February 19, 2011

Eigenfunction expansions: a simple-minded introduction

Supplement to section 1.8 of *Advanced Mathematics for Applications*

by Andrea Prosperetti

Here we present an expanded version of the material contained in sections 1.8 and 2.1 of the text *Advanced Mathematics for Applications*. The purpose is to further clarify the geometric point of view adopted in the book in which "unit vectors" (eigenfunctions) and scalar products play a dominant role. The treatment purposely attempts to help the reader form an intuitive understanding of these very powerful and useful concepts omitting a great deal of mathematical detail.

1 The taut string

As mentioned at the end of section 1.6 of the book, the equation governing the small-amplitude vibrations of a taut string is

$$\mu\frac{\partial^2 u}{\partial t^2} = T\frac{\partial^2 u}{\partial^2 x}\,, \tag{1.1}$$

in which u is the displacement of the string normal to the equilibrium configuration, i.e., the x-axis, μ is the mass per unit length and T is the tension in the string (figure 1). A simple (if far from rigorous) derivation of this equation can be given as follows.

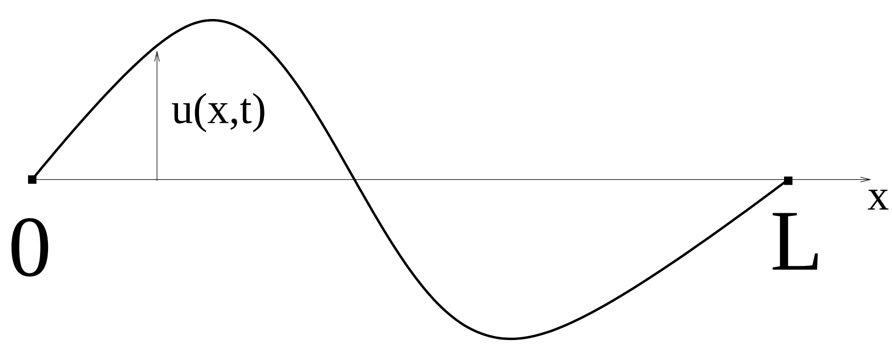

Figure 1: A taut string pinned at the end points

Let us consider a string which extends for $0 \le x \le L$ and focus on a small piece of it of length Δx. The mass of this piece is $\Delta m = \mu\Delta x$, its acceleration normal to the x-axis is $\partial^2 u/\partial t^2$, and the force acting at $x + \Delta x$ in the direction normal to the x-axis, assuming that the string slope is

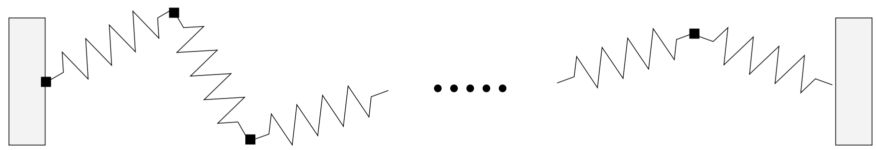

Figure 2: An oscillator chain pinned at the end points

small, is $T \sin\theta \simeq T \tan\theta = T\partial u/\partial x$. So the net force on the short string segment between x and $x + \Delta x$ is

$$T \left.\frac{\partial u}{\partial x}\right|_{x+\Delta x} - T \left.\frac{\partial u}{\partial x}\right|_{x} = \Delta x\, T \frac{\partial^2 u}{\partial x^2}\,. \tag{1.2}$$

The limit of this relation for $\Delta x \to 0$ gives (1.1), which is the wave equation with a speed of propagation

$$c = \sqrt{\frac{T}{\mu}}\,. \tag{1.3}$$

A full mathematical specification of the problem requires also boundary and initial conditions. If, for example, the string is pinned at the end points, then the solution of (1.1) must satisfy

$$u(0,t) = 0\,, \qquad u(L,t) = 0\,. \tag{1.4}$$

As for any mechanical problem, the initial conditions must specify the initial position and initial velocity of the degrees of freedom of the system and, i.e., the points of the string. Thus, they will have the form

$$u(x,0) = u^0(x)\,, \qquad \frac{\partial u}{\partial t}(x,0) = v^0(x)\,. \tag{1.5}$$

Suppose one wanted to solve this numerically. The first step would be to discretize the spatial derivative, i.e., to avoid taking the limit $\Delta x \to 0$ in (1.2) and to replace the first derivatives by finite differences:

$$\begin{aligned}
\left.\frac{\partial^2 u}{\partial x^2}\right|_{x_i} &= \frac{1}{\Delta x}\left(\left.\frac{\partial u}{\partial x}\right|_{i+1/2} - \left.\frac{\partial u}{\partial x}\right|_{i-1/2}\right) \\
&= \frac{1}{\Delta x}\left\{\frac{1}{\Delta x}\,[u(x_{i+1},t) - u(x_i,t)] - \frac{1}{\Delta x}\,[u(x_i,t) - u(x_{i-1},t)]\right\} \\
&= \frac{u(x_{i+1},t) - 2u(x_i,t) + u(x_{i-1},t)}{\Delta x^2}\,.
\end{aligned} \tag{1.6}$$

For simplicity of writing let $u(x_i,t) = u_i(t)$. Then

$$\mu \frac{\partial^2 u_i}{\partial t^2} = \frac{T}{\Delta x^2}(u_{i+1} - 2u_i + u_{i-1})\,, \tag{1.7}$$

or

$$\Delta m \, \frac{\partial^2 u_i}{\partial t^2} \;=\; \frac{T}{\Delta x}\,(u_{i+1} - 2u_i + u_{i-1})\,, \tag{1.8}$$

where $\Delta m = \mu \Delta x$ is the mass of the little piece of string Δx and $T/\Delta x$ is the tension per unit length.

We shall now show that the same equation governs what appears to be, superficially, a very different physical system.

2 A discrete approximation: Oscillator chain

Consider a chain of oscillators of equal mass m connected by equal linear springs of elastic constant K; to simplify, we assume that the equilibrium length of each spring is 0 (figure 2). When the chain is at rest the distance between the masses is Δx and therefore, in the equilibrium state, each spring exerts a force $T = K\,\Delta x$ which is analogous to the tension in the string. Let us now derive an equation for the motion of this system in the direction normal to its equilibrium configuration.

The force in the normal direction exerted on the j-th mass due to the spring at its right is (figure 3)

$$K l \sin\theta \;=\; K l \, \frac{u_{j+1} - u_j}{l} \;=\; K\,(u_{j+1} - u_j) \tag{2.1}$$

and similarly for the spring on the left. Hence the equation of motion of the j-th mass in the

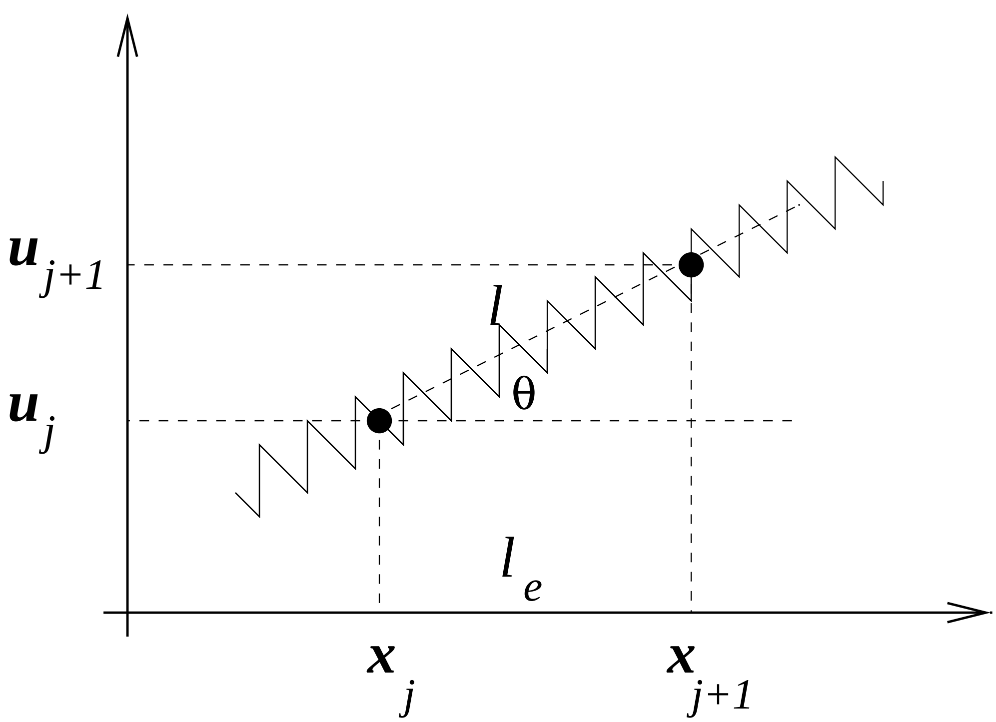

Figure 3: The spring connecting masses j and $j+1$ of the oscillator chain of figure 2.

direction normal to the equilibrium configuration is

$$
\begin{aligned}
m\frac{d^2u_j}{dt^2} &= K\left(u_{j+1}-u_j\right)-K\left(u_j-u_{j-1}\right) \\
&= K(u_{j+1}-2u_j+u_{j-1})\,. \qquad (2.2)
\end{aligned}
$$

If we write $m=\mu\Delta x$, $K\Delta x=T$,[1] this equation becomes identical to the discretized wave equation (1.8). Since the left end of the first string does not move, the general equation of motion (2.2) may be applied to the first oscillator of the chain, $j=1$, by simply setting formally $u_0=0$ to find:

$$
m\frac{d^2u_1}{dt^2} = K(u_2-2u_1)\,. \qquad (2.3)
$$

Similarly, for the last oscillator, $j=N$, we set $u_{N+1}=0$ and we have.

$$
m\frac{d^2u_N}{dt^2} = K(-2u_N+u_{N-1})\,. \qquad (2.4)
$$

The conditions $u_0=0$, $u_{N+1}=0$ are similar to (1.4) for the string pinned at the two ends.[2]

In order to introduce the central idea of the method of eigenfunction expansions, we now show that not only there is a parallelism between the mathematical formulation of the problems for the taut string and for the oscillator chain, but also that there is an *exact correspondence* between the methods of solution.

With the natural frequency of a single mass-spring system defined by

$$
\omega^2 = \frac{K}{m} \qquad (2.5)
$$

we can rewrite the system of equations (2.2), (2.3), (2.4) as

$$
\frac{d^2}{dt^2}\left|\begin{array}{c} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_{N-1} \\ u_N \end{array}\right| + \omega^2 \left|\begin{array}{ccccccc} 2 & -1 & 0 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 2 & -1 & 0 & \dots & 0 \\ & & \dots & & & & \\ 0 & & \dots & -1 & 2 & -1 & \\ 0 & & \dots & & -1 & 2 & -1 \\ 0 & & \dots & & & -1 & 2 \end{array}\right| \left|\begin{array}{c} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_{N-1} \\ u_N \end{array}\right| = 0\,. \qquad (2.6)
$$

Written in this form, this equations suggests to introduce the matrix

$$
\mathsf{M} = \left|\begin{array}{ccccccc} 2 & -1 & 0 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 2 & -1 & 0 & \dots & 0 \\ & & \dots & & & & \\ 0 & & \dots & -1 & 2 & -1 & \\ 0 & & \dots & & -1 & 2 & -1 \\ 0 & & \dots & & & -1 & 2 \end{array}\right|. \qquad (2.7)
$$

[1] Since the springs have been taken to have zero equilibrium length, each one of them exerts a force equal to the separation of the bodies Δx times the elastic constant K.

[2] One may think of a chain consisting of $N+2$ oscillators, the first and last of which are constrained to have a zero displacement, so that $u_0=0$, $u_{N+1}=0$. The dynamics of such a system can therefore be described in a N-dimensional subspace of the $N+2$ dimensional space, and Eq. (2.2) is applicable to all the bodies that can move, $1\le j\le N$.

and the vector

$$\mathbf{u}(t) = \begin{vmatrix} u_1(t) \\ u_2(t) \\ \cdot \\ \cdot \\ \cdot \\ u_{N-1}(t) \\ u_N(t) \end{vmatrix} = u_1(t) \begin{vmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 0 \end{vmatrix} + u_2(t) \begin{vmatrix} 0 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 0 \end{vmatrix} + \ldots + u_N(t) \begin{vmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \end{vmatrix} = \sum_{k=1}^{N} u_k(t)\, \mathbf{i}_k\,, \tag{2.8}$$

where all the elements of each $\mathbf{i}_k$ vanish except the k-th one, which equals 1. Evidently, this set of unit vectors constitutes a basis in the N-dimensional space to which $\mathbf{u}(t)$ belongs, in the same sense in which the unit vectors along the x-, y-, and z-directions are a basis in the usual three-dimensional space. Each component u_k of $\mathbf{u}$ is the projection of $\mathbf{u}$ along one of these basis vectors and we may write

$$u_k = \mathbf{i}_k \cdot \mathbf{u} \tag{2.9}$$

or, with a different notation which will become useful later,

$$u_k = (\mathbf{i}_k, \mathbf{u})\,. \tag{2.10}$$

With these definitions, (2.6) becomes

$$\frac{d^2\mathbf{u}}{dt^2} + \omega^2 \mathsf{M}\, \mathbf{u} = 0\,, \tag{2.11}$$

with initial conditions

$$\mathbf{u}(0) = \mathbf{u}^0\,, \qquad \dot{\mathbf{u}}(0) = \mathbf{v}^0\,. \tag{2.12}$$

To solve this problem we expand the solution $\mathbf{u}$ onto a basis of eigenvectors of M:

$$\mathbf{u} = \sum_{k=1}^{N} U_k(t)\, \mathbf{e}^{(k)}\,, \tag{2.13}$$

where

$$\mathsf{M}\, \mathbf{e}^{(\mathsf{k})} = \lambda_{\mathsf{k}}\, \mathbf{e}^{(\mathsf{k})}\,. \tag{2.14}$$

It is evident that, if we know the U_k's and the $\mathbf{e}^{(k)}$'s, we immediately recover the components of $\mathbf{u}$ in the original representation (2.8) by taking the scalar product of (2.13) with $\mathbf{i}_\ell$:

$$u_\ell(t) = \sum_{k=1}^{N} U_k(t) \left(\mathbf{i}_\ell, \mathbf{e}^{(k)}\right) = \sum_{k=1}^{N} U_k(t)\, e_\ell^{(k)} \qquad \ell = 1, 2, \ldots, N\,, \tag{2.15}$$

in which $e_\ell^{(k)} = \left(\mathbf{i}_\ell, \mathbf{e}^{(k)}\right)$ is the component of $\mathbf{e}^{(k)}$ in the direction of $\mathbf{i}_\ell$.

After substitution of (2.13) into (2.11), by (2.14), we find

$$\sum_{k=1}^{N} \left(\ddot{U}_k + \omega^2 \lambda_k U_k\right) \mathbf{e}^{(k)} = 0\,. \tag{2.16}$$

This is a vector equation from which we generate N scalar equations in the usual way by taking the scalar product in turn with each one of the $\mathbf{v}^{(n)}$'s:

$$\left(\mathbf{e}^{(n)}, \sum_{k=1}^{N} \left(\ddot{U}_k + \omega^2 \lambda_k U_k\right) \mathbf{e}^{(k)}\right) = 0 \qquad n = 1, 2, \ldots, N\,. \tag{2.17}$$

or, since the scalar procudt of a sum is the sum of the scalar products,

$$\sum_{k=1}^{N} \left(\ddot{U}_k + \omega^2 \lambda_k U_k\right) \left(\mathbf{e}^{(n)}, \mathbf{e}^{(k)}\right) = 0 \qquad n = 1, 2, \ldots, N\,. \tag{2.18}$$

We will see later that all the scalar products in this sum vanish except the one for which $n = k$, i.e. that $\left(\mathbf{e}^{(n)}, \mathbf{e}^{(k)}\right) \propto \delta_{kn}$. Thus, upon taking, in turn, $n = 1,\ 2\ \ldots,\ N$, we find from (2.18)[3]

$$\ddot{U}_n + \omega^2 \lambda_n U_n = 0\,, \qquad n = 1, 2, \ldots, N\,, \tag{2.19}$$

the solution of which is of course

$$U_n(t) = C_n \cos \omega\sqrt{\lambda_n}\, t + D_n \sin \omega\sqrt{\lambda_n}\, t\,, \tag{2.20}$$

with the constants C_n, D_n determined from the initial conditions (2.12).

The essential step is therefore to solve the eigenvalue problem (2.14). From the algebraic point of view, this is a linear homogeneous system that possesses non-trivial solutions if and only if the following solvability condition is satisfied:

$$D_N \equiv \det\left(\mathsf{M} - \lambda_\mathsf{k} \mathsf{I}\right) = \mathsf{0}\,, \tag{2.21}$$

where I is the identity matrix. After determining the λ's in this way, one would go back to the system (2.14) and find the eigenvectors. As we show later, this procedure does indeed work. However, having in mind a parallelism with what we will be doing later for the string problem (1.1), we proceed slightly differently, namely (i) we first attempt to solve the algebraic system (2.14) for general λ; (ii) the condition that the first and last equations of the system (2.3) and (2.4) be satisfied will then give an equation which determines λ.

2.1 First step: Eigenvalues

Written out in detail, the eigenvalue problem (2.14) is

$$\left\| \begin{array}{ccccccc} 2-\lambda_k & -1 & 0 & 0 & 0 & \ldots & 0 \\ -1 & 2-\lambda_k & -1 & 0 & 0 & \ldots & 0 \\ 0 & -1 & 2-\lambda_k & -1 & 0 & \ldots & 0 \\ & & \ldots & & & & \\ 0 & & \ldots & -1 & 2-\lambda_k & -1 & \\ 0 & & \ldots & & -1 & 2-\lambda_k & -1 \\ 0 & & \ldots & & & -1 & 2-\lambda_k \end{array} \right\| \left\| \begin{array}{c} e_1^{(k)} \\ e_2^{(k)} \\ \ldots \\ \ldots \\ \ldots \\ e_{N-1}^{(k)} \\ e_N^{(k)} \end{array} \right\| = 0\,, \tag{2.22}$$

[3] More generally, we would find the same result simply provided that all the eigenvectors are linearly independent.

where, in the symbol $e_j^{(k)}$, the upper index k matches the index of λ_k and denotes the particular (eigenvalue, eigenvector) pair that we are considering while the lower index labels the components of this eigenvector.

The generic equation of this system is:

$$-e_{j+1}^{(k)} + (2-\lambda_k)e_j^{(k)} - e_{j-1}^{(k)} \;=\; 0\,. \tag{2.23}$$

This is a *difference equation* determining the dependence of $e_j^{(k)}$ on its variable j.[4] Since this relation links $e_{j+1}^{(k)}$ to the two preceding values $e_j^{(k)}$, $e_{j-1}^{(k)}$, it is classified as a difference equation of the *second order*. Furthermore, since the coefficients do not depend on j, it is a difference equation of the second order with constant coefficients. Just as ordinary differential equations with constant coefficients, such equations have exponential solutions.[5] We assume therefore

$$e_j^{(k)} \propto \exp\left(j\nu_k\right), \tag{2.24}$$

and substitute into (2.23).[6] The result is

$$2\cosh\nu_k \;=\; 2-\lambda_k\,, \tag{2.25}$$

and therefore

$$\nu_k \;=\; \cosh^{-1}\frac{2-\lambda_k}{2}\,, \tag{2.26}$$

Since the hyperbolic cosine is even, $-\nu_k$ also satisfies (2.25). Thus, just as in the case of ordinary differential equations, we write down the general solution of (2.23) by superposing the two possible values of ν_k

$$e_j^{(k)} \;=\; \alpha_k\exp\left(\nu_k j\right) + \beta_k\,\exp\left(-\nu_k j\right). \tag{2.27}$$

2.2 Second step: Eigenvectors

We call (2.27) the general solution of the difference equation (2.23) by analogy with the case of ordinary differential equations. The integration constants α_k, β_k will now be determined by imposing suitable "boundary conditions". These conditions are found by writing out explicitly the first and the last equations of the system (2.22):

$$-e_2^{(k)} + (2-\lambda_k)e_1^{(k)} \;=\; 0\,, \qquad (2-\lambda_k)e_N^{(k)} - e_{N-1}^{(k)} \;=\; 0\,. \tag{2.28}$$

As noted earlier in connection with (2.3) and (2.4), if we were to introduce two fictitious points $e_0^{(k)}$ and $e_{N+1}^{(k)}$ at the beginning and end of the chain, these two equations would have the same general form (2.23) provided that $e_0^{(k)}$ and $e_{N+1}^{(k)}$ were defined to be zero. So, rather than substituting (2.27) into (2.28) (which of course is fine and leads to exactly the same results, but after messier algebra), we can directly impose $e_0^{(k)} = 0$ and $e_{N+1}^{(k)} = 0$ on the general solution (2.27). For $j = 0$ we find

$$\alpha_k + \beta_k \;=\; 0\,, \tag{2.29}$$

[4]As mentioned earlier the index k indicates the particular vector we are focusing on and remains therefore fixed.

[5]This rule for the solution of these equations is derive in section 5.3 of the book.

[6]In the case of an ordinary differential equation we would look for solutions in the form $e(x) \propto exp(\nu x)$. In the present instance, the analog of x is j while ν depends only on the function itself. Since the various e's are labelled by the index k, we affix the same index to ν.

while, for $j = N+1$,

$$\alpha_k \, \exp{(N+1)\nu} + \beta_k \, \exp{-(N+1)\nu} \; = \; 0 \, . \tag{2.30}$$

From the first one β_k = -α_k so that

$$e_j^{(k)} \; = \; 2\alpha_k \sinh{(\nu_k j)} \, . \tag{2.31}$$

Upon setting here $j = N+1$ and requiring the result to vanish as demanded by (2.30) we have

$$2\alpha_k \sinh{(N+1)}\, \nu_k \; = \; 0 \, . \tag{2.32}$$

Hence, either $\alpha_k = 0$, which would give $e_j^{(k)} = 0$ and therefore the trivial solution, or

$$\sinh{(N+1)}\, \nu_k \; = \; 0 \, , \tag{2.33}$$

from which

$$(N+1)\, \nu_k \; = \; i\pi \, k \, . \tag{2.34}$$

Upon redefining the as yet undetermined constant α_k we thus have

$$e_j^{(k)} \; = \; \alpha_k \, \sin{\left(\frac{\pi k}{N+1} \, j \right)} \, . \tag{2.35}$$

Upon recalling (2.25), (2.34) gives

$$\lambda_k \; = \; 2\,(1 - \cosh{\nu_k}) \; = \; 2 \left(1 - \cos{\frac{\pi \, k}{N+1}} \right) \; = \; 4 \, \sin^2{\frac{k\pi}{2(N+1)}} \, . \tag{2.36}$$

Note that $k = 0$ is not acceptable as, from (2.26), it would give $\nu = 0$ and therefore again the trivial solution. If we were to take $k = N+1$, we would find $\nu_{N+1} = i\pi$ and again all the $e_j^{(N+1)}$'s would vanish. Furthermore, for values of k greater than $N+1$, we would be repeating the same values of λ_k that we find by choosing $k = 1, \, 2, \, \ldots, \, N$. Hence we only have N distinct eigenvalues. A similar argument shows that we only have N distinct eigenvectors. Indeed, consider for instance, purely formally, $e_j^{(N+2)}$:

$$e_j^{(N+2)} = \alpha_{N+2} \, \sin{\frac{(N+1+1)\pi j}{N+1}} \; = \alpha_{N+2} \sin{\left(j\pi + \frac{j\pi}{N+1} \right)} \; = \alpha_{N+2} (-1)^j \sin{\frac{j\pi}{N+1}} \, . \tag{2.37}$$

On the other hand

$$\begin{aligned} e_j^{(N)} \quad &= \quad \alpha_N \sin{\frac{Nj\pi}{N+1}} \; = \alpha_N \sin{\frac{(N+1-1)j\pi}{N+1}} \; = \alpha_N \, \sin{\left(j\pi - \frac{j\pi}{N+1} \right)} \\ &= \quad \alpha_N \, (-\cos{j\pi}) \, \sin{\frac{j\pi}{N+1}} \; = \; -\alpha_N (-1)^j \sin{\frac{j\pi}{N+1}} \end{aligned} \tag{2.38}$$

from which we see that $\mathbf{e}^{(N+2)}$ would be proportional to $\mathbf{e}^{(N)}$, and therefore it would not be linearly independent. In a similar way we can convince ourselves that any other choice of k would result in vectors that we have already found by choosing $k = 1, \, 2, \, \ldots, N$.

Written out in detail, the components of the N distinct eigenvectors satisfying (2.14) are therefore

$$\mathbf{e}^{(k)} = \alpha_k \begin{vmatrix} \sin\left(\frac{\pi k}{N+1}\right) \\ \sin\left(2\,\frac{\pi k}{N+1}\right) \\ \sin\left(3\,\frac{\pi k}{N+1}\right) \\ \cdots \\ \sin\left((N-1)\,\frac{\pi k}{N+1}\right) \\ \sin\left(N\,\frac{\pi k}{N+1}\right) \end{vmatrix}. \tag{2.39}$$

The arbitrary constant α_k is present as expected, due to the homogeneity of the problem. Indeed, it is evident from Eq. (2.14) that, given one $\mathbf{e}^{(k)}$ which satisfies the equation, we are at liberty to multiply all its components by the same arbitrary non-zero constant and still have a solution of the equation. We shall choose a convenient value for α_k later by imposing a normalization condition. But before we do this, let us verify that the solvability condition (2.21) is indeed satisfied by our results for the λ_k's. For this purpose we first calculate D_N as a function of λ and then we verify that it vanishes when we replace λ by the values (2.36) just determined.

2.3 Another derivation of the eigenvalues

As noted before, a natural way to solve the eigenvalue problem (2.14) is to find the eigenvalues by setting the determinant of the system to 0. Since we have followed a slightly different route before, it is worth while to show that this alternative procedure leads to the same result.

We need to calculate D_N defined by (2.21), i.e.

$$D_N = \det \begin{vmatrix} 2-\lambda & -1 & 0 & 0 & 0 & \cdots & 0 \\ -1 & 2-\lambda & -1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 2-\lambda & -1 & 0 & \cdots & 0 \\ & & \cdots & & & & \\ 0 & & \cdots & -1 & 2-\lambda & -1 & \\ 0 & & \cdots & & -1 & 2-\lambda & -1 \\ 0 & & \cdots & & & -1 & 2-\lambda \end{vmatrix}. \tag{2.40}$$

We do this by deriving a recurrence relation. Begin by expanding the determinant by the first row. The result is

$$D_N = (2-\lambda)\,D_{N-1} - (-1)\det \begin{vmatrix} -1 & -1 & 0 & 0 & \cdots & 0 \\ 0 & 2-\lambda & -1 & 0 & \cdots & 0 \\ 0 & -1 & 2-\lambda & -1 & \cdots & 0 \\ & & \cdots & & & \\ 0 & \cdots & -1 & 2-\lambda & -1 & \\ 0 & \cdots & & -1 & 2-\lambda & -1 \\ 0 & \cdots & & & -1 & 2-\lambda \end{vmatrix} \tag{2.41}$$

The second determinant can be calculated by expanding by the first column, with which it is found to equal $(-1)\,D_{N-2}$. Thus we are led to

$$D_N - (2-\lambda)\,D_{N-1} + D_{N-2} \,=\, 0\,. \tag{2.42}$$

This is a difference equation identical to the one previously derived for the $e_j^{(k)}$'s, Eq. (2.23), and we now know how to solve it. Let

$$D_N \propto \exp\,\mu N\,, \tag{2.43}$$

and substitute into (2.42) to find

$$2\cosh\mu \,=\, 2-\lambda \qquad \Rightarrow \qquad \lambda \,=\, 2\,(1-\cosh\mu)\;. \tag{2.44}$$

We are seeking D_N as a function of λ. Since cosh is even, for each value of λ, this equation gives us two values of μ and therefore the solution for D_N has the form (cf. the previous Eq. 2.27 and the usual procedure in the case of ordinary differential equations with constant coefficients)

$$D_N \,=\, A\,\exp\,(\mu N) + B\exp\,(-\mu N)\,, \tag{2.45}$$

with A, B to be determined by the "initial conditions" on D_N, i.e. by the values of D_N for $N=1$ and 2. These can be calculated explicitly. For $N=1$

$$D_1 \,=\, \det\,|2-\lambda| \,=\, 2-\lambda\,. \tag{2.46}$$

Similarly, for $N=2$,

$$D_2 \,=\, \det\begin{vmatrix} 2-\lambda & -1 \\ -1 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - 1\,. \tag{2.47}$$

In principle, we could just take (2.45), evaluate it for $N=1$ and $N=2$, and equate the results to the explicit ones just found to evaluate A and B. In practice the algebra is a little simpler if we note that, from (2.42) written for $N=2$, we have, formally,

$$D_0 \,=\, (2-\lambda)\,D_1 - D_2 \,=\, 1\,, \tag{2.48}$$

where the last step follows by substituting the results (2.46), (2.47). Using this we find

$$D_0 \,=\, A+B \,=\, 1\,, \qquad D_1 \,=\, A\,\exp\,\mu + B\exp\,-\mu \,= 2-\lambda\,. \tag{2.49}$$

The solution of this system in A, B is

$$A \,=\, \frac{2-\lambda-\exp\,(-\mu)}{2\sinh\mu}\,, \qquad B \,=\, \frac{\lambda-2+\exp\,\mu}{2\sinh\mu}\,. \tag{2.50}$$

Upon substitution into (2.45), and use of (2.44) to eliminate λ, we find

$$D_N \,=\, \frac{\sinh\mu(N+1)}{2\sinh\mu}\,. \tag{2.51}$$

Upon substitution of the values (2.36) into (2.44) we find

$$\mu_k \,=\, \frac{i\pi\,k}{N+1}\,, \tag{2.52}$$

which, upon substitution into (2.51), indeed gives $D_N = 0$ as expected.

2.4 Third step: Normalization

We noted before that each $\mathbf{e}^{(k)}$ is determined up to an arbitrary non-zero multiplicative constant, which we have denoted by α_k in (2.39). We can fix this constant by imposing any convenient normalization. We will see that it is often useful – although by no means necessary – to render these vectors of unit length:

$$\left(\mathbf{e}^{(k)}, \mathbf{e}^{(k)}\right) = \sum_{j=1}^{N} \left[e_j^{(k)}\right]^2 = 1\,. \tag{2.53}$$

Upon substituting the explicit form (2.35) of $e_j^{(k)}$, we thus have[7]

$$\begin{aligned}
1 &= |\alpha_k|^2 \sum_{j=1}^{N} \left[\sin\left(\frac{\pi k}{N+1}\, j\right)\right]^2 = |\alpha_k|^2 \sum_{j=0}^{N} \left[\sin\left(\frac{\pi k}{N+1}\, j\right)\right]^2 \\
&= \frac{1}{2}|\alpha_k|^2 \sum_{0}^{N} \left(1 - \cos\frac{2k\pi j}{N+1}\right) \\
&= \frac{1}{2}|\alpha_k|^2 \left\{ N + 1 - \mathrm{Re} \sum_{j=0}^{N} \exp\left(\frac{2k\pi i}{N+1}\, j\right)\right] \\
&= \frac{1}{2}|\alpha_k|^2 \left[N + 1 - \mathrm{Re}\, \frac{1 - \exp(2k\pi i)}{1 - \exp i\,\frac{2k\pi}{N+1}}\right] \\
&= \frac{1}{2}|\alpha_k|^2 (N+1)
\end{aligned} \tag{2.54}$$

from which

$$\alpha_k = \sqrt{\frac{2}{N+1}}\,. \tag{2.55}$$

With this choice of α_k, the eigenvectors (2.39) are then

$$\mathbf{e}^{(k)} = \sqrt{\frac{2}{N+1}} \left| \begin{array}{c} \sin\frac{\pi k}{N+1} \\ \sin 2\,\frac{\pi k}{N+1} \\ \sin 3\,\frac{\pi k}{N+1} \\ \cdots \\ \sin\,(N-1)\,\frac{\pi k}{N+1} \\ \sin N\,\frac{\pi k}{N+1} \end{array} \right| \tag{2.56}$$

Upon setting $N = 1$ in this formula we find

$$\lambda_1 = 2\,, \qquad \mathbf{e}^{(1)} = 1\,, \tag{2.57}$$

which are the exact results for a single mass connected to two springs. Similarly, with $N = 2$, we find the results for the two-oscillator system which can be solved by elementary means.

[7] The way in which the summation is calculated is explained in Example 8.2.1 p. 225 of the book.

With a similar calculation one can easily show that the eigenvectors are orthogonal to each other, i.e.

$$\left(\mathbf{e}^{(k)}, \mathbf{e}^{(\ell)}\right) = 0 \qquad \text{if} \quad k \neq \ell, \tag{2.58}$$

so that we can write the *orthonormality relation*

$$\left(\mathbf{e}^{(k)}, \mathbf{e}^{(\ell)}\right) = \delta_{k\ell} \tag{2.59}$$

Clearly, the eigenvectors are linearly independent, as it would be impossible to represent one as a linear combination of the others, given that the others are are all orthogonal to it.

The fact that we have found N distinct, linearly independent vectors justifies our initial idea of expanding the solution of the original problem in the form (2.13). Indeed, in an N-dimensional space, any vector can be written as a linear combination of N linearly independent vectors (whether orthogonal – as here – or not).

2.5 Fourth step: Initial conditions

The last step in the solution of the problem is the determination of the constants C_k, D_k in (2.20). As expected, this step requires the imposition of the initial conditions (2.12).

For this purpose we write $\mathbf{u}(0)$ according to (2.13):

$$\mathbf{u}(0) = \mathbf{u}^0 = \sum_{k=1}^{N} U_k(0)\mathbf{e}^{(k)}, \tag{2.60}$$

and take the scalar product of this equation with $\mathbf{e}^{(\ell)}$, for $\ell = 1,\ 2,\ \ldots, N$. By using (2.59) and (2.20), we find

$$U_\ell(0) = C_\ell = \left(\mathbf{e}^{(\ell)}, \mathbf{u}^0\right) \qquad \ell = 1, 2, \ldots, N\,. \tag{2.61}$$

Similarly, upon taking the time derivative of (2.13) and evaluating it at $t = 0$ we find

$$\dot{\mathbf{u}}(0) = \mathbf{v}^0 = \sum_{k=1}^{N} \dot{U}_k(0)\mathbf{e}^{(k)}, \tag{2.62}$$

from which, as before,

$$\dot{U}_\ell(0) = \omega\sqrt{\lambda_\ell}\, D_\ell = \left(\mathbf{e}^{(\ell)}, \mathbf{v}^0\right) \qquad \ell = 1, 2, \ldots, N\,. \tag{2.63}$$

3 The continuous problem

Let us try to proceed formally in the same way for the case of a taut string. The equation is (1.1), which we rewrite using the definition (1.3) of the speed of propagation:

$$\frac{\partial^2 u}{\partial t^2} - c^2\,\frac{\partial^2 u}{\partial x^2} = 0\,, \tag{3.1}$$

with boundary conditions

$$u(0,t) = 0 \qquad\qquad u(L,t) = 0\,, \tag{3.2}$$

and initial conditions (1.5).

Let us blind ourselves to the fact that the operator in the second term is a derivative instead of a matrix as in (2.11) and simply write

$$\frac{\partial^2 u}{\partial t^2} + c^2 \mathcal{M}\, u = 0\,, \tag{3.3}$$

where

$$\mathcal{M}\, u \equiv -\frac{\partial^2 u}{\partial x^2}\,. \tag{3.4}$$

Next, as in (2.13), we write

$$u(x,t) = \sum_k U_k(t) e_k(x)\,. \tag{3.5}$$

Note that we have not indicated the range of the summation because we don't know how many "eigenvectors" we will find. For the time being we proceed formally and blindly.

Given the definition of $\mathcal{M}$, the analog of the eigenvalue equation (2.14) would be

$$-\frac{\partial^2 e_k}{\partial x^2} = \lambda_k\, e_k\,. \tag{3.6}$$

A differential equation needs boundary conditions. If we want u represented in the form (3.5) to vanish at 0 and L *for all times*, as required by (3.2), the only option is to impose the same conditions on the $e_k(x)$:

$$e_k(0) = 0\,, \qquad\qquad e_k(L) = 0\,, \tag{3.7}$$

for all k's. The solution of (3.6) satisfying the first one is clearly

$$e_k(x) = A_k \sin \sqrt{\lambda_k}\; x\,. \tag{3.8}$$

Satisfaction of the second condition requires either that $A_k = 0$, which is clearly unacceptable if we want non-zero e_k's, or

$$\sin \sqrt{\lambda_k} L = 0\,, \tag{3.9}$$

from which

$$\sqrt{\lambda_k} L = k\pi\,, \qquad \Rightarrow \qquad \lambda_k = \frac{k^2\pi^2}{L^2}\,. \tag{3.10}$$

With this result, the eigenvectors (3.8) become

$$e_k(x) = A_k \sin \frac{k\pi x}{L}\,. \tag{3.11}$$

If $k = 0$, e_0 vanishes identically and therefore we do not need to consider this value of k. If we consider the pair $\pm k$, it is obvious the the corresponding e_k's are one the negative of the other, i.e. they are proportional to each other. Given one of them, we can obtain the other one by multiplying by a constant. Hence we do not miss any e_k by restricting k to be positive. Aside from these two cases, none of the e_k's can be expressed as a linear combination of the others. The index k is therefore unrestricted with all positive integer values allowed. This suggests that we write the expansion (3.5) more explicitly as follows:

$$u(x,t) = \sum_{k=1}^{\infty} U_k(t) e_k(x)\,. \tag{3.12}$$

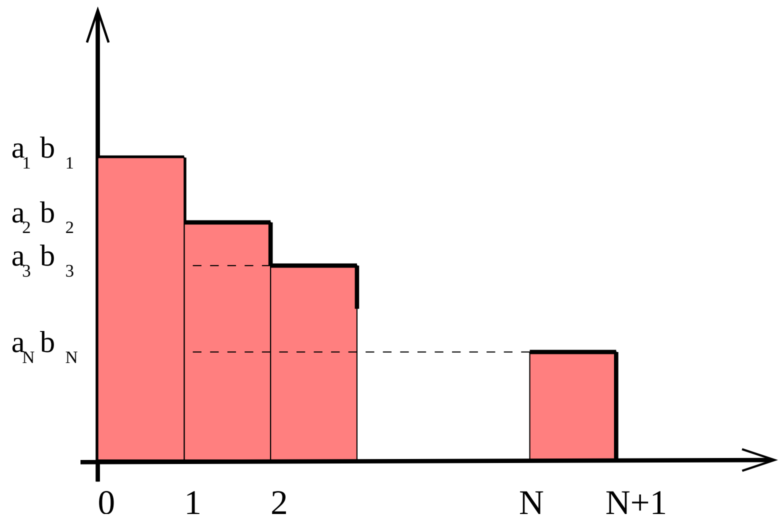

Figure 4: Graphical representation of the summation (3.13)

Since we have an infinite sum, at some point we shall have to worry about issues of convergence. This aspect is treated chapters 8 and 9 of the book.

We have thus reached a point corresponding to Eq. (2.39). The next steps would be to impose some normalization condition, to derive the equation for the U_k's, and to impose the initial conditions. All these steps require that we introduce a concept akin to that of scalar product on which we have relied to execute these steps in the discrete case. How shall we do this? Again, let us be guided by analogy with the discrete case. Since we recover the continuous case as the number of masses of the oscillator chain becomes large, we should look into what happens to the previous idea of scalar product in this limit.

3.1 Scalar product

Given two real N-dimensional vectors $\mathbf{a}$, $\mathbf{b}$, their scalar product is defined by

$$(\mathbf{a}, \mathbf{b}) = \sum_{k=1}^{N} a_k b_k \,. \tag{3.13}$$

With reference to figure 4, where each rectangle has height $a_k b_k$ and width $(k+1)-k=1$, this equals the area under the curve. Suppose now that a_k, b_k, are quantities referring to the k-th oscillator of the chain and we want to draw this figure in such a way that the rectangle corresponding to this term is on top of the k-th oscillator. In this way we will be able to fit the graph onto the segment 0 to L rather than 0 to N (figure 5). (The motivation for this step is that we eventually want to take the limit $N \to \infty$.) If we want to preserve the value of the scalar product and its interpretation as an area, we need to make the rectangles taller by $1/\Delta x$, because now their width has become Δx. Hence we write

$$(\mathbf{a}, \mathbf{b}) = \sum_{k=1}^{N} \frac{a_k}{\sqrt{\Delta x}} \frac{b_k}{\sqrt{\Delta x}} \Delta x \,. \tag{3.14}$$

Suppose now that, as $N \to \infty$,

$$\frac{a_k}{\sqrt{\Delta x}} \to a(x) \,, \qquad \frac{b_k}{\sqrt{\Delta x}} \to b(x) \,, \tag{3.15}$$

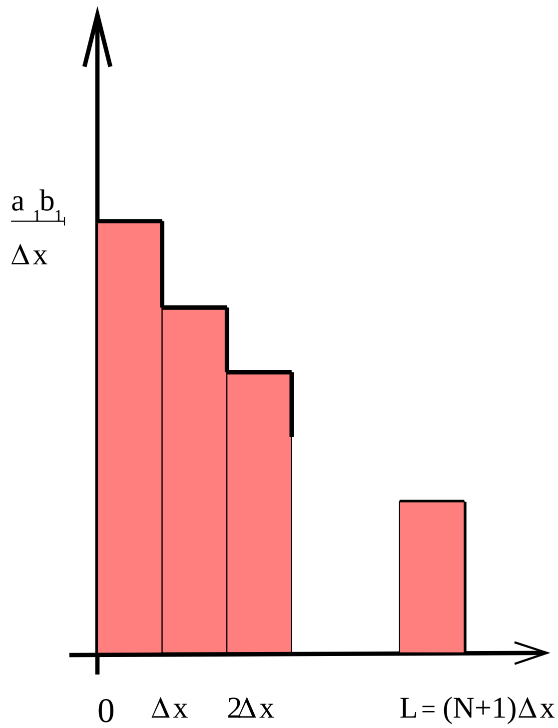

Figure 5: Graphical representation of the summation (3.14)

with $a(x)$, $b(x)$ defined for $0 \leq x \leq L$. As a (partial) justification note that, as $N \to \infty$, the points x_ℓ where the discrete quantities $a_k(x_\ell)$, $b_k(x_\ell)$ are defined get closer and closer. Then it follows from the definition of the Riemann integral[8] that, in this limit,

$$\sum_{k=1}^{N} \frac{a_k}{\sqrt{\Delta x}} \frac{b_k}{\sqrt{\Delta x}} \Delta x \;\to\; \int_0^L a(x)\, b(x)\, dx\,. \tag{3.16}$$

This argument suggests that we define the *scalar product* between two (real) functions by

$$(a, b) \;=\; \int_0^L a(x)\, b(x)\, dx\,. \tag{3.17}$$

We can be led to the same result by another argument. When there are very many oscillators, each spring is short and, therefore, the values of $a_k b_k$ corresponding to neighboring oscillators are not very different from each other otherwise the springs would be enormously extended with respect to their equilibrium length. In other words, there is some sort of “continuity”, albeit in a “discrete” sense. Then, if N is large, instead of doing a sum of N terms, we can proceed approximately and divide the terms in groups each consisting of all the oscillators positioned in a certain interval of the x-axis of width Δ_p. The number of terms in group p is clearly $\Delta_p/\Delta x$ and we can therefore write, approximately,

$$\sum_{k=1}^{N} a_k b_k \;\simeq\; \sum_{\text{all } p} \frac{\Delta_p}{\Delta x}\,(ab)_p\,, \tag{3.18}$$

[8]See section A.4.1 p. 679 for a definition of the Riemann integral.

where $(ab)_p$ is a representative value for the p-th group. Clearly the error in this approximation decreases as the Δ_p's are taken smaller (but, of course, still large enough that each Δ_p contains a significant number of oscillators). With the same assumption (3.15) as before, this is essentially the argument that leads to the definition of the Riemann integral and therefore, again, we are led to (3.17).

To see if all this makes sense, we need to check the assumption (3.15). Since any vector in the N-dimensional space can be written as a superposition of the eigenvectors (2.56), it is sufficient to check it for these vectors. For this purpose, note that

$$e_\ell^{(k)} = \sqrt{\frac{2}{N+1}}\,\sin\left(k\,\frac{\pi\ell}{N+1}\right) = \sqrt{\frac{2\Delta x}{(N+1)\Delta x}}\,\sin\left[k\,\frac{\pi(\ell\Delta x)}{(N+1)\Delta x}\right]. \tag{3.19}$$

But $(N+1)\Delta x$ is precisely equal to L,[9] and $\ell\Delta x$ is the position x_ℓ of the ℓ-th oscillator, so that we can write

$$e_\ell^{(k)} = \sqrt{\frac{2\Delta x}{L}}\,\sin\frac{k\pi x_\ell}{L}, \tag{3.20}$$

from which we see that (3.15) is satisfied with

$$\frac{e^{(k)}}{\sqrt{\Delta x}} \to \sqrt{\frac{2}{L}}\,\sin\frac{k\pi x}{L}, \tag{3.21}$$

which suggests that a good choice for the residual integration constant A_k in (3.11) should be $A_k = \sqrt{2/L}$. We will be further reassured of the reasonableness of our procedure if, with this choice of A_k, the eigenvectors (functions) $e_k(x)$ have "length" 1 according to the scalar product (3.17). Let's check:

$$\left(e^{(k)}, e^{(k)}\right) = \int_0^L \left[\sqrt{\frac{2}{L}}\,\sin\frac{k\pi x}{L}\right]^2 dx = \frac{2}{L}\,\frac{L}{2} = 1. \tag{3.22}$$

It is also readily checked that two different e's are orthogonal (in the sense of the integral) so that we can write the orthonormality relation (2.59) also for these generalized eigenvectors.

3.2 Equation for the "expansion coefficients"

Now that we have found a set of orthogonal eigenvectors (eigenfunctions) and the corresponding eigenvalues we can find the equation governing the "expansion coefficients" $U_k(t)$ in (3.12). We substitute the right-hand side of (3.12) into the wave equation (3.1) and differentiate term by term again assuming, without proof for now, the legitimacy of this step. The result, upon using the eigenfunction relation (3.6), is

$$\sum_{k=1}^{\infty}\left(\ddot{U}_k + c^2\lambda_k U_k\right) e_k(x) = 0. \tag{3.23}$$

[9] Note for example that, for $N = 1$, there are two springs, and therefore $L = 2\Delta x$

which is the perfect analog of (2.16) of the discrete case. As in that case, we now take scalar products by $e_\ell(x)$ with $\ell = 1, 2, \ldots$ assuming that integration and summation can be interchanged to find

$$\ddot{U}_\ell + c^2\lambda_\ell U_\ell \;=\; 0\,, \qquad \ell \;=\; 1,2,\ldots\,, \tag{3.24}$$

which, of course, has a solution analogous to (2.20):

$$\begin{aligned} U_\ell(t) &= C_\ell \cos c\sqrt{\lambda_\ell}\; t + D_\ell \sin c\sqrt{\lambda_\ell}\; t \\ &= C_\ell \cos \frac{\ell\pi c t}{L} + D_\ell \sin \frac{\ell\pi c t}{L}\,. \end{aligned} \tag{3.25}$$

Again, the integration constants C_ℓ and D_ℓ will be found by imposing the initial conditions.

3.3 Initial conditions

The initial conditions are given by formulae similar to (2.61), e.g.

$$C_\ell \;=\; \left(\mathbf{e}^{(\ell)}, \mathbf{u}^0\right) \;=\; \int_0^L \sqrt{\frac{2}{L}}\,\sin\left(\frac{\ell\pi x}{L}\right)\, u^0(x)\,dx\,, \tag{3.26}$$

and similarly for D_ℓ. In the language of the elementary "separation of variables" procedure we would say that we first expand $u^0(x)$ in a Fourier series of sines

$$u^0(x) \;=\; \sum_{k=1}^{\infty} u_k^0 \sin \frac{k\pi x}{L}\,, \tag{3.27}$$

where

$$u_k^0 \;=\; \frac{2}{L}\int_0^L \sin\frac{\ell\pi x}{L}\, u^0(x)\,dx\,, \tag{3.28}$$

and then equate corresponding terms.

3.4 Correspondence with the discrete case

In conclusion, let us see to what extent the oscillator chain results can be considered an approximation to the continuous case of a taut string. We have already checked the eigenvectors. Let us consider the eigenfrequencies. For the discrete oscillator chain these can be read off e.g. from (2.19) and are

$$\omega_\ell^2 \;=\; \omega^2\lambda_\ell \;=\; 4\frac{K}{m}\,\sin^2\,\frac{\ell\pi}{2(N+1)}\,; \tag{3.29}$$

they are plotted in figure 6. For the continuous case we have instead, from (3.24) or (3.25),

$$\omega_\ell^2 \;=\; c^2\lambda_\ell \;=\; \frac{T}{\mu}\,\frac{\ell^2\pi^2}{L^2}\,. \tag{3.30}$$

It is obvious that the discrete results can only be a good approximation to the lower modes.[10] For these, ℓ/N is small and we can approximate the sine in (3.29) by its argument:

$$\omega_\ell^2 \;\simeq\; 4\frac{K}{m}\,\frac{\ell^2\pi^2}{4(N+1)^2} \;=\; \frac{(K\Delta x)}{(m/\Delta x)}\,\frac{\ell^2\pi^2}{(N+1)^2\Delta x^2}\,. \tag{3.31}$$

[10] From the point of view of the numerical solution of the differential equation (1.1), this is equivalent to stating that the discretized approximation can only faithfully reproduce features of the exact solution having a characteristic length much longer than Δx.

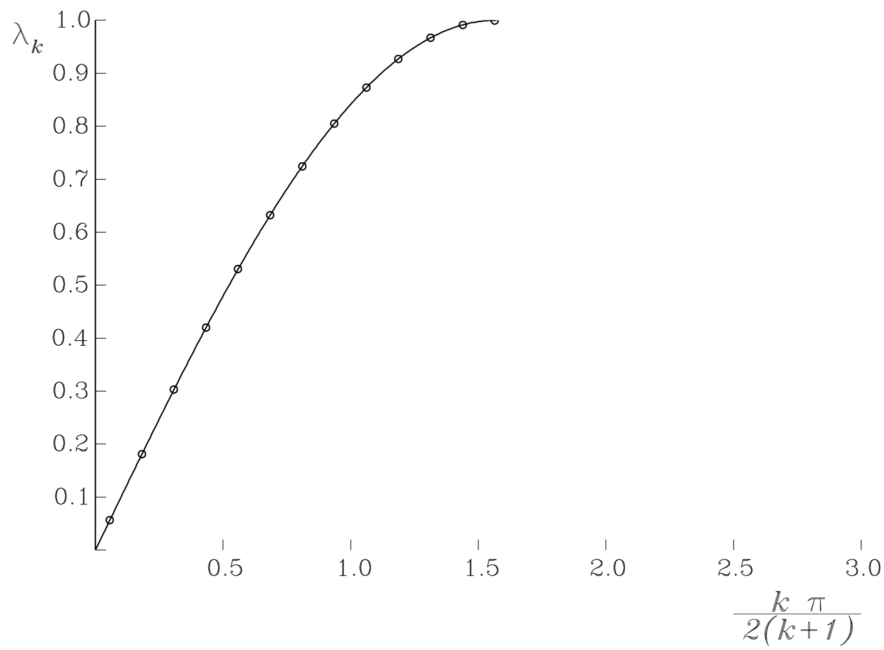

Figure 6: Eigenfrequencies of the oscillator chain according to (3.29)

As already remarked at the beginning in connection with (1.8), $k\Delta x$ and $m/\Delta x$ are the exact analogs of T and μ and, furthermore, $(N+1)\Delta x = L$. We thus conclude that (3.31) is in exact correspondence with (3.30) as expected.