Notes on Continuum Mechanics, CMI, Spring 2018
Govind S. Krishnaswami, March 1, 2018, corrections: 13 May, 2020

Contents

1 Dynamics of continuous media/deformable bodies: Lagrangian and Eulerian descriptions

• Continuum mechanics begins by dealing with the non-relativistic classical dynamics of continuous media such as oscillations of stretched strings, heat conduction in rods, elastic motion of solids (rods/beams), motion of fluids (air, water) and plasmas (ionized gases) in roughly increasing order of complexity. All of these systems involve a very large number of molecules and we will treat them as continuous mass/charge distributions with an infinite number of degrees of freedom. Thus, unlike particle mechanics, continuum mechanics deals with fields (height of string, temperature, elastic displacement, mass density, velocity, pressure, internal energy, entropy, charge density, current density, electric and magnetic fields). Thus continuum mechanics is a collection of (primarily non-relativistic, classical) field theories. Electromagnetism and gravitation are other examples of field theories, though they often involve relativistic and/or quantum effects.

• There are two principal formalisms for treating mechanics of continuous media, the so-called Lagrangian and Eulerian descriptions. The former is closer to our treatment of systems of particles, we follow the motion of each molecule or fluid element/bit of string. For example, if a fluid element occupied the location $\vec{a}$ at $t = 0$, then we seek the trajectory $\vec{r}(\vec{a}, t)$ of this fluid element, which should be determined by Lagrange's equations (ironically, this treatment was originally attempted by Euler). The Lagrangian description is particularly useful if we have some way of keeping track of which material element is where. This is usually not possible in a flowing liquid or gas, but is possible in a vibrating string since the bits of string are ordered and may be labeled by their location along the string or by their horizontal coordinate x for small vertical vibrations of a string that does not 'bend over'. For an elastic solid, the corresponding variable is the local displacement field $\mathbf{s}(\mathbf{r}, t)$ or $\vec{\xi}(\mathbf{r}, t)$ which encodes the departure from equilibrium

location of the element that was originally at $\mathbf{r}$. In a fluid like air or water, it is difficult to follow the motion of individual fluid elements. So Euler developed the so-called Eulerian description, which attempts to understand the dynamics of quantities (Eulerian variables) such as density $\rho(\vec{r},t)$, pressure $p(\vec{r},t)$, velocity $\vec{v}(\vec{r},t)$ and temperature $T(\mathbf{r},t)$ in a fluid at a specified *observation point* $\vec{r}$ at time t. However, it must be emphasized that the laws of mechanics (Newton's laws) apply to material particles or fluid elements, not to points of observation, so one must reformulate the equations of motion so that they apply to the Eulerian variables. The equations of motion in continuum mechanics tend to be expressed as partial differential equations for fields (such as the density of fluid or height of string at a given location at a given time). Thus we are dealing with the classical dynamics of fields. We will begin with the Lagrangian description of the vibration of a stretched string.

2 Vibrations of a stretched string

2.1 Wave equation for transverse vibrations of a stretched string

• Perhaps the simplest physically interesting mechanical system with a continuously infinite number of degrees of freedom is a vibrating stretched string. We will consider the special case where, in equilibrium, the string is stretched between two clamps located at $x = 0$ and $x = L$. We ignore the effects of gravity for now since the tensional forces in the string often dominate (especially when the string is stretched, in a limp string tension can be weaker). We shall call the direction in which the string is stretched, the 'horizontal' direction. The string is free to move only in one direction (vertical) transverse to the direction in which the string is stretched. We assume the string has a mass per unit length of ρ. The instantaneous configuration of the string is specified by giving the height $u(x,t)$ of the string above the horizontal position x at time t. Since the string is stretched, there are tension forces that act on any segment of the string, tangentially at either end of the segment, tending to elongate the segment. It is usually assumed that the tension in the string is a constant τ, though we will allow it to vary slowly with location, so $\tau = \tau(x)$. When the string is horizontal, the tensions at either end of any segment are horizontal, equal and opposite in direction so that the string is in equilibrium. At the end points, the tension is balanced by the force applied by the clamps.

• Note that the length of the string is not fixed, it can stretch to a length more than L, for instance when it is plucked as in a Veena. When the string is displaced from equilibrium by small vertical displacements, tensional forces on the ends of a small segment are not necessarily horizontal. But to a good approximation, the horizontal components of tension are equal and opposite, ensuring that there is no longitudinal/horizontal movement of the string. On the other hand, the vertical components of tension are in general unequal and result in a vertical acceleration of the segment. Let us estimate this. Consider a small segment of string between horizontal locations x and $x + dx$ with corresponding heights $u(x)$ and $u(x + dx) \approx u + du$. We suppose that the tangent to the string at any point x makes a counter-clockwise angle $\theta(x)$ with respect to the horizontal. Draw a diagram of a string bit that is inclined upwards! Then since we assume the inclination angles are small,

$$\cos\theta(x) \approx 1 - \frac{\theta(x)^2}{2} \approx 1 \quad \text{and} \quad \sin\theta(x) \approx \tan\theta(x) \approx \frac{\partial u}{\partial x} \equiv u'(x). \tag{1}$$

Then the horizontal components of tension at the right and left ends of the segment are $\tau(x +$

$dx)\cos\theta(x+dx)$ and $-\tau(x)\cos\theta(x)$. Since we are assuming that the string does not move horizontally, these must be equal and opposite (this is possible if τ is independent of x. More generally we must account for horizontal motion as well. We allow for non-constant τ to see how it affects the equation for transverse motion. The vertical components of tension at the right and left ends of the segment are

$$\tau(x+dx)\sin\theta(x+dx)\hat{z} \approx (\tau u')(x+dx)\hat{z} \quad \text{and} \quad -\tau(x)\sin\theta(x)\hat{z} \approx -(\tau u')(x)\hat{z}. \tag{2}$$

Thus the net upward force on the segment is

$$F_{\text{up}} = (\tau u')(x+dx) - (\tau u')(x) \approx \frac{\partial(\tau(x)u'(x))}{\partial x}\, dx \tag{3}$$

Note that the length of the string segment is $|dx|$ upto terms quadratic in small quantities:

$$\sqrt{dx^2+du^2} \approx |dx|\left[1+(1/2)(\partial u/\partial x)^2\right] \approx |dx|. \tag{4}$$

So Newton's second law for the segment, whose mass is $\rho(x)\ dx$ is (subscripts denote partial derivatives)

$$F_{\text{up}} = (\tau u_x)_x\, dx = \rho\, dx\, u_{tt}. \tag{5}$$

Thus the equation of motion for small transverse (1D) vibrations of the stretched string is $(\tau u_x)_x = \rho u_{tt}$. If the tension τ is a constant, then we get the standard form of d'Alembert's wave equation:

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} \quad \text{where} \quad c = \sqrt{\frac{\tau}{\rho}} = \sqrt{\frac{\text{tension}}{\text{mass per unit length}}}. \tag{6}$$

c has dimensions of a speed and will be seen to be the speed at which waves propagate on the string. The wave equation $u_{tt} = c^2 u_{xx}$ is a linear, homogeneous partial differential equation second order in both space and time derivatives. PDEs involve derivatives with respect to more than one independent variable (x, t here) while ordinary differential equations (ODEs) involve derivatives only in one independent variable (t in Newton's equation for a point particle). The wave equation is linear since it involves only the first power of the unknown function (dependent quantity) u. The wave equation is hyperbolic in a sense to be described later (by contrast with Laplace's equation which is elliptic and Fourier's heat equation which is parabolic). The wave equation also appears in modelling electromagnetic waves, gravitational waves, shear waves in elastic media, sound waves, free massless scalar (Klein-Gordon) fields etc.

• We are interested in solving the initial-boundary value problem for the string. The wave equation is second order in time and requires two initial conditions (say at $t = 0$), just like Newton's equation. These are the initial height $u(x, t = 0)$ and the initial velocity of the string $\dot{u}(x, t = 0)$. In addition, we need to specify what happens at the boundaries. The boundary conditions corresponding to a string clamped at the end points are $u(x = 0, t) = u(x = L, t) = 0$. This is called Dirichlet boundary conditions. Other boundary conditions are also of interest. For example, we might have an end (say at $x = 0$) of the string free to move up and down (though not horizontally), so that the slope of the string vanishes at the end point. This could be implemented by attaching the left end of the string to a massless ring free to move vertically with out friction on a pole. This means u has no slope at the left end point, one cannot apply a vertical force on the ring since it yields, it has no inertia. This leads to the free/open/Neumann

boundary condition $\frac{\partial u}{\partial x} = 0$ at $x = 0$. We could also consider an infinite string with decaying BCs: $u(x,t) \to 0$ as $|x| \to \infty$. All of the above are homogeneous boundary conditions. On the other hand a BC such as $u(0,t) = 1$ is inhomogeneous. In particular, the sum of two solutions with inhomogeneous BCs would not satisfy the BCs and the superposition principle for the linear wave equation would not apply. Sometimes we consider mixed or Robin BCs where a linear combination of u and u_x is specified at the boundaries.

• The nature of solutions and the physics they model depend crucially on boundary conditions. It is important to recognize that without suitable BCs, the equations are in general underdetermined and may have multiple solutions.

• If we were to include the weight of string segments (gravitational force due to constant acceleration due to gravity) show that the equation for small vibrations of a horizontally stretched string becomes the linear inhomogeneous PDE

$$\frac{1}{c^2}u_{tt} = u_{xx} - \frac{g}{c^2}. \tag{7}$$

• We note in passing that the wave equation for small oscillations of a stretched string is Lorentz invariant (with c interpreted as the speed of light) even though the underlying system is non-relativistic. This is associated with the linear dispersion relation for small oscillations. This is an artefact of considering small oscillations and ignoring non-linearities.

• As a consequence of considering small vibrations and small angles θ the equation of motion is linear, however, it is a partial differential equation unlike Newton's ordinary differential equations encountered in the mechanics of finitely many particles. Above, $u(x)$ is the analogue of the generalised coordinate and x labels the particles in the string. The configuration space is the set of possible instantaneous locations of the string segments, i.e. the space of twice differentiable functions $u(x)$ on the interval $[0, L]$ that vanish at the end-points. This is an infinite dimensional space reflecting the fact that a string has infinitely many degrees of freedom. The equations of continuum mechanics (e.g. fluid mechanics, electrodynamics, general relativity, elasticity) are typically systems of partial differential equations and the wave equation is perhaps the simplest prototype.

2.2 Finite differences: Wave equation as a system of ODEs

We may regard a partial differential equation such as the wave equation as a large (infinite) system of coupled ODEs, one ODE for each value of x. To see this we consider the wave equation, for definiteness, on the real line $-\infty < x < \infty$ and imagine discretizing the spatial coordinate $\ldots x_{-2}, x_{-1}, x_0 = 0, x_1, x_2, x_3, \ldots$ with small uniform spacing δx. The first derivative may be approximate as a forward or backward difference:

$$u'_{FD}(x_i) = \frac{1}{\delta x}(u(x_{i+1}) - u(x_i)) \quad \text{or} \quad u'_{BD}(x_i) = \frac{1}{\delta x}(u(x_i) - u(x_{i-1})). \tag{8}$$

If we assemble the values of u in a vector $(u(x_i))_{i\in\mathbb{Z}}$, then we may write the forward and backward difference approximants to the first derivative as upper and lower triangular matrices

with entries along the diagonal and the first super/sub diagonal

$$\left(\frac{d}{dx}\right)_{FD} \approx \frac{1}{\delta x}\begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & -1 & 1 & 0 & 0 & 0 & \dots \\ \dots & 0 & -1 & 1 & 0 & 0 & \dots \\ \dots & 0 & 0 & -1 & 1 & 0 & \dots \\ \dots & 0 & 0 & 0 & -1 & 1 & \dots \\ \dots & 0 & 0 & 0 & 0 & -1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \quad \text{and} \quad \left(\frac{d}{dx}\right)_{BD} \approx \frac{1}{\delta x}\begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & 1 & 0 & 0 & 0 & 0 & \dots \\ \dots & -1 & 1 & 0 & 0 & 0 & \dots \\ \dots & 0 & -1 & 1 & 0 & 0 & \dots \\ \dots & 0 & 0 & -1 & 1 & 0 & \dots \\ \dots & 0 & 0 & 0 & -1 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \tag{9}$$

With decaying BC, d/dx is an anti-hermitian operator $(u, v') = -(u', v)$. This is reflected in the relation $(d/dx)_{BD} = -(d/dx)^t_{FD}$. Now the finite difference approximation to the second derivative can be obtained in several ways. We could apply the forward difference twice or the backward difference twice or compose the forward and backward difference operators (in either order). We prefer the third option as the resulting matrix is hermitian, as is the operator d^2/dx^2 with vanishing/decaying BCs. Thus we have

$$\frac{d^2}{dx^2} \approx \frac{1}{(\delta x)^2}\begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & -2 & 1 & 0 & 0 & 0 & \dots \\ \dots & 1 & -2 & 1 & 0 & 0 & \dots \\ \dots & 0 & 1 & -2 & 1 & 0 & \dots \\ \dots & 0 & 0 & 1 & -2 & 1 & \dots \\ \dots & 0 & 0 & 0 & 1 & -2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \tag{10}$$

Now we may write the wave equation as an infinite collection of coupled ODEs:

$$\frac{1}{c^2}\frac{\partial^2 u(x_i,t)}{\partial t^2} = \frac{1}{(\delta x)^2}(u(x_{i+1},t) - 2u(x_i,t) + u(x_{i+1},t)) \quad \text{for} \quad i = \dots, -2, -1, 0, 1, 2, \dots. \tag{11}$$

This system of ODEs can be solved with suitable initial and BCs.

Remarkably, there is *another* way of converting the wave equation into an infinite array of ODEs, this time uncoupled! This is by the method of separation of variables and Fourier series, which we turn to next.

2.3 Separation of variables, normal modes and solution by Fourier series

• On account of the linearity of the wave equation (subject to homogeneous boundary conditions, say Dirichlet or Neumann), the superposition principle applies. Linear combinations of solutions are again solutions. The solution space forms a linear vector space. This suggests that if we can find a sufficiently large set of linearly independent solutions (called normal modes of oscillation), we may be able to express a solution of interest as a linear combination of the normal modes of oscillation.

• The wave equation $u_{tt} = c^2 u_{xx}$ is a partial differential equation for an unknown height function u dependent on two independent variables t, x. Let us look for solutions which are a product of a function of t alone and a function of x alone: $u(x,t) = X(x)T(t)$. We hope that solutions of this separable type form a basis for the space of all solutions of interest. We also hope that X and T will be determined by simpler ODEs compared to the PDE for u. Indeed, we find, wherever the quotients make sense (remarkably we will see that these quotients are well-defined in a limiting sense even when the denominators vanish),

$$X(x)\ddot{T}(t) = c^2 T(t) X''(x) \quad \Rightarrow \quad \frac{\ddot{T}(t)}{T(t)} = c^2 \frac{X''(x)}{X(x)} = -\omega^2. \tag{12}$$

Now LHS is a function of t alone while RHS is a function of x alone. Thus, both must equal the same constant which we called $-\omega^2$. We anticipate that the constant must be negative for physically interesting vibrational motion. This is because $-\omega^2 = \frac{\ddot{u}}{u}$ is the ratio of the acceleration of the string element to its displacement from the mean position. As in Hooke's law, this quotient must be negative for a restoring force. Thus, our PDE has reduced to a pair of ODEs for each allowed value of ω.

$$\ddot{T}(t) = -\omega^2 T(t) \quad \text{and} \quad X''(x) = -k^2 X(x) \quad \text{where} \quad k = \frac{\omega}{c} \text{ is the angular wave number.} \tag{13}$$

Remarkably, the ODEs for one value of ω are not coupled to those for any other value of ω. Contrast this with the ODEs we got by finite differences in the pervious section.

These ODEs are in fact eigenvalue problems. For example, the first is the eigenvalue problem for the (infinite dimensional) operator d^2/dt^2 with $T(t)$ the eigenvector and $-\omega^2$ the eigenvalue. To make the connection to finite dimensional matrix eigenvalue problems more explicit, we could discretize time and represent $T(t)$ by the column vector whose entries are $T(t_i)$ where t_i are a suitable set of times, say $\delta t(\ldots, -3, -2, -0, 1, 2, 3, \ldots)$ where δt is a small time-step. Then we may represent the operator d^2/dt^2 in this basis by a tri-diagonal real symmetric matrix, a few of whose 'middle' rows and columns are

$$\frac{d^2}{dt^2} \approx \frac{1}{(\delta t)^2} \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \ldots & -2 & 1 & 0 & 0 & 0 & \ldots \\ \ldots & 1 & -2 & 1 & 0 & 0 & \ldots \\ \ldots & 0 & 1 & -2 & 1 & 0 & \ldots \\ \ldots & 0 & 0 & 1 & -2 & 1 & \ldots \\ \ldots & 0 & 0 & 0 & 1 & -2 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \quad \text{where} \quad T(t) \approx \begin{pmatrix} \vdots \\ T(-2\delta t) \\ T(-\delta t) \\ T(0) \\ T(\delta t) \\ T(-2\delta t) \\ \vdots \end{pmatrix} \tag{14}$$

corresponding to the discretization of the second derivative

$$\ddot{T} \approx \frac{1}{\delta t}\left(\frac{T(t+\delta t) - T(t)}{\delta t} - \frac{T(t) - T(t-\delta t)}{\delta t}\right) = \frac{T(t+\delta t) - 2T(t) + T(t-\delta t)}{(\delta t)^2}. \tag{15}$$

The first ODE $\ddot{T} = -\omega^2 T$ is the same as Newton's equation for a simple harmonic oscillator[1] and the second is essentially the same, so we can write their general solutions as

$$T(t) = A\cos\omega t + B\sin\omega t \quad \text{and} \quad X(x) = C\cos kx + D\sin kx. \tag{16}$$

The clamping of end points of the string (Dirichlet boundary conditions) implies $X(0) = X(L) = 0$ so we must have $C = 0$ and $\sin kL = 0$. So $kL = n\pi$ where n is an integer. Thus the allowed eigenvalues are $-\omega_n^2 = -(n\pi c/L)^2$ and $-k_n^2 = -(n\pi/L)^2$. It suffices to take $n \geq 1$ since the negative values give (linearly dependent) solutions that only differ by a sign and $n = 0$ gives the trivial solution. So we may write any separable solution of the wave equation as

$$u_n(x,t) = (A_n\cos\omega_n t + B_n\sin\omega_n t)\sin\frac{n\pi x}{L} \quad \text{where} \quad \omega_n = \frac{n\pi c}{L} \quad \text{for some} \quad n = 1, 2, 3, \ldots \tag{17}$$

Each of these solutions for $n = 1, 2, 3, \ldots$ is called a normal mode of oscillation.

[1] It is notable that Newton's 2nd law for the SHO $\ddot{\delta}(t) = -\omega^2\delta(t)$ where $\delta(t)$ is the displacement of a particle from equilibrium is not an eigenvalue problem since $\omega = \sqrt{k/m}$ here is a fixed constant determined by the spring constant k and mass of the particle m.

The mode $n = 1$ is called the fundamental or first harmonic $n = 2$ the second harmonic or first overtone etc. Normal modes are periodic functions of both x and t. At any fixed time, a normal mode of oscillation has a definite angular wave number k_n and spatial wave length $\lambda_n = 2\pi/k_n = 2L/n$. It has a definite angular frequency $\omega_n = ck_n$ and also a definite time period of oscillation $T_n = 2\pi/\omega_n$. $\nu_n = \omega/2\pi$ is the frequency at which every point along the string vibrates about its mean position. As opposed to a normal mode, a more general motion of a stretched string will not have such a definite wave length and time period, indeed it need not even be periodic in time! Moreover, these normal modes do not necessarily satisfy the prescribed initial conditions. But since the wave equation is linear, we can take linear combinations of normal modes to produce new solutions. The most general such linear combination is a Fourier series

$$u(x,t) = \sum_{n=1}^{\infty} (A_n \cos\omega_n t + B_n \sin\omega_n t) \sin\frac{n\pi x}{L} \tag{18}$$

The 'Fourier coefficients' A_n, B_n must decay sufficiently fast as $n \to \infty$ to ensure that the sum converges ($|A_n|, |B_n| \sim 1/n^2$ will do). The theorems of Fourier series tell us that we can represent any continuous function of x that vanishes at the end points of the interval $[0, L]$ as a Fourier sine series (at any given instant of time). So we may try to fit the initial conditions by a suitable choice of constants A_n, B_n for $1 \le n \le \infty$. They are fixed by the initial height and velocity of the string

$$u(x,0) = \sum_{n=1}^{\infty} A_n \sin\frac{n\pi x}{L} \quad \text{and} \quad \dot{u}(x,0) = \sum_{n=1}^{\infty} B_n\omega_n \sin\frac{n\pi x}{L} \quad \text{where} \quad \omega_n = \frac{n\pi c}{L}. \tag{19}$$

The initial height and initial velocity of the string must vanish at the end points $x = 0, L$ so that the BCs are satisfied.

Using the orthogonality of $\sin(n\pi x/L)$ on the interval $[0, L]$ for $n = 1, 2, 3, \ldots$ and the fact that the average value (over a period) of the square of the sine function is a half, we find[2] the Euler-Fourier formulas

$$A_n = \frac{2}{L}\int_0^L u(x,0) \sin\Big(\frac{n\pi x}{L}\Big)\, dx \quad \text{and} \quad B_n = \frac{2}{n\pi c}\int_0^L \dot{u}(x,0) \sin\Big(\frac{n\pi x}{L}\Big)\, dx. \tag{20}$$

Thus we have solved the initial-boundary value problem for the motion of a stretched string clamped at the end points. It is instructive to plot a movie of the time evolution of one such solution on a computer by choosing a simple initial condition such as an isosceles triangular initial height, zero initial velocity and truncating the Fourier series after a few terms.

• We see that a general vibration of a stretched string involves a superposition of several normal modes and does not possess a definite wave number or time period. However, we will see that in general, higher harmonics cost more energy to excite. We might anticipate this since the restoring force was found to be proportional to $u''(x)$. Higher harmonics $\sin(n\pi x/L)$, for $n \gg 1$ are rapidly oscillating functions with large second derivatives, so they involve significant forces on the string segments. We would expect much energy to be stored in the oscillatory motion of a higher harmonic.

[2] We multiply both sides by $\sin(m\pi x/L)$ and integrate over $x \in [0, L]$ for integer m and consider the cases $m = n$ and $m \neq n$. Use $2\sin\alpha\sin\beta = \cos[\alpha - \beta] - \cos[\alpha + \beta]$.

2.4 Right- and left-moving waves and d'Alembert's solution

• By playing with a stretched string, we may discover the phenomenon of a transverse wave that moves along a string: a traveling wave. A vertical disturbance that is set up somewhere along a string can propagate elsewhere. This is because the vertical component of tension causes the neighboring string element to move vertically, and the process goes on. Indeed, such a traveling wave can reach a boundary (clamp) and get reflected and come back. Two such traveling waves moving in opposite directions can collide and superpose.

• More precisely, by a traveling wave, we mean a wave which maintains its profile as it moves at constant speed (say $c > 0$) along the string (wither to the right or left). For example, if $u(x,0) = f(x)$ is the initial height profile, then $u(x,t) = f(x-ct)$ is a right-moving wave and $u(x,t) = f(x+ct)$ is a left-moving wave.

• Now we investigate whether it is possible to describe the solution of the wave equation in terms of traveling waves.

• Recall that the height $u(x,t)$ (measured relative to the equilibrium height) of a stretched string executing small transverse vibrations must satisfy the wave equation $\Box u = (\frac{1}{c^2}\partial_t^2 - \partial_x^2)u = 0$. In other words, it must be annihilated by the wave operator or d'Alembertian $\Box$. d'Alembert's approach to solving the wave equation arises from factorizing the wave operator $\Box$ into a pair of first order operators. Let us consider the wave equation on an infinite interval $-\infty < x < \infty$ subject to the initial height and initial velocity

$$u(x, t=0) = h(x) \quad \text{and} \quad \dot{u}(x,0) = v(x). \tag{21}$$

The wave equation may be factorized as

$$\left(c^{-2}\partial_t^2 - \partial_x^2\right) u = \left(c^{-1}\partial_t - \partial_x\right)\left(c^{-1}\partial_t + \partial_x\right) u = \left(c^{-1}\partial_t + \partial_x\right)\left(c^{-1}\partial_t - \partial_x\right) u = 0 \tag{22}$$

We assume that we are considering functions for which mixed partials commute $\partial_t\partial_x u = \partial_x\partial_t u$; this is true provided the mixed partials exist and are continuous (Clairaut's Theorem).

• It follows that if u is annihilated by either $\partial_- = c^{-1}\partial_t - \partial_x$ or $\partial_+ = c^{-1}\partial_t + \partial_x$, then it will satisfy the wave equation[3]. Let us consider these first order equations. We notice that any differentiable function $u(x,t) = f(x-ct)$ satisfies $(c^{-1}\partial_t + \partial_x)u = 0$ while any differentiable function $u(x,t) = g(x+ct)$ is annihilated by $c^{-1}\partial_t - \partial_x$. Thus, for any say, twice differentiable functions f and g,

$$u(x,t) = f(x-ct) + g(x+ct) \tag{23}$$

is a solution of the wave equation. A little thought shows that for $c > 0$, $f(x-ct)$ is a right-moving wave with speed c and initial profile (at $t = 0$) given by the function $f(x)$. The shape of the wave $f(x-ct)$ is unaltered as it travels to the right. So $f(x-ct)$ is called a right-moving wave. Similarly, for $c > 0$, $g(x+ct)$ is a left-moving wave. Thus we have found that any superposition of a right- and left-moving wave is a solution of the wave equation.

• One wonders whether such superpositions of right and left moving waves are adequate to solve the initial value problem for a stretched string[4]. We will see that this is indeed the case on an

[3] These are not necessary conditions for solving the wave equation, only sufficient. But functions of these special sorts can be used to obtain the complete solution to the initial value problem as we will soon see.

[4] It can be shown (try!) that the initial value problem for the wave equation has a unique solution. So the solution we find here in terms of left- and right-moving waves and expressed in terms of initial height and initial velocity is the only one.

infinite domain. To solve the IVP, we wish to fix f and g in terms of the initial data.

$$u(x,0) = f(x)+g(x) = h(x) \;\text{ and }\; \dot{u}(x,0) = -cf'(x)+cg'(x) = v(x) \;\text{ or }\; -f'(x)+g'(x) = \frac{1}{c}v(x).$$

Integrating the latter equation with integration constant K we get

$$f(x) + g(x) = h(x) \quad \text{and} \quad -f(x) + g(x) = \frac{1}{c}\int_{x_0}^{x} v(\xi)\, d\xi + K. \tag{24}$$

Adding and subtracting we solve for f, g in terms of initial data

$$f(x) = \frac{1}{2}\left(h(x) - \frac{1}{c}\int_{x_0}^{x} v(\xi)\, d\xi - K\right) \quad \text{and} \quad g(x) = \frac{1}{2}\left(h(x) + \frac{1}{c}\int_{x_0}^{x} v(\xi)\, d\xi + K\right) \tag{25}$$

K and x_0 are not part of the initial data, so we hope to get rid of them. Fortunately, we are not interested in f and g separately, but only $u(x,t) = f(x-ct) + g(x+ct)$. Indeed, adding f, g, we express the solution of the wave equation entirely in terms of initial height and velocity

$$u(x,t) = \frac{1}{2}\left[h(x-ct) + h(x+ct) + \frac{1}{c}\int_{x-ct}^{x+ct} v(\xi)\, d\xi\right]. \tag{26}$$

It is instructive to plot a movie of this solution, for instance in the case of zero initial velocity and a simple initial height profile such as $h(x) = e^{-x^2/2}$. One finds two little waves moving away from $x = 0$. The height at x_o at time t_o depends on the initial ($t = 0$) height at points $x_o - ct_o$ and $x_o + ct_o$. So the initial height only at points a distance ct_o from the observation point x_o can affect the height at the point of observation. This indicates that these 'signals' travel at the speed c[5]. The initial velocity $v(x)$ only at points within a distance ct_o from the observation point can affect the height at the observation point.

2.5 Conserved energy of small oscillations of a stretched string

• Since we have not incorporated any dissipative effects and are not supplying any energy or applying external forces on the string at any time $t > 0$, we expect the energy of the vibrating string to be conserved. Let us derive an expression for the conserved energy in the same way as we did for Newton's equation. Recall that we multiplied $m\ddot{q}_i + \frac{\partial V}{\partial q_i} = 0$ by the integrating factor $\dot{q}_i$ and summed over the degrees of freedom i. The resulting expression was the statement that the time derivative of energy is zero.

• So let us begin with Newton's equation for a string in its pristine form and multiply by u_t

$$\rho u_{tt}\, dx = (\tau u_x)_x\, dx \quad \Rightarrow \quad \rho u_t u_{tt}\, dx - u_t(\tau u_x)_x\, dx = 0 \quad \Rightarrow \quad \frac{1}{2}\rho(u_t^2)_t\, dx - u_t(\tau u_x)_x\, dx = 0. \tag{27}$$

Now we sum over the degrees of freedom by integrating over $x \in [a, b]$

$$\partial_t \int_a^b \frac{1}{2}\rho u_t^2\, dx - \int_a^b u_t(\tau u_x)_x = 0. \tag{28}$$

[5]Note that the speed at which these (transverse) signals travel is quite distinct from the instantaneous vertical velocity $\dot{u}$ of a point on the string.

The first term is the time derivative of what looks like a kinetic energy by analogy with a point particle

$$\frac{m}{2}\sum_i \dot{q}_i^2 \to \int_a^b \frac{1}{2}\,\rho\, u_t^2. \tag{29}$$

So we would like to express the second term as the time derivative of a potential energy. To do so we first integrate by parts (we assume $\tau(x)$ is independent of time)

$$\partial_t \int_a^b \frac{1}{2}\rho u_t^2\, dx - [\tau u_t u_x]_a^b + \int_a^b \tau u_x u_{tx}\, dx = 0 \tag{30}$$

The boundary term vanishes if we use Dirichlet or free boundary conditions ($u = 0$ or $u_x = 0$ at $x = a, b$) or even periodic boundary conditions ($\tau(a) = \tau(b)$, $u(a,t) = u(b,t)$, $u_x(a,t) = u_x(b,t)$) and we get

$$\partial_t \int_a^b \rho\frac{u_t^2}{2}\, dx + \partial_t \int_a^b \frac{1}{2}\tau(u_x^2)\, dx = 0. \tag{31}$$

Thus the conserved energy is a sum of kinetic and potential energies (check the dimensions!)

$$E = T + V = \int_a^b \left[\frac{1}{2}\rho u_t^2 + \frac{1}{2}\tau u_x^2\right]\, dx = \int_a^b \mathcal{E}\, dx \quad \text{with} \quad \frac{dE}{dt} = 0. \tag{32}$$

The kinetic energy T is proportional to the sum of squares of speeds of the bits of string as expected. T and V are separately non-negative and so $E \geq 0$ with equality iff the stretched string is in equilibrium (say $u(x,t) = 0$). The integrand is called the energy density $E = \int \mathcal{E}(x,t)dx$. In general, the energy density 'moves around the string' in such a way that the total energy is conserved.

• The potential energy V is a gradient energy, it is proportional to the square of the gradient (slope) of the string profile. However, the string cannot have a non-zero constant slope if it is clamped at the same height at either end, it must bend. In fact, the potential energy can be regarded as an energy stored in the curvature/bending of the string. By an integration by parts we write $PE = -\int \frac{1}{2}\tau u u_{xx}\, dx$. Here u_{xx} measures the curvature of the string profile. We also see that for fixed A, B, higher ($n \gg 1$) normal modes of oscillation $\sin(n\pi x/L)$ store more potential energy, since higher Fourier modes are more undulatory and have higher second spatial derivatives.

• We verify that the energy is conserved using the eom $\rho u_{tt} = (\tau u_x)_x$ and integration by parts

$$\frac{dE}{dt} = \int \left[\rho u_t u_{tt} + \tau u_x u_{xt}\right]\, dx = \int \left[\rho u_t \frac{1}{\rho}(\tau u_x)_x - (\tau u_x)_x u_t + \partial_x(\tau u_x u_t)\right]\, dx = [\tau u_x u_t]_0^L = 0. \tag{33}$$

We assumed the boundary term $[\tau u_t u_x]_0^L$ vanishes. This is automatic if u or u_x vanish at the end points, which is the case for a clamped string (Dirichlet b.c.) or a string with free boundary conditions ($u_x = 0$). Thus the energy of the string is conserved. The energy was initially supplied to the string when it was set in motion through the initial gradients in the string profile $u_x(t = 0)$ and initial velocity of the string $u_t(t = 0)$.

2.6 Three local conservation laws for the wave equation

1. The total energy $E = \int \mathcal{E}\, dx$ is globally conserved $\dot{E} = 0$. In addition, it is locally conserved in the sense that the energy density satisfies a continuity equation $\partial_t \mathcal{E} + \partial_x j = 0$ for an energy

current density $j = -\tau u_x u_t$. Let us see why

$$\partial_t \mathcal{E} = \rho u_t u_{tt} + \tau u_x u_{xt} = u_t(\tau u_x)_x + (\tau u_x u_t)_x - (\tau u_x)_x u_t = (\tau u_x u_t)_x. \tag{34}$$

The presence of a local conservation law implies that the energy density flows like a fluid, it can move from place to place on the string due to the flux of the energy current, it cannot jump from one place to another discontinuously. This is analogous to how the movement of electric charges is associated to an electric current. It also ensures that the total energy is 'globally' conserved

$$\dot{E} = \partial_t \int_a^b \mathcal{E}\, dx = -\int_0^L \partial_x j\, dx = j(0) - j(L) = 0 \tag{35}$$

if the current vanishes at the boundaries or is equal at the boundaries. This is the case for clamped or open boundaries with periodic boundary conditions (u and u_x periodic with period L) as well as for decaying BC on an infinite domain.

• More generally, the rate of increase of energy in any fixed region (say $[a, b]$) is given by the flux of energy current across its boundary:

$$\frac{d}{dt} \int_a^b \mathcal{E}\, dx = j(a) - j(b). \tag{36}$$

$j(a)$ is the inward flux of energy across the left boundary and $j(b)$ the outward flus of energy at the right boundary.

• In general, local conservation laws imply global conservation laws subject to suitable BCs.

2. The wave equation possesses two other simple conserved quantities for suitable boundary conditions. For free ($u_x(a) = u_x(b) = 0$) or periodic boundary conditions the quantity $Q = \int_a^b \rho u_t\, dx$ is a constant of motion. This is checked by integrating by parts. (Note that a, b could either or both be infinite.)

$$\dot{Q} = \int_a^b \rho u_{tt}\, dx = \int_a^b (\tau u_x)_x\, dx = [\tau u_x]_a^b = 0. \tag{37}$$

ρu_t has the physical meaning of the vertical component of momentum of a string element. So Q is the total vertical momentum of the string. We expect it to be conserved provided there is no external vertical force, like gravity. Notice that Q is in general not conserved for Dirichlet b.c. This is to be expected since there would be a vertical force on the string at the clamps in that case.

• Like energy, $Q = \int \rho u_t\, dx$ too is locally conserved. What we mean is that we can find a current j such that a local continuity equation $\partial_t(\rho u_t) + \partial_x j = 0$ is satisfied. In fact, using the wave equation $\rho u_{tt} = (\tau u_x)_x$ we find that the required current density is $j = -\tau u_x$:

$$\partial_t(\rho u_t) - \partial_x(\tau u_x) = 0. \tag{38}$$

We say that the wave equation can be expressed in conservation form.

3. There is another conserved momentum if the total horizontal force (x-component) vanishes, and the density and tension are uniform (independent of x). This 'field' momentum is $P = \int_a^b \rho u_t u_x\, dx$. It is conserved with periodic b.c. on (a, b) or decaying b.c. on $(-\infty, \infty)$.

$$\dot{P} \;=\; \int_a^b \rho\,(u_{tt}u_x + u_t u_{xt})\; dx = \int_a^b \left(\tau u_{xx}u_x + \frac{1}{2}\rho(u_t^2)_x\right)\; dx = \int_a^b \left[\frac{1}{2}\tau(u_x^2)_x + \frac{1}{2}\rho(u_t^2)_x\right]\; dx$$

$$= \int_a^b \left[\frac{1}{2}\tau u_x^2 + \frac{1}{2}\rho u_t^2\right]_x dx = 0 \tag{39}$$

assuming ρ, τ are constants and using periodic b.c. P is in general *not conserved* for Dirichlet b.c. P too arises from a local conservation law, with the current given by the negative of the energy density:

$$\begin{aligned} \partial_t(\rho u_t u_x) &= \rho u_{tt} u_x + \rho u_t u_{xt} = \tau u_{xx} u_x + \rho u_t u_{tx} = \left(\frac{\tau}{2} u_x^2 + \frac{\rho}{2} u_t^2\right)_x \\ &\Rightarrow \partial_t(\rho u_t u_x) - \mathcal{E}_x = 0. \end{aligned} \tag{40}$$

• Q and P are the conserved quantities from Noether's theorem applied to translation invariance of the wave equation and its Lagrangian in u and in x respectively for suitable boundary conditions. To understand this we need a Lagrangian for the wave equation.

2.7 Lagrangian and Hamiltonian for stretched string

• The possible instantaneous configurations of a vibrating stretched string are the heights $u(x)$ for $0 \le x \le L$. So the configuration space is a space of functions, it is not finite dimensional. The generalised coordinates are the values of the function $u(x,t)$ for $0 \le x \le L$ at a given time t. The generalised velocities at time t are $\dot{u}(x,t)$. We will show that a Lagrangian for small transverse oscillations of the stretched string is

$$L = \int_0^L \frac{1}{2}\left[\rho(u_t)^2 - \tau(u_x)^2\right] dx \equiv \int \mathcal{L}\, dx \tag{41}$$

where $\mathcal{L} = \frac{1}{2}\left[\rho u_t^2 - \tau u_x^2\right]$ is called the Lagrangian density. This formula is simply $L = T - V$ obtained from our earlier formula for the conserved energy $E = T + V$.

• The action is defined as $S = \int_{t_1}^{t_2} L dt$. The principle of extremal action states that the equations of motion are the conditions for S to be stationary with respect to arbitrary small variations $u(x,t) \to u(x,t) + \delta u(x,t)$ holding the initial and final sting profile $u(x,t_1)$ and $u(x,t_2)$ fixed. In other words, we ask for the 'trajectory' $u(x,t)$ of the string for times $t_1 \le t \le t_2$ for prescribed initial and final string profiles. The trajectory is the $u(x,t)$ that extremizes S, i.e., δS should be zero to first order in small variations $\delta u(x,t)$. If we view $u(x,t)$ as defining a surface over a rectangle ($0 \le x \le L$, $t_1 \le t \le t_2$) in the $x-t$ plane, then $u + \delta u$ represents a nearby surface which coincides with the original surface along the edges $t = t_1, t_2$.

• In general, for a Lagrangian density $\mathcal{L}$ that depends on u and its time and space derivatives u_t, u_x, u_{xx}, the Euler-Lagrange conditions for S to be stationary are (subject to suitable boundary conditions such as clamped or decaying etc.)

$$\frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial u_t} = \frac{\partial \mathcal{L}}{\partial u} - \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial u_x}\right) + \partial_x^2\left(\frac{\partial \mathcal{L}}{\partial u_{xx}}\right). \tag{42}$$

To see this, we simply set the first variation of the action $S = \int \mathcal{L} dx\, dt$ to zero after integrating by parts in t in the 2nd term, x in the 3rd term and twice in x in the 4th term:

$$\delta S = \int \left[\frac{\partial \mathcal{L}}{\partial u}\delta u + \frac{\partial \mathcal{L}}{\partial u_t}\delta u_t + \frac{\partial \mathcal{L}}{\partial u_x}\delta u_x + \frac{\partial \mathcal{L}}{\partial u_{xx}}\delta u_{xx}\right] dx\, dt$$

$$
\begin{aligned}
&= \int \left[\frac{\partial \mathcal{L}}{\partial u} - \partial_t\left(\frac{\partial \mathcal{L}}{\partial u_t}\right) - \partial_x\left(\frac{\partial \mathcal{L}}{\partial u_x}\right) + \partial_x^2\left(\frac{\partial \mathcal{L}}{\partial u_{xx}}\right)\right] \delta u \, dx \, dt \\
\delta S = 0 \quad &\Rightarrow \quad \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial u_t} = \frac{\partial \mathcal{L}}{\partial u} - \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial u_x}\right) + \partial_x^2\left(\frac{\partial \mathcal{L}}{\partial u_{xx}}\right). \qquad (43)
\end{aligned}
$$

The boundary terms from the first two integrations by parts are

$$
\int_0^L dx [\frac{\partial \mathcal{L}}{\partial u_t}\delta u]_{t_1}^{t_2} \quad \text{and} \quad \int_{t_1}^{t_2} dt [\frac{\partial \mathcal{L}}{\partial u_x}\delta u]_0^L \qquad (44)
$$

The integrand in the first case vanishes since $\delta u(x,t_1) = \delta u(x,t_2) = 0$ by assumption. The second integrand vanishes if, say, the variations respect clamped ($\delta u(0,t) = \delta u(L,t) = 0$) or decaying BCs or even Neumann BCs as $\frac{\partial \mathcal{L}}{\partial u_x} \propto u_x$ for the Lagrangian proposed. You may try to show that the boundary terms coming from two integrations by parts in the 4th term also vanish subject to suitable BCs.

• For the above Lagrangian density we get $\partial_t(\rho u_t) = \partial_x(\tau u_x)$ or $\rho u_{tt} = \partial_x(\tau u_x)$ since ρ is not explicitly time dependent. When ρ, τ are constants, this reduces to the familiar form of the wave equation $u_{tt} = c^2 u_{xx}$ with $c^2 = \tau/\rho$.

• The momentum conjugate to the coordinate $u(x,t)$ is defined as

$$
\pi(x,t) = \frac{\delta L}{\delta u_t(x)} = \rho u_t(x,t). \qquad (45)
$$

The Hamiltonian is the Legendre transform of the Lagrangian

$$
H[u,\pi] = \text{ext}_{u_t} \int_0^L \left[\pi u_t - \mathcal{L}\right] \, dx = \int \left[\frac{\pi^2}{\rho} - \frac{1}{2}\rho\frac{\pi^2}{\rho^2} + \frac{1}{2}\tau u_x^2\right] \, dx = \int_0^L \left[\frac{\pi(x)^2}{2\rho} + \frac{\tau u_x^2}{2}\right] \, dx. \qquad (46)
$$

The extremization is with respect to the velocities $u_t(x,t)$ at all positions. The conditions for extrema recover the definition of the conjugate momentum $\pi = \delta L/\delta u_t$. Evidently, the Hamiltonian is the same as the energy, but now expressed in terms of u and π instead of u and u_t. Integrating by parts assuming clamped or free boundaries, we may also express the Hamiltonian as

$$
H[u,\pi] = \int \left[\frac{1}{2\rho}\pi^2 - \frac{1}{2}u(\tau u_x)_x\right] \, dx \quad \text{assuming} \quad [\tau u u_x]_0^L = 0. \qquad (47)
$$

• Hamilton's equations are the first order equations

$$
u_t(x) = \frac{\delta H}{\delta \pi(x)} = \frac{\pi(x)}{\rho} \quad \text{and} \quad \pi_t(x) = -\frac{\delta H}{\delta u(x)} = (\tau u_x)_x \qquad (48)
$$

Combining these two 1st order equations, we get the 2nd order wave equation $\rho u_{tt} = (\tau u_x)_x$ or $u_{tt} = c^2 u_{xx}$ for constant tension, as desired.

• The Poisson brackets between canonically conjugate variables (at a common time t) are

$$
\{u(x), \pi(x')\} = \delta(x - x') \quad \text{and} \quad \{u(x), u(x')\} = \{\pi(x), \pi(x')\} = 0. \qquad (49)
$$

The equations of motion written using PBs take the form $\partial_t f = \{f, H\}$ for any dynamical variable function f. Taking $f = u(x,t)$ and $f = \pi(x,t)$ verify that one recovers Hamilton's equations and the wave equation.

2.8 Conserved quantities from Noether's Theorem

• Recall that if the Lagrangian or Hamiltonian is independent of a particular coordinate, then its conjugate momentum is conserved. This is clear from Hamilton's/Lagrange's equation $\dot{p}_i = -\frac{\partial H}{\partial q_i} = \frac{\partial L}{\partial q_i}$. Noether's theorem is a generalization of this.

• The simplest version of Noether's theorem in mechanics states that if the Lagrangian is invariant under the infinitesimal coordinate transformation $q_i \to q_i + \delta q_i$ (called an infinitesimal symmetry), then the quantity $Q = \sum_i p_i \delta q_i$ is conserved, $\dot{Q} = 0$. The generalization to our system is that if $u \to u + \delta u$ is an infinitesimal symmetry of the Lagrangian, then $Q = \int \pi(x,t)\delta u(x,t)dx$ is conserved.

• The Lagrangian $L = \frac{1}{2}\int \left(\rho u_t^2 - \tau u_x^2\right) dx$ is invariant under translations in u and x. So we may use Noether's theorem to find the corresponding conserved quantities. They turn out to the constants of motion Q, P found earlier.

• Under a translation $x \to x + \epsilon$, $u(x) \to u(x+\epsilon) \approx u(x) + \epsilon u_x$. The Lagrangian is translation invariant in x provided the string is homogeneous (τ, ρ independent of x) and we have either periodic or decaying b.c. Dirichlet b.c. would violate translation invariance due to the positions of the clamps. Noether's theorem then gives us the conserved quantity $P = \epsilon \int \rho u_t u_x \, dx$ which we had discovered earlier.

• Under a translation (constant shift) of u, $u \to u + \epsilon$, the Lagrangian does not change since it only involves derivatives of u. If, in addition, the b.c. are also invariant (this is the case for free boundary conditions or periodic b.c., but not for Dirichlet b.c.), then we may apply Noether's theorem to deduce that $Q = \int \pi(x)\delta u(x)\, dx$ is conserved, where $\pi(x) = \rho u_t$. For translations of u, $\delta u = \epsilon$ so $Q = \epsilon \int \rho u_t \, dx$. Omitting the constant ϵ we recover the conserved vertical momentum introduced earlier.

• Q may in fact be interpreted as the conserved momentum conjugate to a cyclic coordinate in the Lagrangian. Indeed, suppose we have periodic b.c. (by this we mean that u and u_x are both periodic with period l), then we may expand the height $u(x,t)$ in a Fourier series

$$u(x,t) = a_0(t) + \sum_{n\geq 1} a_n(t) \cos\left(\frac{2n\pi x}{l}\right) + \sum_{n\geq 1} b_n(t) \sin\left(\frac{2n\pi x}{l}\right). \tag{50}$$

Note that $\cos(\pi x/l)$ is not periodic while the derivative of $\sin(\pi x/l)$ is not periodic with period l which is the reason for the factor of 2 in the arguments.

The zeroth Fourier mode $a_0 = \frac{1}{l}\int_0^l u(x,t)\, dx$ is the average height. Using orthogonality of sines and cosines and the fact that the average value of their squares over a period is a half we get

$$a_n = \frac{2}{l}\int_0^l u(x) \cos\left(\frac{2n\pi x}{l}\right) \quad \text{and} \quad b_n = \frac{2}{l}\int_0^l u(x) \sin\left(\frac{2n\pi x}{l}\right). \tag{51}$$

The infinite collection of Fourier coefficients a_0, a_n and b_n furnish coordinates on the configuration space of the vibrating string. Thus we are changing variables from the values of u at various x to the Fourier coefficients. We find

$$u_t = \dot{a}_0 + \sum_{n\geq 1} \dot{a}_n \cos\left(\frac{n\pi x}{l}\right) + \dot{b}_n \sin\left(\frac{n\pi x}{l}\right) \quad \text{and} \quad u_x = \frac{2\pi}{l}\sum_{n\geq 1} n\left(-a_n \sin\left(\frac{n\pi x}{l}\right) + b_n \cos\left(\frac{n\pi x}{l}\right)\right). \tag{52}$$

The new velocities are the time derivatives $\dot{a}_0, \dot{a}_n, \dot{b}_b$. The Lagrangian is expressed in terms of

the Fourier coefficients and their velocities using orthogonality of the sines and cosines:

$$L = \int \left[\frac{1}{2}\rho u_t^2 - \frac{1}{2}\tau u_x^2\right] dx = \frac{\rho l}{2}\left[\dot{a}_0^2 + \frac{1}{2}\sum_{n\geq 1}\left(\dot{a}_n^2 + \dot{b}_n^2\right)\right] - \frac{\tau\pi^2}{l}\sum_{n=1}^{\infty} n^2(a_n^2 + b_n^2). \tag{53}$$

All Fourier coefficients would be cyclic if $\tau = 0$. For non-zero tension, we see that a_0 is the only cyclic coordinate, it does not appear in L, though its velocity $\dot{a}_0$ does. It follows that the momentum conjugate to the 'zero mode' a_0 is

$$\pi_0 = \frac{\partial L}{\partial \dot{a}_0} = \frac{\partial T}{\partial \dot{a}_0} = \rho l \dot{a}_0 = \rho l \frac{1}{l}\int_0^l u_t\, dx = \int_0^l \rho u_t\, dx = Q. \tag{54}$$

So we see that the Noether conserved charge Q corresponding to symmetry under translations in the height $u \to u + \epsilon$ is the same as the conserved momentum conjugate to the cyclic coordinate a_0.

• Notice that we can find additional cyclic coordinates by going to 'polar-coordinates' in the a_n, b_n plane. Let $a_n = r_n \cos\theta_n$ and $b_n = r_n \sin\theta_n$ or $r_n = \sqrt{a_n^2 + b_n^2}$ and $\tan\theta_n = b_n/a_n$. Write the Lagrangian in terms of a_0, r_n, θ_n and their velocities and try to identify infinitely many cyclic coordinates. What are the corresponding conserved conjugate momenta?

2.9 Dispersion relation, phase and group speeds

• The wave equation in 1D $u_{tt} = c^2 u_{xx}$ admits traveling wave solutions of the form $e^{i(kx-\omega t)}$ provided the angular frequency ω and wave number $k > 0$ satisfy the dispersion relation $\omega^2 = c^2k^2$. The two solutions $\omega = \pm ck$ represent right and left moving waves. The waves move at the phase velocities $c_p = \omega/k = \pm c$.

• The first order wave equation $u_t = cu_x$ has the dispersion relation $\omega = -ck$ and admits only left moving waves (for $k > 0$). Similarly, $u_t = -cu_x$ admits only right moving waves with dispersion relation $\omega = ck$.

• More generally, the 3D wave equation $u_{tt} = c^2\nabla^2 u$ admits plane wave solutions $u_{\mathbf{k}}(\mathbf{r}, t) = e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$ for any wave vector $\mathbf{k}$ provided the angular frequency satisfies the dispersion relation $\omega^2 = c^2|\mathbf{k}|^2$. We call this a plane wave as $u_{\mathbf{k}}(\mathbf{r}, t)$ is constant on any plane orthogonal to the wave vector $\mathbf{k}$. These planes are the wave fronts. If we take $\omega = c|\mathbf{k}|$, this wave travels in the direction of $\mathbf{k}$ at a phase speed $\omega/k = c$.

• In all these cases, the phase speeds of all Fourier modes are the same, they are independent of k. Such an equation is called non-dispersive. More general solutions of these equations may be obtained by superposing plane waves with different wave numbers k: $u(x,t) = \int A(k)e^{i(kx-\omega(k)t)}dk$. If $A(k)$ is a function peaked around some $k = k_0$, the resulting $u(x,t)$ can look like a wave packet peaked around some $x_0(t)$. Since the individual modes travel at the same speed, they will tend to stay together and the wave packet retains its shape as it moves with the same speed c_p.

• On the other hand, for a dispersive equation such as the Schrodinger wave equation, modes with different wave numbers k move at different phase speeds and a wave packet tends to spread out. See Problem Set 1. It can be shown that the packet as a whole moves at the group speed $c_g = \frac{d\omega}{dk}$ evaluated at $k = k_0$. This is obtained using a stationary phase approximation to evaluate the above integral to find that $x_0(t) \approx c_g t$.

3 Fourier's heat conduction/diffusion equation

• Suppose we have a body with some initial (absolute) temperature distribution $T(\mathbf{r}, t = 0)$. We wish to know how this temperature distribution evolves with time.

• Empirically it is found that the heat flux between bodies or parts of a body grows with the temperature difference. Fourier's law of heat diffusion states that the heat flux density vector (with units of energy per unit time crossing unit area normal to the heat flux vector) is proportional to the negative gradient in temperature

$$\mathbf{q} = -k\nabla T \quad \text{where} \quad k = \text{thermal conductivity}. \tag{55}$$

For a perfect thermal insulator, k would be zero.

• If $d\mathbf{S} = \hat{n}dS$ is a small area (vector), then the heat flux across it (energy crossing it per unit time) is given by $\mathbf{q} \cdot d\mathbf{S}$.

• Consider gas in a fixed volume V. The increase in internal energy $U = \int_V \rho c_v T(\mathbf{r}, t)\, d\mathbf{r}$ (here T is the absolute temperature) must be due to the influx of heat across its surface ∂V. This is a consequence of the first law of thermodynamics if we assume no work is done on the gas and that there are no sources/sinks of energy inside the gas. Thus

$$\int_V \partial_t(\rho c_v T) d\mathbf{r} = -\int_{\partial V} \mathbf{q} \cdot \hat{n}\, dS = \int_{\partial V} k\nabla T \cdot \hat{n}\, dS = k\int_V \nabla \cdot \nabla T\, d\mathbf{r}. \tag{56}$$

We used Gauss' divergence theorem to convert the surface integral to a volume integral taking $\hat{n}$ to be the outward pointing normal to the surface. c_v = specific heat/mass (at constant volume, no work) and ρ = density. V is arbitrary, so the integrands must be equal and Fourier's heat diffusion equation follows:

$$\frac{\partial T}{\partial t} = \alpha\nabla^2 T \quad \text{where} \quad \alpha = \frac{k}{\rho c_v} \quad \text{is thermal diffusivity}. \tag{57}$$

$\alpha \geq 0$ has dimensions of area per unit time. Since the heat equation is linear and only involves derivatives of temperature, we are free to translate or rescale temperature.

• Interesting boundary conditions for the heat equation arise when the temperature on the boundary of the body is held fixed or the body is insulated $(\hat{n} \cdot \nabla T = 0)$.

• For definiteness, consider heat conduction on a slender 1D rod that extends from $x = 0$ to $x = L$. If the end points are held at fixed temperatures T_0 and T_1, then we have Dirichlet BCs: $T(0, t) = T_0$ and $T(L, t) = T_1$ at all times t. If the ends are insulated, it means the heat flux vanishes at the ends. This leads to Neumann boundary conditions $T'(0, t) = T'(L, t) = 0$.

• Unlike the wave equation the heat equation is first order in time and requires only one piece of initial data, which may be taken as the initial temperature distribution $T(x, 0)$. We are then interested in solving the initial-boundary value problem for the heat equation.

• Show that the heat equation is not time-reversal invariant: i.e., it is not invariant under $t \to -t$. In other words, the movie of a solution of the heat equation played backwards is generally *not* a possible solution of the heat equation.

• The heat equation does not describe a conservative system, it is an example of a dissipative system. Useful energy stored in the initial temperature distribution is lost with an attendant

rise in entropy. Unlike the time-reversal invariant wave equation, the heat equation does not admit a Hamiltonian or Lagrangian formulation.

• However, dissipative does not mean nothing is conserved. Show that the average temperature of an insulated rod of length L, $\bar{T}(t) = (1/L)\int_0^L T(x,t)dx$ is independent of time! In fact, this is a consequence of the fact that the heat equation is in local conservation form $T_t - \alpha\nabla\cdot(\nabla T) = 0$.

• Heat diffusion tends to smooth the temperature distribution. By Fourier's law, temperature differences tend to even out (unless they are maintained by external agents). Even a discontinuous initial temperature distribution such as $\theta(x)$ (Heaviside step function) becomes a smooth temperature distribution after an arbitrarily short time. As $t \to \infty$, T approaches an equilibrium temperature distribution as $t \to \infty$ when net heat flow stops. The approach to equilibrium is typically exponentially fast, as we shall see. By contrast, the wave equation does not smooth the initial wave height.

• Moreover, the backward time evolution of the heat equation is often ill-posed. This means that while we can specify practically any initial temperature distribution and evolve it forwards in time, it cannot always be evolved backwards, such a solution may not exist. For instance, we can take a Heaviside step function in temperate $T(x,0) = \theta(x)$ and evolve it forward from $t = 0$ to $t = \infty$, the sharp jump in temperature will get smoothed out. On the other hand, there is no temperature distribution that upon evolution, from $t = -\infty$ (or $t = -t_0 < 0$) reaches the above step function. Physically, heat evolution describes a process with increasing entropy and tendency to smooth sharp temperature differences. $T(x,0)$ is discontinuous and it is plausible that $T(x,0)$ is a rather specially arranged state with low entropy. It is not possible that a rod whose temperature distribution has been evolving for infinitely long while increasing its entropy and smoothing $T(x)$ can arrive at the special discontinuous state $T(x,0)$.

• Note, however, that there are initial states that can be evolved backwards in time for some duration. For instance, suppose we perform the forward evolution of the heat equation from $t = -10$ to $t = 0$ starting from some $T(x,-10)$ and ending at some $T(x,0)$. In this case, it plausible that we *can* evolve the state $T(x,0)$ backwards in time, at least till $t = -10$.

• To be a good model for the conduction of heat, it is important that the heat equation preserve the non-negativity of the initial absolute temperature distribution, after all, absolute temperature must be non-negative. Later, we will argue why this is the case.

• Let us argue that the solution to the initial value problem to the heat equation is unique (subject to suitable boundary conditions). Suppose $u(x,t)$ and $v(x,t)$ are both solutions to the heat equation with the same initial condition $u(x,0) = v(x,0) = f(x)$ and the same (consistent) boundary conditions. We wish to show that $u(x,t) = v(x,t)$ at all $t > 0$ or equivalently that $w = u - v \equiv 0$ at all x and $t > 0$. By linearity, $w = u - v$ is also a solution to the heat equation (with BCs that follow: $w = 0$ at the ends or $w' = 0$ at the ends in the Dirichlet and Neumann cases) satisfying the initial condition $w(x,0) \equiv 0$. Since there is no temperature gradient for w initially, the corresponding heat flux density vector vanishes everywhere at $t = 0$ and there cannot be any re-distribution of the temperature w. Thus $w(x,t)$ should remain zero at all subsequent times.

3.1 Solution of IVP for heat equation on an interval by Fourier's series

• Fourier developed his series to solve the heat equation. Let us consider the heat equation $u_t = \alpha u_{xx}$ on the interval $0 \le x \le L$ with the ends insulated so that $u_x(0,t) = u_x(L,t) = 0$ (Neumann BC). We wish to solve the IVP to determine how the temperature distribution evolves with time from a given initial distribution $T(x,0) = f(x)$ satisfying $f'(0) = f'(L) = 0$. Intuitively we expect inhomogeneities in the initial temperature distribution to be washed out and for the temperature all along the rod to settle down at a single temperature $\bar{T} = (1/L)\int f(x)dx$ determined by the conservation of the mean temperature mentioned earlier. How does this equilibration occur?

• As with the wave equation, we make the separation of variables ansatz $u(x,t) = X(x)T(t)$ and obtain

$$\frac{\dot{T}}{T} = \alpha \frac{X''}{X} = \text{constant} = -\kappa < 0. \tag{58}$$

Here κ is a separation constant chosen positive so that the boundary conditions may be non-trivially imposed and so that the time dependence is a decaying rather than growing exponential: we expect equilibration to a steady state rather than a blowing up of the temperature distribution as time evolves.

• The ODEs for X and T have the general solutions

$$X(x) = A\cos\sqrt{\kappa/\alpha}x + B\sin\sqrt{\kappa/\alpha}x \quad \text{and} \quad T(t) = Ce^{-\kappa t}. \tag{59}$$

The Neumann BCs $X'(0)T(t) = 0$ and $X'(L)T(t) = 0$ for all times imply $X'(0) = X'(L) = 0$. The left BC implies $B = 0$ while

$$X'(L) = -A\sqrt{\kappa/\alpha}\,\sin(\sqrt{\kappa/\alpha}L) = 0 \quad \Rightarrow \quad \kappa = \frac{n^2\pi^2}{L^2}\alpha. \tag{60}$$

where n is an integer. Thus the linearly independent separable solutions to the heat equation with insulated end points are the 'normal modes'

$$u_n(x,t) = A_n e^{-\kappa_n t}\cos(\frac{n\pi x}{L}) \quad \text{for} \quad n = 0, 1, 2, 3 \ldots. \tag{61}$$

• Unlike for the wave equation, these modes are not oscillatory. Since $\kappa_n = (n^2\pi^2/L^2)\alpha$ grow quadratically with n, the higher modes decay faster with time. The $n = 0$ mode dominates at large times t while the $n = 1$ mode controls the rate of approach to equilibrium. We may now form a linear combination of these modes to arrive at a Fourier series solution of the heat equation with Neumann boundary conditions

$$u(x,t) = \sum_{n=0}^{\infty} A_n e^{-\kappa_n t}\cos(n\pi x/L) \tag{62}$$

The Fourier coefficients A_n are determined by imposing the initial condition

$$u(x,0) = f(x) = \sum_n A_n \cos(n\pi x/L). \tag{63}$$

As before we get the Euler-Fourier formulae

$$A_0 = \frac{1}{L}\int_0^L f(x)\,dx \quad \text{and} \quad A_n = \frac{2}{L}\int_0^L f(x)\cos(n\pi x/L)\,dx \quad \text{for} \quad n = 1, 2, 3 \ldots. \tag{64}$$

• What if the ends are maintained at a fixed temperature u_0, namely the temperature of the heat bath or surroundings in which the rod is immersed? Can we use the Fourier method to solve the corresponding IVP?

3.2 Evolution kernel for heat equation on a line

• The heat diffusion equation for the temperature distribution $u(x,t)$ on an infinite rod is

$$\frac{\partial u(x,t)}{\partial t} = \alpha \nabla^2 u(x,t) \quad \text{with initial temperature distribution} \quad u(x,0). \tag{65}$$

The thermal diffusivity $\alpha > 0$. The solution may be expressed in terms of the integral kernel of the heat evolution operator $h_t = e^{\alpha t \nabla^2}$. The latter is defined via the convergent exponential series which produces a differential operator of infinite order

$$h_t = e^{\alpha t \nabla^2} = I + \alpha t \nabla^2 + \frac{\alpha^2 t^2 \nabla^4}{2!} + \cdots . \tag{66}$$

Differentiating term by term, h_t is seen to satisfy the heat equation (h_t and the Laplacian commute)

$$\partial_t h_t = \alpha \nabla^2 h_t \quad \text{with initial condition} \quad h_{t \to 0^+} = I. \tag{67}$$

• **Dirac-notation for vectors and operators:** If $|e_i\rangle$ for $i \in I$ is a basis for a (finite-dimensional) vector space, we may write a vector as a linear combination $|v\rangle = \sum_i |e_i\rangle v_i$ and thereby define the components v_i of v. If the space is an inner product space and the basis is orthonormal $\langle e_i | e_j \rangle = \delta_{ij}$, then taking an inner product with e_j we may express $v_j = \langle e_j | v \rangle$. Since the basis is orthonormal, we also have the completeness relation $\sum_i |e_i\rangle\langle e_i| = I$ which expresses the identity as a sum of outer-products: this is easily checked for any basis vector and by linearity holds in general: $\sum_i |e_i\rangle\langle e_i | e_k\rangle = |e_k\rangle = I|e_k\rangle$. Similarly, we define the components of a linear transformation by the way it acts on the basis vectors $A|e_j\rangle = \sum_i |e_i\rangle A_{ij}$. Again taking an inner product with e_k and using orthonormality, we express the components of A as the inner products ('matrix elements') $\langle e_k | A | e_j \rangle = A_{kj}$. This allows us to write A as a linear combination of outer products of basis vectors $A = \sum_{ij} A_{ij} |e_i\rangle\langle e_j|$. Combining this expression with $|v\rangle = \sum_k |e_k\rangle v_k$ we may write $A|v\rangle = \sum_{ij} |e_i\rangle A_{ij} v_j$ or $(Av)_i = \sum_j A_{ij} v_j$.

• These formulae can be formally extended to infinite dimensional inner product spaces (Hilbert spaces of functions such as $L^2(\mathbb{R})$). A convenient basis is that of position eigenstates $|x'\rangle$ which are orthogonal, Delta normalized $\langle x' | x'' \rangle = \delta(x' - x'')$ and satisfy the completeness relation $\int dx |x\rangle\langle x| = I$. They are eigenstates of the multiplication operator $(Mf)(x) = xf(x)$, i.e., $x|x'\rangle = x'|x'\rangle$ or $x\delta(x-x') = x'\delta(x-x')$. Another basic operator is differentiation $(Df)(x) = f'(x)$ or more generally the partial derivative on $L^2(\mathbb{R}^n)$. The second derivative D^2 is similarly defined, as is the Laplacian ∇^2.

• Given an operator A on a function space we have the notion of its integral kernel $A(x,y)$ defined via $(Af)(x) = \int A(x,y) f(y) dy$. In Dirac notation $A(x,y) = \langle x | A | y \rangle$. Check that the integral kernel of the multiplication operator is $M(x,y) = x\delta(x-y) = y\delta(x-y)$. Show that the integral kernel of the differentiation operator is $D(x,y) = \partial_x \delta(x-y) = -\partial_y \delta(x-y)$. Similarly, show that $D^2(x,y) = \partial_x^2 \delta(x-y) = \partial_y^2 \delta(x-y)$ and that the integral kernel of the Laplacian in 3D is $\langle \mathbf{x} | \nabla^2 | \mathbf{y} \rangle = \nabla^2_{\mathbf{x}} \delta^3(\mathbf{x} - \mathbf{y})$

• Let us define the heat kernel $h_t(x,y)$ as the integral kernel of the heat evolution operator. By this we mean that h_t acts on functions via

$$(h_t f)(x) = \int h_t(x,y) f(y)\, dy. \tag{68}$$

Alternatively, $h_t(x,y)$ is the matrix element of h_t between position eigenstates. In Dirac notation, $h_t(x,y) = \langle x|h_t|y\rangle$.

• Then

$$\partial_t h_t(x,y) = \alpha \nabla_x^2 h_t(x,y) \quad \text{with} \quad \lim_{t\to 0^+} h_t(x,y) = \delta(x-y). \tag{69}$$

The first statement is verified using the matrix elements of the laplacian $\langle x|\nabla^2|z\rangle = \nabla_x^2 \delta(x-z)$. Thus taking matrix elements of (67) we get $\langle x|\partial_t h_t|y\rangle = \partial_t h_t(x,y)$ while

$$\langle x|\alpha\nabla^2 h_t|y\rangle = \alpha \int \langle x|\nabla^2|z\rangle\langle z|h_t|y\rangle dz = \alpha \int \nabla_x^2 \delta(x-z) h_t(z,y) dz = \alpha \nabla_x^2 h_t(x,y). \tag{70}$$

The second statement uses the fact that the matrix elements of the identity between position eigenstates is the Dirac delta function.

• Note that $h_t(x,y)$ has dimensions of $1/L^d$ where d is the spatial dimension.

• Then the solution of the heat equation with initial condition $u(x,0)$ may be expressed as

$$u(x,t) = \int_{-\infty}^{\infty} h_t(x,y) u(y,0)\, dy. \tag{71}$$

To see this, we check that the RHS satisfies the heat equation (since $h_t(x,y)$ does) with initial condition $u(x,0)$. By uniqueness, it must be the desired solution of the heat equation. In other words, integration against the heat kernel transforms the initial temperature distribution into the one at time t.

• Taking $u(y,0) = \delta(y)$, we see that $h_t(x) \equiv h_t(x,0)$ may be interpreted as the temperature at x at time t if the initial temperature distribution is a Dirac delta function supported at $x = 0$. [Of course, this $u(y,0)$ and $h_t(x)$ do not have dimensions of temperature. One could use the linearity of the heat equation to multiply by a dimensional constant to make u have dimensions of temperature, but this constant plays no role in what follows.]

• Before trying to determine $h_t(x)$ by solving the heat equation let us try to guess its behaviour. Physically we heat the center of an (infinitely) long rod to a high temperature while keeping the rest of it at zero temperature, then withdraw the heat source and watch the temperature redistribute. We expect the heat to flow out symmetrically in either direction from $x = 0$. The temperature at $x = 0$ should gradually decrease and the temperature distribution should gradually widen. Suppose we guess that the distribution is a gaussian, $h_t(x) = N(t)e^{-x^2/a(t)^2}$. The width $a(t)$ of the gaussian can depend only on time and α and must have dimensions of length. Putting $a^2 \propto t^{2\nu}\alpha^\gamma$ we get $\nu = 1/2$ and $\gamma = 1$ by dimensional analysis, so $h_t(x) \sim N(t)e^{-x^2/c\alpha t}$ for some dimensionless numerical constant c.

• Now $N(t)$ must be a monotonically decreasing function of time, so we might suppose $N(t) \propto 1/t^\mu$. To determine μ we note that the 'total temperature' $\int h_t(x)dx$ is conserved (assuming $h_t(x) \to 0$ as $|x| \to \infty$) and by the initial condition its value is unity. So h_t must have dimensions of an inverse length. Consequently $N(t) = N_0/\sqrt{\alpha t}$ for some positive numerical

constant N_0. In fact, you may determine N_0 in terms of c by requiring $\int h_t(x)\,dx = 1$ at all $t \geq 0$. Thus on physical grounds, we guess that as $t \to \infty$ the heat escapes to infinity and the temperature at all points drops to zero: this is the expected equilibrium state.

• To find a formula for the heat kernel, we note that the heat equation and initial condition for $h_t(x,y)$ is invariant under the translations $x \to x+a, y \to y+a$. In other words, if $h_t(x,y)$ is a solution with IC $\delta(x-y)$ then so is $h_t(x-a,y-a)$. Taking $a=y$ we see that $h_t(x-y,0) \equiv h_t(x-y)$ is a solution of the heat equation with IC $\delta(x-y)$. By uniqueness, it follows that the heat kernel is a function only of the difference in positions $h_t(x,y) = h_t(x-y)$. Thus we have reduced the job of finding the heat kernel to finding a time-dependent function $h_t(x)$ of one rather than two arguments. To solve the equation

$$\partial_t h_t(x) = \alpha \nabla_x^2 h_t(x), \quad \text{with} \quad h_0(x) = \delta(x), \tag{72}$$

we go to momentum space since ∂_x^2 is diagonal in momentum space, differentiation becomes multiplication by ip and this will allow us to convert the PDE to a system of ODEs. Let

$$h_t(x) = \int_{-\infty}^{\infty} \tilde{h}_t(p) e^{ipx}\,[dp], \quad \tilde{h}_t(p) = \int_{-\infty}^{\infty} h_t(x) e^{-ipx}\,dx. \tag{73}$$

Then the heat equation for $h_t(x)$ implies

$$\int_{-\infty}^{\infty} \left[(\partial_t - \alpha p^2)\tilde{h}_t(p)\right] e^{ipx} \frac{dp}{2\pi} = 0 \quad \forall \;\; x,t>0 \tag{74}$$

This can be true for all x only if the quantity in parentheses vanishes, so the heat equation in momentum space becomes

$$\partial_t \tilde{h}_t(p) = -\alpha p^2 \tilde{h}_t(p) \quad \text{and has solution} \quad \tilde{h}_t(p) = A e^{-\alpha p^2 t} \tag{75}$$

The constant of integration $A = \tilde{h}_0(p) = 1$ by the initial condition $h_0(x) = \delta(x)$.

• Now we go back to position space

$$h_t(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dp\; e^{-\alpha t\left(p^2 - \frac{ipx}{\alpha t}\right)} = \frac{1}{2\pi} e^{-\frac{x^2}{4\alpha t}} \int_{-\infty}^{\infty} dp e^{-\alpha t (p - \frac{ix}{2\alpha t})^2} = \frac{1}{\sqrt{4\pi\alpha t}} e^{-\frac{x^2}{4\alpha t}}. \tag{76}$$

Thus the heat kernel is

$$h_t(x,y) = \frac{1}{\sqrt{4\pi\alpha t}} e^{-\frac{(x-y)^2}{4\alpha t}}. \tag{77}$$

The formula says that if the initial temperature is Delta distributed at $x=0$, then as time elapses, the temperature $h_t(x,0)$ at any fixed point $x \neq 0$ first grows with time, reaches a maximum and then decays to zero as a power law $t^{-1/2}$. For a fixed time $t>0$, the temperature exponentially decays to zero as we move x away from the origin. In d spatial dimensions, the heat kernel is similar (it is just the product of d one-dimensional heat kernels)

$$h_t(x,y) = \frac{1}{(4\pi\alpha t)^{d/2}} e^{-\frac{(\mathbf{x}-\mathbf{y})^2}{4\alpha t}}. \tag{78}$$

Since the heat kernel $h_t(x,y)$ is positive, it follows that if $u(x,0)$ is non negative, $u(x,t)$ will also be non-negative. Thus the (absolute) temperature cannot become negative, as we expected on physical grounds.

• Notice also that even an infinitely discontinuous initial temperature distribution $\delta(x)$ evolves into a smooth distribution $h_t(x)$ at any $t>0$.