpha = \sum_{\mu=0}^{3} \Lambda^\alpha_{\ \mu} x^\mu \tag{4.39}$$

with

$$\begin{aligned} \Lambda^\alpha_{\ \mu} &= \begin{pmatrix} \Lambda^0_{\ 0} & \Lambda^0_{\ 1} & \Lambda^0_{\ 2} & \Lambda^0_{\ 3} \\ \Lambda^1_{\ 0} & \Lambda^1_{\ 1} & \Lambda^1_{\ 2} & \Lambda^1_{\ 3} \\ \Lambda^2_{\ 0} & \Lambda^2_{\ 1} & \Lambda^2_{\ 2} & \Lambda^2_{\ 3} \\ \Lambda^3_{\ 0} & \Lambda^3_{\ 1} & \Lambda^3_{\ 2} & \Lambda^3_{\ 3} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{-\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ \frac{-\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{aligned} \tag{4.40}$$

for a Lorentz transformation along the positive x direction with speed v. Other Lorentz transformations are implemented similarly. The rotations which are a subgroup of the Lorentz transformations are the usual rotation elements operating in the bottom three by three spaces in this four by four object. There is some ambiguity in the identification of matrix elements and the two indexed objects, $\Lambda^\mu_{\ \nu}$. The location of indexes will be clarified later but for now the interpretation of Equation 4.40 and Equation 4.39 is consistent with Equation 2.2 and the Einstein convention. There is a broadly accepted convention that simplifies the notation considerably called the Einstein convention which eliminates the summation symbol for cases in which the same index appears up and down in the same term of an equation. In this notation, Equation 4.39 appears simply as

$$x'^\alpha = \Lambda^\alpha_{\ \mu} x^\mu. \tag{4.41}$$

In fact, it might appear that the lower second index on $\Lambda^\mu_{\ \nu}$ is placed there only to accommodate the Einstein convention. We will find that there is a more important significance to the placement of an index in the upper or lower position.

Given two events we can talk about the interval between them. In this language, there is a four vector interval

$$s^\mu \equiv \Delta x^\mu = \left(c\left(t_2 - t_1\right), \left(x_2 - x_3\right), \left(y_2 - y_3\right), \left(z_2 - z_3\right)\right). \tag{4.42}$$

The invariant interval squared, Equation 4.17, is now expressed as

$$\Delta x^2 = s^\alpha g_{\alpha\mu} s^\mu, \tag{4.43}$$

where

$$g_{\alpha\mu} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{4.44}$$

In this case, the identification of the matrix elements with the first index being the matrix row designator and the second index the column indicator. This is more direct than in Equation 4.40 for the Lorentz transformation. $g_{\alpha\mu}$ is called the metric in Minkowski space. This terminology and the index operations are a special case of the general formalism called indexology in Appendix C. We will not need the full power of the indexology formalism until we get to gravitation as geometry in Chapter 9. For our purposes now, it will be sufficient to deal with four vectors defined as any quartet of numbers that transform under Lorentz transformations in the same pattern as Equation 4.41. Lorentz scalars are quadratic forms of four vectors such as Equation 4.43 that are invariant under Lorentz transformations. We will also have use of higher order forms made from the four vectors such as second rank tensors. There are many examples of four vector quantities such as a four velocity, defined below, and the relativistic energy and momentum defined in Chapter 7. In this section, for quantities other than the event coordinates, we will take the transformation properties of these objects for granted and leave for the later definitions the proof of their transformation properties.

The condition that the interval squared be an invariant places a condition on the form of the Lorentz transformations,

$$\Delta x^2 = \Delta x'^2$$

which implies

$$\Delta x^\alpha g_{\alpha\mu} \Delta x^\mu = \Delta x'^\rho g_{\rho\gamma} \Delta x'^\gamma = \Delta x^\delta \Lambda^\rho{}_\delta g_{\rho\gamma} \Lambda^\gamma{}_\omega \Delta x^\omega$$

for all Δx^μ. This relationship is also true for any four vector. The reader must realize that all the indices in this expression are summed and are thus dummies and can take any convenient label as long as the set runs through the values 0,1,2,3.

Thus we have the condition that

$$g_{\alpha\rho} = \Lambda^{\mu}{}_{\alpha} g_{\mu\gamma} \Lambda^{\gamma}{}_{\rho}. \tag{4.45}$$

This condition can be interpreted in several ways. It can be used as the defining equation for the Lorentz transformations. The sixteen numbers, $\Lambda^{\mu}{}_{\nu}$, are a Lorentz transformation if they satisfy Equation 4.45. Equation 4.45 is not sixteen equations since $g_{\mu\nu}$ is symmetric. This is ten independent equations which leaves six free parameters. That is just what is needed – three parameters to label a velocity and three parameters to label rotations in a three space. Another interpretation of Equation 4.45 is that all Lorentz observers have the same metric. In other words, Equation 4.45 is a condition on how the object $g_{\mu\nu}$ transforms; it transforms into itself. It is an invariant tensor.

An alternative notational approach to invariant forms is to define a new set of coordinate four vectors defined by

$$x_{\mu} \equiv g_{\mu\nu} x^{\nu}. \tag{4.46}$$

Using this four vector the invariant takes the form $\Delta x_{\mu} \Delta x^{\mu}$. This is just the statement that for any linear vector space there exists a dual vector and it is linearly related to the original vector space. To facilitate the manipulation of both the upper and lower cased index objects, it is relevant to introduce the inverse metric tensor,

$$\left(g^{-1}\right)_{\mu\nu} \equiv g^{\mu\nu}, \tag{4.47}$$

or

$$g^{\mu\nu} g_{\nu\mu'} = \delta^{\mu}{}_{\mu'}. \tag{4.48}$$

where

$$\delta^{\mu}{}_{\mu'} \equiv \begin{cases} 1 & : \quad \mu = \mu' \\ 0 & : \quad \text{otherwise} \end{cases}. \tag{4.49}$$

For our case of the Minkowski metric,

$$\left(g^{-1}\right)_{\mu\nu} \equiv g^{\mu\nu} = g_{\mu\nu}. \tag{4.50}$$

This greatly simplifies the interpretation of the lower indexed four vector entities. It will not be the case when we deal with relativity as geometry.

The x_{μ} transform as

For a time-like trajectory, we can define a four vector velocity by using the proper time over the trajectory to calculate a rate of change. In other words, a trajectory which is a connected set of events which would be coordinatized by some inertial observer as $(\vec{x}(t), t)$, can be parametrized by the elapsed proper time of the time-like trajectory, $(\vec{x}(\tau), t(\tau))$, where

$$\begin{aligned} \tau\left[trajectory : (\vec{x}_0, t_0; \vec{x}, t)\right] &\equiv \int_{traj., \vec{x}_0, t_0}^{\vec{x}, t} \sqrt{dt^2 - \frac{d\vec{x} \cdot d\vec{x}}{c^2}} \\ &= \int_{traj., \vec{x}_0, t_0}^{\vec{x}, t} \sqrt{1 - \frac{\frac{d\vec{x}}{dt} \cdot \frac{d\vec{x}}{dt}}{c^2}} dt \\ &= \int_{traj., \vec{x}_0, t_0}^{\vec{x}, t} \sqrt{1 - \frac{\vec{v} \cdot \vec{v}}{c^2}} dt. \end{aligned} \tag{4.51}$$

where $\vec{v} \equiv \frac{d\vec{x}}{dt}$ is called the coordinate velocity and is the usual definition of velocity. This may look like a rather complex object but this construction is much like the parametrizing of a curve in two space with distance along the curve, see Section 4.2. The notation is also the same as that in Section 4.2. The elapsed proper time is a functional of the trajectory but is a function of the labels of the events at the end points of the integral. Since it is a function of the time on the trajectory we can derive a differential form for Equation 4.51. Differentiating with respect to the time of the event on the worldline,

$$\frac{d\tau}{dt} = \sqrt{1 - \frac{\vec{v} \cdot \vec{v}}{c^2}} \tag{4.52}$$

Putting all this together, we can construct a four vector velocity,

$$u^\mu \equiv \frac{dx^\mu}{d\tau} = \frac{\frac{dx^\mu}{dt}}{\frac{d\tau}{dt}}, \tag{4.53}$$

which, since the Lorentz transforms are linear and constant, transforms the same way as x^μ in Equation 4.41. In addition, note that all for components of u^μ have the dimensions of a velocity.

The construction of other kinematic four vectors such as a four acceleration follows the same pattern,

$$a^\mu \equiv \frac{du^\mu}{d\tau} = \frac{d^2 x^\mu}{d\tau^2}. \tag{4.54}$$

By construction of the proper time, it follows that the four velocity vector length is always the same,

$$u^\mu g_{\mu\nu} u^\nu = \frac{\frac{dx^\mu}{dt} g_{\mu\nu} \frac{dx^\nu}{dt}}{\left(\frac{d\tau}{dt}\right)^2} = \frac{\left(c^2 - \vec{v}\cdot\vec{v}\right)}{\left(1 - \frac{\vec{v}\cdot\vec{v}}{c^2}\right)} = c^2. \tag{4.55}$$

Any four vector with a positive length squared such as the four velocity is called time-like four vector; there always exists a Lorentz frame in which the four vector takes the form $\left(c, \vec{0}\right)$. In the general frame, the four velocity takes the form

$$u^\mu = \left(\frac{c}{\sqrt{1 - \frac{v^2}{c^2}}}, \frac{\vec{v}}{c\sqrt{1 - \frac{v^2}{c^2}}} \right). \tag{4.56}$$

Differentiating Equation 4.55,

$$\frac{d}{d\tau}\left(u^\mu g_{\mu\nu} u^\nu\right) = 0 = 2a^\mu g_{\mu\nu} u^\nu. \tag{4.57}$$

In the frame in which $u^\mu = \left(c, \vec{0}\right)$, the four acceleration must take the form $a^\mu = (0, \vec{a})$, where $\vec{a}$ is the usual acceleration. We derive this result below in 4.59.

Since the length squared is a Lorentz invariant,

$$a^\mu g_{\mu\nu} a^\nu > 0. \tag{4.58}$$

The acceleration is a space-like four vector. Differentiating Equation 4.56, the general form of the acceleration four vector is

$$\begin{aligned}
\frac{du^\mu}{d\tau} &= \left(\frac{\vec{v}\cdot\frac{d\vec{v}}{d\tau}}{\left\{1 - \frac{v^2}{c^2}\right\}^{\frac{3}{2}}}, \frac{\frac{d\vec{v}}{d\tau}}{\sqrt{1 - \frac{v^2}{c^2}}} + \vec{v}\frac{\vec{v}\cdot\frac{d\vec{v}}{d\tau}}{\left\{1 - \frac{v^2}{c^2}\right\}^{\frac{3}{2}}} \right) \\
&= \left(\frac{\vec{v}\cdot\frac{d\vec{v}}{dt}}{\left\{1 - \frac{v^2}{c^2}\right\}^{\frac{3}{2}}}, \frac{\frac{d\vec{v}}{dt}}{\sqrt{1 - \frac{v^2}{c^2}}} + \vec{v}\frac{\vec{v}\cdot\frac{d\vec{v}}{dt}}{\left\{1 - \frac{v^2}{c^2}\right\}^{\frac{3}{2}}} \right) \frac{dt}{d\tau} \\
&= \left(\frac{\frac{\vec{v}}{c}\cdot\frac{d\vec{v}}{dt}}{\left(1 - \frac{v^2}{c^2}\right)^2}, \frac{\frac{d\vec{v}}{dt}}{\left(1 - \frac{v^2}{c^2}\right)} + \frac{\vec{v}}{c}\frac{\vec{v}\cdot\frac{d\vec{v}}{dt}}{\left(1 - \frac{v^2}{c^2}\right)^2} \right)
\end{aligned} \tag{4.59}$$

In the frame in which $u^\mu = \left(c, \vec{0}\right)$, i. e. $\vec{v} \to 0$, the acceleration four vector is $(0, \vec{a})$, where $\vec{a} = \frac{d\vec{v}}{dt}$ is the coordinate acceleration as measured by a comoving inertial observer. This is the acceleration of Newtonian physics.

4.6.2 Other Four Vector and Four Tensor Elements

Using action as the basic principle of mechanics, see Appendix A, and using Noether's Theorem, Section A.5.4, we can define a relativistic energy and momentum which naturally fall into a four vector. In four vector notation, Noether's Theorem provides that the construction of energy follows from the actions as $E \equiv \frac{\delta S}{\delta t_f}$ and the momentum is $\vec{p} \equiv \frac{\delta S}{\delta \vec{x}_f}$ and, since t_f and $\vec{x}_f$ transform as a four vector and S is Lorentz invariant, the four element object $E^\mu \equiv (E, \vec{p}c)$ will also transform as a four vector. The factor of c on the momentum terms is used to keep the dimensions of the elements of the four vector uniform. For the relativistic free particle, the energy momemtum four vector is simply mcu^μ. The associated invariant is $E^\mu g_{\mu\nu} E^\nu = m^2c^4$. Thus the the energy four vector is time like. The four vector nature of E^μ implies that, in the locally comoving frame, $E^\mu = \left(mc^2, \vec{0}\right)$ and that in general

$$E^\mu = \left(\frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}}, \frac{mc\vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \right) \tag{4.60}$$

Since the Special Theory of Relativity was originally discovered by Einstein from his analysis of the implications of Maxwell's Equations, Equations B.19, B.20, B.21, and B.22. It should be no surprise that the electric and magnetic fields and currents take an especially simple and elegant form in this four vector notation.

Further manipulations of four vectors can be found in Section 6.2.1, Chapter 7, and Appendix C.

4.7 Harry, Dorothy, and Sally Revisited

With the insights gained from the previous discussions, it is worthwhile to revisit Harry, Dorothy and Sally of Section 3.3.5. In that section, we found that Harry, on his return to Sally, had not aged as much as she had and, in addition, he was accelerated. With the definitions of Section 4.3, Equation 4.19, we can see that this effect is a special case of a general result that in space-time, for timelike trajectories, the straight line trajectory, one possessed by an inertial observer, is the longest trajectory and all other timelike trajectories are shorter. The proof of this statement follows from the same reasoning as the proof the in space the straight line is the shortest line. The difference here in the negative contributions of the spatial intervals in the sum of the segments that contribute to the total proper time. In other words, given any two time-like separated events and a connecting time-like trajectory, there always exists an inertial observer for whom the initial and final events are at the same place. Of all the trajectories that can connect these two events, the one with the longest proper time is the straight one since all others will have some contributions from segments with spatial contributions which will reduce the proper time below that of the coordinate time difference.

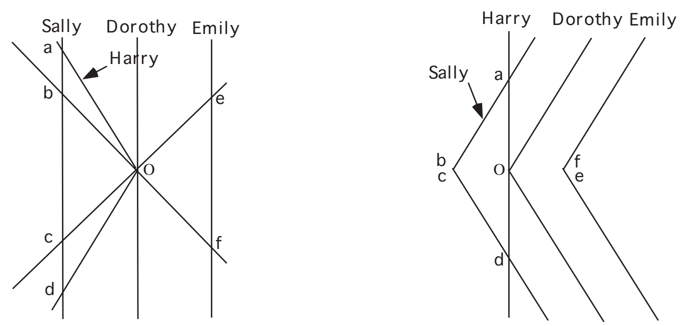

Figure 4.11: **Harry's Turn Around** At event d, Harry leaves Sally moving relative to her to increasing x at a speed of $\frac{3}{5}c$. On reaching Dorothy who is comoving with Sally and one light year away according to them, the event o, Harry turns around and starts back toward to Sally, again, at $\frac{3}{5}c$, meeting her at event a. The line eoc is the line of simultaneity to a Harry co-mover just before turning around and the line bof is the line of simultaneity to a Harry co-mover just after turning around. These co-movers are inertial and thus can develop consistent coordinatizing of space-time. If Harry were to define his coordinate system as that of the co-mover with him before turn around up until turn around and that of the co-mover moving with him after turn around after turn around, since the turn around takes place very quickly all the events within the cones boc and eof are basically at the time that he would label as $\frac{4}{3}$ year. This difficulty is not relieved by spreading the turn-around out in time. Although Sally's events b and c would now be separated and in the correct order, a comover with her such as Emily who is one light year further from her than Dorothy would have her events e and f recorded in a sequence the opposite from that which she records.

Another interesting question is what is the nature of a coordinate system that would be developed by Harry. In other words, how would Harry describe the events around him? Until Harry met Dorothy, he was an inertial observer. The co-mover with him is always inertial and obeys all the requirements to establish a valid coordinate system. This comover could develop a coordinate system that is a Lorentz transform of Sally's, and this would be the same as Harry's at least until he turns around. Once he turns around, he is again inertial and has a new comover who could be used to establish a coordinate system. If we argue that this dual comover coordinate system is the coordinate system that Harry would use, there is an anomaly associated with the events that are between the lines of simultaneity before and after the turn around. If we consider the turn around to be instantaneous, these events must all be at the same time, see Figure 4.11. In other words, events on Sally's world-line that are between the events labeled b and c in Figure 4.11, and are clearly separated in time are all recorded by Harry as at the same instant. This problem is not solved by having the turn-around spread over a small but non-zero interval of time. Although the events on Sally's world-line are now separated and reasonable, the events on a world-line of another inertial observer, say an Emily who was comoving with Sally and Dorothy and further from Sally than Dorothy on the same side as Dorothy, would have events on her world-line that are not in the correct temporal order; those that Emily say occurred earlier are timed later by Harry.

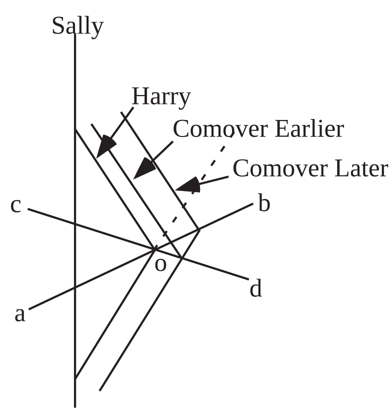

Figure 4.12: **Cohort Coordinatizing by Harry** shinola

Another approach to coordinatizing by Harry could be with comoving confederates. Of course, he if he uses two sets of confederates, he will reproduce the situation described above. If he requires his confederates to actually reproduce his motion, have an accleration, there are several immediate ambiguities. The first is what event on the confederate's worldline should be the turn-around event? Using Figure 4.12. There are two obvious choices; the line of shinola