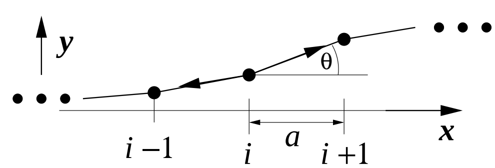
Field Theory

1 String dynamics

In this section we consider two closely related problems, transverse oscillations of a stretched loaded string, and of a stretched heavy string. The latter is is a limiting case of the former. This will provide an introduction to field theory, in which the dynamical degrees of freedom are not a discrete set but are defined at each point in space. Later we will discuss more interesting and involved cases such as the electromagnetic field, where at each point in space we have $\vec{E}$ and $\vec{B}$ as degrees of freedom, though not without constraints. Then we will consider even more interesting fields, transforming under a nonabelian gauged symmetry group.

The loaded string we will consider is a light string under tension τ stretched between two fixed points a distance ℓ apart, say at $x = 0$ and $x = \ell$. On the string, at points $x = a, 2a, 3a, \ldots, na$, are fixed n particles each of mass m, with the first and last a distance a away from the fixed ends. Thus $\ell = (n+1)a$. We will consider only small transverse motion of these masses, using y_i as the transverse displacement of the i'th mass, which is at $x = ia$. We assume all excursions from the equilibrium positions $y_i = 0$ are small, and in particular that the difference in successive displacements $y_{i+1} - y_i \ll a$. Thus we are assuming that the angle made by each segment of the string, $\theta_i = \tan^{-1}[(y_{i+1} - y_i)/a] \ll 1$.



Working to first order in the θ's in the equations of motion, and second order for the Lagrangian, we see that restricting our attention to transverse motions and requiring no horizontal motion implies the tension τ to be constant along the string. The transverse force on the i'th mass is thus

$$F_i = \tau \frac{y_{i+1} - y_i}{a} + \tau \frac{y_{i-1} - y_i}{a} = \frac{\tau}{a}(y_{i+1} - 2y_i + y_{i-1}).$$

The potential energy $U(y_1, \ldots, y_n)$ then satisfies

$$\frac{\partial U}{\partial y_i} = -\frac{\tau}{a}(y_{i+1} - 2y_i + y_{i-1})$$

so

$$
\begin{aligned}
&U(y_1,\ldots,y_i,\ldots,y_n)\\
&\quad=\int_0^{y_i} dy_i\,\frac{\tau}{a}(2y_i-y_{i+1}-y_{i-1})+F(y_1,\ldots,y_{i-1},y_{i+1},\ldots,y_n)\\
&\quad=\frac{\tau}{a}\left(y_i^2-(y_{i+1}+y_{i-1})y_i\right)+F(y_1,\ldots,y_{i-1},y_{i+1},\ldots,y_n)\\
&\quad=\frac{\tau}{2a}\left((y_{i+1}-y_i)^2+(y_i-y_{i-1})^2\right)+F'(y_1,\ldots,y_{i-1},y_{i+1},\ldots,y_n)\\
&\quad=\sum_{i=0}^{n}\frac{\tau}{2a}\left(y_{i+1}-y_i\right)^2+\text{constant}.
\end{aligned}
$$

The F and F' are unspecified functions of all the y_j's except y_i. In the last expression we satisfied the condition for all i, and we have used the convenient definition $y_0=y_{n+1}=0$. We can and will drop the arbitrary constant.

The kinetic energy is simply $T=\frac{1}{2}m\sum_1^n\dot{y}_i^2$. The potential energy $U=\frac{1}{2}y^T\cdot A\cdot y$ has a non-diagonal $n\times n$ matrix

$$
A=-\frac{\tau}{a}\begin{pmatrix}
-2 & 1 & 0 & 0 & \cdots & 0 & 0\\
1 & -2 & 1 & 0 & \cdots & 0 & 0\\
0 & 1 & -2 & 1 & \cdots & 0 & 0\\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\
0 & 0 & 0 & 0 & \cdots & -2 & 1\\
0 & 0 & 0 & 0 & \cdots & 1 & -2
\end{pmatrix}.
$$

The Lagrangian $L=T-U$ and Lagrange's equation tells us $\dfrac{d}{dt}\dfrac{\partial\mathcal{L}}{\partial\dot{y}_i}=\dfrac{\partial L}{\partial y_i}=-(Ay)_i$. While this involves an indefinite number of coupled degrees of freedom, it is not hard to find the general solution,

$$
y(ja,t)=\sum_p \text{Re}\ B_p e^{i\omega_p t}\sin(k_p ja),
$$

with $k_p=p\pi/\ell$, $p=1\ldots n$, with $\omega_p=2\sqrt{\tau/am}\sin(k_p a/2)$. We have arbitrary (complex) amplitudes B_p for each mode p. That is interesting for solid state physics, but we are more interested in the continuum limit, with a view to understanding how to formulate continuum mechanics.

Consider the limit in which the length ℓ is held fixed, but the number of masses $n\to\infty$, $a=\ell/(n+1)\to 0$ with each mass decreasing so that the linear density $\rho=m/a$ is held constant. This constitutes the **continuum**

limit. The function $y(ja)$ which had been defined only at discrete values of $x = ja$ will be assumed to become a continuous function of x[1].

What happens to the kinetic and potential energies in this limit? For the kinetic energy,

$$T = \frac{1}{2}m\sum_i \dot{y}_i^2 = \frac{1}{2}\rho\sum_i a\dot{y}^2(x_i) = \frac{1}{2}\rho\sum_i \Delta x\, \dot{y}^2(x_i) \to \frac{1}{2}\rho\int_0^\ell dx\, \dot{y}^2(x),$$

where the next to last expression is just the definition of a Riemann integral. For the potential energy,

$$U = \frac{\tau}{2a}\sum_i (y_{i+1} - y_i)^2 = \frac{\tau}{2}\sum_i \Delta x \left(\frac{y_{i+1} - y_i}{\Delta x}\right)^2 \to \frac{\tau}{2}\int_0^\ell dx \left(\frac{\partial y}{\partial x}\right)^2.$$

The equation of motion for y_i is

$$m\ddot{y}_i = \frac{\partial L}{\partial y_i} = -\frac{\partial U}{\partial y_i} = \frac{\tau}{a}[(y_{i+1} - y_i) - (y_i - y_{i-1})],$$

or

$$\rho a\ddot{y}(x) = \frac{\tau}{a}([y(x+a) - y(x)] - [y(x) - y_(x-a)]).$$

We need to be careful about taking the limit

$$\frac{y(x+a) - y(x)}{a} \to \frac{\partial y}{\partial x}$$

because we are subtracting two such expressions evaluated at nearby points, and because we will need to divide by a again to get an equation between finite quantities. Thus we note that

$$\frac{y(x+a) - y(x)}{a} = \left.\frac{\partial y}{\partial x}\right|_{x+a/2} + \mathcal{O}(a^2),$$

so

$$\begin{aligned}\rho\ddot{y}(x) &= \frac{\tau}{a}\left(\frac{y(x+a) - y(x)}{a} - \frac{y(x) - y_(x-a)}{a}\right)\\ &\approx \frac{\tau}{a}\left(\left.\frac{\partial y}{\partial x}\right|_{x+a/2} - \left.\frac{\partial y}{\partial x}\right|_{x-a/2}\right) \to \tau\frac{\partial^2 y}{\partial x^2},\end{aligned}$$

[1]This means that the nodes B_p are unrestricted for finite p, but $B_{\alpha n} \to 0$ for fixed nonzero α. The acoustic modes remain but the optical modes don't. Thus $\sin(k_p a)/a \to k_p = k\pi/\ell$ and we have a nondispersive wave with speed $c = \sqrt{\tau/\rho}$.

and we wind up with the wave equation for transverse waves on a massive string

$$\frac{\partial^2 y}{\partial t^2} - c^2 \frac{\partial^2 y}{\partial x^2} = 0,$$

where

$$c = \sqrt{\frac{\tau}{\rho}}.$$

2 Field theory

We now examine how to formulate the continuum limit directly.

2.1 Lagrangian density

We saw in the last section that the kinetic and potential energies in the continuum limit can be written as integrals over x of densities, and so we may also write the Lagrangian as the integral of a **Lagrangian density** $\mathcal{L}(x)$,

$$L = T - U = \int_0^L dx\ \mathcal{L}(x), \qquad \mathcal{L}(x) = \left[\frac{1}{2}\rho\dot{y}^2(x,t) - \frac{1}{2}\tau\left(\frac{\partial y(x,t)}{\partial x}\right)^2\right].$$

This Lagrangian, however, will not be of much use until we figure out what is meant by varying it with respect to each dynamical degree of freedom or its corresponding velocity. In the discrete case we have the canonical momenta $P_i = \partial L/\partial \dot{y}_i$, where the derivative requires holding all $\dot{y}_j$ fixed, for $j \neq i$, as well as all y_k fixed. In the continuum, however, this notion is a bit dubious — how can we vary $\dot{y}(x_0)$ at one point x_0 while holding $\dot{y}(x)$ fixed at all other x? In the discrete case, this variation extracts one term from the sum $\frac{1}{2}\rho\sum a\dot{y}_i^2$, and this would appear to vanish in the limit $a \to 0$. Instead, we define the canonical momentum as a density, $P_i \to aP(x = ia)$, so

$$P(x = ia) = \lim \frac{1}{a}\frac{\partial}{\partial \dot{y}_i}\sum_i a\ \mathcal{L}(y(x), \dot{y}(x), x)\big|_{x=ai}.$$

We may think of the last part of this limit,

$$\lim_{a\to 0}\sum_i a\ \mathcal{L}(y(x), \dot{y}(x), x)\big|_{x=ai} = \int dx\, \mathcal{L}(y(x), \dot{y}(x), x),$$

if we also define a limiting operation

$$\lim_{a\to 0}\frac{1}{a}\frac{\partial}{\partial \dot{y}_i} \to \frac{\delta}{\delta \dot{y}(x)},$$

and similarly for $\frac{1}{a}\frac{\partial}{\partial y_i}$, which act on functionals of $y(x)$ and $\dot{y}(x)$ by

$$\frac{\delta y(x_1)}{\delta y(x_2)} = \delta(x_1 - x_2), \quad \frac{\delta \dot{y}(x_1)}{\delta y(x_2)} = \frac{\delta y(x_1)}{\delta \dot{y}(x_2)} = 0, \quad \frac{\delta \dot{y}(x_1)}{\delta \dot{y}(x_2)} = \delta(x_1 - x_2).$$

where $\delta(x' - x)$ is the **Dirac delta function**[2] Thus

$$P(x) = \frac{\delta}{\delta \dot{y}(x)} \int_0^\ell dx' \frac{1}{2}\rho \dot{y}^2(x', t) = \int_0^\ell dx' \rho \dot{y}(x', t)\delta(x' - x) = \rho \dot{y}(x, t).$$

We also need to evaluate

$$\frac{\delta}{\delta y(x)} L = \frac{\delta}{\delta y(x)} \int_0^\ell dx' \frac{-\tau}{2} \left(\frac{\partial y}{\partial x}\right)^2_{x=x'}.$$

For this we need[3]

$$\frac{\delta}{\delta y(x)} \frac{\partial y(x')}{\partial x'} = \frac{\partial}{\partial x'}\delta(x' - x) := \delta'(x' - x),$$

Thus

$$\frac{\delta}{\delta y(x)} L = -\int_0^\ell dx' \tau \frac{\partial y}{\partial x}(x')\delta'(x' - x) = \tau \frac{\partial^2 y}{\partial x^2},$$

and Lagrange's equations give the wave equation

$$\rho \ddot{y}(x, t) - \tau \frac{\partial^2 y}{\partial x^2} = 0. \tag{1}$$

[2]The Dirac delta function is defined by its integral, $\int_{x_1}^{x_2} f(x')\delta(x' - x)dx' = f(x)$ for any function $f(x)$, provided $x \in (x_1, x_2)$.

[3]which is again defined by its integral,

$$\begin{aligned}
\int_{x_1}^{x_2} f(x')\delta'(x' - x)dx' &= \int_{x_1}^{x_2} f(x') \frac{\partial}{\partial x'}\delta(x' - x)dx' \\
&= f(x')\delta(x' - x)\big|_{x_1}^{x_2} - \int_{x_1}^{x_2} dx' \frac{\partial f}{\partial x'}\delta(x' - x) \\
&= \frac{\partial f}{\partial x}(x),
\end{aligned}$$

where after integration by parts the surface term is dropped because $\delta(x - x') = 0$ for $x \neq x'$, which it is for $x' = x_1, x_2$ if $x \in (x_1, x_2)$.

2.2 Lagrangian Mechanics for 3-D Fields

In sections 1 and 2.1 we considered the continuum limit of a chain of point masses on stretched string. We had a situation in which the potential energy had interaction terms for particle A which depended only on the relative displacements of particles in the neighborhood of A. If we generalize to motion of a three-dimensional material, the displacements from equilibrium will be vectors $\vec{\eta}(\vec{r},t)$, and we expect the potential energy to be integrals over volume of a function of $\vec{\eta}(\vec{r},t)$ and its spatial derivatives. More generally, $\vec{\eta}$ could be some other fields[4]. The dynamics is then determined by a Lagrangian density

$$\mathcal{L} = \mathcal{L}(\eta_i, \frac{\partial \eta_i}{\partial x}, \frac{\partial \eta_i}{\partial y}, \frac{\partial \eta_i}{\partial z}, \frac{\partial \eta_i}{\partial t}, x, y, z, t)$$

with lagrangian $L = \int dx\, dy\, dz\, \mathcal{L}$ and action $I = \int dx\, dy\, dz\, dt\, \mathcal{L}$.

The actual motion of the system will be given by a particular set of functions $\eta_i(x,y,z,t)$, which are functions over the volume in question and of $t \in [t_I, t_f]$. The function will be determined by the laws of dynamics of the system, together with boundary conditions which depend on the initial configuration $\eta_i(x,y,z,t_I)$ and perhaps a final configuration. Generally there are some boundary conditions on the spatial boundaries as well. For example, our stretched string required $y=0$ at $x=0$ and $x=L$, for all values of t.

Before taking the continuum limit we say that the configuration of the system at a given t was a point in a large N dimensional configuration space, and the motion of the system is a path $\Gamma(t)$ in this space. In the continuum limit $N \to \infty$, so we might think of the path as a path in an infinite dimensional space. But we can also think of this path as a mapping $t \to \eta(\cdot,\cdot,\cdot,t)$ of time into the (infinite dimensional) space of functions on ordinary space.

Hamilton's principal says that the actual path is an extremum of the action. If we consider small variations $\delta\eta_i(x,y,z,t)$ which vanish on the boundaries, then

$$\delta I = \int dx\, dy\, dz\, dt\, \delta\mathcal{L} = 0$$

determines the equations of motion.

Note that what is varied here are the functions η_i, not the coordinates (x,y,z,t). x,y,z do not represent the position of some atom — they represent a label which tells us which atom it is that we are talking about. Often they are chosen to be the equilibrium position of that atom, but they are fixed

[4]In the physicist's definition, a function of space and time, not the mathematician's.

labels independent of the motion. It is the $\eta_i(\vec{x})$, for each $\vec{x}$, which are the dynamical degrees of freedom, specifying the configuration of the system. In our discussion of section 2 η_i specified the displacement from equilibrium, but here we generalize to an arbitrary set of dynamical fields.

The variation of the Lagrangian density is

$$\begin{aligned}
&\delta\mathcal{L}(\eta_i, \frac{\partial\eta_i}{\partial x}, \frac{\partial\eta_i}{\partial y}, \frac{\partial\eta_i}{\partial z}, \frac{\partial\eta_i}{\partial t}, x, y, z, t) \\
&\quad = \sum_i \frac{\partial\mathcal{L}}{\partial\eta_i}\delta\eta_i + \sum_i \frac{\partial\mathcal{L}}{\partial(\partial\eta_i/\partial x)}\delta\frac{\partial\eta_i}{\partial x} + \sum_i \frac{\partial\mathcal{L}}{\partial(\partial\eta_i/\partial y)}\delta\frac{\partial\eta_i}{\partial y} \\
&\qquad + \sum_i \frac{\partial\mathcal{L}}{\partial(\partial\eta_i/\partial z)}\delta\frac{\partial\eta_i}{\partial z} + \sum_i \frac{\partial\mathcal{L}}{\partial(\partial\eta_i/\partial t)}\delta\frac{\partial\eta_i}{\partial t}.
\end{aligned}$$

Notice there is no variation of x, y, z, and t, as we discussed.

The notation is getting awkward, so we need to reintroduce the notation $A_{,j} = \partial A/\partial r_j$, for $r_j = (x, y, z)$. In fact, we see that $\partial/\partial t$ enters in the same way as $\partial/\partial x$, so it is time to introduce notation which will become crucial when we consider relativistic dynamics, even though we are not doing so here. So we will consider time to be an additional component of the position, called the zeroth rather than the fourth component. We will also change our notation for coordinates to anticipate needs from relativity, by writing the indices of coordinates as superscripts rather than subscripts. Thus we write $x^0 = ct$, where c will eventually be taken as the speed of light, but for the moment is an arbitrary scaling factor. Until we get to special relativity, one should consider whether an index is raised or lowered as irrelevant, but they are written here in the place which will be correct once we make the distinction between them. In particular the Kronecker delta is now written ${\delta_\mu}^\nu$. For the partial derivatives we now have

$$\partial_\mu := \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{c\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right),$$

for $\mu = 0, 1, 2, 3$, and write $\eta_{,\mu} := \partial_\mu \eta$. If there are several fields η_i, then $\partial_\mu \eta_i = \eta_{i,\mu}$. The comma represents the beginning of differentiation, so we must not use one to separate different ordinary indices.

In this notation, we have

$$\delta\mathcal{L} = \sum_i \frac{\partial\mathcal{L}}{\partial\eta_i}\delta\eta_i + \sum_i \sum_{\mu=0}^{3} \frac{\partial\mathcal{L}}{\partial\eta_{i,\mu}}\delta\eta_{i,\mu},$$

and

$$\delta I = \int \left(\sum_i \frac{\partial \mathcal{L}}{\partial \eta_i} \delta \eta_i + \sum_i \sum_{\mu=0}^{3} \frac{\partial \mathcal{L}}{\partial \eta_{i,\mu}} \delta \eta_{i,\mu} \right) d^4x,$$

where[5] $d^4x = cdx\, dy\, dz\, dt$. Except for the first term, we integrate by parts,

$$\delta I = \int \left[\sum_i \frac{\partial \mathcal{L}}{\partial \eta_i} - \sum_i \sum_{\mu=0}^{3} \left(\partial_\mu \frac{\partial \mathcal{L}}{\partial \eta_{i,\mu}} \right) \right] \delta \eta_i d^4x,$$

where we have thrown away the boundary terms which involve $\delta \eta_i$ evaluated on the boundary, which we assume to be zero. Inside the region of integration, the $\delta \eta_i$ are independent, so requiring $\delta I = 0$ for all functions $\delta \eta_i(x^\mu)$ implies

$$\sum_\mu \frac{d}{dx^\mu} \frac{\partial \mathcal{L}}{\partial \eta_{i,\mu}} - \frac{\partial \mathcal{L}}{\partial \eta_i} = 0. \tag{2}$$

We have written the equations of motion (which are now partial differential equations rather than coupled ordinary differential equations), in a form which looks like we are dealing with a relativistic problem, because t and spatial coordinates are entering in the same way. We have not made any assumption of relativity, however, and our problem will not be relativistically invariant unless the Lagrangian density is invariant under Lorentz transformations (as well as translations).

I have given you this introduction to continuum mechanics chiefly so we can discuss gauge field theories, so I am not pursuing very useful ideas such as the energy-momentum (or stress-energy) tensor and Noether's theorem[6], and the actual description of motion of solid bodies or fluids.

[5] We have also multiplied I by c, which does no harm in finding the extrema.

[6] You can continue this discussion at `http://www.physics.rutgers.edu/~shapiro/507/book9.pdf` page 236, which is where this leaves off.