Chapter 2

Introduction to Field Theory

2.1 Action at a Distance and Field Dynamics

A field is something, generally a measured quantity, that is defined at every point on some manifold. For us the manifold will usually be space. In addition, as you move from one point to a nearby point the value of the something changes smoothly; it varies as you change places. Another essential aspect of the field is that it can also evolve. It is sometimes interesting to discuss static configurations but these are usually only in anticipation of the more important case of situations in which the field changes with time. To appreciate these rather abstract comments let's look at several examples.

There are numerous examples of fields. The temperature in a room is a field. Temperature is measured for instance by a mercury bulb thermometer. As you move the thermometer from point to point, you will get different values for the temperature. If the room is not too drafty, the temperature at nearby points will be similar; the temperature varies smoothly as you move to nearby points. You can even intuit certain rules for how the temperature changes as you move from point to point. For instance, you can guess that a point at the center of a surrounding group of points, the temperature will be the average of the temperatures of the surrounding points. It is because of rules like this that you expect that the temperature varies smoothly as you go among nearby points. In many cases, the temperate distribution in the room will be the same whenever you measure it. If on the other hand you introduce a heat source, the temperature distributions will begin to change. You have a dynamic situation. Other obvious examples of fields are air pressure in a room, height above or below the normal height of water in a pool, or the transverse displacement of a stretched string. With some amount of smoothing you can make a field from such things as population density on the earth. Any system with a dynamic that is defined over a continuous manifold is a field.

In other words, a field is something that is defined over some manifold, usually space, that has a temporal evolution. The rules for the behavior of the field both as regards the spatial and the temporal variation are usually local in the sense that its variation in space and time is determined by what is going on at those points of space at those times. This is the meaning of local causality. It is one of the bedrock principles of modern physics. It ranks with reductionism as one of our fundamental formulating rules. The basic idea is that what happens to an entity happens because of what is going on at the place at which the entity is or the immediate neighborhood. This is in sharp contrast to the situation in theories that are based on what are called "action at a distance" dynamics. Newton's Laws of gravitation are an example of an action at a distance theories. To a large extent, it was the attempt to remove these action at a distance formulation and replace them with locally causal theories that motivated the development of field theories. More importantly, it became apparent that the fields where themselves the fundamental structural elements. In order to appreciate the importance of these ideas let's contrast the field construction with those of action at a distance theories.

2.1.1 Action at a Distance

My former colleague, Johnny Wheeler calls it "spooky" action at a distance. Newton (1642–1727), its inventor, was not comfortable with the concept but could not come up with something better. In a letter to the theologian Robert Bentley, he wrote:

> that gravity should be innate, inherent and essential to Matter, so that one body may act upon another at a Distance thro' a Vacuum, without the Mediation of any thing else, by and through which their Action and Force may be conveyed from one to another, is to me so great an Absurdity that I believe no Man who has in philosophical Matters a competent Faculty of thinking, can ever fall into it. Gravity must be caused by an Agent acting constantly according to certain laws; but whether this Agent be material or immaterial, I have left the consideration of my Readers.

The anomaly of action at a distance force systems is apparent to anyone who tries to teach the concept of force to students for the first time. The usual fall back for a definition of force is to say that it is a "push or a pull" with the connotation that it is a contact force – a hand on the back of a cart, see comments in Appendix A.1.

Regardless of his own reservations and because of the success of the Newtonian approach, physicists became accepting of the anomalous nature of action at a distance and the early formulations of most laws of physics were all in the pattern of action at a distance. Fortunately, Maxwell could not believe these and, for the case of electricity and magnetism, this lead him to the development of the first first-principle field theory. Prior to Maxwell's work there were field theories but these were derivative of an underlying structure. For example, the rules of fluid flow were formulated in a field theory vocabulary. But this was understood to be a consequence of the underlying structure of the fluid. Maxwell's formulation of the nature of the electric and magnetic systems was actually a statement on the intrinsic properties of these entities. In order to understand this important idea, let's review the situation with action at a distance theories and the contrast to field theories.

All the satisfactory theories prior to the 19th century were not what we now call locally causal theories but instead were bases on action at a distance theories, actions resulted from situations that were at a distance from the object of interest. Newton's theory of the gravitational force is a perfect example. In Newton's approach to gravitation, a bodies motion is determined by the separation from a remote other body at the instant under consideration. The moon's acceleration is determined from knowledge of the earth's position which of course is not in contact with the moon but is at a distance and it uses the information about the separation at that instant. It is hard to accept that, if the earth suddenly ceased to exist that, at instant, the moon would instantaneously react by traveling off in a straight line, no longer in orbit. There are two issues here. First the idea that somehow that moon is influenced not by things going on where it is and the fact that the earths disappearance should be realized by the moon instantaneously; it should take some time. Consider the case that I am standing in the front of the lecture hall and announce that I am going to make the clock at the back of the room run differently. If I could do that, you would infer that I had a wire or used sound waves or some other mechanism to communicate the change to the immediate vicinity of the clock. Whatever ultimately changed the clocks running was at the place of the clock not at a distance.

Coulomb's law and all the other laws of electromagnetism that were formulated before the 19th Century were action at distance laws. A charge here effected a charge there and did so instantaneously. Of course, the instantaneous nature of the interaction is especially problematic in light of the modern ideas of Theory of Relativity. Because of difficult underlying conflicts in a theory that mixed particles and fields such as the one that we will develop, there have been many modern attempts to resuscitate a Relativistic action at a distance formulation of particle electromagnetic interactions. A most interesting one that we will deal with is

The solution to the basic philosophical conundrum of action at a distance is in the idea of strict locality for all phenomena and the vehicle is the concept called the field. The ultimate solution is that all interactions are field field interactions and that particles are a derivative concept. This is achieved in the Quantum Theory of Fields. In this course, we will certainly get half way there – a theory with problems of point particles and fields. With luck, we will be able to introduce the entire construction of the Quantum Theory of Fields, Section **??**, and realize a single consistent framework.

Regardless, Maxwell could not deal with an action at a distance description of electromagnetic phenomena. He considered it an impossible description of fundamental processes, see Section 4. Of course, in physics, a philosophic problem is not a good reason for doing something. The idea must be tested experimentally. The proof of the construction is in the testing. Through his treatment of electromagnetic phenomena as a field theory, Maxwell was lead to predict that light was a disturbance of the electromagnetic field. When this prediction was verified by Heinrich Hertz in 1887, there was a general acceptance of Maxwell's approach. Since that time, we have found that all fundamental theories are field theories; the ultimate modern expression of the nature of matter and energy being through the machinery of quantum field theory. For this reason, it is important to understand the idea of the field. For now we will develop the classical field without sources. This is not the usual approach, see Jackson. We will then add the sources and display the problems that develop.

2.1.2 Local Field Theory

Maxwell developed a local field theory to describe the phenomena associated with what is called electricity and magnetism. He reduced all the known laws of electricity and magnetism into four reasonably simple equations. In so doing, he unified the electric and magnetic forces and predicted the fundamental nature of light. These are considerable accomplishments in their own right but also he somewhat inadvertently clarified the idea of the field and the idea of causality. His was not the first field theory; it was the first field theory of a fundamental force system. The first local field theory and the easiest to appreciate was the description of fluid flow. It was the success of a field theory of fluid flow that motivated him to attempt to write the rules of the electricity and magnetism in this field theory form.

How fluids move through space is very complex. At any point in the fluid there are several variables that are necessary to describe the state of the fluid. These variables such as density, velocity, and temperature are all fields, defined at each point in space and subject to change by some set of rules that are determined by the values of these variables at that point and nearby points and by the nature of the fluid. For example, if the temperature at a point is higher than its neighbors, that temperature will tend to decrease because of heat flow to the neighbors. Also depending on the nature of the fluid, the density may increase and this will cause matter to flow away from the point. How much effect each variable has on the magnitude of the the other variables and how fast these variables respond will depend on the fluid. The parameters such as the thermal conductivity and compressibility of the fluid which will control the rates at which these effects can take place are measured phenomenologically for each fluid. It is not hard to understand that the properties of a fluid in motion are controlled by local effects; flow at a point depends on the temperature and pressure and flow at the point and neighboring points not on what is going on some distance away. The rules for the fluid flow are thus local. The difference with the results of Maxwell is that we know there is an underlying structure, the atoms. In the case of the electromagnetic field, it is not made of anything but itself. The inability to associate a reality to the field independent of an underlying structure is the basis for the famous search for an ether, see Section **??**.

In fact, Maxwell suffered from that same problem. He discovered his equations by trying to fill space with a hypothetical something, vortices, that exhibited reasonable mechanical properties and attributing the electric and magnetic forces to the whirling of the vortices that filled the pervasive medium. The idea was that charges produced vortices in this medium and that the whirling of the vortices close to the charge then produced other vortices etc. until space was filled with whirling vortices and the amount of whirling at any place was the electric force. In other words, in order to understand his own equations, he needed an ether, the famous ether that Einstein disposed of later. He also needed to have the vortices properties be determined by the charge or the whirlyness locally. To the modern physicist, the idea of an underlying mechanical system seems out of place and a little weird. In fact, several years back, there was a collection of articles published that were "lighthearted" musings by well known scientists, [Weber 1973]. These articles were written as joke. Among the collected articles was the original paper by Maxwell justifying his vortices in the ether as a mechanism for the electromagnetic field. At the time of the writing, there was nothing lighthearted about it.

2.2 The Stretched String

Since the concept of the field and its dynamical rules are rather hard to grasp in the abstract, let's look at a particularly simple mechanical field system – the transverse displacement of a stretched string. I have to emphasis that this is a field with an obvious underlying mechanical structure – the string, a system with mass and an internal force, the tension. This is in contrast with the fields that we will deal with later. These fields are themselves the fundamental entities. The other thing to realize is that the string that we deal with is an idealized element. It has zero thickness and bends with no resistance. Its only possible displacement is transverse to its alignment. In addition, a real extended string in our three dimensional space has a two plane of transverse directions. For our simple analysis of the field properties of the string, I will use only one of these transverse directions. Later, when we have developed this field system in a more global form, we will add the extra transverse direction and discuss its consequences, see Section **??** This one dimensional string field is much simpler that the electromagnetic field which is a field composed of two vector quantities, the electric and magnetic forces, and with constraints. This one dimensional string is about as simple as you can get, a scalar field in one space one time with a simple dynamical rule. It will turn out that even this simple string system is sufficiently complicated to be interesting.

Like most mechanically based systems the dynamics of the string has two simple sources, energy of the motion of its masses and a potential energy that is due to its configuration. For the case of a string with a tension T_e and with only transverse displacements, the potential energy is the work associated with making the string longer. The displacement of the string in the transverse direction is the field that we will consider and any non-zero displacement causes the string to be longer and thus changes the potential energy. These are global approaches to the behavior of the string and will be useful to us later when use a more universal approach to dynamics based on the concept called action, see Section 3.1. For now because our ultimate goal will be an understanding the conceptual basis of the electromagnetic field, we will use a more local approach for the string and find that the electromagnetic field has many of the same properties as this the simplest of fields. In this approach the electromagnetic field is just a more complex field and the complications do not add much to the understanding of the field nature of the system.

You may also be perplexed by the idea of a stretched string under tension. Our experience is that a string has to be fastened to be under tension. If that is the case, think of the string as tightly stretched between fixed walls. The problem with this is that the walls add complications of their own and for the first pass are not necessary. Here we deal with an idealized situation of an infinite string under tension.

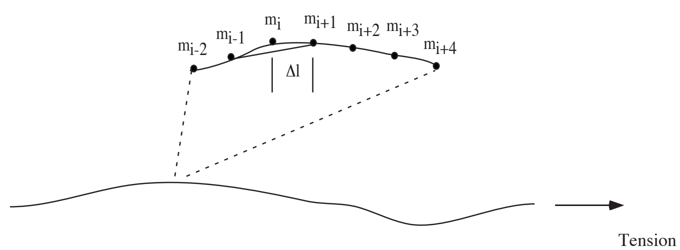

Figure 2.1: **The Stretched String** A string that can move in the transverse direction under tension is a simple example of a local field. In the figure, a section is magnified. In this section, the string is divided into small segments of length Δl and the mass of each segment is concentrated at a point. The dynamic of the string is that the mass segment at location i has a force on it if its transverse displacement is different from the average of its two neighbors. Thus in the case shown, by drawing a straight line between masses at $i-1$ and $i+1$, we can see that at the place of segment i, the neighbors' average is below i's current position. Thus i has a downward force on it.

2.2.1 The Local Stretched String

The local statement of the dynamics of the string are easy to understood; the rule is very simple and intuitive: The force on a segment of the string caused by the transverse displacement of that piece of the string is proportional to the negative of the average of the displacement of that segment of the string from the displacement of its neighbors.

In order to implement this algorithm, divide the string into small segments of length Δl and concentrate the mass in the segment at a point, see Figure 2.1. In the example shown, the segment of string labeled i is above the position of the average of its two neighbors. Thus there is a force to bring it to the position of the average. The proportionality constant for this force has the dimensions of a force per unit length and is thus the twice the tension in the string divided by the length of the segment of string; twice since both neighbors pull. ρ is the mass per unit length of the string and thus the mass of each segment is $\rho\Delta l$. Using $\vec{F} = m_i\vec{a}_i$ and using the position along the string x as label for the piece of string, the transverse displacement of the string at x is $y(x,t)$, the average of the two neighbors of x is $\frac{\{y(x+\Delta l,t)+y(x-\Delta l,t)\}}{2}$, the force equation for the segment at x is

$$\rho\Delta l a_{x,t} = -\frac{2T_e}{\Delta l}\left[y(x,t) - \frac{\{y\left(x+\Delta l,t\right)+y\left(x-\Delta l,t\right)\}}{2}\right], \tag{2.1}$$

where T_e is the tension in the string.

Another way to organize the right side of Equation 2.1, is to note that

$$\frac{2}{\Delta l^2}\left[y(x,t) - \frac{\{y\left(x+\Delta l,t\right)+y\left(x-\Delta l,t\right)\}}{2}\right] = \\ -\left\{\frac{\Delta y}{\Delta l}\left(x+\frac{\Delta l}{2},t\right) - \frac{\Delta y}{\Delta l}\left(x-\frac{\Delta l}{2},t\right)\right\}. \tag{2.2}$$

This last term on the right is the negative of the definition of the second derivative of $y(x,t)$. Note also that the acceleration is the second derivative with respect to time. In the limit that Δl is zero and using partial derivatives because we have both x and t dependence, this force equation becomes

$$\rho\frac{\partial^2 y}{\partial t^2}(x,t) = T_e\frac{\partial^2 y}{\partial x^2}(x,t). \tag{2.3}$$

This is an excellent example of the general form in which the dynamics of fields are expressed. They are generally partial differential equations because we are interested in how the field changes for changes in position and time. Equation 2.3 is second order in the time derivatives because that is how the dynamic operates; it emerged from a mechanical force law based on the idea of force producing acceleration. Other orders of time derivatives are possible and it is not uncommon to have laws that are first order in time. In fact, it is preferable because the interpretation of the evolution is simpler. Maxwell's Equations are an example. The stretched string or any higher order temporal evolution can be reduced to a first order temporal evolution by defining new fields. Defining a new field, the velocity field, $v(x,t) \equiv \frac{\partial y}{\partial t}$, we can get an evolution that has only first time derivatives.

$$\begin{aligned} \frac{\partial y}{\partial t}(x,t) &= v(x,t) \\ \rho \frac{\partial v}{\partial t}(x,t) &= T_e \frac{\partial^2 y}{\partial x^2}(x,t). \end{aligned} \tag{2.4}$$

In a very real sense, you could say the the magnetic part of the electromagnetic system is a manifestation of this kind of substitution. More on this later, see Section **??**. Equation 2.59 is our dynamic for the stretched string.

An important feature of this dynamic is that it is linear in the field variable. This is a very important feature of this field dynamic and allows for many of the properties that we associate with fields – superposition of effects. In this case, it came from the simple formulation of the dynamic. This linearity is a nice feature but as we will see is only approximate. For the case of the electromechanical field system this linearity is a consequence of the formulation of Maxwell and is supported by experiment. When we get to the quantum field theory of electromagnetism, the simplifying linearity will be gone.

The fact that there are only values of the field and spatial derivatives of the field on the right side of the Equation 2.3 is the expression of the locality of the dynamic. This is the definition of local causality: how the field evolves at a place depends only on what is going on at that point.

Also note that the only parameters in the field equation are ρ and T_e. These express the intrinsic properties of the medium in which the field operates. By dividing Equation 2.3 by ρ, we can reduce the effective number of parameters to one, $\frac{T_e}{\rho} \overset{\text{dim}}{=} \frac{L^2}{T^2}$. This has the dimensions of a velocity squared. The fact that there is only this parameter in the dynamic says a great deal about the nature of the evolution of the fields. There are not enough parameters to construct a length or a time. Thus for this field there is no intrinsic size except as it is put in by the starting conditions or put into the problem by boundaries like walls. Thus this particular field system, the stretched rope, is characterized by movement of field configurations. Since the parameter of the medium is a velocity squared, the movement is in both directions with a characteristic speed, $\pm\sqrt{\frac{T_e}{\rho}}$. It is important to remember that the movement of a piece of string is only in the transverse direction whereas the movement of the field configurations is along the direction in which the string is extended.

In order to better understand the operation of field dynamics let's work though the example of the string under tension. Consider our case of a stretched string with mass per unit length ρ and tension T_e. At $t = 0$, we put a distortion in the string as shown in Figure 2.2. Note that at $t = 0$, the string is displaced but no part of the string is moving. It is simplest to interpret the operation of the dynamic in the first order time derivative form, Equation 2.59. In this form, it is clear that a complete description of the initial configuration of the string involves the specification of two fields, the initial velocity field and the initial displacement field. In other words for the case in Figure 2.2 at $t = 0$, the velocity of all parts of the string is zero and there is a simple pulse of displacement in the string. Other starting configurations are possible. You could have the situation in which the string has no displacement and the sting has a distribution of transverse velocity. The difference in the operation of a harpsichord and piano is the the strings are plucked or distorted in the harpsichord and hammered in a piano. You can also have situations with both an initial displacement and velocity.

This is a difficult situation to describe. If you attribute all reality to the hunk of string the only motion is up and down in the transverse direction. Yet the configuration of the string moves along the string. In Section **??**, we will find that there is energy and momentum associated with the configuration of the string and that this moves with the configuration along the string. Thus we have the problem of the 'string' only moving up and down but energy and momentum are flowing along the string.

The converse of the above result that the parameters of the system are not sufficient to determine a size or time scale is that the medium, in the case of the stretched string are ρ and T_e, implies that the disturbances in the string travel with a speed set by the medium, $\sqrt{\frac{T_e}{\rho}}$ and that this speed is independent of the form of the disturbance. In other words disturbances travel with speed $\pm\sqrt{\frac{T_e}{\rho}}$ without distortion. For this reason, systems with this field dynamic are called wavelike. This is the definition of a wavelike medium. Although many systems are wavelike such as sound and light, other field systems may not be. For instance the dynamic for temperature flow in one spatial dimension is

$$\frac{\partial T_{emp}}{\partial t}(x,t) = a^2 \frac{\partial^2 T_{emp}}{\partial x^2}(x,t). \tag{2.5}$$

where a^2 is called the diffusion constant and is the ratio of the heat conductivity to the heat capacity of the material. Notice that $a^2 \stackrel{\text{dim}}{=} \frac{L^2}{T}$ and thus there is no special speed or length or time that is characteristic of the field.

The dynamic of the string requires that all points on the string be at the average of its neighbors. An easy way to compute the average is to pick two

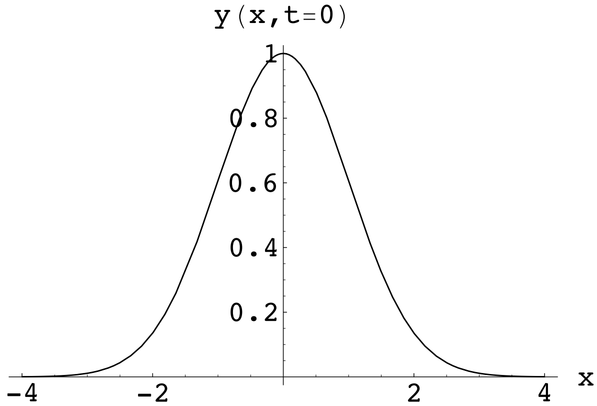

Figure 2.2: **A Simple Displacement Pulse in a String** A simple pulse in a stretched string under tension. At $t = 0$, the string is distorted but no part of the string moving.

neighbors, points on the string close to the point of interest and equidistant from it, and connect the points by a straight line. At the point of interest, x, the point on the line is the average of the two neighbors. Thus from Figure 2.3, we see that the center of the string is pulled strongly down and the edges are pulled up. The points of steepest drop are not pulled at all. This last point is interesting to note. The string is not pulled to the neutral position. Each segment is pulled only by its neighbors. If the string where pulled to the neutral position there would be a force for the entire time of descent and then the string would still have a velocity when it reached the neutral position and thus would overshoot and there would be oscillation at each disturbed point on the string. As we know, the disturbance in the stretched string is removed by the dynamic with the string returning gently to its neutral position.

To make this discussion more quantitative, we look at what goes on in a few small time increments. In a small time, Δt, since the velocity field is initially zero everywhere, we find that the string has not moved.

$$\begin{aligned} y(x, \Delta t) &= v(x,0)\Delta t + y(x,0) \\ &= y(x,0), \end{aligned} \tag{2.6}$$

where $v(x,t)$ is the velocity of the string at the point labeled x at time t. At $t = 0$, the string is not moving and $y(x,0)$ is known.

We will need the velocity of the string at all times and, even in a small

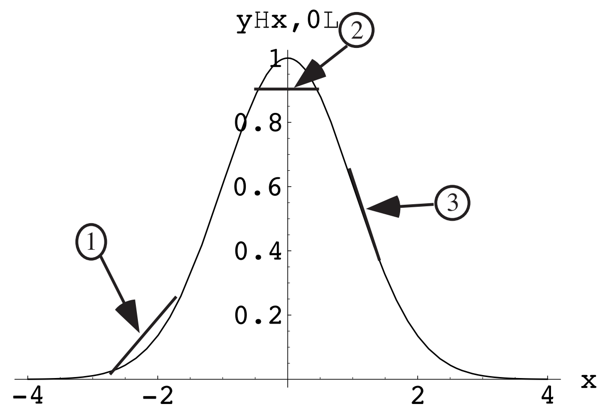

Figure 2.3: **Forces on a Pulsed String** The dynamic of the stretched string require that all points in the string be at the average of its neighbors. A simple rule for finding the force and thus the acceleration of a place on the string is to connect the neighbors with a straight line. If the string at that place is above the line, there will be a downward acceleration with magnitude proportional to the distance above. There are three examples shown. At a point on the edge of the pulse, 1, the string is accelerating upward. At the center, 2, the string is accelerating down. At a point at the midpoint of the side of the pulse, 3, the string has no acceleration.

time, because of the forces from Figure 2.3, the velocity changes.

$$\begin{aligned} v(x, \Delta t) &= a_{t=0}(x)\Delta t + v(x, 0) \\ &= a_{t=0}(x)\Delta t \end{aligned} \tag{2.7}$$

where we find $a_{t=0}(x)$ from an analysis such as that shown in Figures 2.3 for each point on the string. Thus we see that after a time Δt the velocities will have the same pattern as a function of position as the initial accelerations.

Repeating the process for a second Δt using Equations 2.6 and 2.7 but with the time shifted another increment,

$$\begin{aligned} v(x, 2\Delta t) &= a_{t=\Delta t}(x)\Delta t + v(x, \Delta t) \\ &= a_{t=0}(x)\Delta t + a_{t=0}(x)\Delta t \\ &= 2a_{t=0}(x)\Delta t \end{aligned} \tag{2.8}$$

where in the second line, I used the fact that since $y(x, \Delta t) = y(x, 0)$ and the accelerations depend only on $y(x, t)$, then $a_{t=\Delta t}(x) = a_{t=0}(x)$.

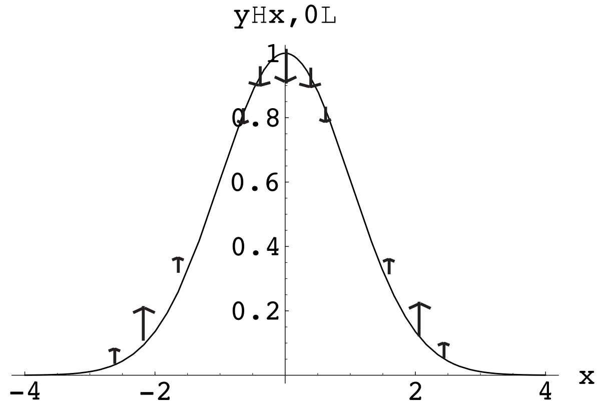

Figure 2.4: **Accelerations on a Pulsed String** Using a technique such as shown in Figure 2.4 for the forces on the string, the algorithm in Equation 2.1 can be applied at each point, x, and find the accelerations shown as arrows above.

The second dynamic is handled similarly,

$$\begin{aligned} y(x, 2\Delta t) &= v(x, \Delta t)\Delta t + y(x, \Delta t) \\ &= a_{t=0}(x)\Delta t^2 + y(x, 0). \end{aligned} \tag{2.9}$$

We now begin to see the string moving.

We can intuit that the pattern shown in Figure 2.5 develops. The region where there is a strong bend at the edge is is pulled up and so has an upward velocity and begins to lift. The middle section is unchanged at first. The center is forced down and after a time step has a downward velocity. Because of the pattern of the upward velocity at the bends and the downward velocity at the center, the two separating pulses appear to be moving along the string away from each other. We have to remember that the all the motion of the string is transverse to its direction.

The general pattern then develops of two distinct transverse displacement pulses of half the original amplitude one moving to the left and one to the right, see figure 2.6. This transverse velocity is patterned so that the two emergent pulses are one moving to the left with speed $-\sqrt{\frac{T_e}{\rho}}$ and one moving to the right with speed $\sqrt{\frac{T_e}{\rho}}$. It is important to realize that these traveling transverse displacement pulses are very different than just taking half the initial transverse displacement pulse and moving one to increasing x and one to decreasing x. These traveling waves, one to the left and one

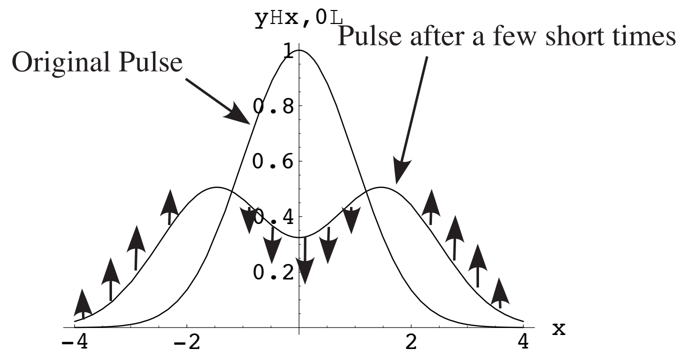

Figure 2.5: **Pulsed String after a Few Short Times** Using appropriate versions of Equations 2.6 and 2.7 to evolve the system, we can see the development of two pulses. Also shown are the velocities by scaled arrows. Remember the parts of the string are only free to move up and down but the pattern of up and down motion conspires to produce the effect that the pulse at negative x is moving toward greater negative x and the pulse at positive x is moving toward greater positive x. The original pulse is shown for comparison.

to the right, each have an accompanying transverse velocity field. It is the pattern of traveling waves that there is both a transverse displacement field and an associated transverse velocity field with the velocity field rising in front of the motion of the traveler and falling behind the traveler. This is a typical pattern for wavelike media. There are two fields that support each other and form the traveling configuration. For sound it is the density of the air and the pressure of the air. For electromagnetic waves, it is the electric and magnetic force fields.

It is worthwhile to also note that our original configuration of the displacement pulse with no transverse velocity, Figure 2.2, can be considered as the sum of two travelers, one going to the left and one going to the right, each of half the amplitude of the original. The addition of the displacement field gives the correct shape for the pulse at the instant of complete overlap, the initial instant, the two transverse velocity fields that accompany the transverse displacement field add to zero, the initial velocity field is everywhere zero. The ability to treat the original distortion as a sum of two independent distortions is an example of superposition and is traced to the fact that these are linear differential systems.

In addition. the travelers have an interesting relationship between the

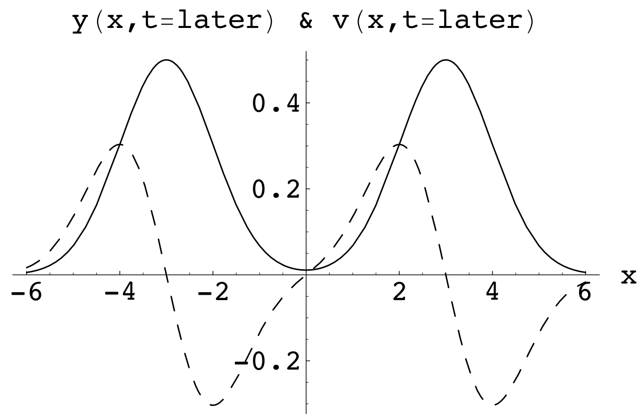

Figure 2.6: **Pulses in String Separating** After a time, the pulse initially placed on a stretched string, see Figure 2.2, separates into two half amplitude pulses. One travels to the left with velocity $v = -\sqrt{\frac{T_e}{\rho}}$ and one travels to the right with velocity $v = \sqrt{\frac{T_e}{\rho}}$. There is also a transverse velocity field that travels along with each pulse shown as the dashed curve instead of using arrows as in Figure 2.5.

displacement field and the velocity field. For a traveler that moves to increasing x, the argument of the displacement field is a single variable, $x - \sqrt{\frac{T_e}{\rho}}t$, instead of x and t as independent variables. This traveler is called a right traveler. For waves that move to decreasing x, called left travelers, the argument is $x + \sqrt{\frac{T_e}{\rho}}t$. This is what makes them travelers; they move to increasing x or decreasing x uniformly without the shape of the disturbance changing. This is a general result and true for all one dimensional wavelike systems. We worked this out for the particular disturbance of Figure 2.2, a simple pulse. It should be clear that this pattern of two separate travelers superposing to produce an initial distortion with no velocity field will hold for any form of distortion for the displacement field. Let's show this analytically. Consider a transverse displacement and transverse velocity configuration that is traveling, say to increasing x. Call the fields $y_{rt}(x,t)$ and $v_{rt}(x,t)$. By definition, since this is a traveler, although the fields are labeled by two variables, x and t, they are really only a function of one variable $x - \sqrt{\frac{T_e}{\rho}}t$. Otherwise they would not travel undistorted. In other words,

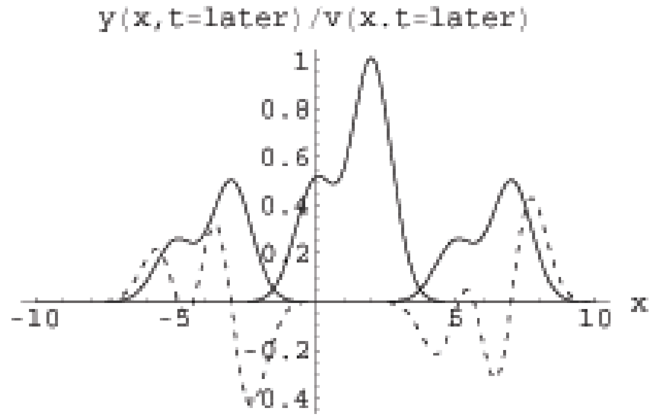

Figure 2.7: **Arbitrary Traveling Waves** Using a more general form for the initial distortion of the string, shown at the center for reference, we see at a later time the two traveling distortions, one moving to increasing x called the right traveler and one moving to decreasing x called the left traveler. The associated velocity profile for each is shown dotted. Because of the special form of the argument of the travelers, the velocity profile for the right traveler is proportional to the negative of the slope of the displacement profile of the right traveler at that instant and the velocity profile of the left traveler is proportional to the slope of the displacement profile of the left traveler at that instant.

$y_{rt}(x,t) = y_{rt}\left(x - \sqrt{\frac{T_e}{\rho}}t, 0\right)$ with a similar construction for the left travelers. Note that $y(x,t=0) = y_{rt}\left(x - \sqrt{\frac{T_e}{\rho}}t, 0\right)|_{t=0} + y_{lt}\left(x + \sqrt{\frac{T_e}{\rho}}t, 0\right)|_{t=0}$ as is required. Because of the first part of Equation 2.59,

$$\frac{\partial y_{rt}}{\partial t}(x,t) \equiv v_{rt}(x,t) = -\sqrt{\frac{T_e}{\rho}}\frac{\partial y_{rt}}{\partial x}(x,t), \tag{2.10}$$

and

$$\frac{\partial y_{lt}}{\partial t}(x,t) \equiv v_{lt}(x,t) = \sqrt{\frac{T_e}{\rho}}\frac{\partial y_{rt}}{\partial x}(x,t), \tag{2.11}$$

In other words, for travelers the transverse velocity field and the transverse displacement field are in a fixed relationship to each other; given the displacement field the velocity field is determined.

In summary, all initial transverse displacement fields constitute traveling systems. Figure 2.7, shows a more general initial configuration and the subsequent travelers. Because to the nature of the relationship between the x and t variables in the travelers, $x - \sqrt{\frac{T_e}{\rho}}t$ for the right traveler and $x + \sqrt{\frac{T_e}{\rho}}t$ for the left traveler, the velocity field in this dynamic is related to the slope of the displacement of the transverse displacement traveler at that point and the general solution for this case is

$$\begin{aligned} y(x,t) &= \frac{1}{2}\left\{ y\left(x + \sqrt{\frac{T_e}{\rho}}t, 0\right) + y\left(x - \sqrt{\frac{T_e}{\rho}}t, 0\right)\right\} \\ v(x,t) &= \frac{1}{2}\left\{ \sqrt{\frac{T_e}{\rho}}\frac{\partial y}{\partial x}\left(x + \sqrt{\frac{T_e}{\rho}}t, 0\right) - \sqrt{\frac{T_e}{\rho}}\frac{\partial y}{\partial x}\left(x - \sqrt{\frac{T_e}{\rho}}t, 0\right)\right\}. \end{aligned} \tag{2.12}$$

The problem with this solution is that it does not include all possible starting configurations. The initial configuration could be an arbitrary configuration of both transverse displacement and transverse velocity fields. We can develop the general solution by using superposition. Treat the initial displacement field and the initial velocity field as if the other field were initially zero and in the end add the resulting fields.

Considering the case of only an initial transverse velocity field, from Equation 2.59, it should be obvious that the velocity field satisfies a wave equation,

$$\rho\frac{\partial^2 v}{\partial t^2}(x,t) = T_e\frac{\partial^2 v}{\partial x^2}(x,t), \tag{2.13}$$

with right and left travelers, and, thus, we can construct the solution for this case for the velocity field in exactly the same way that we constructed the displacement field above:

$$v(x,t) = \frac{1}{2}\left\{ v\left(x + \sqrt{\frac{T_e}{\rho}}t, 0\right) + v\left(x - \sqrt{\frac{T_e}{\rho}}t, 0\right)\right\}. \tag{2.14}$$

From $\frac{\partial y}{\partial t} = v(x,t)$, it follows by direct differentiation that for each of the travelers

$$y_{rt}(x,t) = \frac{1}{\sqrt{\frac{T_e}{\rho}}}\int_{x_0}^{x-\sqrt{\frac{T_e}{\rho}}t} v(x',0)dx' \tag{2.15}$$

where x_0 is anything independent of t such as an arbitrary position. There is obviously a similar construction for y_{lt} and, since these two travelers constitute the configuration of the string, a general solution for the transverse displacement field is

$$
\begin{aligned}
y(x,t) &= \frac{1}{\sqrt{\frac{T_e}{\rho}}}\int_{x_0}^{x+\sqrt{\frac{T_e}{\rho}}t} v(x',0)dx' - \frac{1}{\sqrt{\frac{T_e}{\rho}}}\int_{x_0}^{x-\sqrt{\frac{T_e}{\rho}}t} v(x',0)dx' \\
&= \frac{1}{\sqrt{\frac{T_e}{\rho}}}\int_{x-\sqrt{\frac{T_e}{\rho}}t}^{x+\sqrt{\frac{T_e}{\rho}}t} v(x',0)dx' \qquad (2.16)
\end{aligned}
$$

where the constant at the lower limit of integration for the left and right travelers is the same, so that at $t = 0$, the transverse displacement wave is zero.

Combining our results for the cases of initial displacement field and initial velocity field, the general solution to the homogeneous differential equations are

$$
\begin{aligned}
y(x,t) &= \frac{1}{2}\left(y(x+\sqrt{\frac{T_e}{\rho}}t,0) + y(x-\sqrt{\frac{T_e}{\rho}}t,0)\right) \\
&\quad + \frac{1}{2\sqrt{\frac{T_e}{\rho}}}\int_{x-\sqrt{\frac{T_e}{\rho}}t}^{x+\sqrt{\frac{T_e}{\rho}}t} v(x',0)dx' \\
v(x,t) &= \frac{1}{2}\left(\sqrt{\frac{T_e}{\rho}}\frac{\partial y}{\partial x}(x+\sqrt{\frac{T_e}{\rho}}t,0) - \sqrt{\frac{T_e}{\rho}}\frac{\partial y}{\partial x}(x-\sqrt{\frac{T_e}{\rho}}t,0)\right) \\
&\quad + \frac{1}{2}\left(v(x+\sqrt{\frac{T_e}{\rho}}t,0) + v(x-\sqrt{\frac{T_e}{\rho}}t,0)\right). \qquad (2.17)
\end{aligned}
$$

This is the D'Alembert solution for the configuration for the unforced string given the most general initial field configurations at a given time.

Is this the only case of interest? Instead of determining the two fields at all times from the fields at sometime could we determine the fields at all times from the displacement field at two independent times of from the displacement field and the velocity field at some other time? In the case of Equation 2.17, we have a case of what is called the well posed problem, a differential system and other information to construct a solution. In these other examples, the problem is generally not well posed. Only special cases, admit solutions. Up until now we have used physical insight to construct the solution. It is worthwhile to take a more analytic approach.

2.2.2 Digression on Hyperbolic Equations

In order to better understand the general properties of these solution, we follow the analysis of Somerfeld, [Somerfelf 1949], First return to the single field description of our system, Equation 2.3. This is a second order homogenous linear partial differential equation. For the purposes of our analysis, we will broaden our system and also change notation to allow for the more general analysis. We will also restrict our analysis to two dimensional systems but the ideas that emerge will be generalizable to higher dimensional cases.

Consider the field $u(x, y)$ and the differential system

$$L(u) \equiv A\frac{\partial^2 u}{\partial x^2} + 2B\frac{\partial^2 u}{\partial x \partial y} + C\frac{\partial^2 u}{\partial x^2} + D\frac{\partial u}{\partial x} + E\frac{\partial u}{\partial y} + Fu = 0, \tag{2.18}$$

where $A, B, ..., F$ are possibly functions of u, x, and y. Isolating the leading differential terms,

$$A\frac{\partial^2 u}{\partial x^2} + 2B\frac{\partial^2 u}{\partial x \partial y} + C\frac{\partial^2 u}{\partial x^2} = \Phi\left(u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, x, y\right). \tag{2.19}$$

Φ can be any non-linear non-singular combination of the variables.

Let Γ be a curve in the xy plane. In higher dimensions, $d > 2$, an $d-1$ dimensional surface with both u and $\frac{\partial u}{\partial \hat{n}}$ given where $\hat{n}$ indicates the normal to the curve or surface. Since $u(x, y)$ is given on Γ, $\frac{\partial u}{\partial \hat{s}}$ where $\hat{s}$ is along Γ is known. Since $\frac{\partial u}{\partial \hat{n}}$ and $\frac{\partial u}{\partial \hat{s}}$ are known on Γ, $\frac{\partial u}{\partial x}$ and $\frac{\partial u}{\partial y}$ can also be found on Γ. Define $p \equiv \frac{\partial u}{\partial x}$, $q \equiv \frac{\partial u}{\partial y}$, $r \equiv \frac{\partial^2 u}{\partial x^2}$, $s \equiv \frac{\partial^2 u}{\partial x \partial y}$, and $t \equiv \frac{\partial^2 u}{\partial y^2}$ on Γ. Thus,

$$Ar + 2Bs + Ct = \Phi \tag{2.20}$$

on Γ. In addition,

$$\begin{aligned} d\left(\frac{\partial u}{\partial x}\right) &\equiv dp \\ &= \frac{\partial^2 u}{\partial x^2}dx + \frac{\partial^2 u}{\partial y \partial x}dy \\ &= rdx + sdy \end{aligned} \tag{2.21}$$

and, similarly,

$$dq = sdx + tdy \tag{2.22}$$

again all on Γ.

The set of Equations, 2.20, 2.21, and 2.22, can be solved for r, s, and t if the determinant

$$\Delta \equiv \begin{vmatrix} A & 2B & C \\ dx & dy & 0 \\ 0 & dx & dy \end{vmatrix} = Ady^2 - 2Bdxdy + Cdx^2 \neq 0 \tag{2.23}$$

for points on and along Γ. It is important to realize that the condition $\Delta = 0$ is a condition on directions at a point (x, y),

$$A\left(\frac{dy}{dx}\right)^2 - 2B\frac{dy}{dx} + C = 0. \tag{2.24}$$

At every point, there could be up to two directions for which $\Delta = 0$.

If Γ does not lie along a direction in which $\Delta = 0$, we solve for r, s, and t. Then differentiating Equation 2.20 yields the higher derivatives of $u(x, y)$ again on Γ. For example,

$$\begin{aligned} r_x &\equiv \frac{\partial^3 u}{\partial x^3} \\ s_x &\equiv \frac{\partial^3 u}{\partial y \partial^2 x} = r_y \\ t_x &\equiv \frac{\partial^3 u}{\partial y^2 \partial x} = s_y \\ t_y &\equiv \frac{\partial^3 u}{\partial y^3} \end{aligned} \tag{2.25}$$

can be found by differentiating Equation 2.20,

$$Ar_x + 2Bs_x + Ct_x = \Phi_x + ... \tag{2.26}$$

In addition,

$$\begin{aligned} dr &= r_x dx + s_x dy \\ ds &= s_x dx + t_x dy. \end{aligned} \tag{2.27}$$

The determinant is again Δ which is not zero on and along Γ and thus Equations 2.26 and 2.27 can be solved for r_x, s_x, and t_x. All the other higher derivatives can be calculated similarly. Thus, if Δ is non-zero along Γ, all the derivatives of $u(x, y)$ are known and a complete Taylor expansion of $u(x, y)$ is possible and, thus, $u(x, y)$ is determined in a neighborhood of (x, y) for points on Γ.

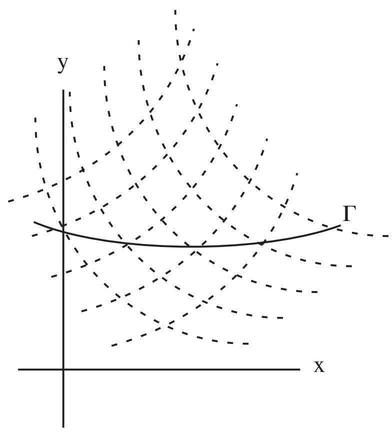

Figure 2.8: **Boundary Information and Characteristics for Hyperbolic Differential System** The relationship of the boundary information curve, Γ, on which $u(x, y)$ and $\frac{\partial u}{\partial \hat{n}}$ are given and the characteristic curves for the case of a hyperbolic differential system which is properly posed. The characteristic curves are shown dashed and Γ is solid.

But, if for any point (x, y), Γ lies along a direction in which $\Delta = 0$, there is no solution for r, s, and t and $u(x, y)$ cannot be constructed. Thus the families of second order partial differential equations are divided into three classes determined by the nature of the coefficients of the highest order derivatives. The directions in which $\Delta = 0$ are called characteristic directions. Since the condition that $\Delta = 0$ is a quadratic in $\frac{dy}{dx}$, how many real solution directions there are is determined by the discriminant of the quadratic. The cases are

$$AC - B^2 < 0 \tag{2.28}$$

for which there are two real directions for any point. In this case, the differential system is said to be hyperbolic. A prominent example of this class is our equation for the stretched string, Equation 2.3. The case

$$AC - B^2 = 0 \tag{2.29}$$

has one real characteristic direction and the case

$$AC - B^2 > 0 \tag{2.30}$$

has no real characteristic directions.

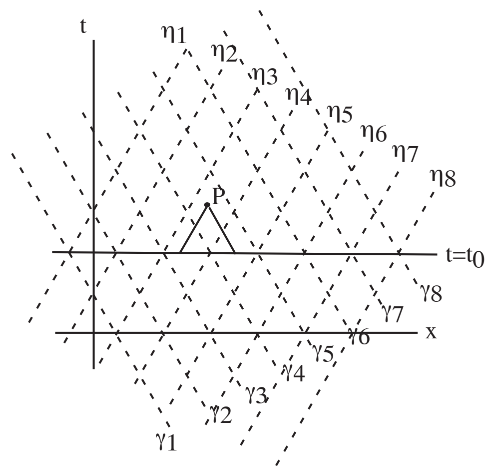

Figure 2.9: **Boundary Information and Characteristics for String Field Theory** The boundary information curve, $t = t_0$, on which $y(x,t)$ and $v(x,t)$ are given and the characteristic curves for the case of the stretched string. The characteristic curves are shown dashed and are labeled by the constant values of η and γ. Given the results of Equation 2.35, the value at the event labeled P, (x,t), is constructed by the superposition of the boundary data propagated over the characteristics, see Figure 2.10 for details. The curves are shown for the case that $\sqrt{\frac{\rho}{T_e}} = \pm\frac{1}{2}$.

The case of the string field theory is hyperbolic and has two characteristic directions. There are several important features of the characteristic curves, surfaces in higher dimensions, which can be illustrated with this simple example, Equation 2.3. Firstly $\Delta = 0$ yields

$$\left(\frac{dt}{dx}\right)^2 - \frac{\rho}{T_e} = 0 \tag{2.31}$$

for the characteristic curves. Integration yields two curves

$$\sqrt{\frac{T_e}{\rho}}\,(t - t_0) - (x - x_0) = 0$$

or

$$\sqrt{\frac{T_e}{\rho}} t - x = \sqrt{\frac{T_e}{\rho}} t_0 - x_0 \equiv \eta \tag{2.32}$$

and

$$\begin{aligned} \sqrt{\frac{T_e}{\rho}}\,(t-t_0)+(x-x_0) &= 0 \\ \sqrt{\frac{T_e}{\rho}}t+x=\sqrt{\frac{T_e}{\rho}}t_0+x_0 &\equiv \gamma \end{aligned} \tag{2.33}$$

where η and γ are the constants of integration and (x_0, t_0) is an arbitrary point in the xy plane. The characteristic curves and our Γ, $t = t_0$, are shown in Figure 2.9

It is clear that our initial data curve, $t = t_0$, is nowhere tangent to a characteristic and thus provide a properly posed boundary value for our partial differential equation.

Another important feature of the characteristic curves is that they provide a useful coordinate grid. In general, for the hyperbolic case, there will always be two directions at each point. Curves generated by integration along these directions, $\phi(x, y)$ and $\psi(x, y)$, can be chosen as coordinates. For our case of the string field theory, choosing as coordinates the two constants of integration, (η, γ), from Equations 2.32 and 2.33 instead of our (x, t), the partial differential equation, Equation 2.3, becomes

$$\frac{\partial^2 y}{\partial\eta\partial\gamma}=0. \tag{2.34}$$

The obvious solution to this equation is

$$y(\eta,\gamma)=F_\gamma(\eta)+F_\eta(\gamma), \tag{2.35}$$

where $F_\gamma(\eta)$ and $F_\eta(\gamma)$ are functions of only one variable and are labeled by the characteristic on which they reside.

Using Figure 2.10, we can construct the string fields at a point (x_1, t_1). The two characteristics that pass through the point are $\eta_1 = ct_1 - x_1$ and $\gamma_1 = ct_1 + x_1$ where I have used $c \equiv \sqrt{\frac{T_e}{\rho}}$ and the choice of c for the label of this factor is fore-shadowing our interest in electromagnetism. Our solution for the string field at (x_1, t_1) is

$$\begin{aligned} y(x_1,t_1) &= F_{\gamma_1}(\eta_1)+F_{\eta_1}(\gamma_1) \\ v(x_1,t_1) &= cF'_{\gamma_1}(\eta_1)-cF'_{\eta_1}(\gamma_1), \end{aligned} \tag{2.36}$$

and, thus, we have to construct the functions $F_{\gamma_1}(\eta)$ and $F_{\eta_1}(\gamma)$. The only boundary information that is relevant is the information at the events common to the boundary, $t = t_0$, and the two characteristics, $ct = \gamma_1 - x$ and

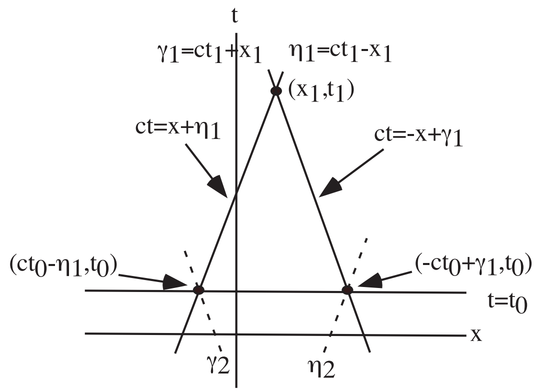

Figure 2.10: **Construction of String Field Using Characteristics** The construction of the string fields at a point (x_1, t_1) from boundary information curve, $t = t_0$, on which $y(x,t)$ and $v(x,t)$ using the characteristic curves is accomplished by extending the two characteristics to the boundary curve and evaluating the relevant single variable function at the boundary. It is then extended back to the event at which the field is being constructed. For simplicity of notation, $c \equiv \sqrt{\frac{T_e}{\rho}}$. The curves are shown for the case that $c = \pm\frac{1}{3}$.

$ct = x + \eta_1$. These are $(\gamma_1 - ct_0, t_0) = (ct_1 + x_1 - ct_0, t_0)$ and $(ct_0 - \eta_1, t_0) = (ct_0 + x_1 - ct_1, t_0)$. The first of these is $(\gamma_1 - ct_0, t_0) \equiv (x_2, t_2) \Rightarrow \eta_2 = ct_2 - x_2 = 2ct_0 - \gamma_1$. Here the field is given as

$$\begin{aligned} y(x_1 + c(t_1 - t_0), t_0) &= F_{\gamma_1}(\eta_2) + F_{\eta_2}(\gamma_1) \\ v(x_1 + c(t_1 - t_0), t_0) &= cF'_{\gamma_1}(\eta_2) - cF'_{\eta_2}(\gamma_1), \end{aligned} \tag{2.37}$$

This is solved for F_{γ_1} as

$$F_{\gamma_1}(\eta_2) = \frac{1}{2}\left\{ y(x_1 + c(t_1 - t_0), t_0) + \frac{1}{c}\int_0^{x_1 + c(t_1 - t_0)} v(x', t_0)dx' \right\},$$

or using the fact that $\eta_2 = 2ct_0 - \gamma_1 \Rightarrow \gamma_1 = 2ct_0 - \eta_2$,

$$F_{\gamma_1}(\eta_2) = \frac{1}{2}\left\{ y(ct_0 - \eta_2, t_0) + \frac{1}{c}\int_0^{ct_0 - \eta_2} v(x', t_0)dx' \right\}.$$

Now that we have isolated the η dependence, replace η_2 with η_1 and then

η_1 by $x_1 + ct_1$. The desired result is,

$$F_{\gamma_1}(\eta_1) = \frac{1}{2}\left\{ y(x_1 - c(t_1 - t_0), t_0) + \frac{1}{c}\int_0^{x_1 - c(t_1 - t_0)} v(x', t_0)dx' \right\}. \quad (2.38)$$

A similar construction for $F_{\eta_1}(\gamma_1)$ yields

$$F_{\eta_1}(\gamma_1) = \frac{1}{2}\left\{ y(x_1 + c(t_1 - t_0), t_0) - \frac{1}{c}\int_0^{x_1 + c(t_1 - t_0)} v(x', t_0)dx' \right\}, \quad (2.39)$$

Combining Equations 2.38, and 2.39, we obtain our most general form of the D'Alembert solutio,

$$\begin{aligned} y(x,t) &= \frac{1}{2}\left(y(x + c(t - t_0), t_0) + y(x - c(t - t_0), t_0)\right) \\ &\quad + \frac{1}{2c}\int_{x - c(t - t_0)}^{x + c(t - t_0)} v(x', t_0)dx' \\ v(x,t) &= \frac{1}{2}\left(v(x + c(t - t_0), t_0) + v(x - c(t - t_0), t_0)\right) \\ &\quad + \frac{c}{2}\left(\frac{\partial y}{\partial x}(x + c(t - t_0), t_0) - \frac{\partial y}{\partial x}(x - c(t - t_0), t_0)\right). \quad (2.40) \end{aligned}$$

Before leaving this section for completeness, we should briefly discuss the other situations that exist for second order partial differential equations. The case in which the discriminant of the quadratic form is zero, Equation 2.29, the parabolic case, there is only one real characteristic direction. In this case, the one characteristic, $\gamma = \phi(x, y) = \psi(x, y)$ and the coordinate $\eta = x$, can be used as coordinates. With these coordinates, the leading terms in the partial differential equation, Equation 2.19 can be written

$$\frac{\partial^2 u}{\partial \eta^2} = X\left(u, \frac{\partial u}{\partial \gamma}, \frac{\partial u}{\partial \eta}, \gamma, \eta\right). \quad (2.41)$$

The case in which the discriminant of the quadratic form is greater than zero, Equation 2.30, the elliptic case, there are no real characteristic directions. At each point, there are instead two complex conjugate directions $\phi(x, y)$ and $\psi(x, y) = \phi^*(x, y)$. In a coordinate system based on coordinates $\gamma = \frac{\phi(x,y)+\psi(x,y)}{2}$ and $\eta = \frac{\phi(x,y)-\psi(x,y)}{2}$, the leading terms in the elliptic partial differential equation, Equation 2.19 can be written

$$\frac{\partial^2 u}{\partial \gamma^2} + \frac{\partial^2 u}{\partial \eta^2} = X\left(u, \frac{\partial u}{\partial \gamma}, \frac{\partial u}{\partial \eta}, \gamma, \eta\right). \quad (2.42)$$

With no real characteristic directions the techniques of using the characteristics for solutions is not available for the elliptic case and other techniques will have to be developed. Elliptic partial differential equations will be important to us in our discussion of potential theory, Section C.2, and the techniques of solution will be discussed there.

2.3 Variations on the Simple String

In this section, we will add important and interesting features to our simple string. These will introduce new field theory concepts and constructions. Some of these variations are motivated by the physically realized object, the string, and the ways that a string can be manipulated. Thus the string can be forced. There can be a mass on the string. The string is suspended between boundaries. The string is free to have its transverse displacement and velocity in a two dimensional space. This leads to a discussion of the symmetries of the string including special relativistic effects. Other extensions of the simple string will be to add features that help elucidate the nature of a simple field theory. The most important of these is to analyze the string in the context of an action based mechanics. This is a necessary first step in the extension to concepts related to a true field theory such energy, momentum, and angular momentum density. These can be used to develop a complete classical field theory of a "string" in interaction with a local source. These will be treated in the next section, Section 3.

2.3.1 Forced Strings

The simplest element to add is an external force on the string. If at each point on the string, there is an external force per unit length, $f(x,t)$, acting, it will add a term to Equation 2.1 and subsequently to Equation 2.3. For our first order in time coupled field equations, Equation 2.59[1], the final form is

$$\frac{\partial y}{\partial t}(x,t) = v(x,t)$$

[1]Although this added term can clearly be identified as a force, for the more general case of a field theory, terms independent of the field on the left side of Equations 2.1 are called sources. Therefore, there could be a source term on the left side of the first of Equations 2.43, say $g(x,t)$. In that case, $g(x,t)$would be a velocity source or, when multiplied by ρ an impulse source. The subsequent analysis of this section is changed by this additional source only by substitution of $f(x,t)$ with $f(x,t)+\rho\frac{\partial g}{\partial t}(x,t)$

$$\rho \frac{\partial v}{\partial t}(x,t) \quad = \quad T_e \frac{\partial^2 y}{\partial x^2}(x,t) + f(x,t)\,. \tag{2.43}$$

It should be clear that the addition of this force term does not change the nature of the partial differential equation; it does not change the leading differential terms. Therefore, this equation is still hyperbolic and has the same characteristics given in Equations 2.32 and 2.33. In fact, Equation 2.43 using the standard form is

$$\frac{\partial^2 y}{\partial \eta \partial \gamma} = \frac{f(\eta,\gamma)}{T_e}. \tag{2.44}$$

Before solving this problem generally, let's consider a simplified special case. We require that the force is local, $f(x,t) = f_{x_0}(t)\delta(x - x_0)$, and not effective before some time t_0, $f_{x_0}(t) = 0$ for $t < t_0$ where t_0 is some finite time. It is also worthwhile to keep track of the dimension of the force terms. $f(x,t)$ in Equation 2.43 is a force per unit length. $f_{x_0}(t)$ is then a force.

We also require that the initial state of the string be quiescent, $y(x,t) = v(x,t) = 0$ for all x and times before the force operates. These boundary conditions are not the only ones available but represent the most reasonable circumstances. There are important implications of this condition. The travelers that develop can only be manifest in the future. This assumption destroys the time reversal symmetry of the system. This restriction on the boundary conditions is called the "Sommerfeld Radiation Condition", [Somerfelf 1949] and it plays a significant role in all classical applications of hyperbolic field theories[2], see Appendix C.

The solution for this force is based on our use of the characteristics and the realization that Equation 2.44 with the localized force is zero almost everywhere and thus follows the form of Equation 2.35 almost everywhere. As a result of the force at $x = x_0$, there is a right moving traveler $x > x_0$ and $t > t_0$ for each time after t_0 on the characteristic $\eta_0 = ct' - x_0 = ct - x$ where $t' > t_0$ is the time of action of the force at x_0 for a given field point, (x,t), and a traveler to decreasing x for $x < x_0$ and $t > t_0$ on the characteristic $\gamma = ct' + x_0 = ct + x$, see Figure 2.11. For this force, integrating the interval $x = x_0 - \epsilon$ to $x = x_0 + \epsilon$ for small ϵ in the second of the Equations 2.35 requires

$$0 = T_e \left\{ \frac{\partial y}{\partial x}(x_0 + \epsilon, t) - \frac{\partial y}{\partial x}(x_0 - \epsilon, t) \right\} + f_{x_0}(t)\,. \tag{2.45}$$

[2] In quantum field theories an alternative condition is applied. This is the Feynman propagator and enables traveler solutions moving into the past. These are identified with the antiparticle contributions to the system.

Thus

$$\frac{\partial y}{\partial x}(x,t) = \begin{cases} -\frac{1}{2T_e} f_{x0}(t - \frac{x-x_0}{c}) & : \quad x > x_0 \quad t > t_0 \\ \frac{1}{2T_e} f_{x0}(t + \frac{x-x_0}{c}) & : \quad x < x_0 \quad t > t_0 \\ 0 & : \quad \text{otherwise} \end{cases} . \tag{2.46}$$

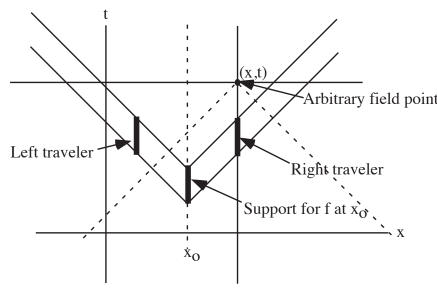

Figure 2.11: **String Fields Causal Diagram** The event plane for the string field with a spatially localized finite time interval force on the string at x_0. There are right and left travelers from the finite time interval force at x_0 providing support inside the rising lines at $\pm\frac{\pi}{4}$ that are the characteristics. An arbitrary event in the plane is influenced only by the events in between and below the characteristic lines falling below the event. A spatial section through the event will have support only in the regions between the rising characteristics from the support at x_0.

For travelers the slope and the velocity field are proportional.

$$v(x,t) = \frac{\partial y}{\partial t}(x,t) = \begin{cases} \frac{c}{2T_e} f_{x0}(t - \frac{x-x_0}{c}) & : \quad x > x_0 \quad t > t_0 \\ \frac{c}{2T_e} f_{x0}(t + \frac{x-x_0}{c}) & : \quad x < x_0 \quad t > t_0 \\ 0 & : \quad \text{otherwise} \end{cases} , \tag{2.47}$$

and therefore

$$y(x,t) = \begin{cases} \frac{c}{2T_e} \int_0^{t-\frac{x-x_0}{c}} f_{x0}(t')dt' & : \quad x > x_0 \quad t > t_0 \\ \frac{c}{2T_e} \int_0^{t+\frac{x-x_0}{c}} f_{x0}(t')dt' & : \quad x < x_0 \quad t > t_0 \\ 0 & : \quad \text{otherwise} \end{cases} . \tag{2.48}$$

where again we have used $c = \sqrt{\frac{T_e}{\rho}}$ throughout.

An especially simple example of this is the case of a localized constant force that is turned on at some time t_0 and turned off at t_f. If the force has magnitude f_{x_0}, the fields are

$$y(x,t) = \frac{f_{x_0}}{2T_e} \begin{cases} 0 & : \quad x > x_0 + c(t-t_0) \\ ct - (x - x_0) & : x_0 + c(t-t_0) > x > x_0 + c(t-t_f) \\ c(t_f - t_0) & : x_0 + c(t-t_f) > x > x_0 - c(t-t_f) \\ ct + (x - x_0) & : x_0 - c(t-t_f) > x > x_0 - c(t-t_0) \\ 0 & : \quad x < x_0 - c(t-t_0) \end{cases}, \tag{2.49}$$

and

$$v(x,t) = \frac{cf_{x_0}}{2T_e} \begin{cases} 0 & : \quad x > x_0 + c(t-t_0) \\ 1 & : x_0 + c(t-t_0) > x > x_0 + c(t-t_f) \\ 0 & : x_0 + c(t-t_f) > x > x_0 - c(t-t_f) \\ 1 & : x_0 - c(t-t_f) > x > x_0 - c(t-t_0) \\ 0 & : \quad x < x_0 - c(t-t_0) \end{cases}. \tag{2.50}$$

Figure 2.12 on page 38 shows the string fields at some time after the force vanishes.

This result can be generalized in two ways. A more general force and more general boundary conditions. The addition of a general force, $f(x,t)$, with support in a finite range in space and time is accomplished with the superposition of a system of localized forces,

$$f(x,t) = \int_{-\infty}^{\infty} f_{x_0}(t)\delta(x - x_0)dx_0. \tag{2.51}$$

This case can be solved by integration Equations 2.49 and 2.50 over x_0. This is not very enlightening since these are messy integrals.

To understand the situation better, lets restrict the form of the force slightly further by using an instantaneous force at the point x_0 at the time t_0. In this case, we define our force as

$$f(x,t) = f_{x_0 t_0}\delta\left(x - x_0\right)\delta\left(t - t_0\right). \tag{2.52}$$

This is an especially important case. It is common practice to define a special function called the Green's function which is the response of the field to the unit instantaneous force at a point and at a given time. With the Green's function, we can generate a very general procedure for the solution of field problems. To get the Green's function for this case, replace $f_{x_0}\left(t'\right)$ with $f_{x_0\, t_0}\delta\left(t' - t_0\right)$ in Equations 2.48 and 2.47. The case with $f_{x_0 t_0} = 1$ generates a solution that is the Green's function. For a general coverage of the use

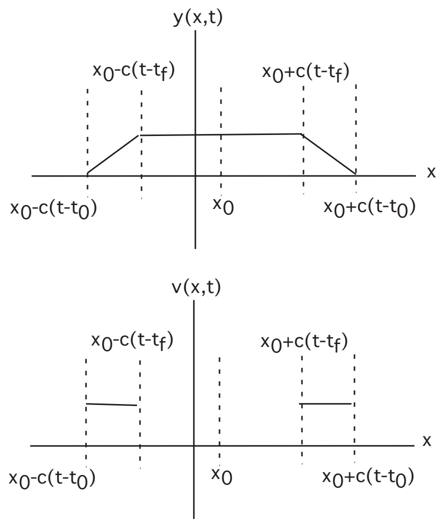

Figure 2.12: **String Fields for Local Step Force** The string displacement field, $y(x,t)$, and the velocity field, $v(x,t)$, at some time $t > t_f$ for a local string force at x_0. The force is zero for times before $t_0 > 0$, constant of magnitude f_{x_0} for times between t_0 and $t_f$and zero again for $t_f > t_0$. For simplicity of notation, $c \equiv \sqrt{\frac{T_e}{\rho}}$. The magnitude of the velocity field for those places at which it is non-zero is $\frac{cf_{x_0}}{2T_e}$. The height of the central plateau of the displacement field is $\frac{cf_{x_0}}{2T_e}\left(t_f - t_0\right)$.

of Green's functions see Appendix C. For a string in quiescent boundary conditions for times before t_0, i. e. Equation 2.48, with $f_{x_0}(t) = \delta(t - t_0)$ is

$$
\begin{aligned}
y(x,t) &= \int_{-\infty}^{\infty} G(x, x_0, t, t_0) f_{x_0 t_0}(t) \delta(t - t_0) \delta(x - x_0) dx_0 dt_0 \\
&= \qquad (2.53)
\end{aligned}
$$

In order, to solve this equation, we will use a method that will seem to be a diversion but has very general application.

2.3.2 The Bounded String

Actually, the string under tension is never infinite in extension. All real strings are confined to a finite region of space and the tension maintained at the boundaries. For the bounded string several new features appear. Simple boundaries are totally absorbing or perfectly reflecting. With reflecting boundaries new phenomena emerges – images and standing waves. Standing waves are repeating patterns in the string. More specifically, the disturbance in the string factors into a separate functions of position and time; each part of the string moves together. For the right conditions, there are distortions of the string that initiate a pattern of standing waves in which each part of the string oscillates with the same frequency. This can be made consistent with our previous analysis that stated that all disturbances in the string are traveling waves by taking advantage of reflections at the boundary.

The simplest type of boundary is the completely absorbing boundary. This case is so simple that it is often not discussed. Consider the case of a string bounded on the left with a perfectly absorbing boundary. Any disturbance that travels toward the boundary passes through the boundary point as if the string were infinite in extent and the traveling disturbance is then lost forever from the string. The analysis of the behavior of the disturbance is easily carried out if an identical string that overlapped the original string in the region in which it exists but the new string extended indefinitely to the left. This idea of extending the finite string through the boundary is the key to understanding the behavior of travelers in the finite string regardless of the nature of the boundary. Here the case is especially trivial. Any traveler moving toward the left boundary merely moves off the finite string. It is absorbed at the boundary. Later we will generalize this case to a case in which a portion of the traveler reflects and a portion is absorbed, but first we treat two important ideal reflecting boundaries.

The next example of a boundary is the case of a region of the string bounded on at least one end at a point at which there can be no transverse displacement and no transverse velocity, $y(0, t) = 0$ and $v(0, t) = 0$ where the end is located. For convenience, choose $x = 0$ as the boundary point. Terminations of this kind are called fixed ends. Consider the string bounded on the left. Again, the string is extended beyond the boundary but with the magic property that for any left moving waves in the original string there is, in addition, a right moving wave in the extension that is the inverse of the original distortion both its displacement wave and velocity wave through the fixed point, see Figure 2.13. The right mover is called the image wave of the original left moving wave. Note that this process must be carried out for all the left travelers that will ever arrive at the fixed end including left travelers that are images of the right travelers generated by a right boundaries of the string. This is clearly a reflecting boundary.

Another termination is a free end. This is a termination that has the

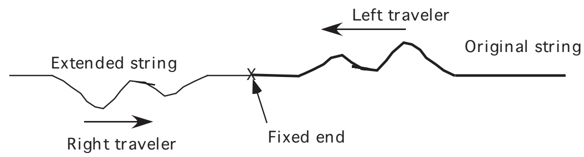

Figure 2.13: **Image Wave in a Fixed End** A fixed end at the left end of a bounded string can be modeled by extending the string through the fixed end and, for each left traveler that can come to the fixed end, inserting an image right traveler that is the inverse amplitude at the same distance from the fixed end of the original left traveler. The figure shows a typical transverse displacement wave. Not shown is the associated transverse velocity wave

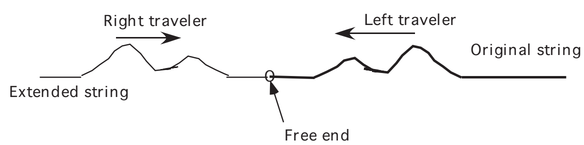

Figure 2.14: **Image Wave in a Free End** A free end at the left end of a bounded string can be modeled by extending the string through the free end and, for each left traveler that can come to the fixed end, inserting an image right traveler that is the same amplitude at the same distance from the fixed end of the original left traveler. The figure shows a typical transverse displacement wave. Not shown is the associated transverse velocity wave

condition that there is no restoring force at the termination. As shown in Section 2.3.1, the absence of force at the termination translates into the statement that the slope of the string at the termination is zero. Also since we are basing our analysis of the string phenomena on travelers and for all travelers the the slope of the transverse displacement and the velocity are proportional, the transverse velocity also vanishes at the termination. For this kind of termination it is simple to construct and image system. For a left traveler coming into a left free termination, instead of inverting image of the transverse displacement wave, the image wave is the same displacement and traveling to the right. Again since the velocity wave of any traveler is determined by the displacement wave, it will turn out that the velocity wave of the image traveler is inverted, see Figure 2.14.

For strings confined to a finite region by two ends, it is simply a matter of finding all the right moving images in the left extension including the

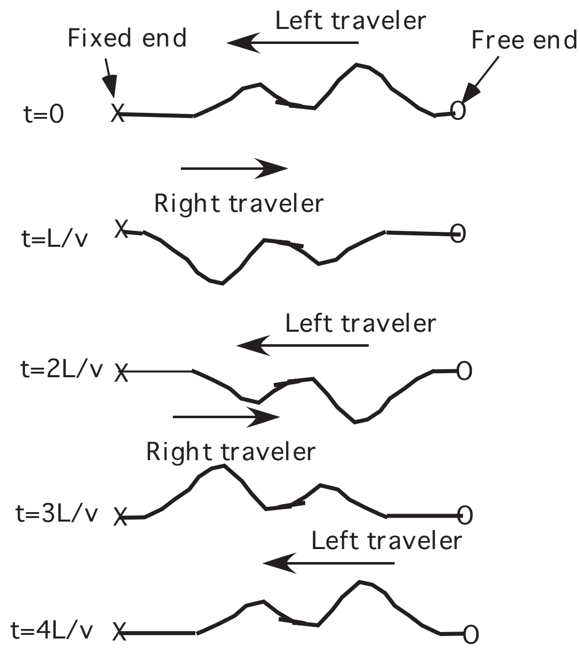

Figure 2.15: **Disturbance in Finite String with Free and Fixed End** A finite string of length L between a fixed and a free end has a left traveler. In time $t = \frac{L}{v}$, the string has an inverted right traveler. In $t = \frac{2L}{v}$, the string has an inverted left traveler. In $t = \frac{3L}{v}$, the string has a right traveler. In $t = \frac{4L}{v}$, the string has a left traveler, the original disturbance.

images from the right end images. Consider the case of the string of length L bounded on the left by a fixed end and on the right by a free end. For a string in which the travelers have speed $v = \sqrt{\frac{T_e}{\rho}}$ disturbances traverse the length of the string in a time $t_0 = \frac{L}{v}$. Thus as shown in Figure 2.15 for an initial situation with a left traveler after one $\frac{L}{v}$, the disturbance inside the string will be a right traveler which is the inverse of the original. In $2\frac{L}{v}$, there will be a left traveler in the string that is the inverse of the original traveler. In $3\frac{L}{v}$, there will be a right traveler in the string that is the image of the original traveler going the other way. In $4\frac{L}{v}$, the system returns to its original configuration with a left traveler in the string that is the same of the original traveler. Obviously, if we had started with a right traveler with minor modifications we can show that the same repeat pattern appears. Since travelers describe all possible distortions, the periodicity of the fixed and free end bounded finite string has period $4\frac{L}{v}$. A similar analysis for the finite string bounded by two fixed ends or two free ends will have a period of $2\frac{L}{v}$. In this sense, all distortions of a finite string with perfectly reflecting boundaries are standing waves.

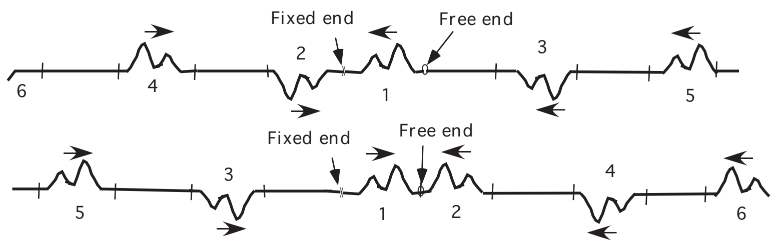

Figure 2.16: **Image Travelers in Finite String with Free and Fixed End** A finite string of length L between a fixed and a free end which initially has a displacement wave. This produces two travelers, one moving to the left, top, and one moving to the right, bottom. All the image waves produced by the boundaries are shown. Shapes 1 are the original traveler waves and their image waves are labeled 2. The images of these in the other boundary are 3. 4 is the image of 3 from the original boundary. This process proceeds through to wave 6. If all the left travelers are combined onto one string, there is a continuous series of waves with period 4L filling the region to the right of the fixed end and moving into the physical region. Similarly for the right travelers. Both cases can be extended into the entire infinite space with no effect in the physical region.

The periodic appearance of the images in the finite string with one fixed end and one free end separated by a distance L and the repeating pattern every $\frac{4L}{v}$ as seen in Figure 2.15 implies that we can model the behavior of the finite string with two infinite strings one of which at $t = 0$ has the region to the right of the fixed boundary with an infinite string of left traveler images with a wavelength of 4L and the region to the right of the free boundary with left traveler images with that same wavelength, see Figure 2.16. Each of these wave systems can be extended into the infinite string in the unphysical region to produce two infinite strings one with only right travelers and one with left travelers and both with wavelength 4L and each aligned with the appropriate boundary point, see Figure 2.17. Obviously, the left travelers to the left of the fixed end location never enter the real segment of the string between 0 and L. Similarly, the right travelers to the right of the free end never enter the physical part of the strings. With this construction, we have two systems of images on the infinite line with wavelengths of 4L which when superposed produce a $t = 0$ periodic displacement wave.

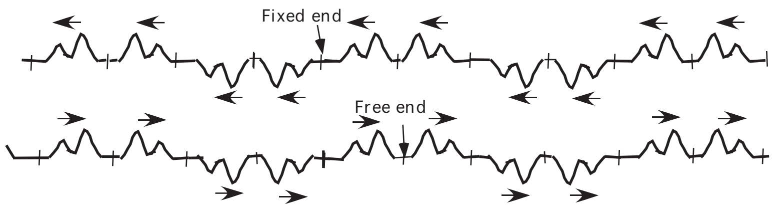

Figure 2.17: **Infinite Series of Image Travelers for a Finite String with Free and Fixed End** A finite string of length L between a fixed and a free end which initially has a displacement wave produces left and right travelersin two infinite systems of traveling waves. Left travelers at top and right travelers at bottom. At $t = 0$, a superposition of these two systems produces an infinite sequence of displacement waves of wavelength 4L.

The infinite system of alternating fixed and free end boundaries has a discrete translation symmetry of the form $y(x \pm 4nL, t) = y(x, t)$. Using the fixed end as the spatial origin, the fields also possess odd inversion symmetry, $y(-x, t) = -y(x, t)$. The irreducible representations of this symmetry are Although we have not analyzed the other perfectly reflecting boundaries, it is easy to do an analysis similar to the one above. For the case of a string bounded by two fixed ends or two free ends the periodicity is $2\frac{L}{v}$ and again Fourier series is appropriate.

An arbitrary displacement wave with null velocity field at $t = 0$, $y(x, 0)$, set in our bounded string of length L with one fixed end and one free end is represented by

$$y(x, 0) = \sum_{n=1}^{\infty} \left(b_n \sqrt{\frac{1}{2L}} \sin \frac{\pi n x}{2L} \right) \tag{2.54}$$

where

$$b_m = \sqrt{\frac{1}{2L}} \int_{-2L}^{2L} y(x', 0) \sin \frac{\pi m x'}{2L} dx', \tag{2.55}$$

see see Appendix B, Equations B.1 and B.3. Since the wave is a superposition of a left traveler and a right traveler, it is a direct substitution to establish the time dependence of the waves. Replace x in Equation 2.54 by $x \pm vt$ or

$$y(x, t) = \sum_{n=1}^{\infty} \frac{b_n}{2} \sqrt{\frac{1}{2L}} \left(\sin \frac{\pi n (x + vt)}{2L} + \sin \frac{\pi n (x - vt)}{2L} \right). \tag{2.56}$$

The velocity field is

$$v(x,t) = \sum_{n=1}^{\infty} \frac{b_n}{2} \sqrt{\frac{1}{2L}} \left(\frac{\pi n v}{2L} \left(\cos \frac{\pi n (x+vt)}{2L} - \cos \frac{\pi n (x-vt)}{2L} \right) \right). \quad (2.57)$$

Using simple trigonometry identities[3], these can be reduced to

$$\begin{aligned} y(x,t) &= \sum_{n=1}^{\infty} b_n \sqrt{\frac{1}{2L}} \left(\sin \frac{\pi n x}{2L} \cos \frac{\pi n v t}{2L} \right) \\ v(x,t) &= -\sum_{n=1}^{\infty} b_n \sqrt{\frac{1}{2L}} \left(\frac{\pi n v}{2L} \left(\sin \frac{\pi n x}{2L} \sin \frac{\pi n v t}{2L} \right) \right). \end{aligned} \quad (2.58)$$

Note that at $x = 2Lm, m = 0. \pm 1, \pm 2, \cdots$ both the displacement and velocity fields are zero, the fixed points. Similarly, the slopes of both the displacement field and velocity field are zero at the free end and the free end images for all time. In this form, Equation 2.58, the position and time parts of the fields are separated. This is the form of the solution in which the dynamics separate. In this case, you there is a distributed oscillator for each n, each part of the string has the same time dependence, and, in each term of the series expansion, that time dependence is that of an oscillator with frequency $\omega = \frac{\pi n v}{4L}$ where this v is as in the above $v = \sqrt{\frac{T_e}{\rho}}$.. The system is a superposition of an infinite number of un coupled distributed oscillators, each with a different natural frequency. Although this analysis was carried through for the case of one fixed and one free end and an initial nonzero displacement field, similar conclusions follow for the other possible end point boundaries of the stretched string and possible initial field configurations.

2.3.3 String-like Fields

There are several different variations that are simple extensions of the stretched string as described here. The most important is the addition of the a spring attachment at each place on the string. For reasons that will be clear shortly that this is sometimes called the massive string. Another extension is to include thickness in the description of the string. We will conclude this section with a discussion of the mechanisms for energy dissipation.

[3] $\sin\gamma + \sin\rho = 2\sin\left(\frac{\gamma+\rho}{2}\right)\cos\left(\frac{\gamma-\rho}{2}\right)$ and $\cos\rho - \cos\gamma = 2\sin\left(\frac{\gamma+\rho}{2}\right)\sin\left(\frac{\gamma-\rho}{2}\right)$

The Massive String

For the massive string, the addition of an elastic force returning the displaced string back to the neutral position modifies Equations 2.59 to

$$\begin{aligned}\frac{\partial y}{\partial t}(x,t) &= v(x,t)\\ \rho\frac{\partial v}{\partial t}(x,t) &= T_e\frac{\partial^2 y}{\partial x^2}(x,t)+ky(x,t).\end{aligned} \tag{2.59}$$

where k is the elastic factor

2.3.4 String in Three Dimensions

In the usual case of the string under tension, the string is free to move in two directions that are transverse to the direction of the string. In most applications, there is no differentiation between the two orthogonal transverse directions and thus the motion is that of two independent identical strings. The equations of motion are simply two copies of Equation 2.3 or better yet Equation 2.59,

$$\begin{aligned}\frac{\partial y_1}{\partial t}(x,t) &= v_1(x,t)\\ \rho\frac{\partial v_1}{\partial t}(x,t) &= T_e\frac{\partial^2 y_1}{\partial x^2}(x,t)\\ \frac{\partial y_2}{\partial t}(x,t) &= v_2(x,t)\\ \rho\frac{\partial v_2}{\partial t}(x,t) &= T_e\frac{\partial^2 y_2}{\partial x^2}(x,t).\end{aligned} \tag{2.60}$$

Obviously, as is the case with the one transverse direction string, the three dimensional string is linear hyperbolic differential system and all disturbances are travelers. Thus there are four boundary conditions for the motions: a displacement wave in 1 field system, a velocity wave in the 1 field system, a displacement wave in the 2 field, and a velocity wave in the 2 field.

We can write this situation more compactly using a vector notation,

$$\begin{aligned}\frac{\partial \vec{y}}{\partial t}(x,t) &= \vec{v}(x,t)\\ \rho\frac{\partial \vec{v}}{\partial t}(x,t) &= T_e\frac{\partial^2 \vec{y}}{\partial x^2}(x,t)\end{aligned} \tag{2.61}$$

where the vector field

$$\vec{y}(x,t) \equiv y_1(x,t)\hat{e}_1 + y_2(x,t)\hat{e}_2, \tag{2.62}$$

with a similar equation for the velocity field. The two orthogonal basis vectors, $\hat{e}_i$ $i = 1, 2$, are chosen to span the transverse plane with the unit vector along the string, $\hat{x}$ and $\hat{e}_1$ and $\hat{e}_2$ forming a right handed three dimensional coordinate system.

We could equally well use a cylindrical basis to describe the string orientation in the transverse plane,

$$\begin{aligned} \hat{\rho} &\equiv \cos\theta\,\hat{e}_1 + \sin\theta\,\hat{e}_2 \\ \hat{\theta} &\equiv -\sin\theta\,\hat{e}_1 + \cos\theta\,\hat{e}_2 \end{aligned} \tag{2.63}$$

with the inverse

$$\begin{aligned} \hat{e}_1 &\equiv \cos\theta\,\hat{\rho} - \sin\theta\,\hat{\theta} \\ \hat{e}_2 &\equiv \sin\theta\,\hat{\rho} + \cos\theta\,\hat{\theta}. \end{aligned} \tag{2.64}$$

2.3.5 Symmetry from Dynamics

Obviously in Section 2.3.4 in the case that no transverse direction is preferref, the choice of the direction of one of the $\hat{e}_i$'s in the transverse plane is arbitrary and the system is symmetric under rotations of that unit vector in the transverse plane, rotations about the axis that is the positive direction of the string. Defining the positive rotation as one in which $\hat{e}_1$ rotates into $\hat{e}_2$, the new coordinate systems basis vectors are related to the old by

$$\begin{aligned} \hat{e}'_1 &= \cos\theta\,\hat{e}_1 + \sin\theta\,\hat{e}_2 \\ \hat{e}'_2 &= -\sin\theta\,\hat{e}_1 + \cos\theta\,\hat{e}_2. \end{aligned} \tag{2.65}$$

In this new coordinate description of the fields,

$$\vec{y}\,'(x,t) = y'_1(x,t)\hat{e}'_1 + y'_2(x,t)\hat{e}'_2 = \vec{y}(x,t) \tag{2.66}$$

or

$$\begin{aligned} y_1(x,t) &= \cos\theta\,y'_1(x,t) + \sin\theta\,y'_2(x,t) \\ y_2(x,t) &= -\sin\theta\,y'_1(x,t) + \cos\theta\,y'_2(x,t). \end{aligned} \tag{2.67}$$

These equations can be inverted with the substitution $\theta = -\theta$.

This is the interpretation of the transformation that is known as the passive view, the field is unchanged only the coordinate system is modified, see Appendix A.4. Note that the transformation, Equation 2.67, is simpler than the usual rotation since it does not effect the field variables. Generally, a transformation will change the field event labels. This is not the case here since the transformation is a rotation about the string direction axis and thus the position label on the string and the time are unaffected. When a transformation effects only the field labels and not the coordinates, such as Equation 2.67, the transformation in called a transformation on the internal degrees of freedom. Thus the string in three dimensions is a system with an intrinsic $O(2)$ symmetry.

In addition, Equation 2.67 can be interpreted in the active view. In this case, the basis vectors are unchanged and the system is rotated. The rotated field is

$$\vec{y}'(x,t) = y_1'(x,t)\,\hat{e}_1 + y_2'(x,t)\,\hat{e}_2, \tag{2.68}$$

where

$$\begin{aligned} y_1'(x,t) &= \cos\theta\, y_1(x,t) - \sin\theta\, y_2(x,t) \\ y_2'(x,t) &= \sin\theta\, y_1(x,t) + \cos\theta\, y_2(x,t). \end{aligned} \tag{2.69}$$

There are also rotated velocity fields, $\vec{v}\,'$, and, obviously, since the dynamic is linear in the coupled fields and the rotation is a linear transformation, the rotated fields satisfy Equation 2.61 in both the primed and unprimed fields. For now, this will serve as our definition of a symmetry: a change that leads to fields with the same dynamical equations. In other words, the two solution spaces have the same set of functions. The transformed solution differing from the original in the boundary configurations. The string in three dimensions is symmetric for rotations about the string axis.

With this definition, there are many other symmetries for the string. In any number of dimension, the inversion of the field variables, $\vec{y}\,'(x,t) = -\vec{y}(x,t)$ and $\vec{v}\,'(x,t) = -\vec{v}(x,t)$, is a discrete symmetry. This is a special case of the more general case of any rescaling of the field amplitudes, $y'(x,t) = \lambda y(x,t)$ and $v'(x,t) = \lambda v(x,t)$. Note that the only dimensional requirement from the dynamic is the $\dim(v) \stackrel{\dim}{=} \frac{\dim(y)}{T}$. Thus a rescaling of $y(x,t)$ implies a rescaling of $v(x,t)$.

There are also coordinate transformations. As implied by the dimensional argument above, time inversion, $t' = -t$ is a symmetry as long as there is an accompanying velocity field inversion, $\vec{v}\,'(x',t') = -\vec{v}(x,-t)$. The only other scaling requirement from dimensional arguments given in the dynamics is that the factor $\frac{T_e}{\rho}$ have the dimensions of a velocity squared which is the case. This analysis implies that the general case of a coordinate transformations will also require some field transformations to bring the full set of dynamical equations into the same form.

2.3.6 Relativity of the String

The simple string system obeys a hyperbolic differential equation with only one dimensional parameter, a velocity, which implies that it is should be relativistically invariant. Before considering the general case, it is worthwhile to rationalize the fields and coordinates. The quantity $\sqrt{\frac{T_e}{\rho}}$ is a velocity and defining

$$c \equiv \sqrt{\frac{T_e}{\rho}}, \tag{2.70}$$

a rationalized time which is dimensionally a length is

$$\tau \equiv ct, \tag{2.71}$$

and a rationalized velocity field

$$\vec{v}_R(\tau, x) \equiv \frac{\vec{v}(\tau, x)}{c}. \tag{2.72}$$

Note that we have also switched the order of the arguments to follow the more usual convention when dealing with relativistic systems. With these substitutions, the dynamic becomes

$$\begin{aligned} \frac{\partial \vec{y}}{\partial \tau}(\tau, x) &= \vec{v}_R(\tau, x) \\ \frac{\partial \vec{v}_R}{\partial \tau}(\tau, x) &= \frac{\partial^2 \vec{y}}{\partial x^2}(\tau, x). \end{aligned} \tag{2.73}$$

Before examining the general coordinate transformation, let's reduce Equations 2.73 back to the one dimensional case and into the usual form of the wave equation, Equation 2.3, but now rationalized[4],

$$\frac{\partial^2 y(\tau, x)}{\partial \tau^2} - \frac{\partial^2 y(\tau, x)}{\partial x^2} = 0 \tag{2.74}$$

[4]In this section dealing with symmetry, we use only the source free wave equation. The addition of source terms to the left hand side of the Equation 2.74, see Equation 2.43, would require that for the system to be symmetric, in addition to the requirements in this section, the source terms be unchanged by the coordinate transformation, i. e. $f'(x', t') = f(x(x', t'), t(x', t')) = f(x', t')$. This is the same as the requirement that coordinates in the source term enter only as form invariants for that family of transformations. Form invariants are combinations of the coordinates such that the combination of coordinates before and after transformation is the same. An example for rotations in three dimensions is the combination $\vec{x}' \cdot \vec{x}' = \vec{x} \cdot \vec{x}$ where $\vec{x}$ is the position coordinate before rotation and $\vec{x}'$ is the position coordinate after, see Section A.5.3.

Consider the general coordinate transformation for our string, again working in the active view,

$$\begin{aligned} x' &= x'(\tau, x) \\ \tau' &= \tau'(\tau, x), \end{aligned} \tag{2.75}$$

and the inverse

$$\begin{aligned} x &= x(\tau', x') \\ \tau &= \tau(\tau', x'). \end{aligned} \tag{2.76}$$

For now, our definition of a symmetry under the coordinate transformation is that the dynamic is recovered in the new description,

$$\frac{\partial^2 y'(\tau', x')}{\partial \tau'^2} - \frac{\partial^2 y'(\tau', x')}{\partial x'^2} = 0, \tag{2.77}$$

where $y'(\tau', x') = y\left(x(\tau', x'), \tau(\tau', x')\right)$.

Although tedious, it is direct to find that

$$\begin{aligned} \frac{\partial^2 y'(x', \tau')}{\partial \tau'^2} = & \frac{\partial y}{\partial x}\frac{\partial^2 x}{\partial \tau'^2} \\ & + \frac{\partial x}{\partial \tau'}\left\{\frac{\partial^2 y}{\partial x^2}\frac{\partial x}{\partial \tau'} + \frac{\partial^2 y}{\partial x \partial \tau}\frac{\partial \tau}{\partial \tau'}\right\} \\ & + \frac{\partial y}{\partial \tau}\frac{\partial^2 \tau}{\partial \tau'^2} \\ & + \frac{\partial \tau}{\partial \tau'}\left\{\frac{\partial^2 y}{\partial x \partial \tau}\frac{\partial x}{\partial \tau'} + \frac{\partial^2 y}{\partial \tau^2}\frac{\partial \tau}{\partial \tau'}\right\} \end{aligned} \tag{2.78}$$

with a similar equation for $\frac{\partial^2 y'(x', \tau')}{\partial x'^2}$. Regrouping, the dynamic in the new description is

$$\begin{aligned} \frac{\partial^2 y'(x', \tau')}{\partial \tau'^2} - \frac{\partial^2 y'(x', \tau')}{\partial x'^2} = & \frac{\partial y}{\partial x}\left\{\frac{\partial^2 x}{\partial \tau'^2} - \frac{\partial^2 x}{\partial x'^2}\right\} + \frac{\partial y}{\partial \tau}\left\{\frac{\partial^2 \tau}{\partial \tau'^2} - \frac{\partial^2 \tau}{\partial x'^2}\right\} \\ & + \frac{\partial^2 y}{\partial \tau^2}\left\{\left(\frac{\partial \tau}{\partial \tau'}\right)^2 - \left(\frac{\partial \tau}{\partial x'}\right)^2\right\} \\ & + \frac{\partial^2 y}{\partial x^2}\left\{\left(\frac{\partial x}{\partial \tau'}\right)^2 - \left(\frac{\partial x}{\partial x'}\right)^2\right\} \\ & + 2\frac{\partial^2 y}{\partial \tau \partial x}\left\{\frac{\partial x}{\partial \tau'}\frac{\partial \tau}{\partial \tau'} - \frac{\partial x}{\partial x'}\frac{\partial \tau}{\partial x'}\right\}. \end{aligned} \tag{2.79}$$

The conditions that the new description has the same dynamic are

$$0 = \frac{\partial^2 x}{\partial \tau'^2} - \frac{\partial^2 x}{\partial x'^2} = \frac{\partial^2 \tau}{\partial \tau'^2} - \frac{\partial^2 \tau}{\partial x'^2} \tag{2.80}$$

$$0 = \frac{\partial x}{\partial \tau'}\frac{\partial \tau}{\partial \tau'} - \frac{\partial x}{\partial x'}\frac{\partial \tau}{\partial x'} \tag{2.81}$$

$$1 = \left(\frac{\partial \tau}{\partial \tau'}\right)^2 - \left(\frac{\partial \tau}{\partial x'}\right)^2 \tag{2.82}$$

$$-1 = \left(\frac{\partial x}{\partial \tau'}\right)^2 - \left(\frac{\partial x}{\partial x'}\right)^2. \tag{2.83}$$

These conditions are easy to interprete, Equations 2.80 would apply to non-linear transformations and imply that the new coordinates satisfy the wave equation. As we now know all the solutions to the wave equations are travelers and thus this requirement basically requires some relation to the special general form for hyperbolic equations, Equation 2.34.

Our interest is the linear transformations which automatically satisfy Equations 2.80. The linear transformations must satisfy the remaining conditions. The case of translations, $x' = x + a$ or $\tau' = \tau + b$, where a and b are arbitrary constants, is trivial since

$$\begin{aligned} \frac{\partial y'}{\partial \tau'} &= \frac{\partial y}{\partial x}\frac{\partial x}{\partial \tau'} + \frac{\partial y}{\partial \tau}\frac{\partial \tau}{\partial \tau'} \\ &= \frac{\partial y}{\partial \tau} \end{aligned} \tag{2.84}$$

with a similar condition for the x coordinate. Since this dynamic contains only derivative terms, the dynamic is clearly symmetric under translations[5] in both space and time.

The most general homogeneous linear transformation is

$$\begin{aligned} x &= \alpha x' + \beta \tau' \\ \tau &= \rho x' + \sigma \tau'. \end{aligned} \tag{2.85}$$

In these terms, Equations 2.81, 2.82, and 2.83 are

$$\begin{aligned} 0 &= \beta\sigma - \alpha\rho \\ 1 &= \rho^2 - \sigma^2 \\ -1 &= \beta^2 - \alpha^2. \end{aligned} \tag{2.86}$$

[5]Note that a common addition to string-like field theories is a term which returns the system to the equilibrium position, $\frac{\partial^2 y}{\partial \tau^2} - \frac{\partial^2 y}{\partial x^2} + m^2 y = 0$. This dynamic does not have translation symmetry.

This is three equations with four unknowns and thus the solution space is a one parameter manifold. A simple labeling of this space is by identifying

$$\begin{aligned} \alpha &= \sigma = \cosh\gamma \\ \beta &= \rho = \sinh\gamma. \end{aligned} \tag{2.87}$$

These transformations are the well known Lorentz transformation for a $(1,1)$ space. Again, this is merely the condition that the transformation does not distort the characteristics of this hyperbolic partial differential equation. In fact, the conditions, Equations 2.86, are nothing more than the conditions that $\tau^2 - x^2$ be a form invariant for the family of transformations which is the usual definition of the Lorentz transformation. This is the form for Lorentz transformations that takes advantage of hangle[6] or rapidity to give a labeling that is additive. It can be identified with the usual labeling by identifying the hangle as $\gamma \equiv \tanh^{-1}\frac{V}{c}$ where V is a relative velocity between the observers used in the Lorentz transformation and c is the speed of the travelers for this system.

It is important to note that the string dynamic in the linearized form, Equation 2.59, does not satisfy this coordinate transformation, and without further development will not have a dynamic that is symmetric under these coordinate transformations. We can recover the symmetry of this linear in time dynamic by expanding the meaning of the transformation to include field transformations. Again, it will be computationally simpler if we deal with more rationalized fields. Also, we can develop a more balanced dynamic by redefining our fields. Instead of starting with Equation 2.73 rewritten for the one spatial dimension, we start with the dynamic for the slope of the string field,

$$s(x,\tau) \equiv \frac{\partial y}{\partial x}(x,\tau) \tag{2.88}$$

and the rationalized velocity field,

$$\begin{aligned} \frac{\partial s}{\partial \tau}(x,\tau) &= \frac{\partial v_R}{\partial x}(x,\tau) \\ \frac{\partial v_R}{\partial \tau}(x,\tau) &= \frac{\partial s}{\partial x}(x,\tau). \end{aligned} \tag{2.89}$$

Following the pattern of vector field transformations under rotations, see Equation 2.68, but in this case also using the coordinate transformations of Equation 2.87

$$\begin{aligned} s'(x',\tau') &= As\left(x(x',\tau'),\tau(x',\tau')\right) + Bv_R\left(x(x',\tau'),\tau(x',\tau')\right) \\ v_R'(x',\tau') &= Cv_R\left(x(x',\tau'),\tau)x',\tau')\right) + Ds\left(x(x',\tau'),\tau(x',\tau')\right) \end{aligned} \tag{2.90}$$

[6] See my notes on Relativity, [Gleeson Relativity]

and requiring that the linear-in-time dynamic, Equation 2.89, be the same in the new coordinates and fields requires

$$\begin{aligned} 0 &= \{A \sinh\gamma - D\cosh\gamma + B\cosh\gamma - C\sinh\gamma\} \\ 0 &= \{A\cosh\gamma - D\sinh\gamma + B\sinh\gamma - C\cosh\gamma\}. \end{aligned} \tag{2.91}$$

The solutions for the transformation coefficients are

$$\begin{aligned} A &= C &= \cosh\gamma \\ B &= D &= \sinh\gamma. \end{aligned} \tag{2.92}$$

Thus the string slope field and velocity field, although both separately satisfy the second order in time wave equation under a Lorentz coordinate transformations, must also transform like the coordinates to maintain the form of the linear time dynamic. This is possible because all these equations are linear in the fields and thus the fields and linear combinations of the fields will satisfy the same second order in time evolution. A similar property of field transformation will obtain in the case of the source free electromagnetic fields, Section 6.

Another feature of the current analysis is that the analogy to the electromagnetic situation goes even further. The displacement field acts as a potential field which is a gauge field for the system of slope and velocity fields, $s(x,\tau) = \frac{\partial y}{\partial x}$ and $v_R(x,\tau) = \frac{\partial y}{\partial \tau}$. It is important to note that the value of the displacement field is not physically significant only the slope and velocity fields are. Of course in this case, the conditions on the gauge field is much simpler. The same slope and velocity fields are obtained by a displacement field that differs by another scalar field $y'(x,\tau) = y(x,\tau) + \lambda(x,\tau)$ such that $\frac{\partial \lambda}{\partial x} = 0 = \frac{\partial \lambda}{\partial \tau}$. This only leaves the condition that $\lambda(x,\tau)$ is the trivial field, constant in space and time. This makes sense physically since the value of the displacement is irrelevant. The other difference with the electromagnetic case is that, in that case, the potential field is introduced to satisfy a constraint condition.

This identification of the slope field and velocity fields as transforming under Lorentz transformations allows writing the system in a manifestly covariant formulation. In fact, had we thought about it, we could have intuited these results. The transverse displacement field, $y(x,\tau)$, is a scalar field under Lorentz transformations along the x axis. Since the coordinates transform as a $(1,1)$ space, the fields $v'_R(x',\tau') \equiv \frac{\partial y'(x',\tau')}{\partial \tau'} = \frac{\partial y}{\partial \tau}\frac{\partial \tau}{\partial \tau'} + \frac{\partial y}{\partial x}\frac{\partial x}{\partial \tau'}$ and $s'(x',\tau') \equiv \frac{\partial y'(x',\tau')}{\partial x'} = \frac{\partial y}{\partial \tau}\frac{\partial \tau}{\partial x'} + \frac{\partial y}{\partial x}\frac{\partial x}{\partial x'}$ or $v'_R(x',\tau') = v_R(x,\tau)\cosh\gamma +$

$s(x,\tau)\sinh\gamma$ and $s'(x',\tau') = v_R(x,\tau)\sinh\gamma + s(x,\tau)\cosh\gamma$ where as usual the coordinates on the right side of these equations are $x(x',\tau')$ and $\tau(x',\tau')$.

Writing the coordinates as a two vector, $x^\mu = (\tau, x)$, where μ ranges over the values of 0 and 1 we can write the system of equations in a more covariant form. As usual, there is a metric, $g_{\mu\nu}$, such that $\tau^2 - x^2 = x^\mu g_{\mu\nu} x^\nu$ where the repeat of an upper and lower index implies summing over the range 0 and 1. All the usual relationships that hold in the usual $(1,3)$ Minkowski space, see Appendix G.3 and G, such as the fact that there is a dual set of coordinates that are covariant, $x_\mu \equiv g_{\mu\nu}x^\nu$, i. e. $x'_\mu = \Lambda_\mu{}^\nu x_\nu$ when $x^{\mu\prime} = \Lambda^\mu{}_\nu x^\nu$. As usual, the coordinate derivatives are also covariant, $\partial_\mu \equiv \frac{\partial}{\partial x^\mu}$ which then implies that the velocity and slope fields can be placed in a covariant two vector $s_\mu \equiv (v_R(\tau,x), s(\tau,x))$. Note that I have switched the order of the variables in the fields to be consistent with our counting. Beside the symmetric invariant $\binom{0}{2}$ tensor, $g'_{\mu\nu} = \Lambda_\mu{}^\gamma \Lambda_\nu{}^\rho g_{\gamma\rho} = g_{\mu\nu}$, in the $(1,1)$ space there is an invariant antisymmetric $\binom{2}{0}$ tensor[7] $\epsilon^{\mu\nu}$ where $\epsilon^{01} = 1$. Note that $\epsilon_{\mu\nu} \equiv g_{\mu\mu'}g_{\nu\nu'}\epsilon^{\mu'\nu'} = -\epsilon^{\mu\nu}$. Using this tensor, we can construct another vector field ${}^*s^\mu = \epsilon^{\mu\nu}s_\nu$. In this notation the dynamic, Equation 2.89, becomes

$$\begin{aligned} \partial_\mu {}^*s^\mu &= 0 \\ \partial_\mu s^\mu &= 0. \end{aligned} \tag{2.93}$$

Using the observation that the displacement field is unknowable to within a constant, we recognize that the first of equation in this dynamic is satisfied for a gauge field[8] $y(\tau,x)$, the original displacement field, with the auxiliary condition that

$$s_\mu(\tau,x) = \partial_\mu y(\tau,x). \tag{2.94}$$

Note that the displacement field is a gauge field only for the simple string. In modified dynamics such as the massive string, the displacement field is measured from a point that is identifiable in the dynamic. In the presence of a force per unit length, $f(\tau,x)$, this dynamic becomes

$$\begin{aligned} \partial_\mu {}^*s^\mu(\tau,x) &= 0 \\ \partial_\mu s^\mu(\tau,x) &= f(\tau,x). \end{aligned} \tag{2.95}$$

[7]In a $(1,3)$ space the corresponding invariant antisymmetric tensor is a $\binom{4}{0}$ form, $\epsilon^{\mu\nu\gamma\rho}$. Both the $(1,3)$ and $(1,1)$ cases derive their invariance from the requirement that the determinant of the Lorentz transformations are ± 1.

[8]This result is similar to the situation in electromagnetism in which the potential functions are gauge fields and the electric and magnetic fields are the measurable entities.

and is partially solved with the identification given in Equation 2.94. Once the gauge field is introduced, the remaining condition is our original condition Equation 2.3 with the force term added.

In addition, all the tools of analysis for the Lorentz group prove useful for the understanding of the bilinear forms that can be formed from these fields. In particular, the reduction of the second rank tenor form $s^{\mu}(\tau,x)s^{\nu}(\tau,x)$ has two singlet forms. One is the inner product, $s^{\mu}(\tau,x)g_{\mu\nu}s^{\nu}(\tau,x)$, is a Lorentz scalar field. Another singlet field is the form $s^{\mu}(\tau,x)s^{\nu}(\tau,x) - s^{\nu}(\tau,x)s^{\mu}(\tau,x)$. There is a doublet field that transforms as a vector

In Appendix A.5 and Section 3.2, we will develop more powerful tools for the analysis of symmetries and apply these to the string system. There will be advantages to using the relativistic identifications developed in this section.