

A Computational Approach to Wave Motion in One and Two Dimensions

1. Introduction

Computational physics has evolved into an equally valid approach to exploring the laws of nature, augmenting the more familiar experimental and theoretical approaches. It affords access to physical regimes hitherto unapproachable. This inaccessibility may be due to limitations of temporal acuity, if the relevant time scales are too short as in explosive reactions, or too long as in climate modeling or the evolution of the Universe. There may be physical limitations due to hostile environmental conditions which prevents direct observation as, for example, in the core of the Sun. Many problems exhibit several limitations such as in a supernova. These computational laboratories are meant to introduce students at the high school level to the computational approach to problem solving. The intent is not to create computational physicists nor to teach computer programming. Therefore the emphasis will not be on creating the programs (although their algorithms will be discussed) but on the underlying concepts. Whenever possible, the derivations have used only algebraic or geometrical arguments, to allow the greatest possible dissemination. Where possible, the relationship to the Calculus is included for those few high schools in which it is taught.

The techniques used in computational physics often appear cumbersome to the student used to the traditional problem solving technique of finding exact analytical solutions. The objections often center around the complaint that the answers are only approximate. This overlooks the many levels of approximations built into the “exact” analytical solutions. The physical equations (often called “laws”) are themselves approximate descriptions of the physical world. There are scale approximations which depend upon the requirements of the problem -- different approximations are required for sub-atomic, molecular, macroscopic, or universal scales. Spatial approximations are often employed as simplifying assumptions reducing the complexity from three dimensions to two dimensions, or only one dimension. Forces having no significant influence are ignored, such friction or gravity, which are always present but unimportant for many applications. This list could continue, but the point is that “approximate solutions” have always been appropriate, assuming that some care was exercised in the approximating. Often these approximations are self evident: sub-atomic physics is not the best way to describe galaxy mergers, for example. In order to make responsible choices concerning which approximations should be made, some general knowledge and experience with the possibilities and possible results is required. It is the same with computational physics, which has its own levels of approximation and different techniques (and resulting complications) of problem solution. An operational definition of problem solution will be used here. “A problem is considered solved if, at any desired time, the required quantities (e.g. position or velocity) are known to within a reasonable tolerance.”

Computational physics is *not* simply plugging values into an analytic solution to get quick numbers. (That is the purpose of pocket calculators.) Rather, one starts with the basic equations and assumptions relevant to the problem to be solved. The computer is then used to advance these equations in space or time (or both) starting from a specified state, called the initial conditions. This is often performed as a series of small steps or iterations. As a simple example of the computational approach, consider a simple harmonic oscillator.

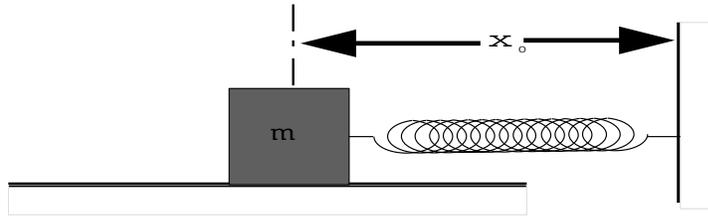
2. Simple Harmonic motion

Figure 2.1 should be familiar to anyone with a knowledge of physics. It consists of a mass m attached to a massless spring. The spring is attached to an immovable wall at the right. The mass is constrained (as in a trough) to one dimensional horizontal motion, without friction. (Note the approximations!) The equilibrium position is x_0 . It is assumed that the spring obeys Hooke’s Law

$$F = -k(x - x_0), \quad (2.1)$$

where F is the force exerted by the spring on the mass and k is the spring constant, expressing the spring’s stiffness. Equation (2.1) is simply a mathematically elegant way of stating what in English can become quite cumbersome: “when the length of a spring is changed by being stretched or compressed beyond its

Figure 2.1 This problem has intrigued physics students for generations. A mass attached to a spring is allowed to slide back and forth on a frictionless surface. The equilibrium spring length is x_0 .



equilibrium (relaxed) length, a force is exerted by the spring to resist this change in length. This restoring force is linearly proportional to the change in length. For example, the force doubles if the change in length doubles.” Newton’s 2nd law is needed to relate the acceleration of the mass to the force produced by the spring. This law states that the force exerted on an object is directly proportional to its observed acceleration, or

$$F = m a, \tag{2.2}$$

where m , the constant of proportionality, is the *inertial mass*. Combining equations (2.1) and (2.2) produces

$$ma = -k(x - x_0). \tag{2.3}$$

This is often called the equation of simple harmonic motion, because the motion this equation describes consists of a periodic sinusoidal movement that has a single frequency, which is

$$f = \frac{1}{2\pi} \sqrt{\frac{k}{m}} \tag{2.4}$$

One computational approach to solving this equation might begin by dividing time into discrete intervals of length Δt . This introduces a standard notation which is used throughout this paper: that a change (usually small) in some quantity q from the one value q_1 to another value q_2 is

$$\Delta q = q_2 - q_1. \tag{2.5}$$

Please take careful note of the order of subtraction. This follows the generally accepted convention of “new” minus “old”, which may give rise to negative changes. Now the equation to be solved (here, eqn 2.3) must be converted into a form compatible with modern digital computers. This may involve expressing other quantities in addition to time in terms of small changes as well. This process is called a *discretization* of the equation.

First consider the general definition of velocity:

$$v = \frac{\Delta(\text{position})}{\Delta(\text{time})}. \tag{2.6}$$

Equation (2.6) states that velocity is a change in position divided by that change in time over which this change in position occurred. Likewise, acceleration is

$$a = \frac{\Delta(\text{velocity})}{\Delta(\text{time})}. \tag{2.7}$$

The form of this discretization should be familiar to those with some knowledge of the Calculus. For example the definition of the derivative is

$$\frac{df(x)}{dx} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}. \tag{2.8}$$

Equation (2.3) now becomes

$$\frac{\Delta v}{\Delta t} = -\frac{k}{m}(x - x_0) \tag{2.9}$$

The change in velocity may be written as

$$\Delta v = v^{n+1} - v^n, \tag{2.10}$$

which is the difference between the velocity at the new time level and the velocity at the old. The number of *timesteps* or iterations taken so far is given by the superscript which, in this standard notation, does *not* imply an involution (exponentiation). The use of “n” rather than some specific number implies the equation is general and valid for any timestep. An expression for the new or updated velocity may now be derived from eqn (2.9):

$$v^{n+1} = v^n - \frac{k}{m} (x^n - x_0)\Delta t \tag{2.11a}$$

Likewise, the updated position is

$$x^{n+1} = x^n + \frac{(v^n + v^{n+1})}{2} \Delta t \tag{2.11b}$$

(The astute reader will have noted that in eqn. (2.11b) the velocity used to advance the position in time was the average of the velocity at the beginning of the timestep, v^n , and the velocity at the end of the timestep, v^{n+1} .)

The procedure is to first select a small timestep Δt and then, beginning with the initial conditions, step forward in time by finding successive values for position and velocity from eqns (2.11). It should be noted that such a solution no longer consists of a continuum of positions or velocities, but of values found only at specific points in time determined by the size of the timestep.

From the discretization of the equations, one would expect the accuracy of this computational solution to increase as the timestep decreases. This can be verified using a pocket calculator. However, even after only a couple of oscillations it should become clear that the answer diverges, that is, the oscillations grow larger with each repetition, regardless of how small the timestep is made (tables 1 and 2, figure 2.2). This behavior, called *numerical instability*, is undesirable.

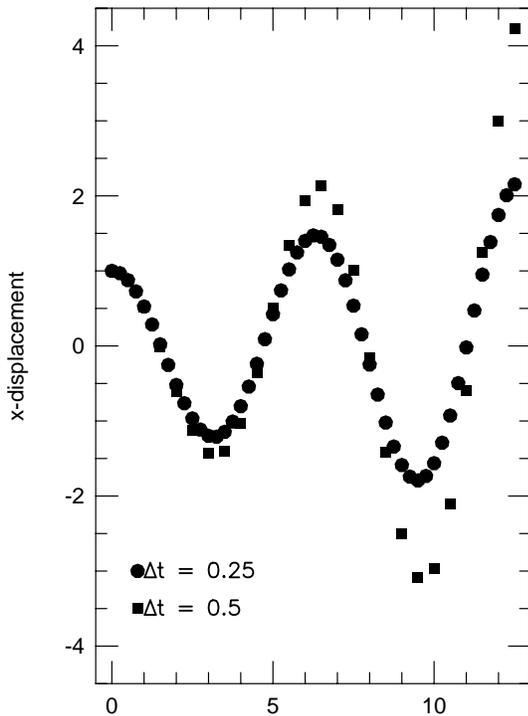


Figure 2.2 This shows results from two different timesteps using the unstable numerical algorithm given by eqns. (2.11) when used for the simple harmonic oscillator shown in figure 2.1.

The initial conditions are:
 $x_0 = 0$, $x^0 = 1$, and $v^0 = 0$.

Results using two different timesteps are displayed. Circles are the results using a timestep of $\Delta t = 0.25$, squares are the results using a timestep of $\Delta t = 0.5$. Note the increasing amplitude with each oscillation, betraying the unstable nature of the algorithm.

The problem is buried within equation (2.11a). The new velocity is computed with a constant acceleration: the acceleration at the beginning of the timestep. This ignores the change in acceleration as the position changes from x^n to x^{n+1} during the timestep Δt . Thus the acceleration is always over-estimated in the velocity calculation. Each time the mass slows and reverses direction, at the left and right

ends of its motion, the deceleration is underestimated by using accelerations calculated only from “old” values. This is a systematic error, and in computational physics systematic errors generally accumulate to unfortunate effect.

A way to correct this is to advance the position using velocities computed halfway through the time step. That is, equations (2.11) now become

$$v^{n+1/2} = v^{n-1/2} + \frac{k}{m} (x^n - x_0) \Delta t \tag{2.12a}$$

$$x^{n+1} = x^n + v^{n+1/2} \Delta t. \tag{2.12b}$$

The resulting scheme is known as a *leapfrog* scheme because of the computational grid pattern (figure 2.3). Numerical grids such as the one shown in figure 2.3, in which some quantities are computed only at certain grid points and other quantities are computed only at points betwixt, are call *staggered* grids.

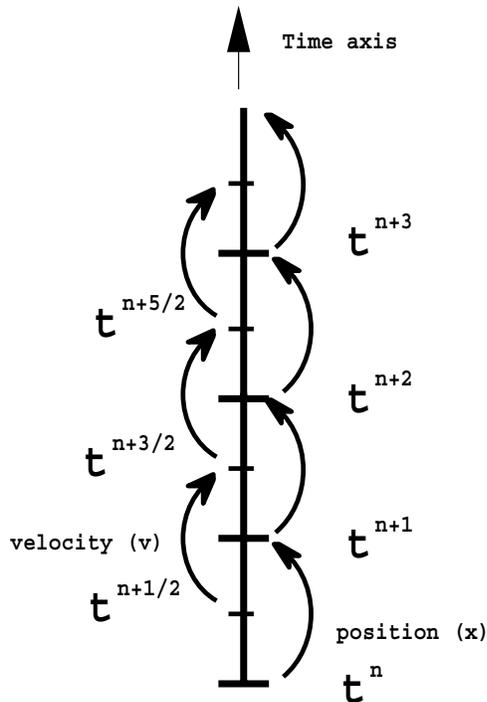


Figure 2.3 In a leapfrog scheme, some quantities are calculated at on integral points in time ($\Delta t, 2\Delta t, 3\Delta t, \dots$), while other quantities are calculated at points in between ($\frac{1}{2}\Delta t, \frac{3}{2}\Delta t, \frac{5}{2}\Delta t, \dots$). The name of this family of numerical schemes is derived from this pattern.

One disadvantage of leapfrog type schemes is the cumbersome initial condition requirement: knowledge of the unknown velocity, $v^{-1/2}$, at the at time $t = -\Delta t/2$, would be required. In practice, the value at time $t = 0$ is used instead. This introduces a small starting error, the magnitude of which is dependent upon the size of the timestep. However, this error does not grow as the calculation progresses. (The technical term for such small errors is *starting glitch*.) The stability of this scheme can be easily demonstrated on a pocket calculator. (Table 3, figure 2.4. Note that the use of $v^{-1/2} = 0$ causes the magnitude of the resulting oscillations to be approximately 1.03 rather than 1.00 as would be expected for the given parameters. Note also that this error does not grow.)

3. Isolated mass between two springs

This problem is a simple extension of the isolated mass on a spring problem. It will prove useful later in constructing a model of a stretched string.

a. Theory

First consider the geometry of figure 3.1. A mass m is suspended between two vertical attachments by two identical massless springs. Gravity is neglected. Each spring exerts a force on the mass. Recall that force is a vector quantity, so that both magnitude and direction are important. At the equilibrium position shown in figure 3.1, the two forces are equal in magnitude, and collinear, with opposite directions. If left undisturbed, the mass will remain in this position.

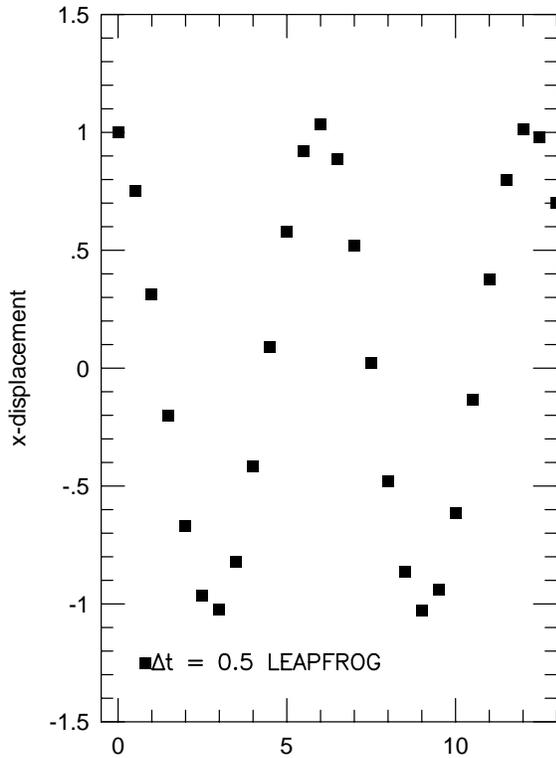


Figure 2.4 This shows results from the stable leapfrog algorithm given by eqns (2.12). The initial conditions are: $x_0 = 0$, $x^0 = 1$, and $v^{-1/2} = 0$. The timestep is $\Delta t = 0.5$.

Note that the error introduced by using $v^{-1/2} = 0$ for the velocity at $t = -1/2$ causes the magnitude of the amplitude to be about 3% larger than the expected 1.0. This starting error does not grow.

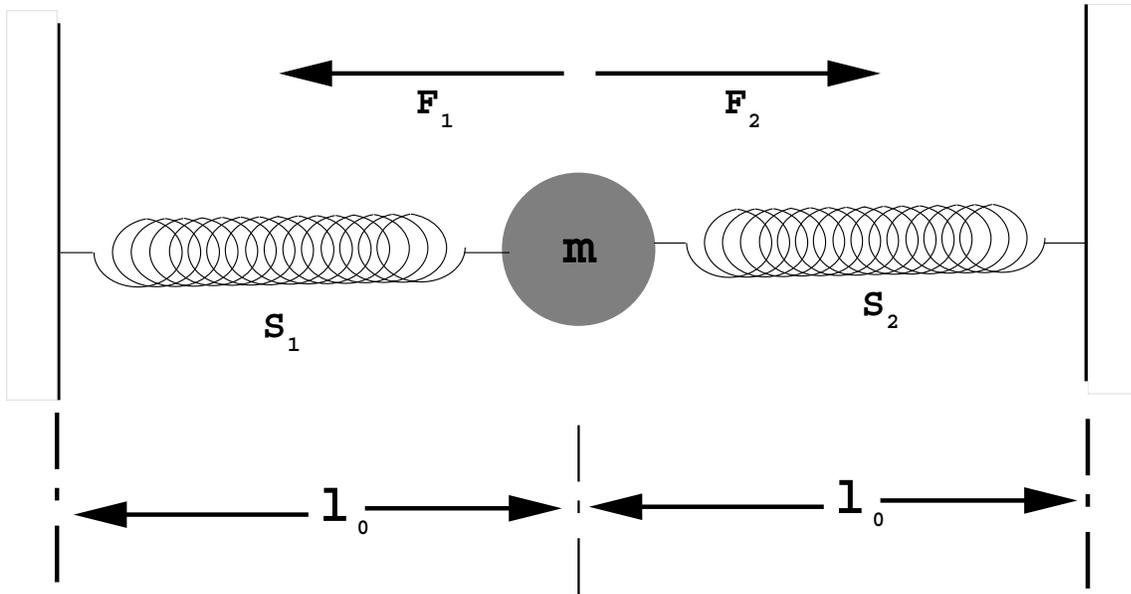


Figure 3.1 This shows a mass suspended between two vertical posts by two massless springs. The springs pull on the weight with sufficient force so that gravity may be neglected, that is, as shown here at equilibrium, there is no appreciable vertical displacement.

Now consider a vertical displacement of the mass by a small amount as shown in figure 3.2. Assume that this displacement is much smaller than the length of the two springs, l_0 . Now the forces exerted on the mass by the two springs are no longer collinear, causing the vector nature of a force to become apparent.

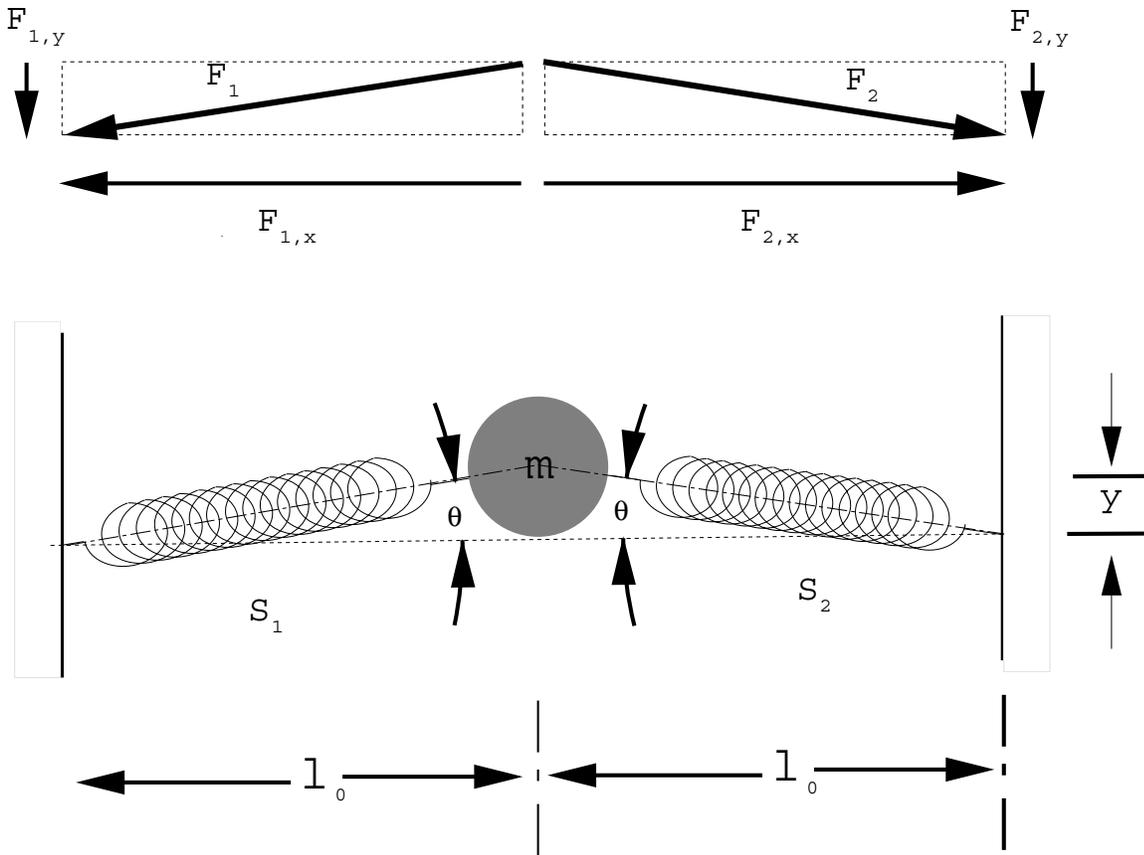


Figure 3.2 This shows the same mass and springs as in figure 3.1, but with a small vertical displacement (greatly exaggerated here). This vertical displacement y is much smaller than the length of either spring, l .

Recall that a vector is a quantity enjoying both magnitude and direction. Vectors will be represented by a **bold** typeface in equations. To determine the resulting motion, these forces must be resolved into their horizontal and vertical vector components, that is, the components parallel to the x-axis and the y-axis, resp. In this way the force created by spring S_1 may be expressed as

$$\mathbf{F}_1 = -F_1 \cos(\theta) \hat{\mathbf{x}} - F_1 \sin(\theta) \hat{\mathbf{y}}, \quad 3.1a$$

and the force created by spring S_2 may be expressed as

$$\mathbf{F}_2 = F_2 \cos(\theta) \hat{\mathbf{x}} - F_2 \sin(\theta) \hat{\mathbf{y}}. \quad 3.1b$$

Here $\hat{\mathbf{x}}$ is a unit vector pointing in the positive x-direction and $\hat{\mathbf{y}}$ is a unit vector pointing in the positive y-direction. The total force acting on the mass is the vector sum of these two forces, or

$$\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2 = (F_2 - F_1) \cos(\theta) \hat{\mathbf{x}} - (F_1 + F_2) \sin(\theta) \hat{\mathbf{y}}. \quad (3.2)$$

When the springs are identical, that is, $F_s = F_1 = F_2$ (where F_s is the magnitude of the force exerted by one of the springs on the mass), the horizontal components are equal in magnitude, but opposite in direction, thus canceling. This leaves only a vertical component

$$F = -2F_s \sin(\theta) \hat{\mathbf{y}}. \quad (3.3)$$

Next, it must be determined by how much each spring is stretched further (if at all) by the small vertical displacement. The length of each spring is now

$$l' = \sqrt{l_0^2 + y^2} = l_0 \sqrt{1 + \left(\frac{y}{l_0}\right)^2}, \quad (3.4)$$

found by using the theory of Pythagoras, which equates the square of the length of the hypotenuse of a right-angled triangle to the sum of the squares of the lengths of its two legs. If the vertical displacement y is small compared to l_0 , then

$$\sqrt{1 + \left(\frac{y}{l_0}\right)^2} \approx 1. \quad (3.5)$$

Therefore any additional stretching of the springs due to small vertical displacements of the mass is negligible. Because there is no appreciable change in length of the strings, then application of Hooke's law, eqn. (2.1), implies no appreciable change in the magnitude of the forces exerted on the mass by the springs. That is, F_s is constant. Using

$$\sin(\theta) = \frac{y}{\sqrt{l_0^2 + y^2}} = \frac{y}{l_0 \sqrt{1 + (y/l_0)^2}} \approx \frac{y}{l_0}, \quad (3.6)$$

in which equation (3.5) has been used to make this approximation, the restoring force felt by the mass is

$$F = -\frac{2F_s}{l_0} y. \quad (3.7)$$

It should be stressed that the restoring force arises *not* from any increase in the magnitude of the forces acting on the mass, but purely from the changes in the *directions* of these forces. That is, the restoring force arises from the vector nature of a force. Understanding this point is crucial for the remaining discussion. The equation of motion for the mass becomes

$$ma_y = -\frac{2F_s}{l_0} y \quad (3.8)$$

This is again the form of the familiar equation of simple harmonic motion, eqn. (2.3). This motion has a single frequency

$$f = \frac{1}{2\pi} \sqrt{\frac{2F_s}{ml_0}} \quad (3.9)$$

b. Numerical Example

The string program was written to simulate the behavior of a continuous vibrating string. Nevertheless, by running it using its coarsest possible representation of the string, it can also be used to simulate the behavior of a single mass suspended by two springs. First, one selects difference scheme number 2, with 1 point mass, from the graphical user interface (gui), figure 3.3.

It may prove useful to realize that the total length is, by construction, unity so the spacing between masses (which is obviously the length of the springs, l_0) is determined by the number of masses requested. Thus, with only one mass, $l_0 = 1/2$. Exploration of the relationship between the various parameters may now proceed. It should be noted that only the tension and mass affect the frequency of the resulting motion. In particular it should be noted that the frequency is independent of the initial displacement and initial velocity. This may be demonstrated by plotting of position vs. time and velocity vs. time for various choices of initial conditions.

It is suggested that the various user-adjustable parameters be chosen such that

$$\frac{2F_s}{ml_0} = 1. \quad (3.10)$$

With this simplification, eqn (3.8) becomes

$$a_y = -y. \quad (3.11)$$

Using this choice of parameters, verify that given an initial displacement in position (a_0) or an initial velocity (v_0) of 0.01, oscillations are produced in both position and velocity (as a function of time) having equal amplitudes but differing phase. Phase is the difference between a particular point or stage of an oscillation relative to a standard position such as an extremum. Or in other words, the oscillations produced by either an initial displacement (but no initial velocity) or by an initial velocity (but no initial displacement) obtain their maxima and minima at different times, but are otherwise the same. This is illustrated by figure 3.4a for the specific initial conditions of $a_0 = 0.01$ and $v_0 = 0.0$, (example A) and figure 3.4b for the specific

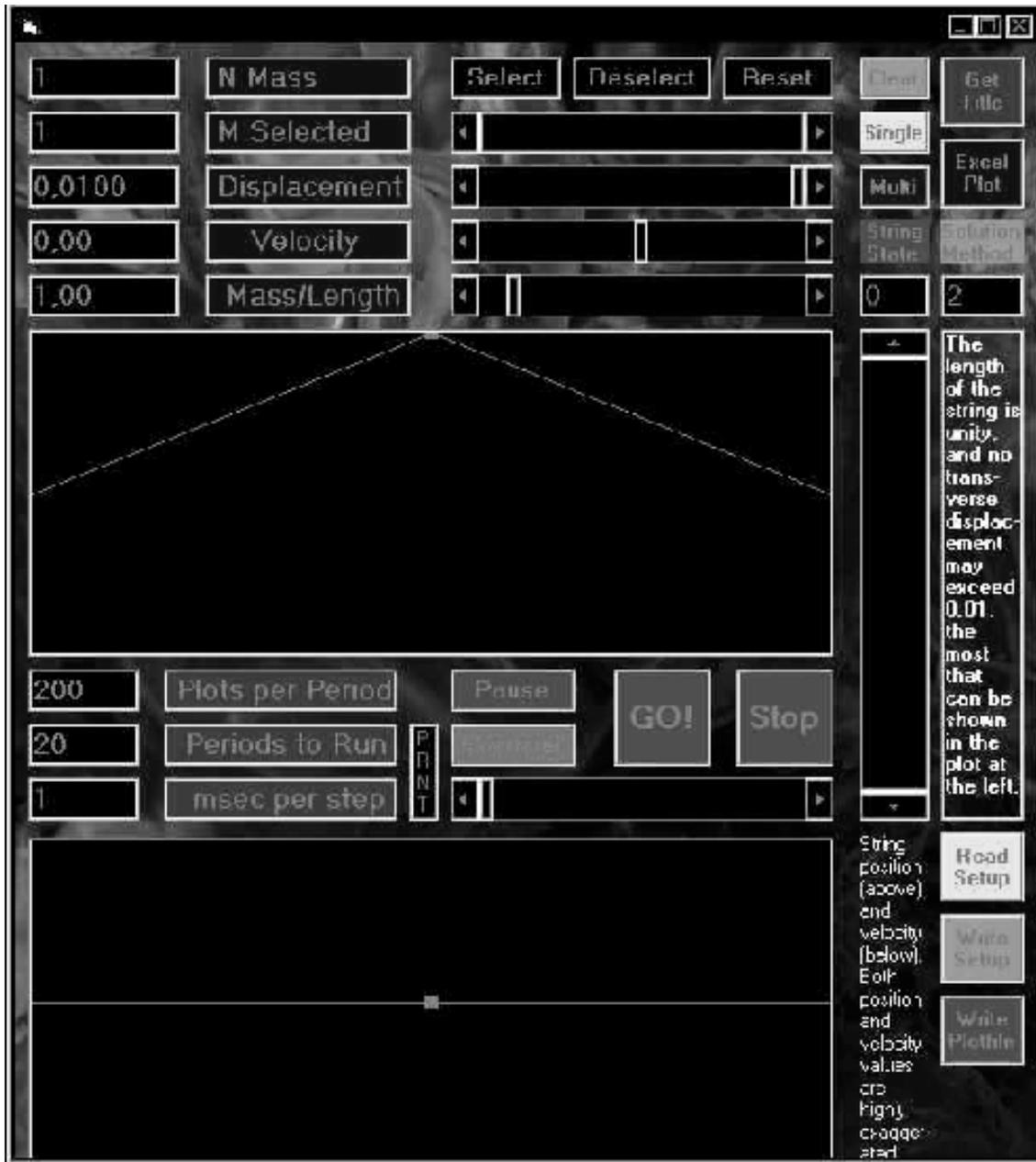


Figure 3.3 This is an example of a typical setup for a single mass. The leapfrog method discussed in the text has clearly been selected.

initial conditions of $a_0 = 0.0$ and $v_0 = 0.01$ (example B). Now measure the resulting amplitudes of the oscillations in position and velocity when the initial position displacement is doubled ($2a_0$) and halved ($\frac{1}{2}a_0$). It should be noted that the resulting oscillation amplitudes double and halve, respectively. This is a consequence of the *linearity* of this problem: the response is directly proportional to the input.

Another consequence of linearity is that when the initial conditions consist of a combination of two simpler initial conditions, the resulting behavior is the sum of the behaviors of these two components. For example, combine the above two examples A and B to produce the initial conditions $a_0 = 0.01$; $v_0 = 0.01$. The amplitude of the resulting oscillations is *not* 0.02, but approximately 0.013, because the solutions must be added in phase, as illustrated by figure 3.5 using position displacement oscillations. In figure 3.5, the transverse displacement as a function of time is plotted at various points for the two examples A and B. Also plotted is the transverse displacement as a function of time resulting from the combination of the initial conditions of examples A and B ($a_0 = 0.01$; $v_0 = 0.01$). At each point in time, the resulting

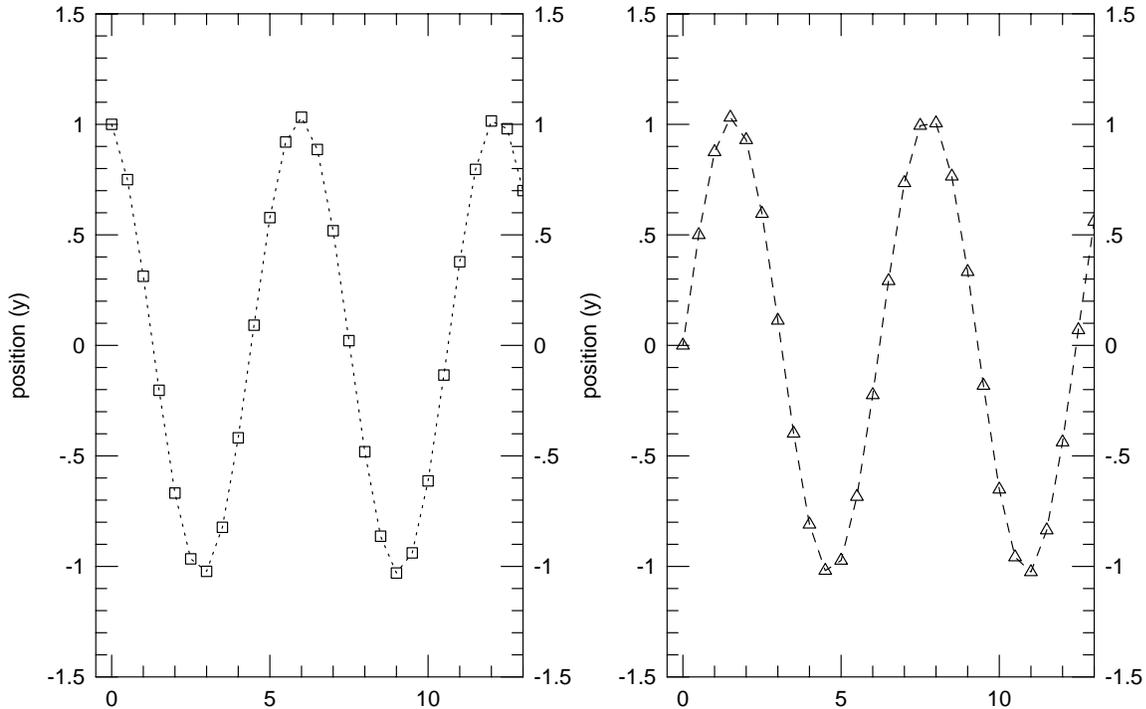


Figure 3.4 Here are two examples showing the sinusoidal behaviour of small oscillations of a mass suspended between two springs. In 3.4A (left) the initial conditions consist of a small displacement in position but no initial velocity, while in 3.4B (right) the initial conditions consist of a small initial velocity but no position displacement. The same behaviour results from either. The two oscillations differ only in phase

displacement is the sum of the displacements of A and B. Because A and B are not in phase, the maximum displacement is not the simple sum of the maxima of A and B. The frequency is, of course, the same. The construction of a similar diagram for the velocity is left to the reader.

4. A Chain of Masses

The example of an isolated mass between two springs from the previous section will now be extended to a one dimensional chain of masses and springs.

a. Theory

Now consider several masses and springs tethered together as in figure 4.1. It is assumed that the spacings are equal and that any *incremental* transverse displacement, $y_i - y_{i-1}$, is small relative to this spacing. Subscripts are used to indicate the spatial relationships among individual objects or measurements. For example, m_i refers to any individual mass, m_{i-1} is the next mass on its left, and m_{i+1} is the next mass on its right. If the vertical displacement is represented by y , then y_i , y_{i-1} , and y_{i+1} are the vertical displacements of the three previously mentioned masses, respectively. The use of i indicates that the relationship being discussed is general and applies to any mass. If a specific mass is desired, it would be indicated by its specific ordinal number: m_1 would indicate the first mass, for example. Note that in a long chain, the absolute displacement of mass j is the sum of all the individual incremental displacements from the first

Figure 3.5 This is an example of linearity. Here, the solid line (stars) results from initial conditions consisting of both a positional displacement and an initial velocity. The resulting oscillations are simply the sum of the oscillations due to either problem in isolation, that is, initial conditions consisting only of a displacement, but no initial velocity (dotted line, squares) and initial conditions consisting only of an initial velocity, but no displacement (dashed line, squares). The resulting oscillation has the same frequency as either of the two individual oscillations, but differing amplitude and phase.

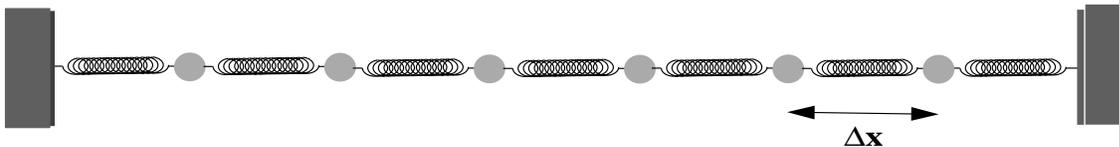
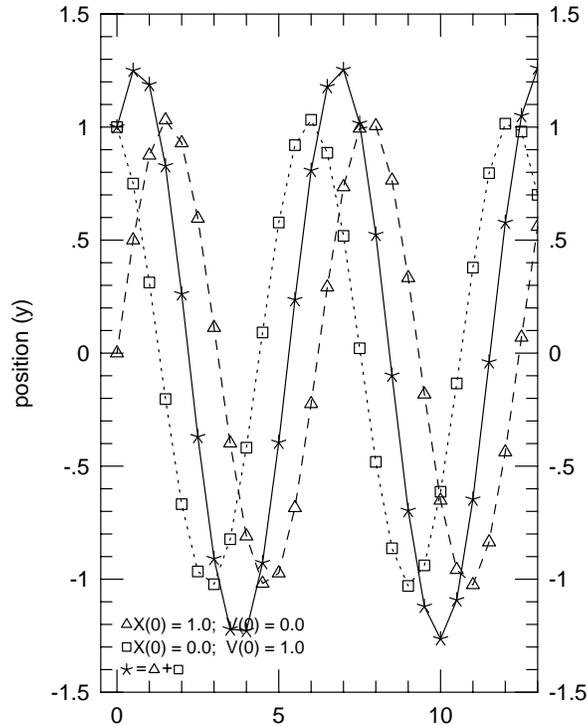


Figure 4.1. This shows a chain of identical masses and springs. The masses of the springs are negligible, and they exert sufficient force on the masses to prevent any appreciable sagging of the chain. This allows gravity to be neglected. The length of each spring at equilibrium (spacing between masses) is everywhere the same: Δx .

mass:

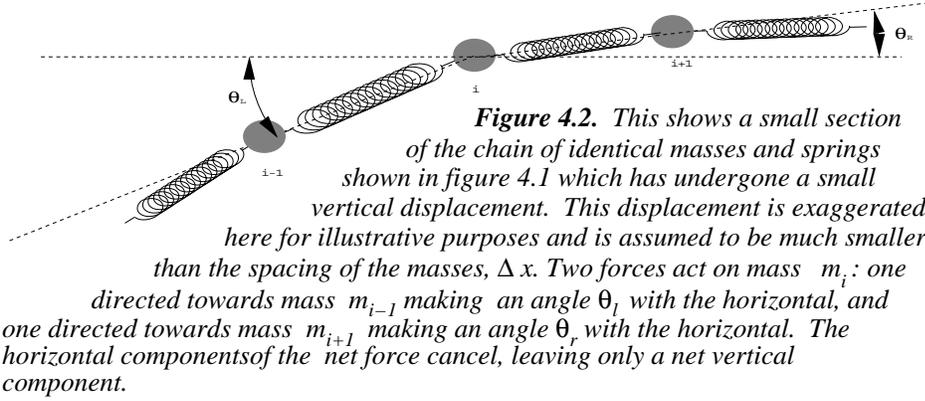
$$y_j - y_0 = \sum_{i=1}^j y_i - y_{i-1}. \tag{4.1}$$

Thus the absolute displacement can be relatively large.

Now the forces affecting the i th mass, m_i , will be constructed using figure 4.2. Figure 4.2 depicts a short segment from fig. 4.1 with a transverse displacement. As with the problem of a single mass suspended between two springs of section 3, the horizontal force components will cancel, assuming small incremental displacements, leaving only a vertical component

$$F = F_s(\sin(\theta_l) - \sin(\theta_r)), \tag{4.2}$$

where F_s is the magnitude of the force each spring exerts on the individual masses, θ_l is the angle the spring to the left of the mass makes with respect to the horizontal, and θ_r is the angle the spring to the right of the mass makes with respect to the horizontal. From figure 4.2 it is clear that



$$\sin(\theta_l) = \frac{y_i - y_{i-1}}{\sqrt{(\Delta x)^2 + (y_i - y_{i-1})^2}} \approx \frac{y_i - y_{i-1}}{\Delta x} \quad (4.3)$$

and

$$\sin(\theta_r) = \frac{y_{i+1} - y_i}{\sqrt{(\Delta x)^2 + (y_{i+1} - y_i)^2}} \approx \frac{y_{i+1} - y_i}{\Delta x}, \quad (4.4)$$

using the approximation of eqn. (3.4). Thus the acceleration on the i th mass becomes

$$a_i = \frac{F_s}{m_i} \left(\frac{y_{i+1} - y_i}{\Delta x} - \frac{y_i - y_{i-1}}{\Delta x} \right). \quad (4.5)$$

A numerical solution may now be constructed using eqns (2.12):

$$v_i^{n+1/2} = v_i^{n-1/2} + \frac{F_s}{m_i \Delta x} (y_{i-1}^n - 2y_i^n + y_{i+1}^n) \Delta t \quad (4.6a)$$

and

$$y_i^{n+1} = y_i^n + v_i^{n+1/2} \Delta t. \quad (4.6b)$$

(The relationship with the single mass of section 3 can be made apparent by holding masses m_{i-1} and m_{i+1} stationary, and setting their offsets y_{i-1} and y_{i+1} to zero. Equation (4.6a) then becomes

$$v_i^{n+1/2} = v_i^{n-1/2} + 2 \frac{F_s}{m_i \Delta x} y_i^n \Delta t. \quad (4.6a')$$

This is recognizable as eqn. (3.8)).

Obviously, the initial conditions must specify the starting displacements and velocities for each mass. To complete *one* timestep, eqns (4.6) must be computed N times for N masses. The advantages of fast computers quickly become evident when considering a problem involving thousands of individual masses.

Most readers will have noticed a problem using eqns (4.6): information for the non-existent 0th mass is required to update m_1 , as well as information for the non-existent $(N+1)$ th mass in order to update the last or N th mass. This additional information, called the *boundary conditions* (for obvious reasons), must also be supplied, defining the positions and velocities of the missing points. For fixed ends (as shown in the illustrations here) these conditions are particularly simple: no motion. These conditions are usually called reflecting, or hard wall, boundary conditions.

b. Numerical Examples

As with the single mass, we use the LCSE string numerical laboratory and difference method 2, but now with 2 masses, figure 4.3. As you may recall, one mass produced only one type of motion. The up-down-up-down motion observed in section 2.b. There was nothing else the single mass could do. A more precise way of saying this is that there existed only one *degree of freedom*, and there was only one *mode* of vibration. Now, with two masses, there are two degrees of freedom which will produce two fundamental modes of vibration.

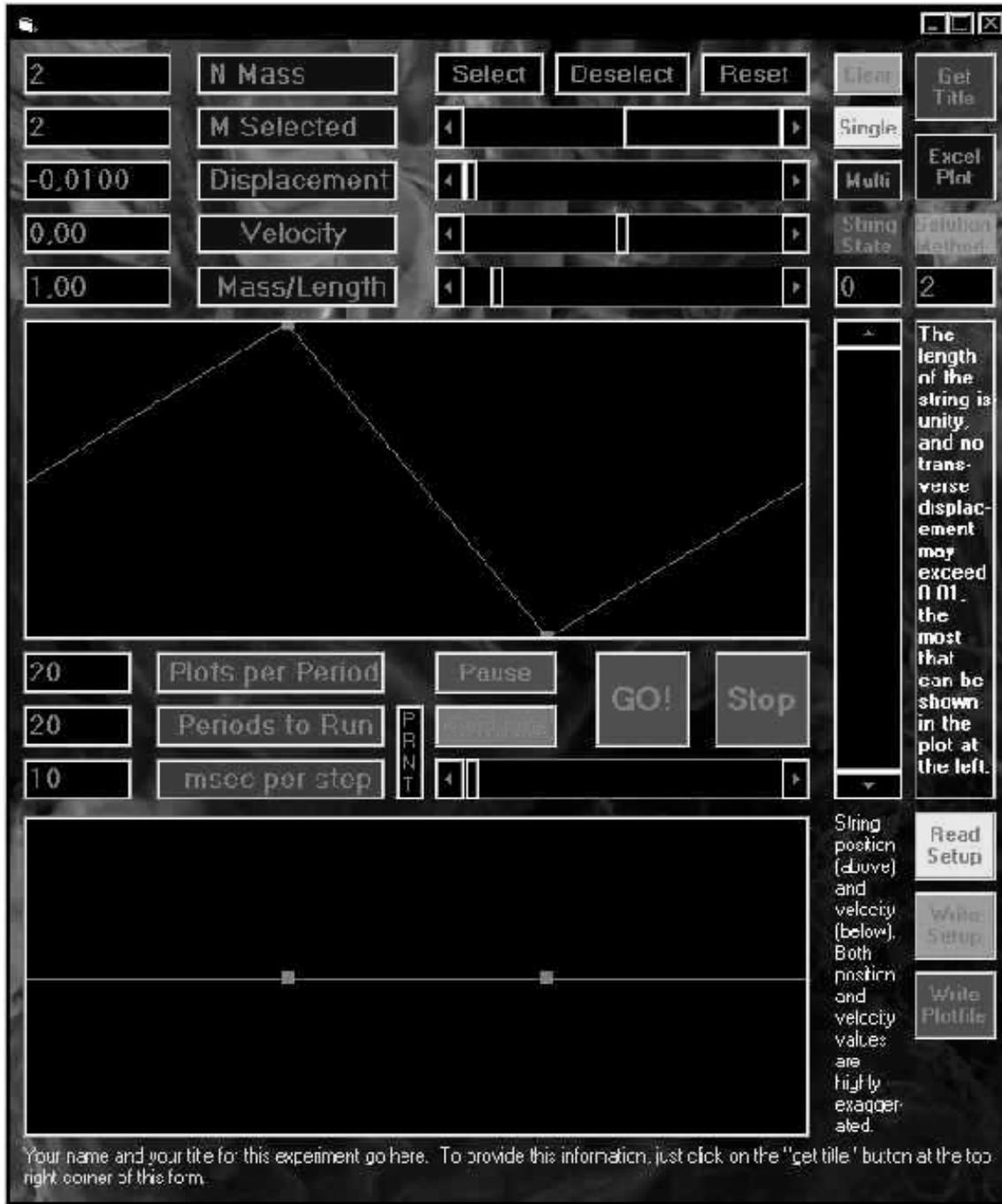


Figure 4.3 This shows the LCSE string program in a typical setup for two masses using the leapfrog method described in the text.

The analysis of these two modes begins with equation (4.5) to provide the acceleration for m_1 :

$$a_1 = \frac{F_s}{m_1} \left(\frac{y_2 - y_1}{\Delta x} - \frac{y_1 - y_0}{\Delta x} \right). \quad (4.7)$$

Using the assumption of reflecting boundary conditions in which the end points are held stationary (that is, $y_0, y_3 = 0$) this becomes

$$a_1 = \frac{F_s}{m_1} \left(\frac{y_2 - 2y_1}{\Delta x} \right). \quad (4.8a)$$

Likewise, the acceleration for m_2 is

$$a_2 = \frac{F_s}{m_2} \left(\frac{y_1 - 2y_2}{\Delta x} \right). \quad (4.8b)$$

These are an example of a coupled set of equations: the motion of m_1 depends upon the position of m_2 , and conversely.

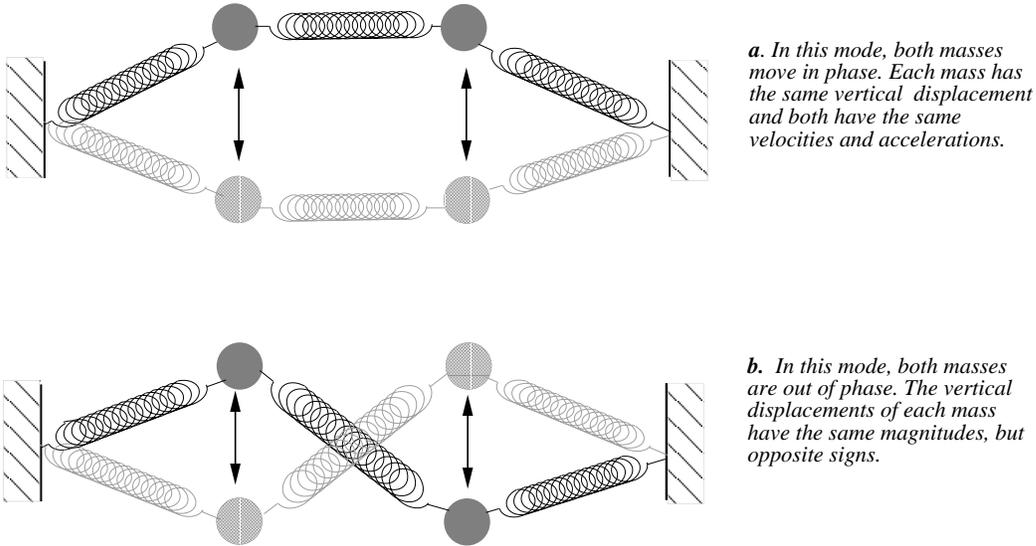


Figure 4.4. With two individual masses, two distinct vibratory modes are possible.

When the masses are identical, there exist two special configurations which can be easily analyzed by most people. The first mode is a motion in which the two masses move in phase (figure 4.4a and figure 4.5, left column). To see this, begin with identical displacements ($y_1 = y_2$) and zero velocities. Then eqns. (4.8) become

$$a_1 = - \left(\frac{F_s}{m \Delta x} \right) y_1 \quad (4.9a)$$

and

$$a_2 = - \left(\frac{F_s}{m \Delta x} \right) y_2 \quad (4.9b)$$

Thus each mass undergoes simple harmonic motion. They begin in phase, and remain in phase since their individual accelerations are always the same. Using eqn. (2.4), this motion has a frequency of

$$f = \frac{1}{2\pi} \sqrt{\frac{F_s}{m \Delta x}}. \quad (4.10)$$

The subscripts on the mass m have been dropped because we have assumed that they are identical.

In the second mode the two masses are out of phase, shown in figure 4.4b and figure 4.5, middle column. This may be initialized using zero initial velocities with displacements which are opposite in sign, but equal in magnitude ($y_1 = -y_2$). Now eqns. (4.8) become

$$a_1 = - \left(\frac{3 F_s}{m \Delta x} \right) y_1 \quad (4.11a)$$

and

$$a_2 = - \left(\frac{3 F_s}{m \Delta x} \right) y_2. \quad (4.11b)$$

Again, each mass undergoes simple harmonic motion. In this mode, they begin out of phase, and remain so since their individual accelerations are always the same. This motion has a frequency of

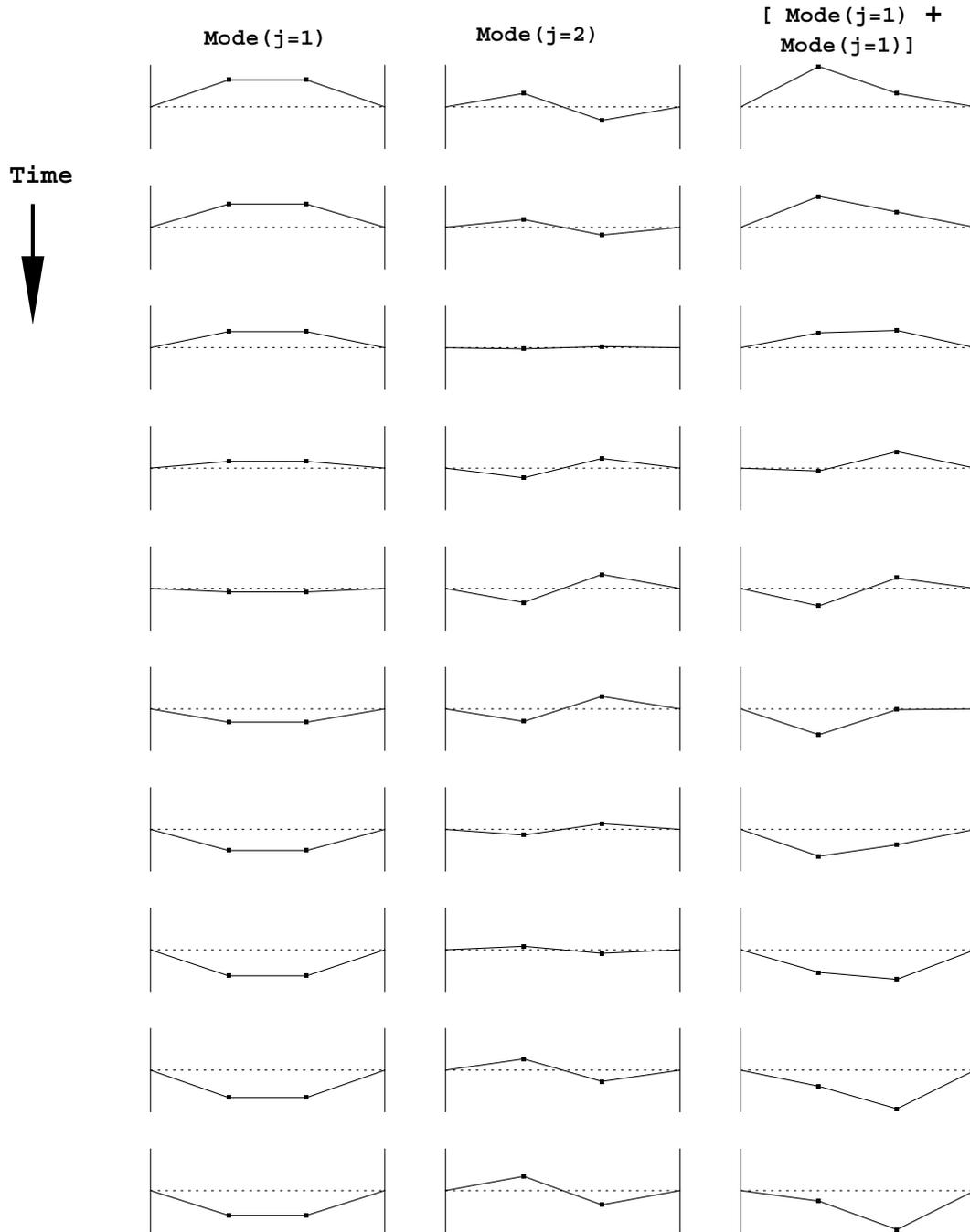


Figure 4.5 These are snapshots of the oscillations of two masses given three different initial conditions. The snapshots are taken at equal times and at uniform intervals, with time increasing from top to bottom. The left and middle columns are the $j=1$ and $j=2$ resonate modes, resp. Note the different oscillation frequencies. The initial amplitude of mode $j=1$ is twice that of mode $j=2$. The initial conditions for the right column is the sum of these two modes. Note the subsequent behaviour is at any one time the sum of the two modes at that time.

$$f = \frac{1}{2\pi} \sqrt{\frac{3F_s}{m \Delta x}}. \quad (4.12)$$

(A similar analysis utilizing velocities rather than displacements is left as a relaxing little exercise for the reader.)

It should be noted that the two different modes have different frequencies: the frequency of the in phase mode being lower than that of the out of phase mode. Furthermore, these two modes are

independent. Initial conditions which are a linear combination of these two modes (as illustrated in figure 4.5) results in motion which is the sum of the two individual oscillations. This is easily verified by the string lab by the interested reader. Any desired initial conditions may be expressed as a linear combination of these two modes. Figure 4.6 shows how the initial conditions in which only one of the masses has a displacement may be constructed.

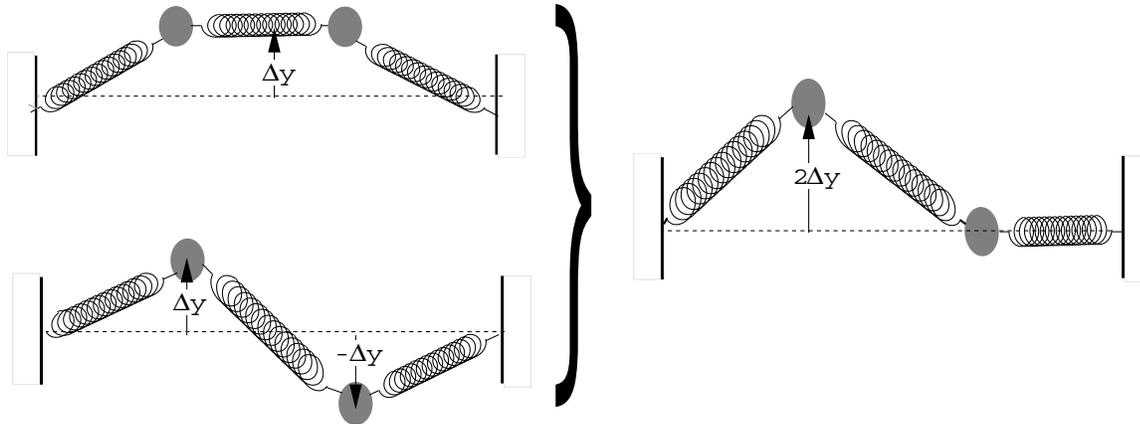


Figure 4.6 Any initial conditions may be created from a linear combination of the two normal modes of vibration. This shows that the initial conditions consisting of only one of the two masses having a positional offset is the result of a simple sum of the symmetric and antisymmetric modes.

An analytical analysis like the above for more than two masses is hopelessly beyond the scope of this text. However, a qualitative approach follows.

Each allowed mode of vibration, called a *normal mode*, corresponds to some integral number of half sine waves corresponding to the equilibrium length of the chain of masses and springs. This integral number of half sine waves, j , is the *mode* number of the vibration. It will be shown, but mercifully *not* derived, that for N particles, there exist only $j = 1, 2, 3, \dots, N$ distinct modes. Assuming the separation of the masses (that is, the length of the springs), Δx is uniform, the initial position of the i th particle for the j th normal mode may be determined by evaluating

$$y_i = A_j \sin \left[\frac{\pi j x_i}{\Delta x (N + 1)} \right], \quad (4.13)$$

where x_i is the x -location of the i th mass, N is the number of masses, A_j is the (small but arbitrary) amplitude of the j th mode, and j is the mode number, $j = 1, 2, \dots, N$. The resulting patterns for two and three masses are shown in figures 4.7 and 4.8. The broken lines are the curves, $f(x)$, resulting from an equation similar to eqn. (4.13) in which the *discrete* locations x_i are replaced by a continuous variable, x :

$$f(x) = A_j \sin \left[\frac{\pi j x}{\Delta x (N + 1)} \right]. \quad (4.14)$$

The only physically meaningful features of these curves are the values at the positions occupied by the masses! Because this has confused so many people, this statement will be repeated: **The only physically meaningful features of these curves are the values at the positions occupied by the masses!** These positions are indicated by large dots, rendering them easily visible to the older reader. Some people become confused by such diagrams, giving the continuous curves more physical meaning than exists. For these systems of masses, there is nothing special about the extrema, or the nodes of these curves. This confusion undoubtedly arises from the fact that the system consists of a *small number of discrete* masses, not a continuous system. Therefore, the extrema and nodes are *not necessarily* occupied by masses as some would assert. Figure 4.7 shows the displacements for the two modes allowed by two masses, and figure 4.8 shows the displacements for the three modes allowed by three masses. It is clear from this that there is

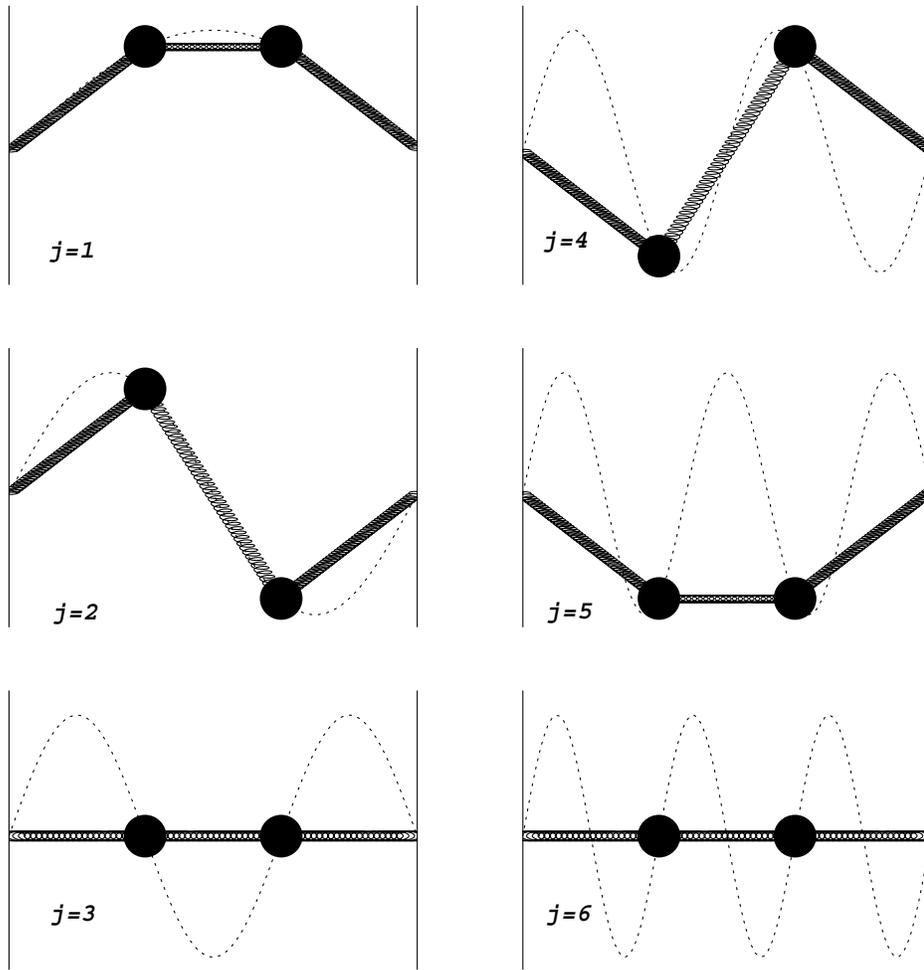


Figure 4.7 This clearly shows that a system made up of a discrete number of particles has a limited number of unique modes of vibration $j: j=1,2,\dots,N$ where N is the number of particles. Increasing j beyond N produces solutions which have already been depicted. For two masses, as shown here, there are two modes, $j=1, j=2$. Mode $j=3$ is the null mode (not meaningful), $j=4$ is the same as $j=2$ (except for a trivial minus sign) as are modes $j=5$, and $j=1$. Mode $j=6$ is again the null mode. The broken line is from eqn. (4.13). Note that the only meaningful features of these curves are the values at the positions occupied by the masses.

an upper limit to the frequency (or equivalently, a lower limit on wavelength) which can be uniquely represented. This limit is reached when the displacements exhibit an alternating up-down-up-down pattern. Such a limit will exist regardless of the number of masses chosen, but of course at differing wavelengths for differing numbers of masses, assuming the total length of the chain is the same in all cases. The longer wavelengths, however are easily represented as the number of masses increases. Figure 4.9 shows the two lowest modes ($j = 1$ and $j = 2$) for nine, ten, and eleven masses. By comparing figure 4.9 with figures 4.7 and 4.8, it is clear that as the number of masses increase, the behaviour approaches that of a continuous string for long wavelength modes. That is, the discrete nature of the model does not appreciably affect the motion.

5. A Stretched String

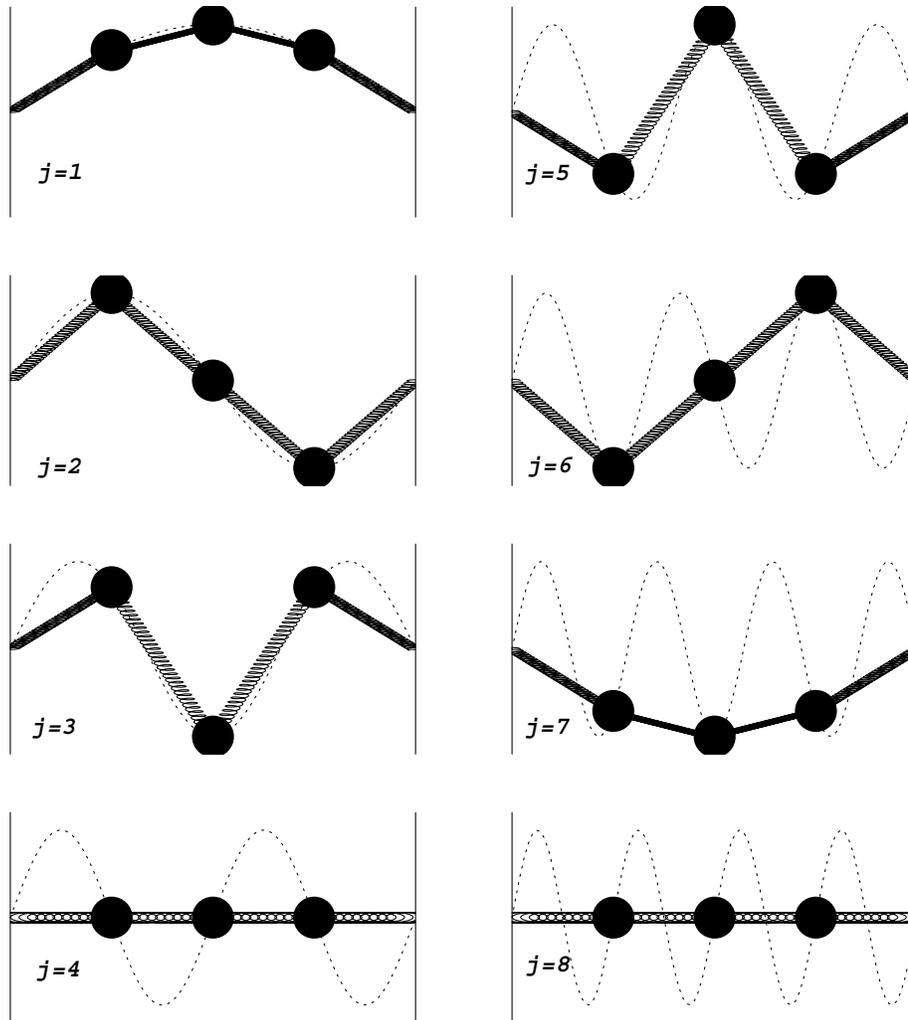


Figure 4.8 Three masses have three distinct modes of vibration. Shown here are initial positions based on eqn. (4.13) The broken line is the number of half sine waves fitting within the length of the system upon which eqn. (4.13) is based. Note that as the mode number j is increased beyond three, no new modes of vibration exist. Mode $j=4$ is the null mode, as is $j=8$. Mode $j=5$ is the same as mode $j=3$, except for a trivial factor of minus one, as are modes $j=6,2$ and $j=7,1$. Note that the only meaningful features of the sinusoidal curves are the values at the positions occupied by the masses.

a. Theory

Now consider a stretched string. A “string” will be any flexible length of material whose diameter is insignificant when compared to its length. It will be assumed to be clamped at each end and “stretched” by forces pulling on the ends. This supplies the tension in the string. It is assumed that enough tension is applied to prevent any significant sagging, so that gravity may be neglected. It should be noted that in a string at equilibrium (that is, at rest), the tension must be everywhere directed along the string. If not, transverse motions would arise. Furthermore, the magnitude of the tension must also be everywhere identical, even at a juncture between two different materials. To verify this, consider such a juncture (point A) as shown in figure 5.1. If the magnitude of the tension on the left, T_l , differed from that on the right, T_r , a net horizontal force would exist at point A, resulting in horizontal motion, which would violate the assumption

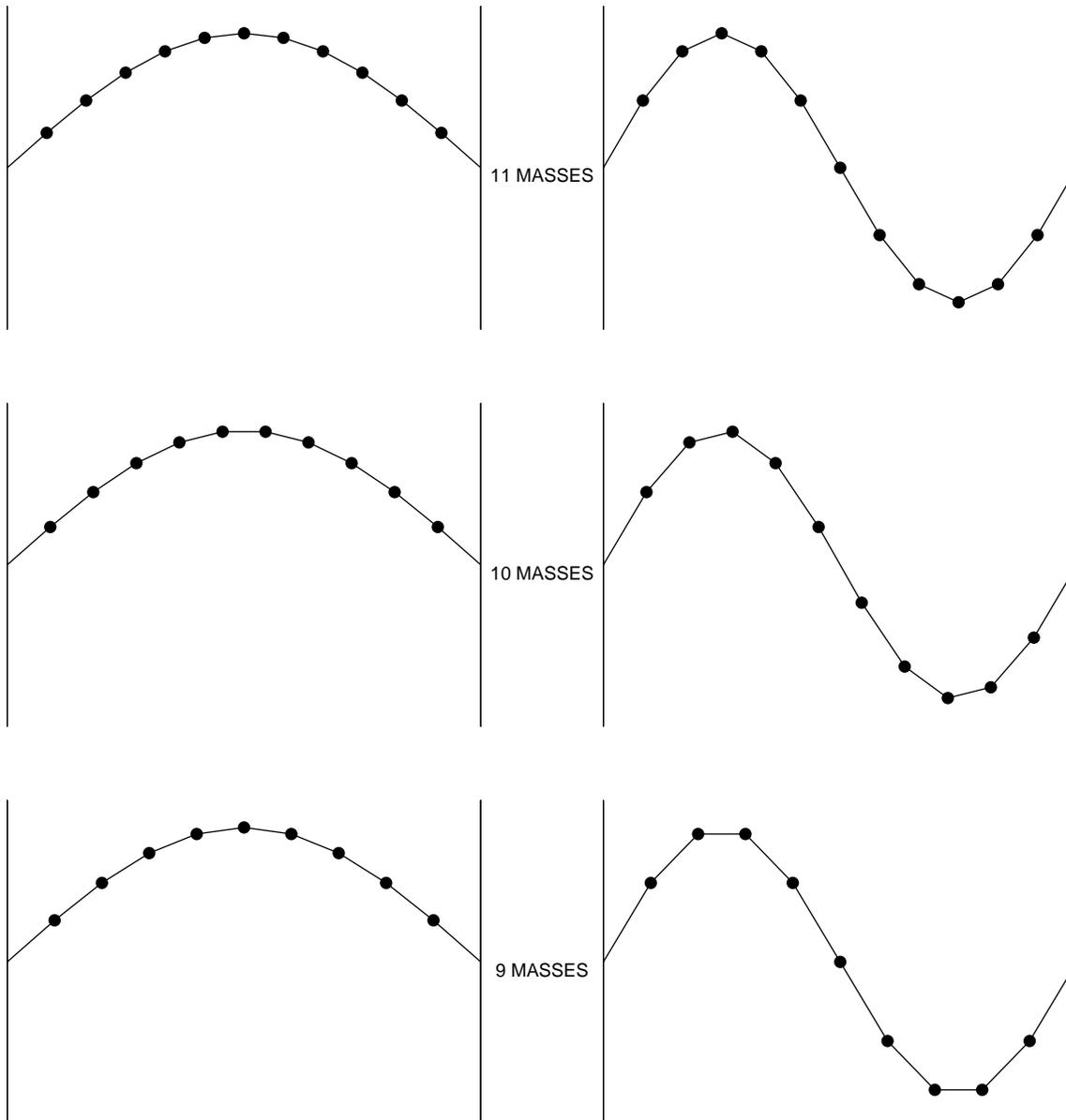


Figure 4.9 Displayed here are the lowest two modes ($j=1$ on the left and $j=2$ on the right) for strings of 9, 10, and 11 masses. Long wavelength vibratory modes are easily represented by even a relatively small number of discrete masses. The discrete nature of the model becomes apparent for modes whose wavelengths are comparable to the spacing of the masses.

that the string was at equilibrium.

For sufficiently long samples, almost any solid substance could be made into a string, thus its sole defining parameter is its *linear density*. Because some readers may not be comfortable with the concept of linear density, a little further explanation may be in order. This terminology comes from the most general definition of density, that is

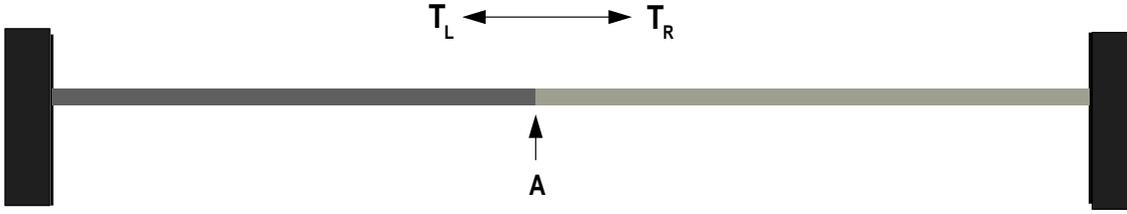


Figure 5.1 At equilibrium, the tension within a string must be uniform, even when the string is a composite. Here a string is composed of two different materials joined at point A. If the magnitude of the tension on the left, T_L , differed from that on the right, T_R , a net horizontal force would exist at point A, resulting in horizontal motion, violating the assumption that the string was at equilibrium.

$$\text{density} = \frac{\text{mass}}{\text{volume}}, \tag{5.1}$$

where volume is determined by the dimensions of the problem. In three dimensions, the volume element, or unit volume, is the unit of length, cubed. Examples are: cubic feet, cubic centimeters, and cubic meters. Thus, in three dimensions, density, or *volume density*, enjoys such familiar units as slugs[†] per cubic foot, grams per cubic centimeter, kilograms per cubic meter, etc.

However, when solving some specific problem, any special geometrical properties of that problem are used to simplify its solution. This simplification often includes reducing the dimensions to two (say for a flat plate, or motion confined to a table top) or to one (as in this case of a string). In two dimensions, the volume element is length squared, producing *surface densities* such as slugs per square foot (at the Earth's surface, force is often used in the English system to produce pounds per square foot), kilograms per square meter, grams per square centimeter. In one dimension density becomes a linear density, examples of which are slugs per foot (or pounds per foot if you really insist), kilograms per meter, or grams per centimeter. Thus eqn (5.1) is really

$$\text{density} = \frac{\text{mass}}{\text{length}^{\text{(number of dimensions)}}}, \tag{5.2}$$

There may be also some confusion with the concept of density and the physical diameter of the wire. It is obvious that when given two wires of the same composition, the wire with the greater linear density will have a larger diameter, but this is not necessarily the case when the composition of the two wires differ, as for example, steel and plastic. It may prove useful to compare the resulting diameters of wires made from various materials with the same linear densities (but differing volume densities). For example two steel wires whose linear densities differ by a factor of two, or a steel wire and a plastic rope.) Next consider the forces on a short segment or *element* of the string as in figure 5.2. The mass of the string element is

$$m_i = \lambda \Delta l \approx \lambda \Delta x_i \tag{5.3}$$

The directions of the forces at each end of this element are tangent to the string at these points. In general these directions are not exactly parallel, giving rise to a net vertical force

$$T_s (\sin(\theta_l) - \sin(\theta_r)). \tag{5.4}$$

Thus if the string is considered to be constructed of a large number of small segments, the tangents may be approximated by drawing lines connecting the centers of adjacent segments, and if each segment is considered to be a mass $\lambda \Delta x$ concentrated at the center of that segment connected to each other by springs, the model from the previous section may be used. Substituting eqn. (5.3) for m_i in eqn. (4.4) produces

$$a_i = \frac{F_s}{\lambda} \frac{1}{\Delta x} \left(\frac{y_{i+1} - y_i}{(\Delta x)} - \frac{y_i - y_{i-1}}{(\Delta x)} \right). \tag{5.5}$$

[†] If you do not know what a slug is, then shame on you.

That a string could be modeled in this manner should not really be surprising. At the molecular level, this is very close to what the problem really is: a set of mass concentrations (atoms or molecules) held together by restoring forces (electrostatic forces), but at a much smaller scale than can be realistically computed.

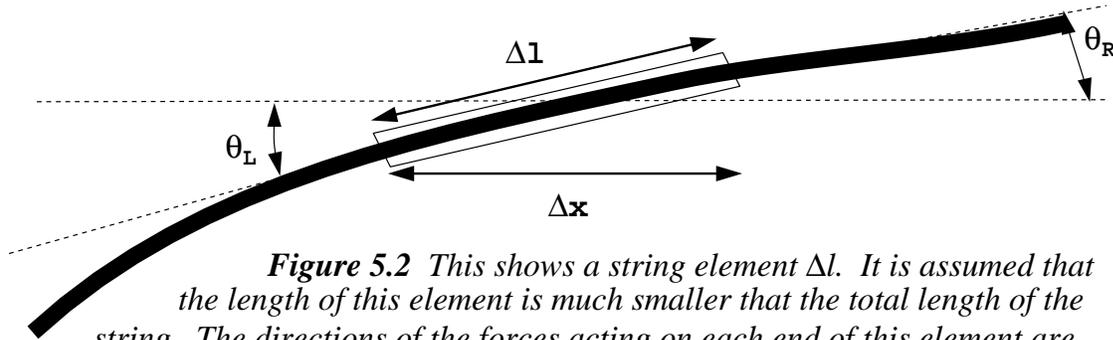


Figure 5.2 This shows a string element Δl . It is assumed that the length of this element is much smaller than the total length of the string. The directions of the forces acting on each end of this element are tangent to the string at these points. In general these directions are not exactly parallel, making an angle θ_l with respect to the horizontal at the left edge of the element and an angle θ_r at the right edge. Both of these angles are small, so that $\Delta l \cong \Delta x$. The vertical scale is magnified in this illustration.

Readers familiar with the calculus will recognize that $\frac{1}{\Delta x} \left(\frac{y_{i+1} - y_i}{(\Delta x)} - \frac{y_i - y_{i-1}}{(\Delta x)} \right)$ from the right-hand side of eqn. (5.5) is really the second derivative of y with respect to x , in the limit that $\Delta x \rightarrow 0$. Using the definition of acceleration:

$$a = \frac{\partial^2 y}{\partial t^2}, \tag{5.6}$$

eqn. (5.5) becomes

$$\frac{\partial^2 y}{\partial t^2} = \frac{F_s}{\lambda} \frac{\partial^2 y}{\partial x^2}. \tag{5.7}$$

This is known as the *wave equation*, a dynamical relationship appearing in many seemingly unrelated circumstances, from waves on strings to light. For light, the wave equation reads

$$\frac{\partial^2 y}{\partial t^2} = \frac{c^2}{\eta^2} \frac{\partial^2 y}{\partial x^2}, \tag{5.8}$$

where η is the index of refraction of the medium through which the light is traveling, and c is the speed of light in vacuum. Therefore, the same numerical algorithm developed to simulate waves in strings will simulate plane light waves if the identification

$$\frac{c}{\eta} = \sqrt{\frac{F_s}{\lambda}} \tag{5.9}$$

is made.

b. Dissipation

There is one addition necessary to complete this numerical string model. After some consideration of the figures in section 4 above, it is apparent that there are allowed some very “unstring-like” modes of vibration: particularly a very high frequency (short wavelength) up-down motion shown in figure 5.3. This is known as a *grid oscillation* since it occurs at the highest resolution of the numerical grid. At the atomic level, the atoms do vibrate in modes such as this, but not in a manner which is coordinated across the entire diameter of the string. Such small-scale motions are eventually dissipated as heat. In the computational model as it now stands, no such dissipation exists, and any such short wavelength (high frequency) vibrations are allowed to remain, which would violate the assumption that the string model was smooth. The model for a smooth string must therefore not allow any vibrations to occur with wavelengths comparable to

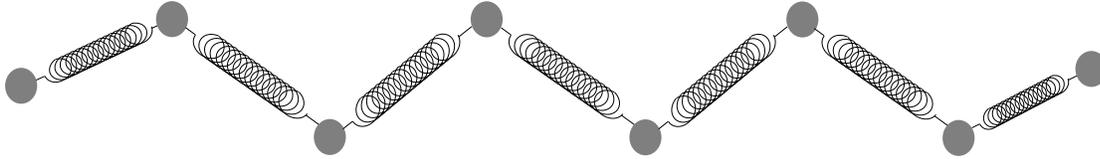


Figure 5.3 Here is an example an allowed up–down motion or grid oscillation. This is the highest frequency (shortest wavelength) allowed on the numerical grid. In actual objects, such vibrations would occur at the molecular scale and would be dissipated as heat. Because the computational resolution cannot extend to the molecular level, additional numerical dissipation must be added to avoid grid oscillations.

the size of the string element or discretization. Therefore dissipation is added to the computation to damp these unwanted vibratory modes.

To add dissipation to the model one must first realize that not only are the positions oscillating in this up–down manner, but the velocities must be oscillating as well to produce such displacements. Thus if a given mass, m_i is in a region undergoing such an oscillation, its velocity will be different than its nearest neighbors. However, many modes are allowable, as was demonstrated in the last section. In the very long wavelength modes, which are well described by this string model, the velocity of m_i is very close to the average of its neighbors. This suggests that the long wavelength signals could be preserved while damping the short wavelength oscillations if the velocities used in eqn (4.6b), $v_i^{n+1/2}$, are replaced by a blend of that velocity and a velocity found from some interpolation using its neighboring masses:

$$v_i' = (1 - f)v_i^{n+1/2} + f \bar{v}_i, \tag{5.10}$$

where f is the fractional weighting, $0 \leq f \leq 1$, and \bar{v}_i , is the interpolated velocity.

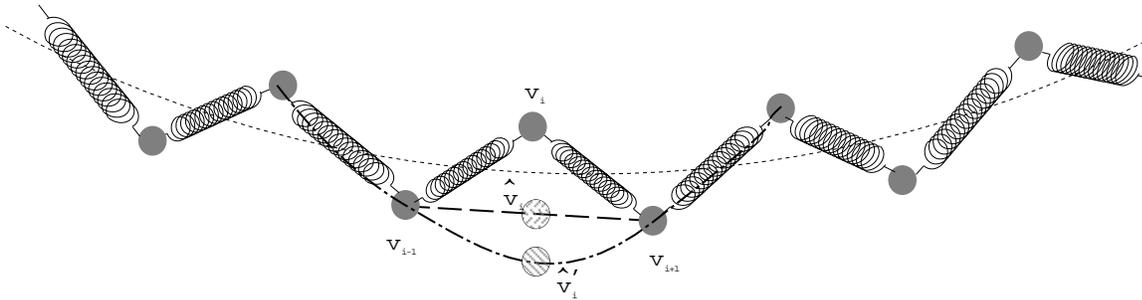


Figure 5.4 Here are shown velocities in part of a string of masses for which the motion has two components. One is a long wavelength component, shown as a dotted line. The other component is the grid oscillation from figure 5.3. Two methods are offered for damping the short wavelength component while leaving the long wavelength component relatively unscathed. One procedure blends the velocity at each point with the value obtained by a simple average of the two nearest neighbors (dashed line). This is method 3 in the LCSE lab. The other procedure blends the velocity at each point with a values obtained by fitting a curve to the four neared neighbors (dot–dash line). This is offered as method 4 in the LCSE lab.

This interpolated velocity, \bar{v}_i , may be determined by a number of procedures. Two are offered in the LCSE string laboratory, figure 5.4. First, a simple average (straight-line interpolation) of the two nearest neighbors:

$$\hat{v} = \frac{v_{i-1} + v_{i+1}}{2}. \tag{5.11}$$

This is incorporated into method 3 in the LCSE string laboratory. The second uses a smooth curve, here a cubic polynomial,

$$g(x) = A + Bx + Cx^2 + Dx^3, \tag{5.12}$$

which passes through the four nearest points, and evaluating that curve at $x = x_i$ to obtain \hat{v}'_i as shown in figure 5.4. This is incorporated into method 4 in the LCSE string laboratory.

For those who really must know, \hat{v}'_i may be found in the following procedure. The point masses are uniformly spaced, so the symmetry of the problem suggests using weighted averages of the neighboring point masses taken in symmetrical pairs:

$$\bar{v}_i = \alpha(v_{i-1}^{n+1/2} + v_{i+1}^{n+1/2}) + \beta(v_{i-2}^{n+1/2} + v_{i+2}^{n+1/2}). \quad (5.13)$$

Some tedious algebra will reveal that $\alpha = 2/3$ and $\beta = -1/6$. This may be verified in the following manner. If eqn (5.13) holds for the special cases: $v_i = x_i$, $v_i = x_i^2$, and $v_i = x_i^3$, then it must work for all linear combinations of them, and hence for all polynomials of the desired form, eqn (5.12). Choose x_i as the central mass, setting $x_i = 0$. Using a spacing Δx of 1 gives

$$x_{i\pm 1} = \pm 1 \quad (5.14a)$$

and

$$x_{i\pm 2} = \pm 2. \quad (5.14b)$$

Grouping the terms as in eqn (5.13) using only two coefficients guarantees that the formula will work for the cases $v_i = x_i$ and $v_i = x_i^3$. The other two cases imply that $A + B = 1/2$ and $A + 4B = 0$. Now, aren't you glad you asked?

The maximum timestep for each iteration can be estimated by looking at the behavior of the grid oscillations. This is a short cut to a result that can be derived more rigorously (and tediously) in general through a more complex (and definitely tedious) analysis than that which confronts the poor reader here. Often when numerical algorithms such as this one fail, the failure usually appears first in the highest frequency mode possible, the grid oscillation mode. Writing c^2 for $\frac{F_s}{\lambda}$ in eqn (5.5), consider a grid oscillation using eqns (4.6) with initial displacement values

$$y_i = (-1)^i y_0 \quad (5.15a)$$

and initial velocities

$$v_i^{-1/2} = (-1)^i 2y_0 \left(\frac{c}{\Delta x} \right)^2 \Delta t. \quad (5.15b)$$

This produces accelerations (from eqn 5.5) of

$$a_i = (-1)^{i+1} 4y_0 \left(\frac{c}{\Delta x} \right)^2. \quad (5.16)$$

It will prove convenient to define a quantity, the *Courant number*, as

$$\sigma = \frac{c \Delta t}{\Delta x}. \quad (5.17)$$

The Courant number, or stability parameter, is named after R. Courant, a mathematician who contributed greatly to this field.

It will be illustrative to determine what happens when the Courant number exceeds unity, using these initial conditions. The first iteration of eqns (5.5) and (4.6) produces

$$a_i = (-1)^i 4y_0 \left(\frac{\sigma}{\Delta t} \right)^2, \quad (5.18a)$$

$$v_i^{1/2} = (-1)^{i+1} 2\sigma^2 \frac{y_0}{\Delta t}, \quad (5.18b)$$

and

$$y_i^1 = (-1)^i y_0 (1 - 2\sigma^2). \quad (5.18c)$$

It is clear that even after one timestep the amplitude of the grid oscillation increases when σ exceeds unity. To verify that this oscillation continues to grow, it is sufficient to perform only one more iteration, using

$$\sigma = 1 + \varepsilon, \quad (5.19)$$

where ε is a small positive value. Performing an additional iteration of the algorithm produces

$$a_i^2 = (-1)^{i+1} 4y_o(1 + 4\varepsilon) \left(\frac{\sigma}{\Delta t} \right)^2 = -(1 + 4\varepsilon)a_i^1, \quad (5.20a)$$

$$v_i^2 = (-1)^{i+1} \sigma^2 \frac{y_0}{\Delta t} [-2 + 4(1 + 4\varepsilon)] = -(1 + 8\varepsilon)v_i^1, \quad (5.20b)$$

and

$$y_i^2 = (-1)^{i+1} y_0^1 [-1 - 4\varepsilon + (1 + 2\varepsilon)(2 + 16\varepsilon)] = (1 + 16\varepsilon)y_i^1, \quad (5.20c)$$

where all terms involving ε raised to powers of two or greater have been dropped. (Recall that it was assumed that ε was very small, so ε^2 is smaller still.) Clearly, the algorithm blows up whenever the Courant number exceeds unity. This stability limit for difference schemes approximating the wave equation was discovered by Courant, and hence bears his name. Thus the limit on the timestep is

$$\Delta t \leq \frac{\Delta x}{c}. \quad (5.21)$$

The blending factor f in eqn (5.10) may now be determined. This factor will be set equal to the Courant number times that value of f to be used when the Courant number is unity. Considering the grid oscillation example above, it is clear that a value of $\frac{1}{2}$ for f will bring all velocities to zero in a single timestep when the Courant number is unity. Such a choice would provide very strong damping for the highest frequency modes, but a softer touch, such as

$$f = \frac{\sigma}{4} \quad (5.22)$$

will suffice.