Contents

List of Figures

List of Tables

Chapter 4

Lagrangian Mechanics

4.1 Snell's Law

Warm-up problem: You are standing at point (x_1, y_1) on the beach and you want to get to a point (x_2, y_2) in the water, a few meters offshore. The interface between the beach and the water lies at $x = 0$. What path results in the shortest travel time? It is not a straight line! This is because your speed v_1 on the sand is greater than your speed v_2 in the water. The optimal path actually consists of two line segments, as shown in fig. 4.1. Let the path pass through the point $(0, y)$ on the interface. Then the time T is a function of y:

$$T(y) = \frac{1}{v_1}\sqrt{x_1^2 + (y - y_1)^2} + \frac{1}{v_2}\sqrt{x_2^2 + (y_2 - y)^2} \quad . \tag{4.1}$$

To find the minimum time, we set

$$\begin{aligned} \frac{dT}{dy} = 0 &= \frac{1}{v_1}\frac{y - y_1}{\sqrt{x_1^2 + (y - y_1)^2}} - \frac{1}{v_2}\frac{y_2 - y}{\sqrt{x_2^2 + (y_2 - y)^2}} \\ &= \frac{\sin\theta_1}{v_1} - \frac{\sin\theta_2}{v_2} \quad . \end{aligned} \tag{4.2}$$

Thus, the optimal path satisfies

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{v_1}{v_2} \quad , \tag{4.3}$$

which is known as *Snell's Law*. Snell's Law is familiar from optics, where the speed of light in a polarizable medium is written $v = c/n$, where n is the index of refraction. Thus $n_1 \sin\theta_1 = n_2 \sin\theta_2$. If there are several interfaces, Snell's law holds at each one, so that

$$n_i \sin\theta_i = n_{i+1} \sin\theta_{i+1} \qquad \Leftrightarrow \qquad \frac{\sin\theta_i}{v_i} = \frac{\sin\theta_{i+1}}{v_{i+1}} \quad , \tag{4.4}$$

at the interface between media i and $i + 1$.

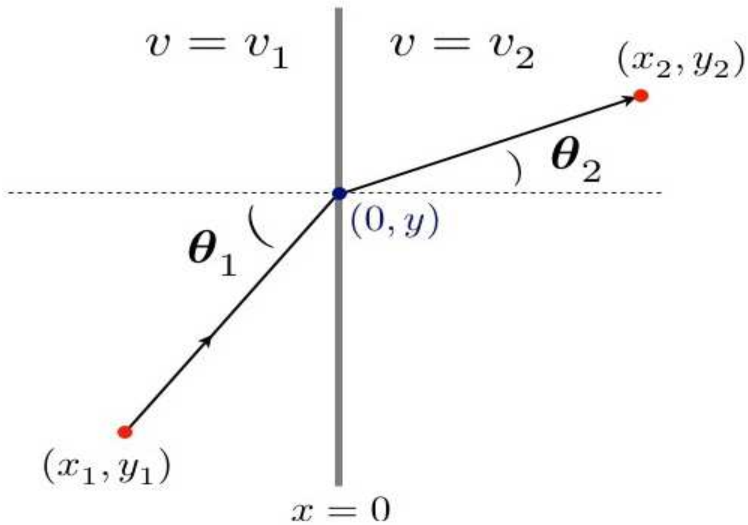

Figure 4.1: The shortest path between (x_1, y_1) and (x_2, y_2) is not a straight line, but rather two successive line segments of different slope.

In the limit where the number of slabs goes to infinity but their thickness is infinitesimal, we can regard n and θ as functions of a continuous variable x. One then has

$$\frac{\sin\theta(x)}{v(x)} = \frac{\sin\theta(x+dx)}{v(x+dx)} \quad , \tag{4.5}$$

which tells us that

$$\frac{d}{dx}\bigg(\frac{\sin\theta}{v}\bigg) = 0 \quad . \tag{4.6}$$

On a differential scale, trigonometry tells us that

$$\sin\theta(x) = \frac{dy}{\sqrt{dx^2 + dy^2}} = \frac{y'}{\sqrt{1+y'^2}} \quad , \tag{4.7}$$

and therefore eqn. 4.6 yields

$$\begin{aligned} \frac{d}{dx}\bigg(\frac{y'}{v\sqrt{1+y'^2}}\bigg) &= \frac{y''}{v\sqrt{1+y'^2}} - \frac{y'^2 y''}{v\big(1+y'^2\big)^{3/2}} - \frac{v' y'}{v^2\sqrt{1+y'^2}} \\ &= \frac{1}{v\big(1+y'^2\big)^{3/2}}\bigg(y'' - \frac{v'}{v}(1+y'^2)\,y'\bigg) = 0 \quad . \end{aligned} \tag{4.8}$$

Thus we arrive at the homogeneous second order nonlinear ODE,

$$y'' - \frac{v'}{v}(1+y'^2)\,y' = 0 \quad , \tag{4.9}$$

This is a differential equation that $y(x)$ must satisfy if the *functional*

$$T\big[y(x)\big] = \int \frac{ds}{v} = \int\limits_{x_1}^{x_2} dx\,\frac{\sqrt{1+y'^2}}{v(x)} \tag{4.10}$$

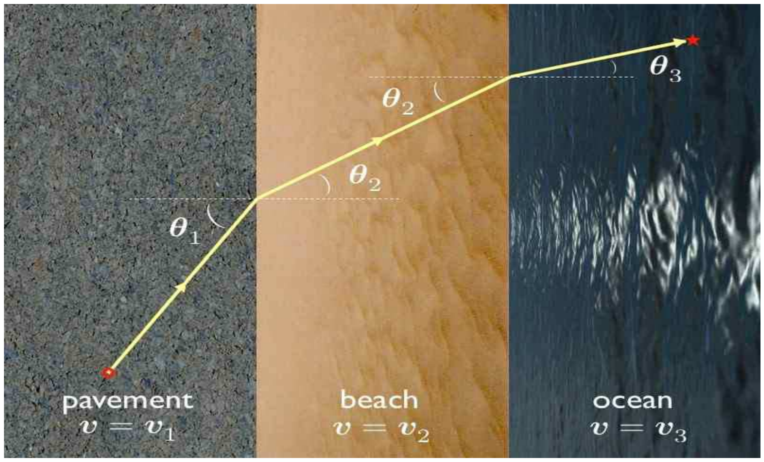

Figure 4.2: The path of shortest length is composed of three line segments. The relation between the angles at each interface is governed by Snell's Law.

is to be minimized. The solution of eqn. 4.9 will require two initial conditions, such as $y(x_0) = y_0$ and $y'(x_0) = y'_0$, or perhaps two boundary conditions $y(x_0) = y_0$ and $y(x_1) = y_1$. Indeed from eqn. 4.6 we can already integrate once, yielding

$$\frac{\sin\theta}{v} = \frac{y'}{v\sqrt{1+y'^2}} = P \quad , \tag{4.11}$$

where P is a constant. Thus, we arrive at a first order ODE, which after isolating y' may be written as

$$\frac{dy}{dx} = \pm\frac{Pv(x)}{\sqrt{1-P^2v^2(x)}} \quad , \tag{4.12}$$

and for which we must supply one initial/boundary condition.

4.2 The Calculus of Variations

4.2.1 Functions and functionals

A *function* is a mathematical object which takes a real (or complex) variable, or several such variables, and returns a real (or complex) number. A *functional* is a mathematical object which takes an entire function and returns a number. In the case at hand, we have

$$T\big[y(x)\big] = \int\limits_{x_1}^{x_2} dx\, L(y, y', x) \quad , \tag{4.13}$$

where the function $L(y, y', x)$ is given by

$$L(y, y', x) = \frac{1}{v(x)}\sqrt{1+y'^2} \quad . \tag{4.14}$$

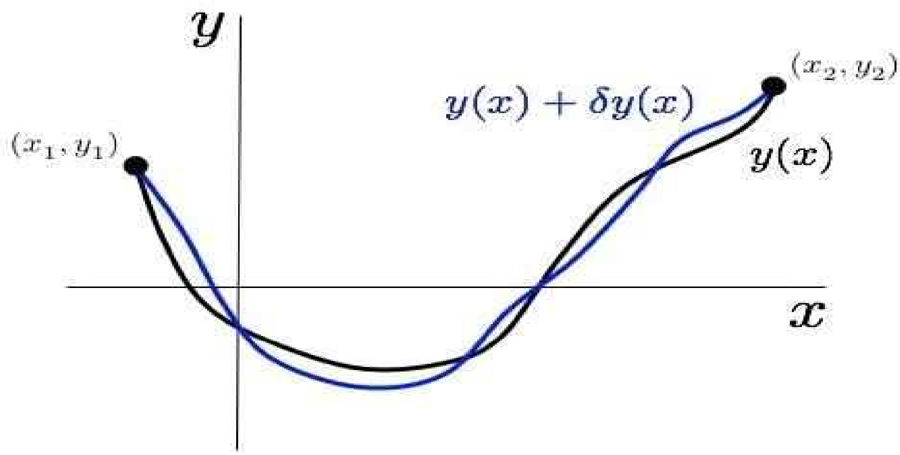

Figure 4.3: A path $y(x)$ and its variation $y(x) + \delta y(x)$.

Here $v(x)$ is a given function characterizing the medium, and $y(x)$ is the path whose time is to be evaluated.

In ordinary calculus, we extremize a function $f(x)$ by demanding that f not change to lowest order when we change $x \to x + dx$:

$$f(x + dx) = f(x) + f'(x)\,dx + \tfrac{1}{2}\,f''(x)\,(dx)^2 + \ldots \quad . \tag{4.15}$$

We say that $x = x^*$ is an extremum when $f'(x^*) = 0$.

For a functional, the first *functional variation* is obtained by sending $y(x) \to y(x) + \delta y(x)$, and extracting the variation in the functional to order δy. Thus, we compute

$$\begin{aligned} T\big[y(x) + \delta y(x)\big] &= \int\limits_{x_1}^{x_2} dx\, L(y + \delta y, y' + \delta y', x) \\ &= \int\limits_{x_1}^{x_2} dx \left\{ L + \frac{\partial L}{\partial y}\,\delta y + \frac{\partial L}{\partial y'}\,\delta y' + \mathcal{O}\big((\delta y)^2\big) \right\} \\ &= T\big[y(x)\big] + \int\limits_{x_1}^{x_2} dx \left\{ \frac{\partial L}{\partial y}\,\delta y + \frac{\partial L}{\partial y'}\,\frac{d}{dx}\,\delta y \right\} \\ &= T\big[y(x)\big] + \int\limits_{x_1}^{x_2} dx \left[\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right)\right] \delta y + \frac{\partial L}{\partial y'}\,\delta y \,\Bigg|_{x_1}^{x_2} \quad . \end{aligned} \tag{4.16}$$

Now one very important thing about the variation $\delta y(x)$ is that it must vanish at the endpoints: $\delta y(x_1) = \delta y(x_2) = 0$. This is because the space of functions under consideration satisfy fixed boundary conditions $y(x_1) = y_1$ and $y(x_2) = y_2$. Thus, the last term in the above equation vanishes, and we have

$$\delta T = \int\limits_{x_1}^{x_2} dx \left[\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right)\right] \delta y \quad . \tag{4.17}$$

We say that the first functional derivative of T with respect to $y(x)$ is

$$\frac{\delta T}{\delta y(x)} = \left[\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right)\right]_x \quad , \tag{4.18}$$

where the subscript indicates that the expression inside the square brackets is to be evaluated at x. The functional $T\big[y(x)\big]$ is *extremized* when its first functional derivative vanishes, which results in a differential equation for $y(x)$,

$$\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) = 0 \quad , \tag{4.19}$$

known as the *Euler-Lagrange* equation.

$L(y, y', x)$ **independent of** y

Suppose $L(y, y', x)$ is independent of y. Then from the Euler-Lagrange equations we have that

$$P \equiv \frac{\partial L}{\partial y'} \tag{4.20}$$

is a constant. In classical mechanics, this will turn out to be a *generalized momentum*. For $L = v^{-1}\sqrt{1+y'^2}$ we have

$$P = \frac{y'}{v\sqrt{1+y'^2}} \quad . \tag{4.21}$$

Setting $dP/dx = 0$, we recover the second order ODE of eqn. 4.9. Solving for y',

$$\frac{dy}{dx} = \pm\frac{v(x)}{\sqrt{v_0^2 - v^2(x)}} \quad , \tag{4.22}$$

where $v_0 = 1/P$.

$L(y, y', x)$ **independent of** x

When $L(y, y', x)$ is independent of x, we can again integrate the Euler-Lagrange equation. Consider the quantity

$$H = y'\,\frac{\partial L}{\partial y'} - L \quad . \tag{4.23}$$

Then

$$\begin{aligned} \frac{dH}{dx} = \frac{d}{dx}\left[y'\,\frac{\partial L}{\partial y'} - L\right] &= y''\,\frac{\partial L}{\partial y'} + y'\,\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) - \frac{\partial L}{\partial y'}\,y'' - \frac{\partial L}{\partial y}\,y' - \frac{\partial L}{\partial x} \\ &= y'\left[\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) - \frac{\partial L}{\partial y}\right] - \frac{\partial L}{\partial x} = -\frac{\partial L}{\partial x} \quad , \end{aligned} \tag{4.24}$$

where we have used the Euler-Lagrange equations to write $\frac{d}{dx}\big(\frac{\partial L}{\partial y'}\big) = \frac{\partial L}{\partial y}$. So if $\partial L/\partial x = 0$, we have $dH/dx = 0$, *i.e.* H is a constant.

4.2.2 Functional Taylor series

In general, we may expand a functional $F[y+\delta y]$ in a *functional Taylor series*,

$$\begin{aligned} F[y+\delta y] = F[y] + \int dx_1\, K_1(x_1)\,\delta y(x_1) + \tfrac{1}{2!}\int dx_1 \int dx_2\, K_2(x_1,x_2)\,\delta y(x_1)\,\delta y(x_2) \\ + \tfrac{1}{3!}\int dx_1 \int dx_2 \int dx_3\, K_3(x_1,x_2,x_3)\,\delta y(x_1)\,\delta y(x_2)\,\delta y(x_3) + \ldots \end{aligned} \tag{4.25}$$

and we write

$$K_n(x_1,\ldots,x_n) \equiv \frac{\delta^n F}{\delta y(x_1)\cdots\delta y(x_n)} \tag{4.26}$$

for the $n^{\rm th}$ functional derivative.

4.2.3 Examples

Here we present three useful examples of variational calculus as applied to problems in mathematics and physics.

Example 1 : minimal surface of revolution

Consider a surface formed by rotating the function $y(x)$ about the x-axis. The area is then

$$A\big[y(x)\big] = \int\limits_{x_1}^{x_2} dx\; 2\pi y \sqrt{1+\left(\frac{dy}{dx}\right)^2} \quad , \tag{4.27}$$

and is a functional of the curve $y(x)$. Thus we can define $L(y,y') = 2\pi y\sqrt{1+y'^2}$ and make the identification $y(x) \leftrightarrow q(t)$. Since $L(y,y',x)$ is independent of x, we have

$$H = y'\,\frac{\partial L}{\partial y'} - L \quad \Rightarrow \quad \frac{dH}{dx} = -\frac{\partial L}{\partial x} \quad , \tag{4.28}$$

and when L has no explicit x-dependence, H is conserved. One finds

$$H = 2\pi y \cdot \frac{y'^2}{\sqrt{1+y'^2}} - 2\pi y\,\sqrt{1+y'^2} = -\frac{2\pi y}{\sqrt{1+y'^2}} \quad . \tag{4.29}$$

Solving for y',

$$\frac{dy}{dx} = \pm\sqrt{\left(\frac{2\pi y}{H}\right)^2 - 1} \quad , \tag{4.30}$$

which may be integrated with the substitution $y = \frac{H}{2\pi}\cosh u$, yielding

$$y(x) = b\,\cosh\!\left(\frac{x-a}{b}\right) \quad , \tag{4.31}$$

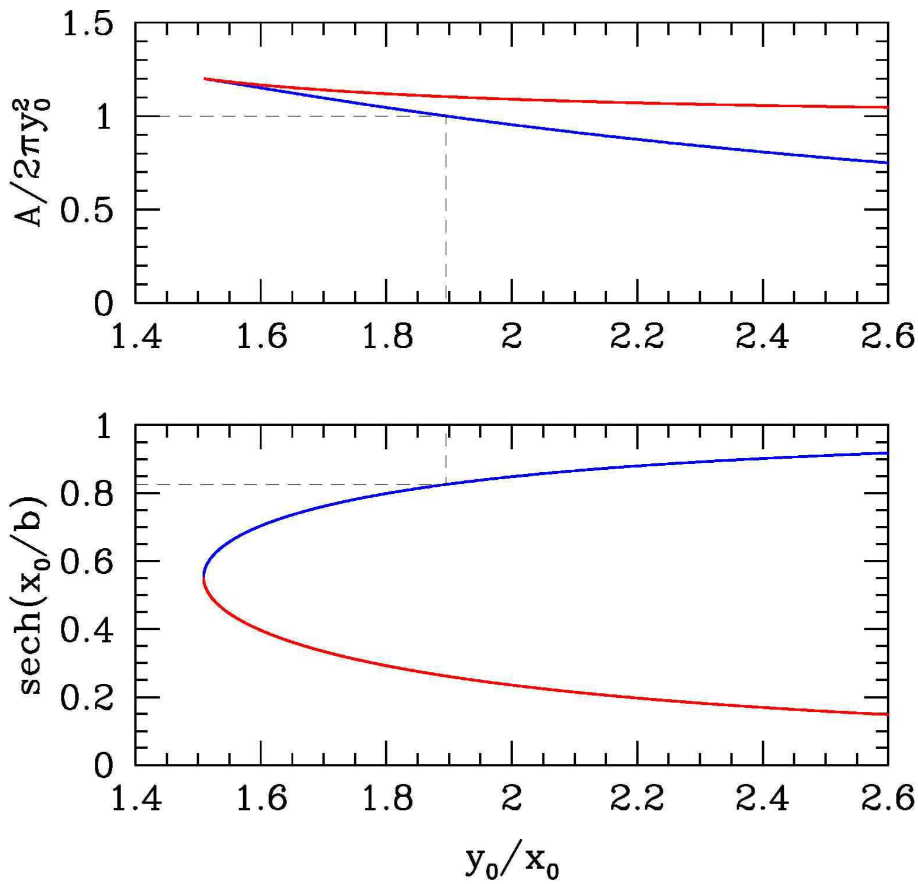

Figure 4.4: Minimal surface solution, with $y(x) = b\cosh(x/b)$ and $y(x_0) = y_0$. Top panel: $A/2\pi y_0^2$ *vs.* y_0/x_0. The discontinuous configuration is shown by the dashed black line. Bottom panel: $\mathrm{sech}(x_0/b)$ *vs.* y_0/x_0. The blue curve corresponds to a global minimum of $A[y(x)]$, and the red curve to a local minimum or saddle point.

where a and $b = \frac{H}{2\pi}$ are constants of integration. Note there are two such constants, as the original equation was second order. This shape is called a *catenary*. As we shall later find, it is also the shape of a uniformly dense rope hanging between two supports, under the influence of gravity. To fix the constants a and b, we invoke the boundary conditions $y(x_1) = y_1$ and $y(x_2) = y_2$.

Consider the case where $-x_1 = x_2 \equiv x_0$ and $y_1 = y_2 \equiv y_0$. Then clearly $a = 0$, and we have

$$y_0 = b\,\cosh\!\left(\frac{x_0}{b}\right) \quad \Rightarrow \quad \gamma = \kappa^{-1}\,\cosh\kappa \quad , \tag{4.32}$$

with $\gamma \equiv y_0/x_0$ and $\kappa \equiv x_0/b$. One finds that for any $\gamma > 1.5089$ there are two solutions, one of which is a global minimum and one of which is a local minimum or saddle of $A[y(x)]$. The solution with the smaller value of κ (*i.e.* the larger value of $\mathrm{sech}\,\kappa$) yields the smaller value of A, as shown in fig. 4.4. Note that

$$\frac{y}{y_0} = \frac{\cosh(x/b)}{\cosh(x_0/b)} \quad , \tag{4.33}$$

so $y(x=0) = y_0\,\mathrm{sech}(x_0/b)$.

When extremizing functions that are defined over a finite or semi-infinite interval, one must take care to evaluate the function at the boundary, for it may be that the boundary yields a global extremum even though the derivative may not vanish there. Similarly, when extremizing functionals, one must investigate the functions at the boundary of function space. In this case, such a function would be the discontinuous solution, with

$$y(x) = \begin{cases} y_1 & \text{if } x = x_1 \\ 0 & \text{if } x_1 < x < x_2 \\ y_2 & \text{if } x = x_2 \quad . \end{cases} \tag{4.34}$$

This solution corresponds to a surface consisting of two discs of radii y_1 and y_2, joined by an infinitesimally thin thread. The area functional evaluated for this particular $y(x)$ is clearly $A = \pi(y_1^2 + y_2^2)$. In fig. 4.4, we plot $A/2\pi y_0^2$ *versus* the parameter $\gamma = y_0/x_0$. For $\gamma > \gamma_c \approx 1.564$, one of the catenary solutions is the global minimum. For $\gamma < \gamma_c$, the minimum area is achieved by the discontinuous solution.

Note that the functional derivative,

$$K_1(x) = \frac{\delta A}{\delta y(x)} = \frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) = \frac{2\pi\big(1 + y'^2 - yy''\big)}{(1 + y'^2)^{3/2}} \quad , \tag{4.35}$$

indeed vanishes for the catenary solutions, but does not vanish for the discontinuous solution, where $K_1(x) = 2\pi$ throughout the interval $(-x_0, x_0)$. Since $y = 0$ on this interval, y cannot be decreased. The fact that $K_1(x) > 0$ means that increasing y will result in an increase in A, so the boundary value for A, which is $2\pi y_0^2$, is indeed a local minimum.

We furthermore see in fig. 4.4 that for $\gamma < \gamma_* \approx 1.5089$ the local minimum and saddle are no longer present. This is the familiar saddle-node bifurcation, here in function space. Thus, for $\gamma \in [0, \gamma_*)$ there are no extrema of $A[y(x)]$, and the minimum area occurs for the discontinuous $y(x)$ lying at the boundary of function space. For $\gamma \in (\gamma_*, \gamma_c)$, two extrema exist, one of which is a local minimum and the other a saddle point. Still, the area is minimized for the discontinuous solution. For $\gamma \in (\gamma_c, \infty)$, the local minimum is the global minimum, and has smaller area than for the discontinuous solution.

Example 2 : geodesic on a surface of revolution

We use cylindrical coordinates (ρ, ϕ, z) on the surface $z = z(\rho)$. Thus,

$$\begin{aligned} ds^2 &= d\rho^2 + \rho^2\, d\phi^2 + dx^2 \\ &= \Big\{1 + \big[z'(\rho)\big]^2\Big\}\, d\rho + \rho^2\, d\phi^2 \quad , \end{aligned} \tag{4.36}$$

and the distance functional $D\big[\phi(\rho)\big]$ is

$$D\big[\phi(\rho)\big] = \int\limits_{\rho_1}^{\rho_2} d\rho\, L(\phi, \phi', \rho) \quad , \tag{4.37}$$

where

$$L(\phi, \phi', \rho) = \sqrt{1 + z'^2(\rho) + \rho^2\, \phi'^2(\rho)} \quad . \tag{4.38}$$

The Euler-Lagrange equation is

$$\frac{\partial L}{\partial \phi} - \frac{d}{d\rho}\left(\frac{\partial L}{\partial \phi'}\right) = 0 \quad \Rightarrow \quad \frac{\partial L}{\partial \phi'} = \text{const.} \tag{4.39}$$

Thus,

$$\frac{\partial L}{\partial \phi'} = \frac{\rho^2\,\phi'}{\sqrt{1+z'^2+\rho^2\,\phi'^2}} = a \quad , \tag{4.40}$$

where a is a constant. Solving for ϕ', we obtain

$$d\phi = \frac{a\sqrt{1+\big[z'(\rho)\big]^2}}{\rho\sqrt{\rho^2-a^2}}\,d\rho \quad , \tag{4.41}$$

which we must integrate to find $\phi(\rho)$, subject to boundary conditions $\phi(\rho_i) = \phi_i$, with $i = 1, 2$.

On a cone, $z(\rho) = \lambda\rho$, and we have

$$d\phi = a\sqrt{1+\lambda^2}\,\frac{d\rho}{\rho\sqrt{\rho^2-a^2}} = \sqrt{1+\lambda^2}\,d\tan^{-1}\sqrt{\frac{\rho^2}{a^2}-1} \quad , \tag{4.42}$$

which yields

$$\phi(\rho) = \beta + \sqrt{1+\lambda^2}\ \tan^{-1}\sqrt{\frac{\rho^2}{a^2}-1} \quad , \tag{4.43}$$

which is equivalent to

$$\rho\,\cos\left(\frac{\phi-\beta}{\sqrt{1+\lambda^2}}\right) = a \quad . \tag{4.44}$$

The constants β and a are determined from $\phi(\rho_i) = \phi_i$.

Example 3 : brachistochrone

Problem: find the path between (x_1, y_1) and (x_2, y_2) which a particle sliding frictionlessly and under constant gravitational acceleration will traverse in the shortest time. To solve this we first must invoke some elementary mechanics. Assuming the particle is released from (x_1, y_1) at rest, energy conservation says

$$\tfrac{1}{2}mv^2 + mgy = mgy_1 \quad . \tag{4.45}$$

Then the time, which is a functional of the curve $y(x)$, is

$$T\big[y(x)\big] =\equiv \int\limits_{x_1}^{x_2} dx\,L(y,y',x) = \int\limits_{x_1}^{x_2} \frac{ds}{v} = \frac{1}{\sqrt{2g}}\int\limits_{x_1}^{x_2} dx\,\sqrt{\frac{1+y'^2}{y_1-y}} \quad , \tag{4.46}$$

with

$$L(y,y',x) = \sqrt{\frac{1+y'^2}{2g(y_1-y)}} \quad . \tag{4.47}$$

Since L is independent of x, eqn. 4.24, we have that

$$H = y' \frac{\partial L}{\partial y'} - L = -\Big[2g\,(y_1 - y)\,\big(1 + y'^2\big)\Big]^{-1/2} \tag{4.48}$$

is conserved. This yields

$$dx = -\sqrt{\frac{y_1 - y}{2a - y_1 + y}}\,dy \quad , \tag{4.49}$$

with $a = (4gH^2)^{-1}$. This may be integrated parametrically, writing

$$y_1 - y = 2a \sin^2\!\big(\tfrac{1}{2}\theta\big) \quad \Rightarrow \quad dx = 2a \sin^2\!\big(\tfrac{1}{2}\theta\big)\,d\theta \quad , \tag{4.50}$$

which results in the parametric equations

$$\begin{aligned} x - x_1 &= a\,\big(\theta - \sin\theta\big) \\ y - y_1 &= -a\,(1 - \cos\theta) \quad . \end{aligned} \tag{4.51}$$

This curve is known as a *cycloid.*

4.2.4 Ocean waves

Surface waves in fluids propagate with a definite relation between their angular frequency ω and their wavevector $k = 2\pi/\lambda$, where λ is the wavelength. The *dispersion relation* is a function $\omega = \omega(k)$. The *group velocity* of the waves is then $v(k) = d\omega/dk$.

In a fluid with a flat bottom at depth h, the dispersion relation turns out to be

$$\omega(k) = \sqrt{gk\,\tanh kh} \approx \begin{cases} \sqrt{gh}\,k & \text{shallow } (kh \ll 1) \\ \\ \sqrt{gk} & \text{deep } (kh \gg 1) \quad . \end{cases} \tag{4.52}$$

Suppose we are in the shallow case, where the wavelength λ is significantly greater than the depth h of the fluid. This is the case for ocean waves which break at the shore. The phase velocity and group velocity are then identical, and equal to $v(h) = \sqrt{gh}$. The waves propagate more slowly as they approach the shore.

Let us choose the following coordinate system: x represents the distance parallel to the shoreline, y the distance perpendicular to the shore (which lies at $y = 0$), and $h(y)$ is the depth profile of the bottom. We assume $h(y)$ to be a slowly varying function of y which satisfies $h(0) = 0$. Suppose a disturbance in the ocean at position (x_2, y_2) propagates until it reaches the shore at $(x_1, y_1 = 0)$. The time of propagation is

$$T\big[y(x)\big] = \int \frac{ds}{v} = \int\limits_{x_1}^{x_2} \! dx\,\sqrt{\frac{1 + y'^2}{g\,h(y)}} \quad . \tag{4.53}$$

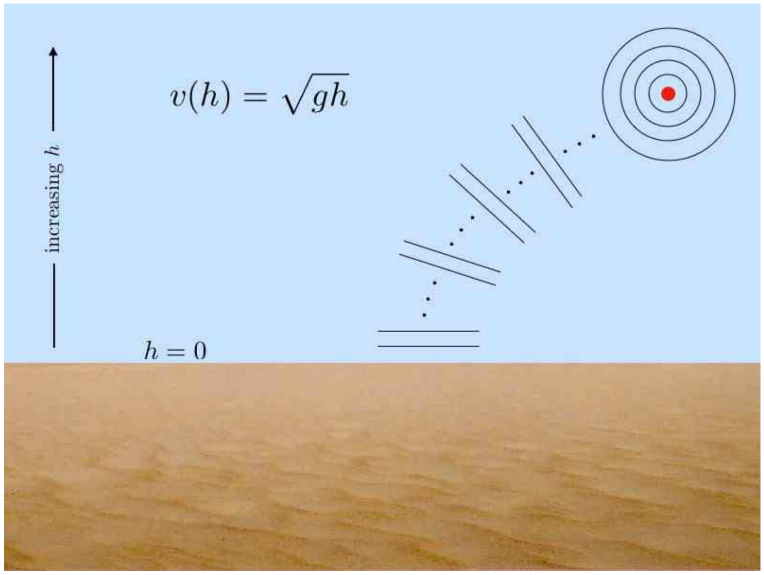

Figure 4.5: For shallow water waves, $v = \sqrt{gh}$. To minimize the propagation time from a source to the shore, the waves break parallel to the shoreline.

We thus identify the integrand

$$L(y, y', x) = \sqrt{\frac{1 + y'^2}{g\,h(y)}} \quad . \tag{4.54}$$

As with the brachistochrone problem, to which this bears an obvious resemblance, L is cyclic in the independent variable x, hence

$$H = y' \frac{\partial L}{\partial y'} - L = -\Big[g\,h(y)\,\big(1 + y'^2\big)\Big]^{-1/2} \tag{4.55}$$

is constant. Solving for $y'(x)$, we have

$$\tan\theta = \frac{dy}{dx} = \sqrt{\frac{a}{h(y)} - 1} \quad , \tag{4.56}$$

where $a = (gH)^{-1}$ is a constant, and where θ is the local slope of the function $y(x)$. Thus, we conclude that near $y = 0$, where $h(y) \to 0$, the waves come in *parallel to the shoreline.* If $h(y) = \alpha y$ has a linear profile, the solution is again a cycloid, with

$$x(\theta) = b\,(\theta - \sin\theta) \quad , \quad y(\theta) = b\,(1 - \cos\theta) \quad , \tag{4.57}$$

where $b = 2a/\alpha$ and where the shore lies at $\theta = 0$. Expanding in a Taylor series in θ for small θ, we may eliminate θ and obtain $y(x)$ as

$$y(x) = \big(\tfrac{9}{2}\big)^{1/3}\, b^{1/3}\, x^{2/3} + \ldots \quad . \tag{4.58}$$

A *tsunami* is a shallow water wave that propagates in deep water. This requires $\lambda > h$, as we've seen, which means the disturbance must have a very long spatial extent out in the open ocean, where $h \sim$

10 km. An undersea earthquake is the only possible source; the characteristic length of earthquake fault lines can be hundreds of kilometers. If we take $h = 10\,\text{km}$, we obtain $v = \sqrt{gh} \approx 310\,\text{m/s}$ or $1100\,\text{km/hr}$. At these speeds, a tsunami can cross the Pacific Ocean in less than a day.

As the wave approaches the shore, it must slow down, since $v = \sqrt{gh}$ is diminishing. But energy is conserved, which means that the amplitude must concomitantly rise. In extreme cases, the water level rise at shore may be 20 meters or more.

4.2.5 More on functionals

We remarked in section 4.2.1 that a function f is an animal which gets fed a real number x and excretes a real number $f(x)$. We say f maps the reals to the reals, or $f\colon \mathbb{R} \to \mathbb{R}$. Of course we also have functions $g\colon \mathbb{C} \to \mathbb{C}$ which eat and excrete complex numbers, multivariable functions $h\colon \mathbb{R}^N \to \mathbb{R}$ which eat N-tuples of numbers and excrete a single number, *etc.*

A *functional* $F[f(x)]$ eats entire functions (!) and excretes numbers. That is,

$$F\colon \Big\{ f(x) \;\Big|\; x \in \mathbb{R} \Big\} \to \mathbb{R} \tag{4.59}$$

We may write $F\colon C(\mathbb{R}) \to \mathbb{R}$, where $C(\mathbb{R})$ is the *space of continuous functions*[1]. This says that F operates on the set of real-valued functions of a single real variable, yielding a real number. Some examples:

$$\begin{aligned}
F[f(x)] &= \tfrac{1}{2}\int\limits_{-\infty}^{\infty} dx\, \big[f(x)\big]^2 \\
F[f(x)] &= \tfrac{1}{2}\int\limits_{-\infty}^{\infty} dx \int\limits_{-\infty}^{\infty} dx'\, K(x,x')\, f(x)\, f(x') \\
F[f(x)] &= \tfrac{1}{2}\int\limits_{-\infty}^{\infty} dx \left[A\, f^2(x) + B\left(\frac{df}{dx}\right)^{\!2}\right] \quad .
\end{aligned} \tag{4.60}$$

In classical mechanics, the action S is a functional of the path $q(t)$:

$$S[q(t)] = \int\limits_{t_a}^{t_b} dt\, \Big\{ \tfrac{1}{2} m \dot{q}^2 - U(q) \Big\} \quad . \tag{4.61}$$

We can also have functionals which feed on functions of more than one independent variable, such as

$$S[y(x,t)] = \int\limits_{t_a}^{t_b} dt \int\limits_{x_a}^{x_b} dx\, \left\{ \tfrac{1}{2}\mu \left(\frac{\partial y}{\partial t}\right)^{\!2} - \tfrac{1}{2}\tau \left(\frac{\partial y}{\partial x}\right)^{\!2} \right\} \quad , \tag{4.62}$$

[1] The notation $C^\infty(\mathbb{R})$ indicates the space of continuous *smooth* (*i.e.* infinitely differentiable) real functions of a real variable.

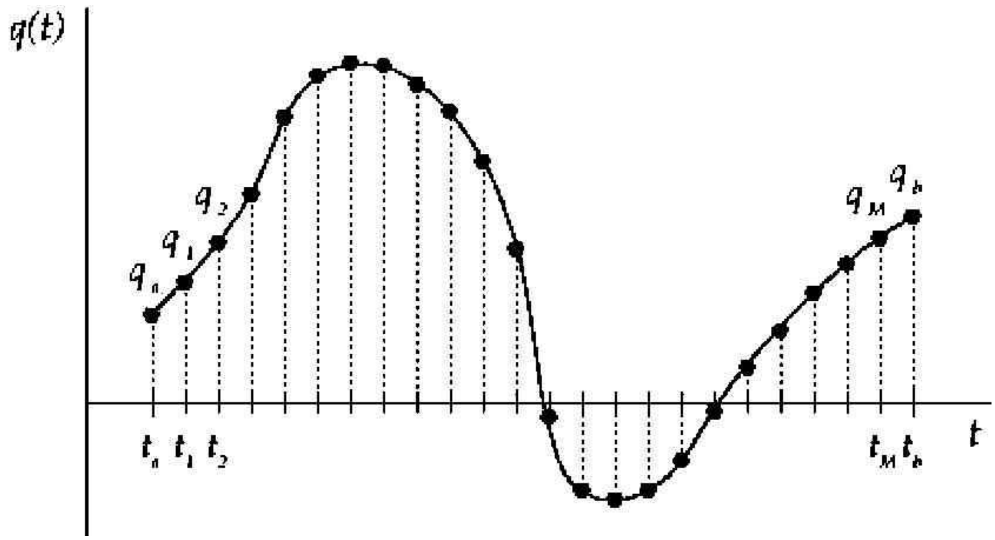

Figure 4.6: A functional $S[q(t)]$ is the continuum limit of a function of a large number of variables, $S(q_1, \ldots, q_M)$.

which happens to be the functional for a string of mass density μ under uniform tension τ. Another example comes from electrodynamics:

$$S[A^\mu(\boldsymbol{x},t)] = -\int dt \int d^3x \left\{ \frac{1}{16\pi} F_{\mu\nu}\, F^{\mu\nu} + \frac{1}{c} j_\mu\, A^\mu \right\} \quad , \tag{4.63}$$

which is a functional of the four fields $\{A^0, A^1, A^2, A^3\}$, where $A^0 = c\phi$. These are the components of the 4-potential, each of which is itself a function of four independent variables (x^0, x^1, x^2, x^3), with $x^0 = ct$. The field strength tensor is written in terms of derivatives of the A^μ: $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, where we use a metric $g_{\mu\nu} = \text{diag}(+,-,-,-)$ to raise and lower indices. The 4-potential couples linearly to the source term J_μ, which is the electric 4-current $(c\rho, \boldsymbol{J})$.

We extremize functions by sending the independent variable x to $x + dx$ and demanding that the variation $df = 0$ to first order in dx. That is,

$$f(x+dx) = f(x) + f'(x)\, dx + \tfrac{1}{2} f''(x)(dx)^2 + \ldots \quad , \tag{4.64}$$

whence $df = f'(x)\, dx + \mathcal{O}\big((dx)^2\big)$ and thus

$$f'(x^*) = 0 \quad \Longleftrightarrow \quad x^* \text{ an extremum.} \tag{4.65}$$

We extremize *functionals* by sending

$$f(x) \to f(x) + \delta f(x) \tag{4.66}$$

and demanding that the variation δF in the functional $F[f(x)]$ vanish to first order in $\delta f(x)$. The variation $\delta f(x)$ must sometimes satisfy certain boundary conditions. For example, if $F[f(x)]$ only operates on functions which vanish at a pair of endpoints, *i.e.* $f(x_a) = f(x_b) = 0$, then when we extremize the functional F we must do so *within the space of allowed functions.* Thus, we would in this case require $\delta f(x_a) = \delta f(x_b) = 0$. We may expand the functional $F[f + \delta f]$ in a *functional Taylor series*,

$$\begin{aligned} F[f+\delta f] = F[f] &+ \int dx_1\, K_1(x_1)\, \delta f(x_1) + \tfrac{1}{2!} \int dx_1 \int dx_2\, K_2(x_1,x_2)\, \delta f(x_1)\, \delta f(x_2) \\ &+ \tfrac{1}{3!} \int dx_1 \int dx_2 \int dx_3\, K_3(x_1,x_2,x_3)\, \delta f(x_1)\, \delta f(x_2)\, \delta f(x_3) + \ldots \end{aligned} \tag{4.67}$$

and we write

$$K_n(x_1, \ldots, x_n) \equiv \frac{\delta^n F}{\delta f(x_1) \cdots \delta f(x_n)} \quad . \tag{4.68}$$

In a more general case, $F = F\big[\{f_i(\boldsymbol{x})\}\big]$ is a functional of several functions, each of which is a function of several independent variables[2]. We then write

$$\begin{aligned} F[\{f_i+\delta f_i\}] = F[\{f_i\}] &+ \int d\boldsymbol{x}_1\, K_1^i(\boldsymbol{x}_1)\, \delta f_i(\boldsymbol{x}_1) + \frac{1}{2!} \int d\boldsymbol{x}_1 \int d\boldsymbol{x}_2\, K_2^{ij}(\boldsymbol{x}_1,\boldsymbol{x}_2)\, \delta f_i(\boldsymbol{x}_1)\, \delta f_j(\boldsymbol{x}_2) \\ &+ \frac{1}{3!} \int d\boldsymbol{x}_1 \int d\boldsymbol{x}_2 \int d\boldsymbol{x}_3\; K_3^{ijk}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)\, \delta f_i(\boldsymbol{x}_1)\, \delta f_j(\boldsymbol{x}_2)\, \delta f_k(\boldsymbol{x}_3) + \ldots \quad , \end{aligned} \tag{4.69}$$

with

$$K_n^{i_1 i_2 \cdots i_n}(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) = \frac{\delta^n F}{\delta f_{i_1}(\boldsymbol{x}_1)\, \delta f_{i_2}(\boldsymbol{x}_2)\, \delta f_{i_n}(\boldsymbol{x}_n)} \quad . \tag{4.70}$$

Another way to compute functional derivatives is to send

$$f(x) \to f(x) + \epsilon_1\, \delta(x - x_1) + \ldots + \epsilon_n\, \delta(x - x_n) \tag{4.71}$$

and then differentiate n times with respect to ϵ_1 through ϵ_n. That is,

$$\frac{\delta^n F}{\delta f(x_1) \cdots \delta f(x_n)} = \frac{\partial^n}{\partial\epsilon_1 \cdots \partial\epsilon_n} \bigg|_{\epsilon_1=\epsilon_2=\cdots\epsilon_n=0} F\big[f(x) + \epsilon_1\, \delta(x-x_1) + \ldots + \epsilon_n\, \delta(x - x_n)\big] \quad . \tag{4.72}$$

Let's see how this works. As an example, we'll take the action functional from classical mechanics,

$$S[q(t)] = \int\limits_{t_a}^{t_b} dt\, \Big\{ \tfrac{1}{2} m\dot{q}^2 - U(q) \Big\} \quad . \tag{4.73}$$

To compute the first functional derivative, we replace the function $q(t)$ with $q(t) + \epsilon\, \delta(t - t_1)$, and expand in powers of ϵ:

$$\begin{aligned} S\big[q(t) + \epsilon\delta(t-t_1)\big] &= S[q(t)] + \epsilon \int\limits_{t_a}^{t_b} dt\, \Big\{ m\dot{q}\, \delta'(t-t_1) - U'(q)\, \delta(t-t_1) \Big\} \\ &= -\epsilon \Big\{ m\ddot{q}(t_1) + U'\big(q(t_1)\big) \Big\} \quad , \end{aligned} \tag{4.74}$$

[2]It may be also be that different functions depend on a different number of independent variables. *E.g.* $F = F[f(x),\, g(x,y),\, h(x,y,z)]$.

hence

$$\frac{\delta S}{\delta q(t)} = -\Big\{ m\ddot{q}(t) + U'\big(q(t)\big)\Big\} \tag{4.75}$$

and setting the first functional derivative to zero yields Newton's Second Law, $m\ddot{q} = -U'(q)$, for all $t \in [t_a, t_b]$. Note that we have used the result

$$\int\limits_{-\infty}^{\infty} dt\, \delta'(t - t_1)\, h(t) = -h'(t_1) \quad , \tag{4.76}$$

which is easily established upon integration by parts.

To compute the second functional derivative, we replace

$$q(t) \to q(t) + \epsilon_1\, \delta(t - t_1) + \epsilon_2\, \delta(t - t_2) \tag{4.77}$$

and extract the term of order $\epsilon_1\, \epsilon_2$ in the double Taylor expansion. One finds this term to be

$$\epsilon_1\, \epsilon_2 \int\limits_{t_a}^{t_b} dt\, \Big\{ m\, \delta'(t - t_1)\, \delta'(t - t_2) - U''(q)\, \delta(t - t_1)\, \delta(t - t_2)\Big\} \quad . \tag{4.78}$$

Note that we needn't bother with terms proportional to ϵ_1^2 or ϵ_2^2 since the recipe is to differentiate once with respect to each of ϵ_1 and ϵ_2 and then to set $\epsilon_1 = \epsilon_2 = 0$. This procedure uniquely selects the term proportional to $\epsilon_1\, \epsilon_2$, and yields

$$\frac{\delta^2 S}{\delta q(t_1)\, \delta q(t_2)} = -\Big\{ m\, \delta''(t_1 - t_2) + U''\big(q(t_1)\big)\, \delta(t_1 - t_2)\Big\} \quad . \tag{4.79}$$

In multivariable calculus, the stability of an extremum is assessed by computing the matrix of second derivatives at the extremal point, known as the Hessian matrix. One has

$$\frac{\partial f}{\partial x_i}\bigg|_{\boldsymbol{x}^*} = 0 \quad \forall\, i \qquad ; \qquad H_{ij} = \frac{\partial^2 f}{\partial x_i\, \partial x_j}\bigg|_{\boldsymbol{x}^*} \quad . \tag{4.80}$$

The eigenvalues of the Hessian H_{ij} determine the stability of the extremum. Since H_{ij} is a symmetric matrix, its eigenvectors η^α may be chosen to be orthogonal. The associated eigenvalues λ_α, defined by the equation

$$H_{ij}\, \eta_j^\alpha = \lambda_\alpha\, \eta_i^\alpha \quad , \tag{4.81}$$

are the respective curvatures in the directions η^α, where $\alpha \in \{1, \ldots, n\}$ where n is the number of variables. The extremum is a local minimum if all the eigenvalues λ_α are positive, a maximum if all are negative, and otherwise is a saddle point. Near a saddle point, there are some directions in which the function increases and some in which it decreases.

In the case of functionals, the second functional derivative $K_2(x_1, x_2)$ defines an eigenvalue problem for $\delta f(x)$:

$$\int\limits_{x_a}^{x_b} dx_2\, K_2(x_1, x_2)\, \delta f(x_2) = \lambda\, \delta f(x_1) \quad . \tag{4.82}$$

In general there are an infinite number of solutions to this equation which form a basis in function space, subject to appropriate boundary conditions at x_a and x_b. For example, in the case of the action functional from classical mechanics, the above eigenvalue equation becomes a differential equation,

$$-\left\{m\frac{d^2}{dt^2} + U''\big(q^*(t)\big)\right\}\delta q(t) = \lambda\,\delta q(t) \quad , \tag{4.83}$$

where $q^*(t)$ is the solution to the Euler-Lagrange equations. As with the case of ordinary multivariable functions, the functional extremum is a local minimum (in function space) if every eigenvalue λ_α is positive, a local maximum if every eigenvalue is negative, and a saddle point otherwise.

Consider the simple harmonic oscillator, for which $U(q) = \frac{1}{2}\,m\omega_0^2\,q^2$. Then $U''\big(q^*(t)\big) = m\,\omega_0^2$; note that we don't even need to know the solution $q^*(t)$ to obtain the second functional derivative in this special case. The eigenvectors obey $m(\delta\ddot{q} + \omega_0^2\,\delta q) = -\lambda\,\delta q$, hence

$$\delta q(t) = A\cos\left(\sqrt{\omega_0^2 + (\lambda/m)}\;t + \varphi\right) \quad , \tag{4.84}$$

where A and φ are constants. Demanding $\delta q(t_a) = \delta q(t_b) = 0$ requires

$$\sqrt{\omega_0^2 + (\lambda/m)}\,\big(t_b - t_a\big) = n\pi \quad , \tag{4.85}$$

where n is an integer. Thus, the eigenfunctions are

$$\delta q_n(t) = A\sin\left(n\pi\cdot\frac{t - t_a}{t_b - t_a}\right) \quad , \tag{4.86}$$

and the eigenvalues are

$$\lambda_n = m\Big(\frac{n\pi}{T}\Big)^2 - m\omega_0^2 \quad , \tag{4.87}$$

where $T = t_b - t_a$. Thus, so long as $T > \pi/\omega_0$, there is at least one negative eigenvalue. Indeed, for $\frac{n\pi}{\omega_0} < T < \frac{(n+1)\pi}{\omega_0}$ there will be n negative eigenvalues. This means the action is generally not a minimum, but rather lies at a *saddle point* in the (infinite-dimensional) function space.

To test this explicitly, consider a harmonic oscillator with the boundary conditions $q(0) = 0$ and $q(T) = Q$. The equations of motion, $\ddot{q} + \omega_0^2\,q = 0$, along with the boundary conditions, determine the motion,

$$q^*(t) = \frac{Q\,\sin(\omega_0 t)}{\sin(\omega_0 T)} \quad . \tag{4.88}$$

The action for this path is then

$$\begin{aligned} S[q^*(t)] &= \int\limits_0^T\! dt\,\Big(\tfrac{1}{2}m\dot{q}^{*2} - \tfrac{1}{2}m\omega_0^2\,q^{*2}\Big) \\ &= \frac{m\,\omega_0^2\,Q^2}{2\sin^2\!\omega_0 T}\int\limits_0^T\! dt\,\Big(\cos^2\!\omega_0 t - \sin^2\!\omega_0 t\Big) = \tfrac{1}{2}m\omega_0\,Q^2\,\operatorname{ctn}(\omega_0 T) \quad . \end{aligned} \tag{4.89}$$

Next consider the path $q(t) = Q\,t/T$ which satisfies the boundary conditions but does not satisfy the equations of motion (it proceeds with constant velocity). One finds the action for this path is

$$S[q(t)] = \tfrac{1}{2} m\omega_0\, Q^2 \left(\frac{1}{\omega_0 T} - \tfrac{1}{3}\omega_0 T \right) \quad . \tag{4.90}$$

Thus, provided $\omega_0 T \neq n\pi$, in the limit $T \to \infty$ we find that the constant velocity path has lower action.

Finally, consider the general mechanical action,

$$S\big[q(t)\big] = \int\limits_{t_a}^{t_b} dt\, L(q, \dot{q}, t) \quad . \tag{4.91}$$

We now evaluate the first few terms in the functional Taylor series:

$$\begin{aligned} S\big[q^*(t) + \delta q(t)\big] = \int\limits_{t_a}^{t_b} dt \Bigg\{ & L(q^*, \dot{q}^*, t) + \frac{\partial L}{\partial q_i}\bigg|_{q^*} \delta q_i + \frac{\partial L}{\partial \dot{q}_i}\bigg|_{q^*} \delta \dot{q}_i \\ & + \frac{1}{2}\, \frac{\partial^2 L}{\partial q_i\, \partial q_j}\bigg|_{q^*} \delta q_i\, \delta q_j + \frac{\partial^2 L}{\partial q_i\, \partial \dot{q}_j}\bigg|_{q^*} \delta q_i\, \delta \dot{q}_j + \frac{1}{2}\, \frac{\partial^2 L}{\partial \dot{q}_i\, \partial \dot{q}_j}\bigg|_{q^*} \delta \dot{q}_i\, \delta \dot{q}_j + \ldots \Bigg\} \quad . \end{aligned} \tag{4.92}$$

To identify the functional derivatives, we integrate by parts. Let $\Phi_{\ldots}(t)$ be an arbitrary function of time. Then

$$\int\limits_{t_a}^{t_b} dt\, \Phi_i(t)\, \delta \dot{q}_i(t) = -\int\limits_{t_a}^{t_b} dt\, \dot{\Phi}_i(t)\, \delta q_i(t) \tag{4.93}$$

and

$$\begin{aligned} \int\limits_{t_a}^{t_b} dt\ \Phi_{ij}(t)\, \delta q_i(t)\, \delta \dot{q}_j(t) &= \int\limits_{t_a}^{t_b} dt \int\limits_{t_a}^{t_b} dt'\ \Phi_{ij}(t)\, \delta(t-t')\, \frac{d}{dt'}\, \delta q_i(t)\, \delta q_j(t') \\ &= \int\limits_{t_a}^{t_b} dt \int\limits_{t_a}^{t_b} dt'\ \Phi_{ij}(t))\, \delta'(t-t')\, \delta q_i(t)\, \delta q_j(t') \quad , \end{aligned} \tag{4.94}$$

and

$$\begin{aligned} \int\limits_{t_a}^{t_b} dt\ \Phi_{ij}(t)\, d\dot{q}_i(t)\, \delta \dot{q}_j(t) &= \int\limits_{t_a}^{t_b} dt \int\limits_{t_a}^{t_b} dt'\ \Phi_{ij}(t)\, \delta(t-t')\, \frac{d}{dt}\, \frac{d}{dt'}\, \delta q_i(t)\, \delta q_j(t') \\ &= -\int\limits_{t_a}^{t_b} dt \int\limits_{t_a}^{t_b} dt' \Big[\dot{\Phi}_{ij}(t)\, \delta'(t-t') + \Phi_{ij}(t)\, \delta''(t-t') \Big]\, \delta q_i(t)\, \delta q_j(t') \quad . \end{aligned} \tag{4.95}$$

Thus, the first two functional derivatives are given by

$$\frac{\delta S}{\delta q_i(t)} = \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\right]_{q^*(t)} \tag{4.96}$$

and

$$\begin{aligned}\frac{\delta^2 S}{\delta q_i(t)\, \delta q_j(t')} = \Bigg\{ &\frac{\partial^2 L}{\partial q_i\, \partial q_j}\bigg|_{q^*(t)} \delta(t-t') - \frac{\partial^2 L}{\partial \dot{q}_i\, \partial \dot{q}_j}\bigg|_{q^*(t)} \delta''(t-t') \\ &+ \left[2\,\frac{\partial^2 L}{\partial q_i\, \partial \dot{q}_j} - \frac{d}{dt}\left(\frac{\partial^2 L}{\partial \dot{q}_i\, \partial \dot{q}_j}\right)\right]_{q^*(t)} \delta'(t-t')\Bigg\} \quad .\end{aligned} \tag{4.97}$$

4.3 Lagrangian Mechanics

4.3.1 Generalized coordinates

A set of *generalized coordinates* $q_1, \ldots, q_n$ completely describes the positions of all particles in a mechanical system. In a system with $d_{\rm f}$ degrees of freedom and k constraints, $n = d_{\rm f} - k$ independent generalized coordinates are needed to completely specify all the positions. A constraint is a relation among coordinates, such as $x^2 + y^2 + z^2 = a^2$ for a particle moving on a sphere of radius a. In this case, $d_{\rm f} = 3$ and $k = 1$. In this case, we could eliminate z in favor of x and y, *i.e.* by writing $z = \pm\sqrt{a^2 - x^2 - y^2}$, or we could choose as coordinates the polar and azimuthal angles θ and ϕ.

For the moment we will assume that $n = d_{\rm f} - k$, and that the generalized coordinates are independent, satisfying no additional constraints among them. Later on we will learn how to deal with any remaining constraints among the $\{q_1, \ldots, q_n\}$.

The generalized coordinates may have units of length, or angle, or perhaps something totally different. In the theory of small oscillations, the normal coordinates are conventionally chosen to have units of $({\rm mass})^{1/2} \times ({\rm length})$. However, once a choice of generalized coordinate is made, with a concomitant set of units, the units of the conjugate momentum and force are determined:

$$\big[p_\sigma\big] = \frac{ML^2}{T} \cdot \frac{1}{\big[q_\sigma\big]} \qquad , \qquad \big[F_\sigma\big] = \frac{ML^2}{T^2} \cdot \frac{1}{\big[q_\sigma\big]} \quad , \tag{4.98}$$

where $\big[A\big]$ means 'the units of A', and where M, L, and T stand for mass, length, and time, respectively. Thus, if q_σ has dimensions of length, then p_σ has dimensions of momentum and F_σ has dimensions of force. If q_σ is dimensionless, as is the case for an angle, p_σ has dimensions of angular momentum (ML^2/T) and F_σ has dimensions of torque (ML^2/T^2).

4.3.2 Hamilton's principle

The equations of motion of classical mechanics are embodied in a variational principle, called *Hamilton's principle.* Hamilton's principle states that the motion of a system is such that the *action functional*

$$S\big[q(t)\big] = \int\limits_{t_1}^{t_2} dt\, L(q, \dot{q}, t) \tag{4.99}$$

is an extremum, *i.e.* $\delta S = 0$. Here, $q = \{q_1, \ldots, q_n\}$ is a complete set of *generalized coordinates* for our mechanical system, and

$$L = T - U \tag{4.100}$$

is the *Lagrangian*, where T is the kinetic energy and U is the potential energy. Setting the first variation of the action to zero gives the Euler-Lagrange equations,

$$\frac{d}{dt}\overbrace{\left(\frac{\partial L}{\partial \dot{q}_\sigma}\right)}^{\text{momentum } p_\sigma} = \overbrace{\frac{\partial L}{\partial q_\sigma}}^{\text{force } F_\sigma} \quad . \tag{4.101}$$

Thus, we have the familiar $\dot{p}_\sigma = F_\sigma$, also known as Newton's second law. Note, however, that the $\{q_\sigma\}$ are *generalized coordinates*, so p_σ may not have dimensions of momentum, nor F_σ of force. For example, if the generalized coordinate in question is an angle ϕ, then the corresponding generalized momentum is the angular momentum about the axis of ϕ's rotation, and the generalized force is the torque.

4.3.3 Invariance of the equations of motion

Suppose

$$\tilde{L}(q, \dot{q}, t) = L(q, \dot{q}, t) + \frac{d}{dt}\, G(q, t) \quad . \tag{4.102}$$

Then

$$\tilde{S}[q(t)] = S[q(t)] + G(q_b, t_b) - G(q_a, t_a) \quad . \tag{4.103}$$

Since the difference $\tilde{S} - S$ is a function only of the endpoint values $\{q_a, q_b\}$, their variations are identical: $\delta\tilde{S} = \delta S$. This means that L and $\tilde{L}$ result in the same equations of motion. Thus, the equations of motion are invariant under a shift of L by a total time derivative of a function of coordinates and time.

4.3.4 Remarks on the order of the equations of motion

The equations of motion are second order in time. This follows from the fact that $L = L(q, \dot{q}, t)$. Using the chain rule,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\sigma}\right) = \frac{\partial^2 L}{\partial \dot{q}_\sigma\, \partial \dot{q}_{\sigma'}}\, \ddot{q}_{\sigma'} + \frac{\partial^2 L}{\partial \dot{q}_\sigma\, \partial q_{\sigma'}}\, \dot{q}_{\sigma'} + \frac{\partial^2 L}{\partial \dot{q}_\sigma\, \partial t} \quad . \tag{4.104}$$

That the equations are second order in time can be regarded as an empirical fact. It follows, as we have just seen, from the fact that L depends on q and on $\dot{q}$, but on no higher time derivative terms. Suppose the Lagrangian did depend on the generalized accelerations $\ddot{q}$ as well. What would the equations of motion look like? Taking the variation of S,

$$\delta\int_{t_a}^{t_b} dt\, L(q,\dot{q},\ddot{q},t) = \left[\frac{\partial L}{\partial \dot{q}_\sigma}\,\delta q_\sigma + \frac{\partial L}{\partial \ddot{q}_\sigma}\,\delta\dot{q}_\sigma - \frac{d}{dt}\bigg(\frac{\partial L}{\partial \ddot{q}_\sigma}\bigg)\,\delta q_\sigma\right]_{t_a}^{t_b} + \int_{t_a}^{t_b} dt\,\bigg\{\frac{\partial L}{\partial q_\sigma} - \frac{d}{dt}\bigg(\frac{\partial L}{\partial \dot{q}_\sigma}\bigg) + \frac{d^2}{dt^2}\bigg(\frac{\partial L}{\partial \ddot{q}_\sigma}\bigg)\bigg\}\,\delta q_\sigma \quad . \tag{4.105}$$

The boundary term vanishes if we require $\delta q_\sigma(t_a) = \delta q_\sigma(t_b) = \delta\dot{q}_\sigma(t_a) = \delta\dot{q}_\sigma(t_b) = 0\ \forall\,\sigma$. The equations of motion would then be *fourth order* in time.

4.3.5 Lagrangian for a free particle

For a free particle, we can use Cartesian coordinates for each particle as our system of generalized coordinates. For a single particle, the Lagrangian $L(\boldsymbol{x},\boldsymbol{v},t)$ must be a function solely of $\boldsymbol{v}^2$. This is because homogeneity with respect to space and time preclude any dependence of L on $\boldsymbol{x}$ or on t, and isotropy of space means L must depend on $\boldsymbol{v}^2$. We next invoke Galilean relativity, which says that the equations of motion are invariant under transformation to a reference frame moving with constant velocity. Let $\boldsymbol{V}$ be the velocity of the new reference frame $\mathcal{K}'$ relative to our initial reference frame $\mathcal{K}$. Then $\boldsymbol{x}' = \boldsymbol{x} - \boldsymbol{V}t$, and $\boldsymbol{v}' = \boldsymbol{v} - \boldsymbol{V}$. In order that the equations of motion be invariant under the change in reference frame, we demand

$$L'(\boldsymbol{v}') = L(\boldsymbol{v}) + \frac{d}{dt}\,G(\boldsymbol{x},t) \quad . \tag{4.106}$$

The only possibility is $L = \frac{1}{2}m\boldsymbol{v}^2$, where the constant m is the mass of the particle. Note:

$$L' = \tfrac{1}{2}m(\boldsymbol{v}-\boldsymbol{V})^2 = \tfrac{1}{2}m\boldsymbol{v}^2 + \frac{d}{dt}\Big(\tfrac{1}{2}m\boldsymbol{V}^2\,t - m\boldsymbol{V}\cdot\boldsymbol{x}\Big) = L + \frac{dG}{dt} \quad . \tag{4.107}$$

For N interacting particles,

$$L = \tfrac{1}{2}\sum_{a=1}^{N} m_a\left(\frac{d\boldsymbol{x}_a}{dt}\right)^2 - U\big(\{\boldsymbol{x}_a\},\{\dot{\boldsymbol{x}}_a\}\big) \quad . \tag{4.108}$$

Here, U is the *potential energy*. Generally, U is of the form

$$U = \sum_a U_1(\boldsymbol{x}_a) + \sum_{a<a'} v(\boldsymbol{x}_a - \boldsymbol{x}_{a'}) \quad , \tag{4.109}$$

however, as we shall see, velocity-dependent potentials appear in the case of charged particles interacting with electromagnetic fields. In general, though,

$$L = T - U \quad , \tag{4.110}$$

where T is the kinetic energy, and U is the potential energy.

4.3.6 Conserved quantities

A conserved quantity $\Lambda(q, \dot{q}, t)$ is one which does not vary throughout the motion of the system. This means

$$\frac{d\Lambda}{dt}\bigg|_{q=q(t)} = 0 \quad . \tag{4.111}$$

We shall discuss conserved quantities in detail in the chapter on Noether's Theorem, which follows.

Momentum conservation

The simplest case of a conserved quantity occurs when the Lagrangian does not explicitly depend on one or more of the generalized coordinates, *i.e.* when

$$F_\sigma = \frac{\partial L}{\partial q_\sigma} = 0 \quad . \tag{4.112}$$

We then say that L is *cyclic* in the coordinate q_σ. In this case, the Euler-Lagrange equations $\dot{p}_\sigma = F_\sigma$ say that the conjugate momentum p_σ is conserved. Consider, for example, the motion of a particle of mass m near the surface of the earth. Let (x, y) be coordinates parallel to the surface and z the height. We then have

$$\begin{aligned} T &= \tfrac{1}{2} m \big(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\big) \\ U &= mgz \\ L = T - U &= \tfrac{1}{2} m \big(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\big) - mgz \quad . \end{aligned} \tag{4.113}$$

Since

$$F_x = \frac{\partial L}{\partial x} = 0 \quad \text{and} \quad F_y = \frac{\partial L}{\partial y} = 0 \quad , \tag{4.114}$$

we have that p_x and p_y are conserved, with

$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x} \quad , \quad p_y = \frac{\partial L}{\partial \dot{y}} = m\dot{y} \quad . \tag{4.115}$$

These first order equations can be integrated to yield

$$x(t) = x(0) + \frac{p_x}{m}\, t \quad , \quad y(t) = y(0) + \frac{p_y}{m}\, t \quad . \tag{4.116}$$

The z equation is of course

$$\dot{p}_z = m\ddot{z} = -mg = F_z \quad , \tag{4.117}$$

with solution

$$z(t) = z(0) + \dot{z}(0)\, t - \tfrac{1}{2} g t^2 \quad . \tag{4.118}$$

As another example, consider a particle moving in the (x, y) plane under the influence of a potential $U(x, y) = U\big(\sqrt{x^2 + y^2}\big)$ which depends only on the particle's distance from the origin $\rho = \sqrt{x^2 + y^2}$. The Lagrangian, expressed in two-dimensional polar coordinates (ρ, ϕ), is

$$L = \tfrac{1}{2} m \big(\dot{\rho}^2 + \rho^2 \dot{\phi}^2\big) - U(\rho) \quad . \tag{4.119}$$

We see that L is cyclic in the angle ϕ, hence

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = m\rho^2 \dot{\phi} \tag{4.120}$$

is conserved. p_ϕ is the angular momentum of the particle about the $\hat{\boldsymbol{z}}$ axis. In the language of the calculus of variations, momentum conservation is what follows when the integrand of a functional is independent of the *independent variable*.

Energy conservation

When the integrand of a functional is independent of the *dependent* variable, another conservation law follows. For Lagrangian mechanics, consider the expression

$$H(q, \dot{q}, t) = \sum_{\sigma=1}^{n} p_\sigma \, \dot{q}_\sigma - L \quad . \tag{4.121}$$

Now we take the total time derivative of H:

$$\frac{dH}{dt} = \sum_{\sigma=1}^{n} \left\{ p_\sigma \, \ddot{q}_\sigma + \dot{p}_\sigma \, \dot{q}_\sigma - \frac{\partial L}{\partial q_\sigma} \, \dot{q}_\sigma - \frac{\partial L}{\partial \dot{q}_\sigma} \, \ddot{q}_\sigma \right\} - \frac{\partial L}{\partial t} \quad . \tag{4.122}$$

We evaluate $\dot{H}$ along the motion of the system, which entails that the terms in the curly brackets above cancel for each σ:

$$p_\sigma = \frac{\partial L}{\partial \dot{q}_\sigma} \quad , \quad \dot{p}_\sigma = \frac{\partial L}{\partial q_\sigma} \quad . \tag{4.123}$$

Thus, we find

$$\frac{dH}{dt} = -\frac{\partial L}{\partial t} \quad , \tag{4.124}$$

which means that H is conserved *whenever the Lagrangian contains no explicit time dependence*. For a Lagrangian of the form

$$L = \sum_a \tfrac{1}{2} \, m_a \dot{\boldsymbol{r}}_a^2 - U(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \quad , \tag{4.125}$$

we have that $\boldsymbol{p}_a = m_a \, \dot{\boldsymbol{r}}_a$, and

$$H = T + U = \sum_a \tfrac{1}{2} \, m_a \dot{\boldsymbol{r}}_a^2 + U(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \quad . \tag{4.126}$$

However, it is not always the case that $H = T + U$ is the total energy, as we shall further on below.

4.3.7 Choosing generalized coordinates

Any choice of generalized coordinates will yield an equivalent set of equations of motion. However, some choices result in an apparently simpler set than others. This is often true with respect to the form of the potential energy. Additionally, certain constraints that may be present are more amenable to treatment using a particular set of generalized coordinates.

The kinetic energy T is always simple to write in Cartesian coordinates, and it is good practice, at least when one is first learning the method, to write T in Cartesian coordinates and then convert to generalized coordinates. In Cartesian coordinates, the kinetic energy of a single particle of mass m is

$$T = \tfrac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) \quad . \tag{4.127}$$

If the motion is two-dimensional, and confined to the plane $z = \text{const.}$, one of course has $T = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right)$.

Two other commonly used coordinate systems are the cylindrical and spherical systems. In cylindrical coordinates (ρ, ϕ, z), ρ is the radial coordinate in the (x, y) plane and ϕ is the azimuthal angle:

$$x = \rho\,\cos\phi \quad , \quad y = \rho\,\sin\phi \quad , \quad \dot{x} = \cos\phi\,\dot{\rho} - \rho\sin\phi\,\dot{\phi} \qquad \dot{y} = \sin\phi\,\dot{\rho} + \rho\cos\phi\,\dot{\phi} \quad , \tag{4.128}$$

and the third, orthogonal coordinate is of course z. The kinetic energy is

$$T = \tfrac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + \dot{x}^2\right) = \tfrac{1}{2}m\left(\dot{\rho}^2 + \rho^2\,\dot{\phi}^2 + \dot{z}^2\right) \quad . \tag{4.129}$$

When the motion is confined to a plane with $z = \text{const.}$, this coordinate system is often referred to as 'two-dimensional polar' coordinates.

In spherical coordinates (r, θ, ϕ), r is the radius, θ is the polar angle, and ϕ is the azimuthal angle. On the globe, θ would be the 'colatitude', which is $\theta = \frac{\pi}{2} - \lambda$, where λ is the latitude. *I.e.* $\theta = 0$ at the north pole. In spherical polar coordinates,

$$x = r\,\sin\theta\,\cos\phi \qquad \dot{x} = \sin\theta\,\cos\phi\,\dot{r} + r\,\cos\theta\,\cos\phi\,\dot{\theta} - r\,\sin\theta\sin\phi\,\dot{\phi} \tag{4.130}$$

$$y = r\,\sin\theta\,\sin\phi \qquad \dot{y} = \sin\theta\,\sin\phi\,\dot{r} + r\,\cos\theta\,\sin\phi\,\dot{\theta} + r\,\sin\theta\cos\phi\,\dot{\phi} \tag{4.131}$$

$$z = r\,\cos\theta \qquad \dot{z} = \cos\theta\,\dot{r} - r\,\sin\theta\,\dot{\theta} \quad . \tag{4.132}$$

The kinetic energy is

$$T = \tfrac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) = \tfrac{1}{2}m\left(\dot{r}^2 + r^2\,\dot{\theta}^2 + r^2\,\sin^2\!\theta\,\dot{\phi}^2\right) \quad . \tag{4.133}$$

4.4 How to Solve Mechanics Problems

Here are some simple steps you can follow toward obtaining the equations of motion:

1. Choose a set of generalized coordinates $\{q_1, \ldots, q_n\}$.
2. Find the kinetic energy $T(q, \dot{q}, t)$, the potential energy $U(q, t)$, and the Lagrangian $L(q, \dot{q}, t) = T - U$. It is often helpful to first write the kinetic energy in Cartesian coordinates for each particle before converting to generalized coordinates.
3. Find the canonical momenta $p_\sigma = \frac{\partial L}{\partial \dot{q}_\sigma}$ and the generalized forces $F_\sigma = \frac{\partial L}{\partial q_\sigma}$.

4. Identify any conserved quantities (more about this later).
5. Evaluate the time derivatives $\dot{p}_\sigma$ and write the equations of motion $\dot{p}_\sigma = F_\sigma$. Be careful to differentiate properly, using the chain rule and the Leibniz rule where appropriate.
6. Integrate the equations of motion to obtain $\{q_\sigma(t)\}$ (easier said than done).

We not consider several examples:

4.4.1 One-dimensional motion

For a one-dimensional mechanical system with potential energy $U(x)$,

$$L = T - U = \tfrac{1}{2}m\dot{x}^2 - U(x) \quad . \tag{4.134}$$

The canonical momentum is

$$p = \frac{\partial L}{\partial \dot{x}} = m\dot{x} \tag{4.135}$$

and the equation of motion is

$$\frac{d}{dt}\bigg(\frac{\partial L}{\partial \dot{x}}\bigg) = \frac{\partial L}{\partial x} \quad \Rightarrow \quad m\ddot{x} = -U'(x) \quad , \tag{4.136}$$

which is of course $F = ma$.

Note that we can multiply the equation of motion by $\dot{x}$ to get

$$0 = \dot{x}\,\Big\{m\ddot{x} + U'(x)\Big\} = \frac{d}{dt}\Big\{\tfrac{1}{2}m\dot{x}^2 + U(x)\Big\} = \frac{dE}{dt} \quad , \tag{4.137}$$

where $E = T + U$.

4.4.2 Central force in two dimensions

Consider next a particle of mass m moving in two dimensions under the influence of a potential $U(\rho)$ which is a function of the distance from the origin $\rho = \sqrt{x^2 + y^2}$. Clearly cylindrical (2d polar) coordinates are called for:

$$L = \tfrac{1}{2}m\left(\dot{\rho}^2 + \rho^2\,\dot{\phi}^2\right) - U(\rho) \quad . \tag{4.138}$$

The equations of motion are

$$\begin{aligned} \frac{d}{dt}\bigg(\frac{\partial L}{\partial \dot{\rho}}\bigg) = \frac{\partial L}{\partial \rho} \quad &\Rightarrow \quad m\ddot{\rho} = m\rho\,\dot{\phi}^2 - U'(\rho) \\ \frac{d}{dt}\bigg(\frac{\partial L}{\partial \dot{\phi}}\bigg) = \frac{\partial L}{\partial \phi} \quad &\Rightarrow \quad \frac{d}{dt}\left(m\rho^2\,\dot{\phi}\right) = 0 \quad . \end{aligned} \tag{4.139}$$

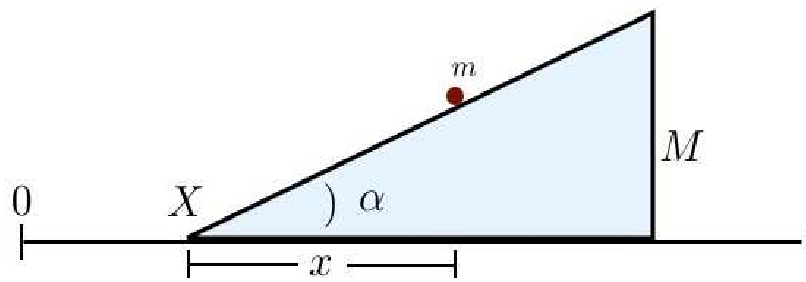

Figure 4.7: A wedge of mass M and opening angle α slides frictionlessly along a horizontal surface, while a small object of mass m slides frictionlessly along the wedge.

Note that the canonical momentum conjugate to ϕ, which is to say the angular momentum, is conserved:

$$p_\phi = m\rho^2\,\dot{\phi} = \text{const.} \tag{4.140}$$

We can use this to eliminate $\dot{\phi}$ from the first Euler-Lagrange equation, obtaining

$$m\ddot{\rho} = \frac{p_\phi^2}{m\rho^3} - U'(\rho) \quad . \tag{4.141}$$

We can also write the total energy as

$$\begin{aligned} E &= \tfrac{1}{2}m\left(\dot{\rho}^2 + \rho^2\,\dot{\phi}^2\right) + U(\rho) \\ &= \tfrac{1}{2}m\dot{\rho}^2 + \frac{p_\phi^2}{2m\rho^2} + U(\rho) \quad , \end{aligned} \tag{4.142}$$

from which it may be shown that E is also a constant:

$$\frac{dE}{dt} = \left(m\ddot{\rho} - \frac{p_\phi^2}{m\rho^3} + U'(\rho)\right)\dot{\rho} = 0 \quad . \tag{4.143}$$

We shall discuss this case in much greater detail in the coming weeks.

4.4.3 A sliding point mass on a sliding wedge

Consider the situation depicted in fig. 4.7, in which a point object of mass m slides frictionlessly along a wedge of opening angle α. The wedge itself slides frictionlessly along a horizontal surface, and its mass is M. We choose as generalized coordinates the horizontal position X of the left corner of the wedge, and the horizontal distance x from the left corner to the sliding point mass. The vertical coordinate of the sliding mass is then $y = x\,\tan\alpha$, where the horizontal surface lies at $y = 0$. With these generalized coordinates, the kinetic energy is

$$\begin{aligned} T &= \tfrac{1}{2}M\dot{X}^2 + \tfrac{1}{2}m\,(\dot{X} + \dot{x})^2 + \tfrac{1}{2}m\dot{y}^2 \\ &= \tfrac{1}{2}(M+m)\dot{X}^2 + m\dot{X}\dot{x} + \tfrac{1}{2}m\,(1+\tan^2\!\alpha)\,\dot{x}^2 \quad . \end{aligned} \tag{4.144}$$

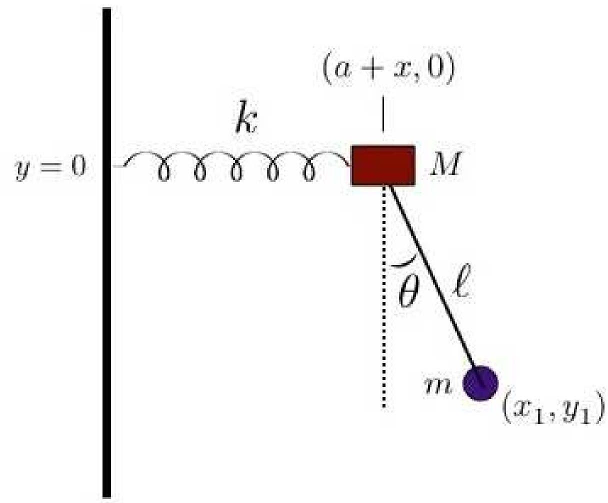

Figure 4.8: The spring–pendulum system.

The potential energy is simply

$$U = mgy = mg\,x\,\tan\alpha \quad . \tag{4.145}$$

Thus, the Lagrangian is

$$L = \tfrac{1}{2}(M+m)\dot{X}^2 + m\dot{X}\dot{x} + \tfrac{1}{2}m\,(1+\tan^2\!\alpha)\,\dot{x}^2 - mg\,x\,\tan\alpha \quad , \tag{4.146}$$

and the equations of motion are

$$\begin{aligned} \frac{d}{dt}\bigg(\frac{\partial L}{\partial \dot{X}}\bigg) &= \frac{\partial L}{\partial X} \quad \Rightarrow \quad (M+m)\ddot{X} + m\ddot{x} = 0 \\ \frac{d}{dt}\bigg(\frac{\partial L}{\partial \dot{x}}\bigg) &= \frac{\partial L}{\partial x} \quad \Rightarrow \quad m\ddot{X} + m\,(1+\tan^2\!\alpha)\,\ddot{x} = -mg\tan\alpha \quad . \end{aligned} \tag{4.147}$$

At this point we can use the first of these equations to write

$$\ddot{X} = -\frac{m}{M+m}\,\ddot{x} \quad . \tag{4.148}$$

Substituting this into the second equation, we obtain the constant accelerations

$$\ddot{x} = -\frac{(M+m)g\,\sin\alpha\,\cos\alpha}{M+m\sin^2\!\alpha} \quad , \quad \ddot{X} = \frac{mg\,\sin\alpha\,\cos\alpha}{M+m\sin^2\!\alpha} \quad . \tag{4.149}$$

4.4.4 A pendulum attached to a mass on a spring

Consider next the system depicted in fig. 4.8 in which a mass M moves horizontally while attached to a spring of spring constant k. Hanging from this mass is a pendulum of arm length ℓ and bob mass m.

A convenient set of generalized coordinates is (x, θ), where x is the displacement of the mass M relative to the equilibrium extension a of the spring, and θ is the angle the pendulum arm makes with respect to the vertical. Let the Cartesian coordinates of the pendulum bob be (x_1, y_1). Then

$$x_1 = a + x + \ell \, \sin\theta \quad , \quad y_1 = -l\cos\theta \quad . \tag{4.150}$$

The kinetic energy is

$$\begin{aligned} T &= \tfrac{1}{2}M\dot{x}^2 + \tfrac{1}{2}m\,(\dot{x}^2 + \dot{y}^2) \\ &= \tfrac{1}{2}M\dot{x}^2 + \tfrac{1}{2}m\left[(\dot{x} + \ell\cos\theta\,\dot{\theta})^2 + (\ell\sin\theta\,\dot{\theta})^2\right] \\ &= \tfrac{1}{2}(M+m)\,\dot{x}^2 + \tfrac{1}{2}m\ell^2\,\dot{\theta}^2 + m\ell\cos\theta\,\dot{x}\,\dot{\theta} \quad , \end{aligned} \tag{4.151}$$

and the potential energy is

$$\begin{aligned} U &= \tfrac{1}{2}kx^2 + mgy_1 \\ &= \tfrac{1}{2}kx^2 - mg\ell\cos\theta \quad . \end{aligned} \tag{4.152}$$

Thus,

$$L = \tfrac{1}{2}(M+m)\,\dot{x}^2 + \tfrac{1}{2}m\ell^2\,\dot{\theta}^2 + m\ell\cos\theta\,\dot{x}\,\dot{\theta} - \tfrac{1}{2}kx^2 + mg\ell\cos\theta \quad . \tag{4.153}$$

The canonical momenta are

$$\begin{aligned} p_x &= \frac{\partial L}{\partial \dot{x}} = (M+m)\,\dot{x} + m\ell\cos\theta\,\dot{\theta} \\ p_\theta &= \frac{\partial L}{\partial \dot{\theta}} = m\ell\cos\theta\,\dot{x} + m\ell^2\,\dot{\theta} \quad , \end{aligned} \tag{4.154}$$

and the canonical forces are

$$\begin{aligned} F_x &= \frac{\partial L}{\partial x} = -kx \\ F_\theta &= \frac{\partial L}{\partial \theta} = -m\ell\sin\theta\,\dot{x}\,\dot{\theta} - mg\ell\,\sin\theta \quad . \end{aligned} \tag{4.155}$$

The equations of motion then yield

$$\begin{aligned} (M+m)\,\ddot{x} + m\ell\,\cos\theta\,\ddot{\theta} - m\ell\,\sin\theta\,\dot{\theta}^2 &= -kx \\ m\ell\,\cos\theta\,\ddot{x} + m\ell^2\,\ddot{\theta} &= -mg\ell\,\sin\theta \quad . \end{aligned} \tag{4.156}$$

Small Oscillations : If we assume both x and θ are small, we may write $\sin\theta \approx \theta$ and $\cos\theta \approx 1$, in which case the equations of motion may be linearized to

$$\begin{aligned} (M+m)\,\ddot{x} + m\ell\,\ddot{\theta} + kx &= 0 \\ m\ell\,\ddot{x} + m\ell^2\,\ddot{\theta} + mg\ell\,\theta &= 0 \quad . \end{aligned} \tag{4.157}$$

If we define

$$u \equiv \frac{x}{\ell} \quad , \quad \alpha \equiv \frac{m}{M} \quad , \quad \omega_0^2 \equiv \frac{k}{M} \quad , \quad \omega_1^2 \equiv \frac{g}{\ell} \quad , \tag{4.158}$$

then may be linearized to

$$\begin{aligned} (1+\alpha)\,\ddot{u} + \alpha\,\ddot{\theta} + \omega_0^2\,u &= 0 \\ \ddot{u} + \ddot{\theta} + \omega_1^2\,\theta &= 0 \quad . \end{aligned} \tag{4.159}$$

We can solve by writing

$$\begin{pmatrix} u(t) \\ \theta(t) \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} e^{-i\omega t} \quad , \tag{4.160}$$

in which case

$$\begin{pmatrix} \omega_0^2 - (1+\alpha)\,\omega^2 & -\alpha\,\omega^2 \\ -\omega^2 & \omega_1^2 - \omega^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad . \tag{4.161}$$

In order to have a nontrivial solution (*i.e.* without $a = b = 0$), the determinant of the above 2×2 matrix must vanish. This gives a condition on ω^2, with solutions

$$\omega_\pm^2 = \tfrac{1}{2}\big[\,\omega_0^2 + (1+\alpha)\,\omega_1^2\,\big] \pm \tfrac{1}{2}\sqrt{\big[\,\omega_0^2 - (1+\alpha)\omega_1^2\,\big]^2 + 4\alpha\,\omega_0^2\,\omega_1^2} \quad . \tag{4.162}$$

4.4.5 The double pendulum

As yet another example of the generalized coordinate approach to Lagrangian dynamics, consider the double pendulum system, sketched in fig. 4.9. We choose as generalized coordinates the two angles θ_1 and θ_2. In order to evaluate the Lagrangian, we must obtain the kinetic and potential energies in terms of the generalized coordinates $\{\theta_1, \theta_2\}$ and their corresponding velocities $\{\dot{\theta}_1, \dot{\theta}_2\}$.

In Cartesian coordinates,

$$\begin{aligned} T &= \tfrac{1}{2}m_1\,(\dot{x}_1^2 + \dot{y}_1^2) + \tfrac{1}{2}m_2\,(\dot{x}_2^2 + \dot{y}_2^2) \\ U &= m_1\,g\,y_1 + m_2\,g\,y_2 \quad . \end{aligned} \tag{4.163}$$

We therefore express the Cartesian coordinates $\{x_1, y_1, x_2, y_2\}$ in terms of the generalized coordinates $\{\theta_1, \theta_2\}$:

$$\begin{aligned} x_1 &= +\ell_1\,\sin\theta_1 \quad , & x_2 &= \ell_1\,\sin\theta_1 + \ell_2\,\sin\theta_2 \\ y_1 &= -\ell_1\,\cos\theta_1 \quad , & y_2 &= -\ell_1\,\cos\theta_1 - \ell_2\,\cos\theta_2 \quad . \end{aligned} \tag{4.164}$$

Thus, the velocities are

$$\begin{aligned} \dot{x}_1 &= \ell_1\,\dot{\theta}_1\,\cos\theta_1 \quad , & \dot{x}_2 &= \ell_1\,\dot{\theta}_1\,\cos\theta_1 + \ell_2\,\dot{\theta}_2\,\cos\theta_2 \\ \dot{y}_1 &= \ell_1\,\dot{\theta}_1\,\sin\theta_1 \quad , & \dot{y}_2 &= \ell_1\,\dot{\theta}_1\,\sin\theta_1 + \ell_2\,\dot{\theta}_2\,\sin\theta_2 \quad . \end{aligned} \tag{4.165}$$

Thus,

$$\begin{aligned} T &= \tfrac{1}{2}m_1\,\ell_1^2\,\dot{\theta}_1^2 + \tfrac{1}{2}m_2\Big\{\ell_1^2\,\dot{\theta}_1^2 + 2\ell_1\,\ell_2\,\cos(\theta_1 - \theta_2)\,\dot{\theta}_1\,\dot{\theta}_2 + \ell_2^2\,\dot{\theta}_2^2\Big\} \\ U &= -m_1\,g\,\ell_1\,\cos\theta_1 - m_2\,g\,\ell_1\,\cos\theta_1 - m_2\,g\,\ell_2\,\cos\theta_2 \quad , \end{aligned} \tag{4.166}$$

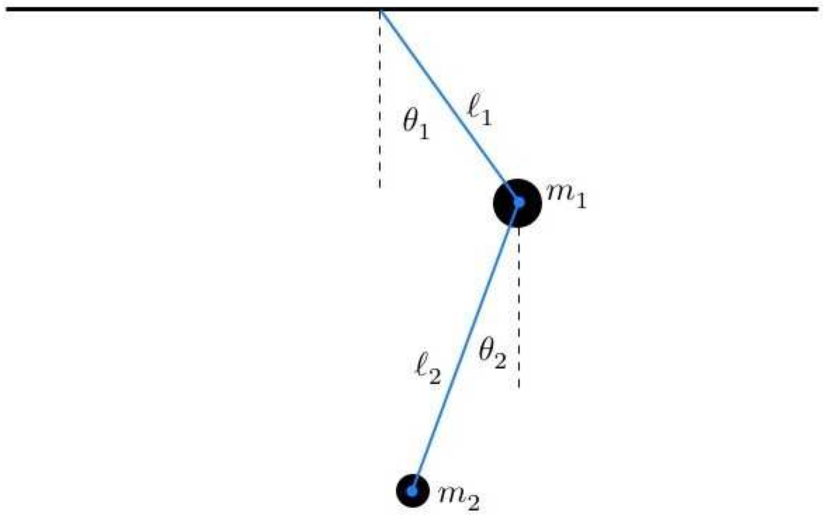

Figure 4.9: The double pendulum, with generalized coordinates θ_1 and θ_2. All motion is confined to a single plane.

and

$$\begin{aligned} L = T - U = \tfrac{1}{2}(m_1 + m_2)\,\ell_1^2\,\dot\theta_1^2 + m_2\,\ell_1\,\ell_2\,\cos(\theta_1 - \theta_2)\,\dot\theta_1\,\dot\theta_2 + \tfrac{1}{2} m_2\,\ell_2^2\,\dot\theta_2^2 \\ + (m_1 + m_2)\,g\,\ell_1\,\cos\theta_1 + m_2\,g\,\ell_2\,\cos\theta_2 \quad . \end{aligned} \tag{4.167}$$

The generalized (canonical) momenta are

$$\begin{aligned} p_1 &= \frac{\partial L}{\partial\dot\theta_1} = (m_1 + m_2)\,\ell_1^2\,\dot\theta_1 + m_2\,\ell_1\,\ell_2\,\cos(\theta_1 - \theta_2)\,\dot\theta_2 \\ p_2 &= \frac{\partial L}{\partial\dot\theta_2} = m_2\,\ell_1\,\ell_2\,\cos(\theta_1 - \theta_2)\,\dot\theta_1 + m_2\,\ell_2^2\,\dot\theta_2 \quad , \end{aligned} \tag{4.168}$$

and the equations of motion are

$$\begin{aligned} \dot p_1 &= (m_1 + m_2)\,\ell_1^2\,\ddot\theta_1 + m_2\,\ell_1\,\ell_2\,\cos(\theta_1 - \theta_2)\,\ddot\theta_2 - m_2\,\ell_1\,\ell_2\,\sin(\theta_1 - \theta_2)\,(\dot\theta_1 - \dot\theta_2)\,\dot\theta_2 \\ &= -(m_1 + m_2)\,g\,\ell_1\,\sin\theta_1 - m_2\,\ell_1\,\ell_2\,\sin(\theta_1 - \theta_2)\,\dot\theta_1\,\dot\theta_2 = \frac{\partial L}{\partial\theta_1} \end{aligned} \tag{4.169}$$

and

$$\begin{aligned} \dot p_2 &= m_2\,\ell_1\,\ell_2\,\cos(\theta_1 - \theta_2)\,\ddot\theta_1 - m_2\,\ell_1\,\ell_2\,\sin(\theta_1 - \theta_2)\,(\dot\theta_1 - \dot\theta_2)\,\dot\theta_1 + m_2\,\ell_2^2\,\ddot\theta_2 \\ &= -m_2\,g\,\ell_2\,\sin\theta_2 + m_2\,\ell_1\,\ell_2\,\sin(\theta_1 - \theta_2)\,\dot\theta_1\,\dot\theta_2 = \frac{\partial L}{\partial\theta_2} \quad . \end{aligned} \tag{4.170}$$

We therefore find

$$\begin{aligned}\ell_1\,\ddot\theta_1 + \frac{m_2\,\ell_2}{m_1+m_2}\,\cos(\theta_1-\theta_2)\,\ddot\theta_2 + \frac{m_2\,\ell_2}{m_1+m_2}\,\sin(\theta_1-\theta_2)\,\dot\theta_2^2 + g\,\sin\theta_1 &= 0\\ \ell_1\,\cos(\theta_1-\theta_2)\,\ddot\theta_1 + \ell_2\,\ddot\theta_2 - \ell_1\,\sin(\theta_1-\theta_2)\,\dot\theta_1^2 + g\,\sin\theta_2 &= 0 \quad .\end{aligned} \tag{4.171}$$

Small Oscillations : The equations of motion are coupled, nonlinear second order ODEs. When the system is close to equilibrium, the amplitudes of the motion are small, and we may expand in powers of the θ_1 and θ_2. The linearized equations of motion are then

$$\begin{aligned}\ddot\theta_1 + \alpha\beta\,\ddot\theta_2 + \omega_0^2\,\theta_1 &= 0\\ \ddot\theta_1 + \beta\,\ddot\theta_2 + \omega_0^2\,\theta_2 &= 0 \quad ,\end{aligned} \tag{4.172}$$

where we have defined

$$\alpha \equiv \frac{m_2}{m_1+m_2} \quad , \quad \beta \equiv \frac{\ell_2}{\ell_1} \quad , \quad \omega_0^2 \equiv \frac{g}{\ell_1} \quad . \tag{4.173}$$

We can solve this coupled set of equations by a nifty trick. Let's take a linear combination of the first equation plus an undetermined coefficient, r, times the second:

$$(1+r)\,\ddot\theta_1 + (\alpha+r)\beta\,\ddot\theta_2 + \omega_0^2\,(\theta_1 + r\,\theta_2) = 0 \quad . \tag{4.174}$$

We now demand that the ratio of the coefficients of θ_2 and θ_1 is the same as the ratio of the coefficients of $\ddot\theta_2$ and $\ddot\theta_1$:

$$\frac{(\alpha+r)\,\beta}{1+r} = r \quad \Rightarrow \quad r_\pm = \tfrac{1}{2}(\beta-1) \pm \tfrac{1}{2}\sqrt{(1-\beta)^2 + 4\alpha\beta} \tag{4.175}$$

When $r = r_\pm$, the equation of motion may be written

$$\frac{d^2}{dt^2}\,\big(\theta_1 + r_\pm\,\theta_2\big) = -\frac{\omega_0^2}{1+r_\pm}\,\big(\theta_1 + r_\pm\,\theta_2\big) \tag{4.176}$$

and defining the (unnormalized) *normal modes* $\xi_\pm \equiv \big(\theta_1 + r_\pm\,\theta_2\big)$ we find $\ddot\xi_\pm + \omega_\pm^2\,\xi_\pm = 0$ with

$$\omega_\pm = \frac{\omega_0}{\sqrt{1+r_\pm}} \quad . \tag{4.177}$$

Thus, by switching to the normal coordinates, we have decoupled the equations of motion, and identified the two *normal frequencies of oscillation*. We shall have much more to say about small oscillations in chapter 6.

For example, with $\ell_1 = \ell_2 = \ell$ and $m_1 = m_2 = m$, we have $\alpha = \frac{1}{2}$, and $\beta = 1$, in which case

$$r_\pm = \pm\tfrac{1}{\sqrt{2}} \quad , \quad \xi_\pm = \theta_1 \pm \tfrac{1}{\sqrt{2}}\,\theta_2 \quad , \quad \omega_\pm = \sqrt{2 \mp \sqrt{2}}\,\sqrt{\frac{g}{\ell}} \quad . \tag{4.178}$$

Note that the oscillation frequency for the 'in-phase' mode ξ_+ is low, and that for the 'out of phase' mode ξ_- is high.

4.4.6 The thingy

Four massless rods of length L are hinged together at their ends to form a rhombus. A particle of mass M is attached to each vertex. The opposite corners are joined by springs of spring constant k. In the square configuration, the strings are unstretched. The motion is confined to a plane, and the particles move only along the diagonals of the rhombus. Introduce suitable generalized coordinates and find the Lagrangian of the system. Deduce the equations of motion and find the frequency of small oscillations about equilibrium.

Solution

The rhombus is depicted in figure 4.10. Let a be the equilibrium length of the springs; clearly $b = \frac{1}{\sqrt{2}}\,a$. Let ϕ be half of one of the opening angles, as shown. Then the masses are located at $(\pm X, 0)$ and $(0, \pm Y)$, with $X = \frac{1}{\sqrt{2}}\,a\cos\phi$ and $Y = \frac{1}{\sqrt{2}}\,a\sin\phi$. The spring extensions are $\delta X = 2X - a$ and $\delta Y = 2Y - a$. The

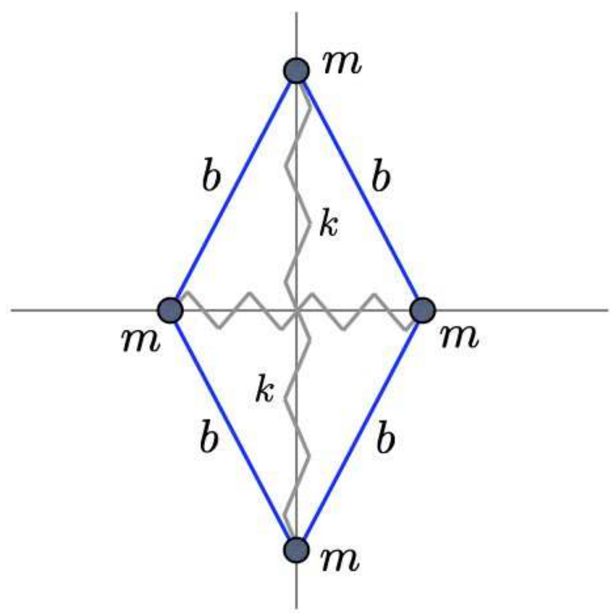

Figure 4.10: The thingy: a rhombus with opening angles 2ϕ and $\pi - 2\phi$.

kinetic and potential energies are therefore

$$T = M\big(\dot{X}^2 + \dot{Y}^2\big) = \tfrac{1}{2}Ma^2\,\dot{\phi}^2 \tag{4.179}$$

and

$$\begin{aligned} U &= \tfrac{1}{2}k\big(\delta X\big)^2 + \tfrac{1}{2}k\big(\delta Y\big)^2 \\ &= \tfrac{1}{2}ka^2\left\{\big(\sqrt{2}\,\cos\phi - 1\big)^2 + \big(\sqrt{2}\,\sin\phi - 1\big)^2\right\} \\ &= \tfrac{1}{2}ka^2\left\{3 - 2\sqrt{2}(\cos\phi + \sin\phi)\right\} \quad . \end{aligned} \tag{4.180}$$

Note that minimizing $U(\phi)$ gives $\sin\phi = \cos\phi$, *i.e.* $\phi_{\rm eq} = \frac{\pi}{4}$. The Lagrangian is then

$$L = T - U = \tfrac{1}{2} M a^2\, \dot\phi^2 + \sqrt{2}\, k a^2 \big(\cos\phi + \sin\phi\big) + {\rm const.} \tag{4.181}$$

The equations of motion are

$$\frac{d}{dt}\frac{\partial L}{\partial\dot\phi} = \frac{\partial L}{\partial\phi} \qquad \Rightarrow \qquad M a^2\, \ddot\phi = \sqrt{2}\, k a^2\, (\cos\phi - \sin\phi) \tag{4.182}$$

It's always smart to expand about equilibrium, so let's write $\phi = \frac{\pi}{4} + \delta$, which leads to

$$\ddot\delta + \omega_0^2\, \sin\delta = 0 \quad , \tag{4.183}$$

with $\omega_0 = \sqrt{2k/M}$. This is the equation of a pendulum! Linearizing gives $\ddot\delta + \omega_0^2\, \delta = 0$, so the small oscillation frequency is just ω_0.

4.5 The Virial Theorem

The virial theorem is a statement about the time-averaged motion of a mechanical system. Define the *virial*,

$$G(q,p) = \sum_\sigma p_\sigma\, q_\sigma \quad . \tag{4.184}$$

Then

$$\begin{aligned} \frac{dG}{dt} &= \sum_\sigma \big(\dot p_\sigma\, q_\sigma + p_\sigma\, \dot q_\sigma\big) \\ &= \sum_\sigma q_\sigma\, F_\sigma + \sum_\sigma \dot q_\sigma\, \frac{\partial L}{\partial \dot q_\sigma} \quad . \end{aligned} \tag{4.185}$$

Now suppose that $T = \frac{1}{2}\sum_{\sigma,\sigma'} {\rm T}_{\sigma\sigma'}(q)\, \dot q_\sigma\, \dot q_{\sigma'}$ is homogeneous of degree $k=2$ in $\dot q$, and that U is homogeneous of degree zero in $\dot q$. Then

$$\sum_\sigma \dot q_\sigma\, \frac{\partial L}{\partial \dot q_\sigma} = \sum_\sigma \dot q_\sigma\, \frac{\partial T}{\partial \dot q_\sigma} = 2\, T, \tag{4.186}$$

which follows from Euler's theorem on homogeneous functions.

Now consider the time average of $\dot G$ over a period τ:

$$\Big\langle \frac{dG}{dt} \Big\rangle = \frac{1}{\tau}\int\limits_0^\tau \! dt\, \frac{dG}{dt} = \frac{1}{\tau}\Big[G(\tau) - G(0)\Big] \quad . \tag{4.187}$$

If $G(t)$ is bounded, then in the limit $\tau\to\infty$ we must have $\langle \dot G\rangle = 0$. Any bounded motion, such as the orbit of the earth around the Sun, will result in $\langle \dot G\rangle_{\tau\to\infty} = 0$. But then

$$\Big\langle \frac{dG}{dt} \Big\rangle = 2\,\langle T\rangle + \Big\langle \sum_\sigma q_\sigma\, F_\sigma \Big\rangle = 0 \quad , \tag{4.188}$$

which implies

$$\langle T \rangle = -\frac{1}{2} \Big\langle \sum_\sigma q_\sigma \, F_\sigma \Big\rangle = \Big\langle \frac{1}{2} \sum_i \boldsymbol{x}_i \cdot \boldsymbol{\nabla}_i \, U\big(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\big) \Big\rangle = \tfrac{1}{2} \, k \, \langle U \rangle \quad , \tag{4.189}$$

where above equation pertains to homogeneous potentials of degree k in the Cartesian coordinates[3]. Finally, since $T + U = E$ is conserved, we have

$$\langle T \rangle = \frac{k\,E}{k+2} \quad , \quad \langle U \rangle = \frac{2\,E}{k+2} \quad . \tag{4.190}$$

4.6 Noether's Theorem

4.6.1 Continuous symmetry implies conserved charges

Consider a particle moving in two dimensions under the influence of an external potential $U(r)$. The potential is a function only of the magnitude of the vector $\boldsymbol{r}$. The Lagrangian is then

$$L = T - U = \tfrac{1}{2} m \left(\dot{r}^2 + r^2 \, \dot{\phi}^2 \right) - U(r) \quad , \tag{4.191}$$

where we have chosen generalized coordinates (r, ϕ). The momentum conjugate to ϕ is $p_\phi = m \, r^2 \, \dot{\phi}$. The generalized force F_ϕ clearly vanishes, since L does not depend on the coordinate ϕ. (One says that L is 'cyclic' in ϕ.) Thus, although $r = r(t)$ and $\phi = \phi(t)$ will in general be time-dependent, the combination $p_\phi = m \, r^2 \, \dot{\phi}$ is constant. This is the conserved angular momentum about the $\hat{\boldsymbol{z}}$ axis.

If instead the particle moved in a potential $U(y)$, independent of x, then writing

$$L = \tfrac{1}{2} m \left(\dot{x}^2 + \dot{y}^2 \right) - U(y) \quad , \tag{4.192}$$

we have that the momentum $p_x = \partial L / \partial \dot{x} = m\dot{x}$ is conserved, because the generalized force $F_x = \partial L / \partial x = 0$ vanishes. This situation pertains in a uniform gravitational field, with $U(x, y) = mgy$, independent of x. The horizontal component of momentum is conserved.

In general, whenever the system exhibits a *continuous symmetry*, there is an associated *conserved charge*. (The terminology 'charge' is from field theory.) Indeed, this is a rigorous result, known as *Noether's Theorem*. Consider a one-parameter family of transformations,

$$q_\sigma \longrightarrow \tilde{q}_\sigma(q, \zeta) \quad , \tag{4.193}$$

where ζ is the continuous parameter. Suppose further (without loss of generality) that at $\zeta = 0$ this transformation is the identity, *i.e.* $\tilde{q}_\sigma(q, 0) = q_\sigma$. The transformation may be nonlinear in the generalized

[3]Note that $-\sum_\sigma q_\sigma F_\sigma = -\sum_\sigma q_\sigma \, (\partial L / \partial q_\sigma) \neq \sum_\sigma q_\sigma \, (\partial U / \partial q_\sigma)$ in general because $T = \frac{1}{2} \sum_{\sigma\sigma'} \mathrm{T}_{\sigma\sigma'}(q) \, \dot{q}_\sigma \, \dot{q}_{\sigma'}$, and so the inequality holds whenever $\mathrm{T}_{\sigma\sigma'}(q)$ is q-dependent. In a Cartesian coordinate system, however, we have $T = \frac{1}{2} \sum_j m_j \dot{\boldsymbol{x}}_j^2$ and therefore eqn. 4.189 holds

coordinates. Suppose further that the Lagrangian L is invariant under the replacement $q \to \tilde{q}$. Then we must have

$$\begin{aligned} 0 = \frac{d}{d\zeta}\bigg|_{\zeta=0} L(\tilde{q}, \dot{\tilde{q}}, t) &= \sum_{\sigma=1}^{n} \left\{ \frac{\partial L}{\partial q_\sigma} \frac{\partial \tilde{q}_\sigma}{\partial \zeta}\bigg|_{\zeta=0} + \frac{\partial L}{\partial \dot{q}_\sigma} \frac{\partial \dot{\tilde{q}}_\sigma}{\partial \zeta}\bigg|_{\zeta=0} \right\} \\ &= \sum_{\sigma=1}^{n} \left\{ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\sigma}\right) \frac{\partial \tilde{q}_\sigma}{\partial \zeta}\bigg|_{\zeta=0} + \frac{\partial L}{\partial \dot{q}_\sigma} \frac{d}{dt}\left(\frac{\partial \tilde{q}_\sigma}{\partial \zeta}\right)_{\zeta=0} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\sigma}\frac{\partial \tilde{q}_\sigma}{\partial \zeta}\right)_{\zeta=0} \right\} \quad . \end{aligned} \tag{4.194}$$

Thus, there is an associated conserved charge

$$\Lambda = \sum_{\sigma=1}^{n} \frac{\partial L}{\partial \dot{q}_\sigma} \frac{\partial \tilde{q}_\sigma}{\partial \zeta}\bigg|_{\zeta=0} \quad . \tag{4.195}$$

4.6.2 Examples of one-parameter families of transformations

Consider the Lagrangian

$$L = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) - U\big(\sqrt{x^2+y^2}\big) \quad . \tag{4.196}$$

In two-dimensional polar coordinates, we have

$$L = \tfrac{1}{2}m(\dot{\rho}^2 + \rho^2\dot{\phi}^2) - U(\rho) \quad , \tag{4.197}$$

and we may now define

$$\tilde{\rho}(\zeta) = \rho \qquad , \qquad \tilde{\phi}(\zeta) = \phi + \zeta \quad . \tag{4.198}$$

Note that $\tilde{\rho}(0) = \rho$ and $\tilde{\phi}(0) = \phi$, *i.e.* the transformation is the identity when $\zeta = 0$. We now have

$$\Lambda = \sum_\sigma \frac{\partial L}{\partial \dot{q}_\sigma} \frac{\partial \tilde{q}_\sigma}{\partial \zeta}\bigg|_{\zeta=0} = \frac{\partial L}{\partial \dot{\rho}} \frac{\partial \tilde{\rho}}{\partial \zeta}\bigg|_{\zeta=0} + \frac{\partial L}{\partial \dot{\phi}} \frac{\partial \tilde{\phi}}{\partial \zeta}\bigg|_{\zeta=0} = m\rho^2\dot{\phi} \quad . \tag{4.199}$$

Another way to derive the same result which is somewhat instructive is to work out the transformation in Cartesian coordinates. We then have

$$\begin{aligned} \tilde{x}(\zeta) &= x\,\cos\zeta - y\,\sin\zeta \\ \tilde{y}(\zeta) &= x\,\sin\zeta + y\,\cos\zeta \quad . \end{aligned} \tag{4.200}$$

Thus,

$$\frac{\partial \tilde{x}}{\partial \zeta} = -\tilde{y} \quad , \quad \frac{\partial \tilde{y}}{\partial \zeta} = \tilde{x} \tag{4.201}$$

and

$$\Lambda = \frac{\partial L}{\partial \dot{x}} \frac{\partial \tilde{x}}{\partial \zeta}\bigg|_{\zeta=0} + \frac{\partial L}{\partial \dot{y}} \frac{\partial \tilde{y}}{\partial \zeta}\bigg|_{\zeta=0} = m(x\dot{y} - y\dot{x}) \quad . \tag{4.202}$$

But

$$m(x\dot{y} - y\dot{x}) = m\hat{\boldsymbol{z}} \cdot \boldsymbol{\rho} \times \dot{\boldsymbol{\rho}} = m\rho^2\dot{\phi} \quad . \tag{4.203}$$

As another example, consider the potential

$$U(\rho,\phi,z) = V(\rho, a\phi + z) \quad , \tag{4.204}$$

where (ρ,ϕ,z) are cylindrical coordinates for a particle of mass m, and where a is a constant with dimensions of length. The Lagrangian is

$$\tfrac{1}{2}m\big(\dot\rho^2 + \rho^2\dot\phi^2 + \dot z^2\big) - V(\rho, a\phi + z) \quad . \tag{4.205}$$

This model possesses a helical symmetry, with a one-parameter family

$$\tilde\rho(\zeta) = \rho \qquad , \qquad \tilde\phi(\zeta) = \phi + \zeta \qquad , \qquad \tilde z(\zeta) = z - \zeta a \quad . \tag{4.206}$$

Note that

$$a\tilde\phi + \tilde z = a\phi + z \quad , \tag{4.207}$$

so the potential energy, and the Lagrangian as well, is invariant under this one-parameter family of transformations. The conserved charge for this symmetry is

$$\Lambda = \frac{\partial L}{\partial\dot\rho}\,\frac{\partial\tilde\rho}{\partial\zeta}\bigg|_{\zeta=0} + \frac{\partial L}{\partial\dot\phi}\,\frac{\partial\tilde\phi}{\partial\zeta}\bigg|_{\zeta=0} + \frac{\partial L}{\partial\dot z}\,\frac{\partial\tilde z}{\partial\zeta}\bigg|_{\zeta=0} = m\rho^2\dot\phi - ma\dot z \quad . \tag{4.208}$$

We can check explicitly that Λ is conserved, using the equations of motion

$$\begin{aligned} \frac{d}{dt}\bigg(\frac{\partial L}{\partial\dot\phi}\bigg) &= \frac{d}{dt}\big(m\rho^2\dot\phi\big) = \frac{\partial L}{\partial\phi} = -a\frac{\partial V}{\partial z} \\ \frac{d}{dt}\bigg(\frac{\partial L}{\partial\dot z}\bigg) &= \frac{d}{dt}\big(m\dot z\big) = \frac{\partial L}{\partial z} = -\frac{\partial V}{\partial z} \quad . \end{aligned} \tag{4.209}$$

Thus,

$$\dot\Lambda = \frac{d}{dt}\big(m\rho^2\dot\phi\big) - a\,\frac{d}{dt}\big(m\dot z\big) = 0 \quad . \tag{4.210}$$

4.6.3 Conservation of linear and angular momentum

Suppose that the Lagrangian of a mechanical system is invariant under a uniform translation of all particles in the $\hat{\boldsymbol n}$ direction. Then our one-parameter family of transformations is given by

$$\tilde{\boldsymbol x}_a = \boldsymbol x_a + \zeta\,\hat{\boldsymbol n} \quad , \tag{4.211}$$

and the associated conserved Noether charge is

$$\Lambda = \sum_a \frac{\partial L}{\partial\dot{\boldsymbol x}_a}\cdot\hat{\boldsymbol n} = \hat{\boldsymbol n}\cdot\boldsymbol P \quad , \tag{4.212}$$

where $\boldsymbol P = \sum_a \boldsymbol p_a$ is the *total momentum* of the system.

If the Lagrangian of a mechanical system is invariant under rotations about an axis $\hat{\boldsymbol{n}}$, then

$$\begin{aligned}\tilde{\boldsymbol{x}}_a &= R(\zeta,\hat{\boldsymbol{n}})\,\boldsymbol{x}_a \\ &= \boldsymbol{x}_a + \zeta\,\hat{\boldsymbol{n}}\times\boldsymbol{x}_a + \mathcal{O}(\zeta^2) \quad ,\end{aligned} \tag{4.213}$$

where we have expanded the rotation matrix $R(\zeta,\hat{\boldsymbol{n}})$ in powers of ζ. The conserved Noether charge associated with this symmetry is

$$\Lambda = \sum_a \frac{\partial L}{\partial \dot{\boldsymbol{x}}_a}\cdot\hat{\boldsymbol{n}}\times\boldsymbol{x}_a = \hat{\boldsymbol{n}}\cdot\sum_a \boldsymbol{x}_a\times\boldsymbol{p}_a = \hat{\boldsymbol{n}}\cdot\boldsymbol{L} \quad , \tag{4.214}$$

where $\boldsymbol{L}$ is the *total angular momentum* of the system.

4.6.4 Invariance of L *vs.* invariance of S

Observant readers might object that demanding invariance of L is too strict. We should instead be demanding invariance of the action S[4]. Suppose S is invariant under

$$t \to \tilde{t}(q,t,\zeta) \qquad , \qquad q_\sigma(t) \to \tilde{q}_\sigma(q,t,\zeta) \quad . \tag{4.215}$$

Then invariance of S means

$$S = \int\limits_{t_a}^{t_b} dt\, L(q,\dot{q},t) = \int\limits_{\tilde{t}_a}^{\tilde{t}_b} dt\, L(\tilde{q},\dot{\tilde{q}},t) \quad . \tag{4.216}$$

Note that t is a dummy variable of integration, so it doesn't matter whether we call it t or $\tilde{t}$. The end-points of the integral, however, do change under the transformation. Now consider an infinitesimal transformation, for which $\delta t = \tilde{t} - t$ and $\delta q = \tilde{q}(\tilde{t}) - q(t)$ are both small. Thus,

$$S = \int\limits_{t_a}^{t_b} dt\, L(q,\dot{q},t) = \int\limits_{t_a+\delta t_a}^{t_b+\delta t_b} dt\,\bigg\{ L(q,\dot{q},t) + \frac{\partial L}{\partial q_\sigma}\,\bar{\delta} q_\sigma + \frac{\partial L}{\partial \dot{q}_\sigma}\,\bar{\delta}\dot{q}_\sigma + \ldots \bigg\} \quad , \tag{4.217}$$

where

$$\begin{aligned}\bar{\delta} q_\sigma(t) &\equiv \tilde{q}_\sigma(t) - q_\sigma(t) \\ &= \tilde{q}_\sigma(\tilde{t}) - \tilde{q}_\sigma(\tilde{t}) + \tilde{q}_\sigma(t) - q_\sigma(t) \\ &= \delta q_\sigma - \dot{q}_\sigma\,\delta t + \mathcal{O}(\delta q\,\delta t)\end{aligned} \tag{4.218}$$

Subtracting eqn. 4.217 from eqn. 4.216, we obtain

$$\begin{aligned}0 &= L_b\,\delta t_b - L_a\,\delta t_a + \frac{\partial L}{\partial \dot{q}_\sigma}\bigg|_b \bar{\delta} q_{\sigma,b} - \frac{\partial L}{\partial \dot{q}_\sigma}\bigg|_a \bar{\delta} q_{\sigma,a} + \int\limits_{t_a+\delta t_a}^{t_b+\delta t_b} dt\,\bigg\{ \frac{\partial L}{\partial q_\sigma} - \frac{d}{dt}\bigg(\frac{\partial L}{\partial \dot{q}_\sigma}\bigg)\bigg\}\,\bar{\delta} q_\sigma(t) \\ &= \int\limits_{t_a}^{t_b} dt\,\frac{d}{dt}\bigg\{\bigg(L - \frac{\partial L}{\partial \dot{q}_\sigma}\,\dot{q}_\sigma\bigg)\,\delta t + \frac{\partial L}{\partial \dot{q}_\sigma}\,\delta q_\sigma\bigg\} \quad ,\end{aligned} \tag{4.219}$$

---

[4] Indeed, we should be demanding that S only change by a function of the endpoint values.

where $L_{a,b}$ is $L(q,\dot{q},t)$ evaluated at $t = t_{a,b}$. Thus, if $\zeta \equiv \delta\zeta$ is infinitesimal, and

$$\delta t = A(q,t)\,\delta\zeta \qquad , \qquad \delta q_\sigma = B_\sigma(q,t)\,\delta\zeta \quad , \tag{4.220}$$

then the conserved charge is

$$\begin{aligned} \Lambda &= \left(L - \frac{\partial L}{\partial \dot{q}_\sigma}\,\dot{q}_\sigma \right) A(q,t) + \frac{\partial L}{\partial \dot{q}_\sigma}\,B_\sigma(q,t) \\ &= -\,H(q,p,t)\,A(q,t) + p_\sigma\,B_\sigma(q,t) \quad . \end{aligned} \tag{4.221}$$

Thus, when $A = 0$, we recover our earlier results, obtained by assuming invariance of L. Note that conservation of H follows from time translation invariance: $t \to t + \zeta$, for which $A = 1$ and $B_\sigma = 0$. Here we have written

$$H = p_\sigma\,\dot{q}_\sigma - L \quad , \tag{4.222}$$

and expressed it in terms of the momenta p_σ, the coordinates q_σ, and time t. H is called the *Hamiltonian*.

4.7 The Hamiltonian

4.7.1 From Lagrangian to Hamiltonian

The Lagrangian is a function of generalized coordinates, velocities, and time. The canonical momentum conjugate to the generalized coordinate q_σ is

$$p_\sigma = \frac{\partial L}{\partial \dot{q}_\sigma} \quad . \tag{4.223}$$

The Hamiltonian is a function of coordinates, *momenta*, and time. It is defined as the Legendre transform[5] of L:

$$H(q,p,t) = \sum_\sigma p_\sigma\,\dot{q}_\sigma - L \quad . \tag{4.224}$$

Let's examine the differential of H:

$$\begin{aligned} dH &= \sum_\sigma \left(\dot{q}_\sigma\,dp_\sigma + p_\sigma\,d\dot{q}_\sigma - \frac{\partial L}{\partial q_\sigma}\,dq_\sigma - \frac{\partial L}{\partial \dot{q}_\sigma}\,d\dot{q}_\sigma \right) - \frac{\partial L}{\partial t}\,dt \\ &= \sum_\sigma \left(\dot{q}_\sigma\,dp_\sigma - \frac{\partial L}{\partial q_\sigma}\,dq_\sigma \right) - \frac{\partial L}{\partial t}\,dt \quad , \end{aligned} \tag{4.225}$$

where we have invoked the definition of p_σ to cancel the coefficients of $d\dot{q}_\sigma$. Since $\dot{p}_\sigma = \partial L/\partial q_\sigma$, we have *Hamilton's equations of motion*,

$$\dot{q}_\sigma = \frac{\partial H}{\partial p_\sigma} \quad , \quad \dot{p}_\sigma = -\frac{\partial H}{\partial q_\sigma} \quad . \tag{4.226}$$

[5]See the appendix in §4.12 for more on Legendre transformations.

Thus, we can write

$$dH = \sum_\sigma \Big(\dot{q}_\sigma \, dp_\sigma - \dot{p}_\sigma \, dq_\sigma \Big) - \frac{\partial L}{\partial t} \, dt \quad . \tag{4.227}$$

Dividing by dt, we obtain

$$\frac{dH}{dt} = -\frac{\partial L}{\partial t} \quad , \tag{4.228}$$

which says that the Hamiltonian is *conserved* (*i.e.* it does not change with time) whenever there is no *explicit* time dependence to L.

Example #1 : For a simple $d = 1$ system with $L = \frac{1}{2} m \dot{x}^2 - U(x)$, we have $p = m\dot{x}$ and

$$H = p\,\dot{x} - L = \tfrac{1}{2} m \dot{x}^2 + U(x) = \frac{p^2}{2m} + U(x) \quad . \tag{4.229}$$

Example #2 : Consider now the mass point – wedge system analyzed above, with

$$L = \tfrac{1}{2}(M+m)\dot{X}^2 + m\dot{X}\dot{x} + \tfrac{1}{2} m \sec^2\!\alpha \, \dot{x}^2 - mg\tan(\alpha)\, x \quad , \tag{4.230}$$

The canonical momenta are

$$\begin{aligned} P &= \frac{\partial L}{\partial \dot{X}} = (M+m)\,\dot{X} + m\dot{x} \\ p &= \frac{\partial L}{\partial \dot{x}} = m\dot{X} + m\sec^2\!\alpha \, \dot{x} \quad . \end{aligned} \tag{4.231}$$

The Hamiltonian is given by

$$\begin{aligned} H &= P\,\dot{X} + p\,\dot{x} - L \\ &= \tfrac{1}{2}(M+m)\dot{X}^2 + m\dot{X}\dot{x} + \tfrac{1}{2} m \sec^2\!\alpha \, \dot{x}^2 + mg\tan(\alpha)\, x \\ &= \frac{1}{2} \begin{pmatrix} \dot{X} & \dot{x} \end{pmatrix} \overbrace{\begin{pmatrix} M+m & m \\ m & m\sec^2\!\alpha \end{pmatrix}}^{A} \begin{pmatrix} \dot{X} \\ \dot{x} \end{pmatrix} + mg\tan(\alpha)\, x \quad . \end{aligned} \tag{4.232}$$

However, this is not quite H, since $H = H(X, x, P, p, t)$ must be expressed in terms of the coordinates and the *momenta* and not the coordinates and velocities. So we must eliminate $\dot{X}$ and $\dot{x}$ in favor of P and p. We do this by inverting the relations

$$\begin{pmatrix} P \\ p \end{pmatrix} = \begin{pmatrix} M+m & m \\ m & m\,\sec^2\!\alpha \end{pmatrix} \begin{pmatrix} \dot{X} \\ \dot{x} \end{pmatrix} = A \begin{pmatrix} \dot{X} \\ \dot{x} \end{pmatrix} \tag{4.233}$$

to obtain

$$\begin{pmatrix} \dot{X} \\ \dot{x} \end{pmatrix} = \frac{1}{M\sec^2\!\alpha + m\tan^2\!\alpha} \begin{pmatrix} \sec^2\!\alpha & -1 \\ -1 & \frac{M}{m} + 1 \end{pmatrix} \begin{pmatrix} P \\ p \end{pmatrix} = A^{-1} \begin{pmatrix} P \\ p \end{pmatrix} \quad . \tag{4.234}$$

Substituting into 4.232, we obtain

$$
\begin{aligned}
H &= \frac{1}{2}\begin{pmatrix} P & p \end{pmatrix} A^{-1} \begin{pmatrix} P \\ p \end{pmatrix} + mg\tan(\alpha)\,x \\
&= \frac{P^2}{2(M+m\sin^2\!\alpha)} - \frac{2Pp\cos^2\!\alpha}{2(M+m\sin^2\!\alpha)} + \frac{\big(\frac{M}{m}+1\big)\,p^2\cos^2\!\alpha}{2(M+m\sin^2\!\alpha)} + mg\tan(\alpha)x \quad .
\end{aligned}
\tag{4.235}
$$

Notice that $\dot{P}=0$ since $\frac{\partial L}{\partial X}=0$. P is the total horizontal momentum of the system (wedge plus particle) and it is conserved. As a sanity check, consider the limit $M\to\infty$ with P and p finite. The wedge then has infinite inertia and remains fixed. Accordingly, we find

$$
H(X,x,P,p,t)\big|_{M\to\infty} = \frac{p^2\cos^2\!\alpha}{2m} + mg\tan(\alpha)\,x \quad .
\tag{4.236}
$$

4.7.2 Is $H = T + U$?

The most general form of the kinetic energy is

$$
\begin{aligned}
T &= T_2 + T_1 + T_0 \\
&= \tfrac{1}{2} T_2^{\sigma\sigma'}(q,t)\,\dot{q}_\sigma\,\dot{q}_{\sigma'} + T_1^\sigma(q,t)\,\dot{q}_\sigma + T_0(q,t) \quad ,
\end{aligned}
\tag{4.237}
$$

where $T_n(q,\dot{q},t)$ is homogeneous of degree n in the velocities[6]. We assume a potential energy of the form

$$
\begin{aligned}
U &= U_1 + U_0 \\
&= U_1^\sigma(q,t)\,\dot{q}_\sigma + U_0(q,t) \quad ,
\end{aligned}
\tag{4.238}
$$

which allows for velocity-dependent forces, as we have with charged particles moving in an electromagnetic field. The Lagrangian is then

$$
L = T - U = \tfrac{1}{2} T_2^{\sigma\sigma'}(q,t)\,\dot{q}_\sigma\,\dot{q}_{\sigma'} + T_1^\sigma(q,t)\,\dot{q}_\sigma + T_0(q,t) - U_1^\sigma(q,t)\,\dot{q}_\sigma - U_0(q,t) \quad .
\tag{4.239}
$$

The canonical momentum conjugate to q_σ is

$$
p_\sigma = \frac{\partial L}{\partial \dot{q}_\sigma} = T_2^{\sigma\sigma'}\,\dot{q}_{\sigma'} + T_1^\sigma(q,t) - U_1^\sigma(q,t)
\tag{4.240}
$$

which is inverted to give

$$
\dot{q}_\sigma = \big(T_2^{-1}\big)^{\sigma\sigma'}\Big(p_{\sigma'} - T_1^{\sigma'} + U_1^{\sigma'}\Big) \quad .
\tag{4.241}
$$

The Hamiltonian is then

$$
\begin{aligned}
H &= p_\sigma\,\dot{q}_\sigma - L \\
&= \tfrac{1}{2}\big(T_2^{-1}\big)^{\sigma\sigma'}\Big(p_\sigma - T_1^\sigma + U_1^\sigma\Big)\Big(p_{\sigma'} - T_1^{\sigma'} + U_1^{\sigma'}\Big) - T_0 + U_0 \\
&= T_2 - T_0 + U_0 \quad .
\end{aligned}
\tag{4.242}
$$

[6] A homogeneous function of degree k satisfies $f(\lambda x_1,\ldots,\lambda x_n) = \lambda^k f(x_1,\ldots,x_n)$. It is then easy to prove *Euler's theorem*, $\sum_{i=1}^n x_i \frac{\partial f}{\partial x_i} = kf$.

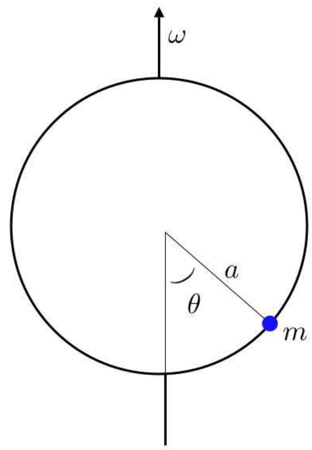

Figure 4.11: A bead of mass m on a rotating hoop of radius a.

If T_0, T_1, and U_1 vanish, *i.e.* if $T(q, \dot{q}, t)$ is a homogeneous function of degree two in the generalized velocities, and $U(q, t)$ is velocity-independent, then $H = T + U$. But if T_0 or T_1 is nonzero, or the potential is velocity-dependent, then $H \neq T + U$.

4.7.3 Example: a bead on a rotating hoop

Consider a bead of mass m constrained to move along a hoop of radius a. The hoop is further constrained to rotate with angular velocity $\dot{\phi} = \omega$ about the $\hat{\boldsymbol{z}}$-axis, as shown in fig. 4.11.

The most convenient set of generalized coordinates is spherical polar (r, θ, ϕ), in which case

$$\begin{aligned} T &= \tfrac{1}{2} m \big(\dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin^2\theta \, \dot{\phi}^2\big) \\ &= \tfrac{1}{2} m a^2 \big(\dot{\theta}^2 + \omega^2 \sin^2\theta\big) \quad . \end{aligned} \tag{4.243}$$

Thus, $T_2 = \frac{1}{2} m a^2 \dot{\theta}^2$ and $T_0 = \frac{1}{2} m a^2 \omega^2 \sin^2\theta$. The potential energy is $U(\theta) = mga(1 - \cos\theta)$. The momentum conjugate to θ is $p_\theta = m a^2 \dot{\theta}$, and thus

$$\begin{aligned} H(\theta, p_\theta) &= T_2 - T_0 + U \\ &= \tfrac{1}{2} m a^2 \dot{\theta}^2 - \tfrac{1}{2} m a^2 \omega^2 \sin^2\theta + mga(1 - \cos\theta) \\ &= \frac{p_\theta^2}{2ma^2} - \tfrac{1}{2} m a^2 \omega^2 \sin^2\theta + mga(1 - \cos\theta) \quad . \end{aligned} \tag{4.244}$$

For this problem, we can define the *effective potential*

$$\begin{aligned} U_{\text{eff}}(\theta) \equiv U - T_0 &= mga(1 - \cos\theta) - \tfrac{1}{2} m a^2 \omega^2 \sin^2\theta \\ &= mga\Big(1 - \cos\theta - \frac{\omega^2}{2\omega_0^2} \sin^2\theta\Big) \quad , \end{aligned} \tag{4.245}$$

where $\omega_0^2 \equiv g/a$. The Lagrangian may then be written

$$L = \tfrac{1}{2} m a^2 \dot{\theta}^2 - U_{\text{eff}}(\theta) \quad , \tag{4.246}$$

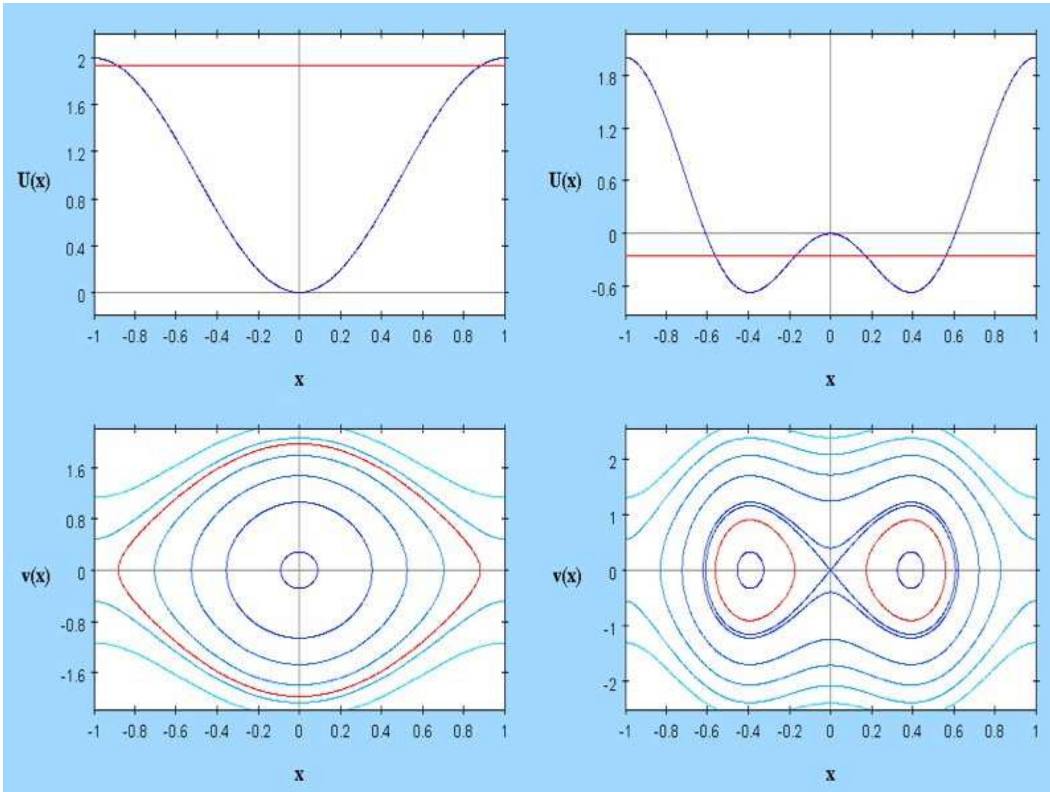

Figure 4.12: The effective potential $U_{\text{eff}}(\theta) = mga\big[1 - \cos\theta - \frac{\omega^2}{2\omega_0^2}\sin^2\theta\big]$. (The dimensionless potential $\tilde{U}_{\text{eff}}(x) = U_{\text{eff}}/mga$ is shown, where $x = \theta/\pi$.) Left panels: $\omega = \frac{1}{2}\sqrt{3}\,\omega_0$. Right panels: $\omega = \sqrt{3}\,\omega_0$.

and thus the equations of motion are

$$ma^2\ddot{\theta} = -\frac{\partial U_{\text{eff}}}{\partial\theta} \quad . \tag{4.247}$$

Equilibrium is achieved when $U'_{\text{eff}}(\theta) = 0$, which gives

$$\frac{\partial U_{\text{eff}}}{\partial\theta} = mga\sin\theta\left\{1 - \frac{\omega^2}{\omega_0^2}\cos\theta\right\} = 0 \quad , \tag{4.248}$$

i.e. $\theta^* = 0$, $\theta^* = \pi$, or $\theta^* = \pm\cos^{-1}(\omega_0^2/\omega^2)$, where the last pair of equilibria are present only for $\omega^2 > \omega_0^2$. The stability of these equilibria is assessed by examining the sign of $U''_{\text{eff}}(\theta^*)$. We have

$$U''_{\text{eff}}(\theta) = mga\left\{\cos\theta - \frac{\omega^2}{\omega_0^2}\big(2\cos^2\theta - 1\big)\right\} \quad . \tag{4.249}$$

Thus,

$$U''_{\text{eff}}(\theta^*) = \begin{cases} +mga\left(1-\frac{\omega^2}{\omega_0^2}\right) & \text{at } \theta^*=0 \\ \\ -mga\left(1+\frac{\omega^2}{\omega_0^2}\right) & \text{at } \theta^*=\pi \\ \\ +mga\left(\frac{\omega^2}{\omega_0^2}-\frac{\omega_0^2}{\omega^2}\right) & \text{at } \theta^*=\pm\cos^{-1}\left(\frac{\omega_0^2}{\omega^2}\right) \quad . \end{cases} \tag{4.250}$$

Thus, $\theta^*=0$ is stable for $\omega^2<\omega_0^2$ but becomes unstable when the rotation frequency ω is sufficiently large, *i.e.* when $\omega^2>\omega_0^2$. In this regime, there are two new equilibria, at $\theta^*=\pm\cos^{-1}(\omega_0^2/\omega^2)$, which are both stable. The equilibrium at $\theta^*=\pi$ is always unstable, independent of the value of ω. The situation is depicted in fig. 4.12.

4.7.4 Charged particle in an electromagnetic field

Consider next the case of a charged particle moving in the presence of an electromagnetic field. The particle's potential energy is

$$U(\boldsymbol{r},\dot{\boldsymbol{r}}) = q\,\phi(\boldsymbol{r},t) - \frac{q}{c}\,\boldsymbol{A}(\boldsymbol{r},t)\cdot\dot{\boldsymbol{r}} \quad , \tag{4.251}$$

which *is* velocity-dependent. The kinetic energy is $T=\frac{1}{2}m\dot{\boldsymbol{r}}^2$, as usual, and $L=T-U$. Here $\phi(\boldsymbol{r},t)$ is the scalar potential and $\boldsymbol{A}(\boldsymbol{r},t)$ the vector potential. The electric and magnetic fields are given by

$$\boldsymbol{E} = -\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{\partial\boldsymbol{A}}{\partial t} \quad , \quad \boldsymbol{B} = \boldsymbol{\nabla}\times\boldsymbol{A} \quad . \tag{4.252}$$

The canonical momenta and forces are

$$\begin{aligned} \boldsymbol{p} &= \frac{\partial L}{\partial\dot{\boldsymbol{r}}} = m\dot{\boldsymbol{r}} + \frac{q}{c}\,\boldsymbol{A} \\ \boldsymbol{F} &= \frac{\partial L}{\partial\boldsymbol{r}} = -q\boldsymbol{\nabla}\phi + \frac{q}{c}\,\boldsymbol{\nabla}\big(\boldsymbol{A}\cdot\dot{\boldsymbol{r}}\big) \quad . \end{aligned} \tag{4.253}$$

The Euler-Lagrange equations are

$$\dot{\boldsymbol{p}} = \frac{d}{dt}\left(\frac{\partial L}{\partial\dot{\boldsymbol{r}}}\right) = \frac{\partial L}{\partial\boldsymbol{r}} = \boldsymbol{F} \tag{4.254}$$

which is to say

$$m\ddot{\boldsymbol{r}} + \frac{q}{c}\,\frac{d\boldsymbol{A}}{dt} = -q\,\boldsymbol{\nabla}\phi + \frac{q}{c}\,\boldsymbol{\nabla}(\boldsymbol{A}\cdot\dot{\boldsymbol{r}}) \quad , \tag{4.255}$$

or, in component notation,

$$m\ddot{x}_i + \frac{q}{c}\overbrace{\left(\frac{\partial A_i}{\partial x_j}\,\dot{x}_j + \frac{\partial A_i}{\partial t}\right)}^{dA_i/dt} = -q\,\frac{\partial\phi}{\partial x_i} + \frac{q}{c}\,\frac{\partial A_j}{\partial x_i}\,\dot{x}_j \quad . \tag{4.256}$$

Here we are using the Einstein convention of summing over repeated indices. Thus,

$$m\ddot{x}_i = -q\,\frac{\partial\phi}{\partial x_i} - \frac{q}{c}\,\frac{\partial A_i}{\partial t} + \frac{q}{c}\left(\frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j}\right)\dot{x}_j \quad . \tag{4.257}$$

It is convenient to express the cross product in terms of the completely antisymmetric tensor of rank three, ϵ_{ijk}:

$$B_k = \epsilon_{klm}\,\frac{\partial A_m}{\partial x_l} \quad , \tag{4.258}$$

and using the result

$$\epsilon_{kij}\,\epsilon_{klm} = \delta_{il}\,\delta_{jm} - \delta_{im}\,\delta_{jl} \quad , \tag{4.259}$$

we have

$$\epsilon_{kij}\,B_k = \frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j} \quad . \tag{4.260}$$

and therefore

$$m\ddot{x}_i = -q\,\frac{\partial\phi}{\partial x_i} - \frac{q}{c}\,\frac{\partial A_i}{\partial t} + \frac{q}{c}\,\epsilon_{ijk}\,\dot{x}_j\,B_k \quad . \tag{4.261}$$

In vector notation, using $\epsilon_{ijk}\,v_j\,B_k = (\boldsymbol{v}\times\boldsymbol{B})_i$, we have

$$m\ddot{\boldsymbol{r}} = q\boldsymbol{E} + \frac{q}{c}\,\dot{\boldsymbol{r}}\times\boldsymbol{B} \quad , \tag{4.262}$$

which is, of course, the Lorentz force law.

Next, we compute the Hamiltonian:

$$\begin{aligned} H(\boldsymbol{r},\boldsymbol{p},t) &= \boldsymbol{p}\cdot\dot{\boldsymbol{r}} - L \\ &= m\dot{\boldsymbol{r}}^2 + \frac{q}{c}\,\boldsymbol{A}\cdot\dot{\boldsymbol{r}} - \tfrac{1}{2}m\dot{\boldsymbol{r}}^2 - \frac{q}{c}\,\boldsymbol{A}\cdot\dot{\boldsymbol{r}} + q\phi \\ &= \tfrac{1}{2}m\dot{\boldsymbol{r}}^2 + q\phi \\ &= \frac{1}{2m}\Big(\boldsymbol{p} - \frac{q}{c}\,\boldsymbol{A}(\boldsymbol{r},t)\Big)^2 + q\phi(\boldsymbol{r},t) \quad . \end{aligned} \tag{4.263}$$

If $\boldsymbol{A}$ and ϕ are time-independent, then $dH/dt = -\partial L/\partial t = 0$ and $H(\boldsymbol{r},\boldsymbol{p})$ is conserved.

4.8 Motion in Rapidly Oscillating Fields

4.8.1 Slow and fast dynamics

Consider a free particle moving under the influence of an oscillating force $F(t) = F_0\,\cos(\omega t)$. Newton's second law is then $m\ddot{q} = F\cos\omega t$, the solution to which is

$$q(t) = a + b\,t - \frac{F_0\,\cos\omega t}{m\omega^2} \quad . \tag{4.264}$$

where $q_{\rm h}(t) \equiv a + b\,t$ is the solution to the homogeneous (unforced) equation of motion. Note that the amplitude of the response $q - q_{\rm h}$ goes as ω^{-2} and is therefore small when ω is large.

Now consider a general $n = 1$ system, with

$$H(q,p,t) = H^0(q,p) + \widetilde{V}(q)\cos(\omega t) \quad , \tag{4.265}$$

where we will assume $\widetilde{V}(q)$ is small. We also assume that ω is much greater than any natural oscillation frequency associated with H_0. We separate the motion $q(t)$ and $p(t)$ into slow and fast components:

$$\begin{aligned} q(t) &= Q(t) + \zeta(t) \\ p(t) &= P(t) + \pi(t) \quad , \end{aligned} \tag{4.266}$$

where $\zeta(t)$ and $\pi(t)$ oscillate with the driving frequency ω. Since ζ and π will be small, we expand Hamilton's equations in these quantities:

$$\begin{aligned} \dot{Q} + \dot{\zeta} &= \frac{\partial H^0}{\partial P} + \frac{\partial^2 H^0}{\partial P^2}\,\pi + \frac{\partial^2 H^0}{\partial Q\,\partial P}\,\zeta + \frac{1}{2}\,\frac{\partial^3 H^0}{\partial Q^2\,\partial P}\,\zeta^2 + \frac{\partial^3 H^0}{\partial Q\,\partial P^2}\,\zeta\pi + \frac{1}{2}\,\frac{\partial^3 H^0}{\partial P^3}\,\pi^2 + \ldots \\ \dot{P} + \dot{\pi} &= -\frac{\partial H^0}{\partial Q} - \frac{\partial^2 H^0}{\partial Q^2}\,\zeta - \frac{\partial^2 H^0}{\partial Q\,\partial P}\,\pi - \frac{1}{2}\,\frac{\partial^3 H^0}{\partial Q^3}\,\zeta^2 - \frac{\partial^3 H^0}{\partial Q^2\,\partial P}\,\zeta\pi - \frac{1}{2}\,\frac{\partial^3 H^0}{\partial Q\,\partial P^2}\,\pi^2 \\ &\qquad - \frac{\partial \widetilde{V}}{\partial Q}\,\cos(\omega t) - \frac{\partial^2 \widetilde{V}}{\partial Q^2}\,\zeta\cos(\omega t) - \ldots \quad . \end{aligned} \tag{4.267}$$

We now average over the fast degrees of freedom to obtain an equation of motion for the slow variables Q and P, which we here carry to lowest nontrivial order in averages of fluctuating quantities:

$$\begin{aligned} \dot{Q} &= H^0_P + \tfrac{1}{2} H^0_{QQP}\,\langle\zeta^2\rangle + H^0_{QPP}\,\langle\zeta\pi\rangle + \tfrac{1}{2} H^0_{PPP}\,\langle\pi^2\rangle \\ \dot{P} &= -H^0_Q - \tfrac{1}{2} H^0_{QQQ}\,\langle\zeta^2\rangle - H^0_{QQP}\,\langle\zeta\pi\rangle - \tfrac{1}{2} H^0_{QPP}\,\langle\pi^2\rangle - \widetilde{V}_{QQ}\,\langle\zeta\cos\omega t\rangle \quad , \end{aligned} \tag{4.268}$$

where we now adopt the shorthand notation $H^0_{QQP} = \frac{\partial^3 H^0}{\partial^2 Q\,\partial P}$, *etc.* The fast degrees of freedom obey

$$\begin{aligned} \dot{\zeta} &= H^0_{QP}\,\zeta + H^0_{PP}\,\pi \\ \dot{\pi} &= -H^0_{QQ}\,\zeta - H^0_{QP}\,\pi - \widetilde{V}_Q\,\cos(\omega t) \quad . \end{aligned} \tag{4.269}$$

We can solve these by replacing $\widetilde{V}_Q\,\cos\omega t$ with $\widetilde{V}_Q\,e^{-i\omega t}$, and writing $\zeta(t) = \zeta_0\,e^{-i\omega t}$ and $\pi(t) = \pi_0\,e^{-i\omega t}$, resulting in

$$\begin{pmatrix} H^0_{QP} + i\omega & H^0_{PP} \\ -H^0_{QQ} & -H^0_{QP} + i\omega \end{pmatrix} \begin{pmatrix} \zeta_0 \\ \pi_0 \end{pmatrix} = \begin{pmatrix} 0 \\ \widetilde{V}_Q \end{pmatrix} \quad . \tag{4.270}$$

We now invert the matrix to obtain ζ_0 and π_0, then take the real part, which yields

$$\begin{aligned} \zeta(t) &= \omega^{-2}\,H^0_{PP}\,\widetilde{V}_Q\,\cos\omega t + \mathcal{O}\big(\omega^{-4}\big) \\ \pi(t) &= -\omega^{-2}\,H^0_{QP}\,\widetilde{V}_Q\,\cos\omega t - \omega^{-1}\,\widetilde{V}_Q\,\sin\omega t + \mathcal{O}\big(\omega^{-3}\big) \quad . \end{aligned} \tag{4.271}$$

Invoking $\langle\cos^2(\omega t)\rangle = \langle\sin^2(\omega t)\rangle = \frac{1}{2}$ and $\langle\cos(\omega t)\sin(\omega t)\rangle = 0$, we substitute into eqns. 4.268 to obtain

$$\begin{aligned}\dot{Q} &= H^0_P + \tfrac{1}{4}\,\omega^{-2}\,H^0_{PPP}\,\widetilde{V}^2_Q + \mathcal{O}(\omega^{-4}) \\ \dot{P} &= -H^0_Q - \tfrac{1}{4}\,\omega^{-2}\,H^0_{QPP}\,\widetilde{V}^2_Q - \tfrac{1}{2}\,\omega^{-2} H^0_{PP}\,\widetilde{V}_Q\,\widetilde{V}_{QQ} + \mathcal{O}(\omega^{-4}) \quad .\end{aligned} \tag{4.272}$$

These equations may be written compactly as

$$\dot{Q} = \frac{\partial K}{\partial P} \quad , \quad \dot{P} = -\frac{\partial K}{\partial Q} \quad , \tag{4.273}$$

where

$$K(Q,P) = H^0(Q,P) + \frac{1}{4\omega^2}\,\frac{\partial^2 H^0}{\partial P^2}\bigg(\frac{\partial \widetilde{V}}{\partial Q}\bigg)^2 + \ldots \quad . \tag{4.274}$$

4.8.2 Example : pendulum with oscillating support

Consider a pendulum with a vertically oscillating point of support. The coordinates of the pendulum bob are

$$x = \ell\,\sin\theta \qquad , \qquad y = a(t) - \ell\,\cos\theta \quad . \tag{4.275}$$

The Lagrangian is easily obtained:

$$\begin{aligned}L &= \tfrac{1}{2}m\ell^2\,\dot{\theta}^2 + m\ell\dot{a}\,\dot{\theta}\sin\theta + mg\ell\cos\theta + \tfrac{1}{2}m\dot{a}^2 - mga \\ &= \tfrac{1}{2}m\ell^2\,\dot{\theta}^2 + m(g+\ddot{a})\ell\cos\theta + \overbrace{\tfrac{1}{2}m\dot{a}^2 - mga - \frac{d}{dt}\big(m\ell\dot{a}\,\cos\theta\big)}^{\text{these may be dropped}} \quad .\end{aligned} \tag{4.276}$$

Thus we may take the Lagrangian to be

$$L = \tfrac{1}{2}m\ell^2\,\dot{\theta}^2 + m(g+\ddot{a})\,\ell\cos\theta \quad , \tag{4.277}$$

from which we derive the Hamiltonian

$$\begin{aligned}H(\theta,p_\theta) &= \frac{p_\theta^2}{2m\ell^2} - mg\ell\cos\theta - m\ell\ddot{a}\,\cos\theta \\ &= H_0(\theta,p_\theta,t) + \widetilde{V}(\theta)\,\sin\omega t \quad .\end{aligned} \tag{4.278}$$

We have assumed $a(t) = a_0\sin\omega t$, so

$$\widetilde{V}(\theta) = m\ell a_0\,\omega^2\cos\theta \quad . \tag{4.279}$$

Writing $\theta \equiv \Theta + \zeta$ and $p_\theta \equiv L + \pi$, the effective Hamiltonian, per eqn. 4.274, is

$$K(\Theta,L) = \frac{L^2}{2m\ell^2} - mg\ell\cos\Theta + \tfrac{1}{4}m\,a_0^2\,\omega^2\sin^2\!\Theta \quad . \tag{4.280}$$

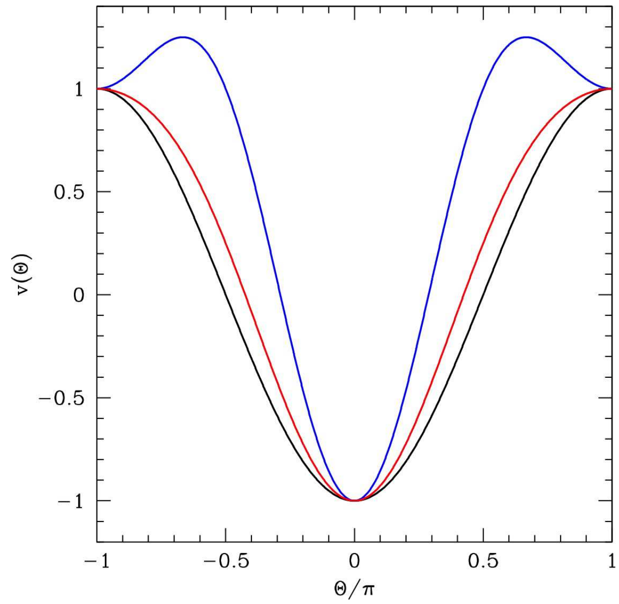

Figure 4.13: Dimensionless potential $v(\Theta)$ for $r=0$ (black curve), $r=0.5$ (red), and $r=2$ (blue).

Let's define the dimensionless parameter $r \equiv \omega^2 a_0^2/2g\ell$. The slow variable Θ then executes motion in the *effective potential* $V_{\text{eff}}(\Theta) = mg\ell\, v(\Theta)$, with

$$v(\Theta) = -\cos\Theta + \frac{r}{2}\sin^2\!\Theta \quad . \tag{4.281}$$

Differentiating, we find that $V_{\text{eff}}(\Theta)$ is stationary when

$$v'(\Theta) = 0 \quad \Rightarrow \quad r\sin\Theta\,\cos\Theta = -\sin\Theta \quad . \tag{4.282}$$

Thus, $\Theta = 0$ and $\Theta = \pi$, where $\sin\Theta = 0$, are equilibria. When $r > 1$ (note $r > 0$ always), there are two new solutions, given by the roots of $\cos\Theta = -r^{-1}$.

To assess stability of these equilibria, we compute the second derivative:

$$v''(\Theta) = \cos\Theta + r\,\cos 2\Theta \quad . \tag{4.283}$$

From this, we see that $\Theta = 0$ is stable, *i.e.* $v''(\Theta = 0) > 0$, always, but $\Theta = \pi$ is stable for $r > 1$ and unstable for $r < 1$. When $r > 1$, two new solutions appear, at $\cos\Theta = -r^{-1}$, for which

$$v''\big(\cos^{-1}(-1/r)\big) = r^{-1} - r \quad , \tag{4.284}$$

which is always negative since $r > 1$ in order for these equilibria to exist. The situation is sketched in fig. 4.13, showing $v(\Theta)$ for three representative values of the parameter r. For $r < 1$, the equilibrium at $\Theta = \pi$ is unstable, but as r increases, a subcritical pitchfork bifurcation is encountered at $r = 1$, and $\Theta = \pi$ becomes stable, while the outlying $\Theta = \cos^{-1}(-1/r)$ solutions are unstable.

4.9 Field Theory: Systems with Several Independent Variables

4.9.1 Equations of motion and Noether's theorem

Suppose $\phi_a(x)$ depends on several independent variables: $x = \{x^1, x^2, \ldots, x^n\}$. Furthermore, suppose

$$S\big[\{\phi_a(x)\}\big] = \int\limits_\Omega d^n\!x\, \mathcal{L}(\phi_a\, \partial_\mu \phi_a, x) \quad , \tag{4.285}$$

i.e. the *Lagrangian density* $\mathcal{L}$ is a function of the fields ϕ_a; their partial derivatives $\partial\phi_a/\partial x^\mu$, and possibly the independent variables x^μ as well. Here Ω is a region in $\mathbb{R}^n$. In dynamical field theories, we write $x = (x^0, x^1, \ldots, x^d)$ where d is the dimension of space and $x^0 = ct$, where t is time and c is a constant with dimensions of speed. In such cases $n = d + 1$ and we can identify $x^0 \equiv x^n$.

Then the first variation of S is

$$\begin{aligned} \delta S &= \int\limits_\Omega d^n\!x \left\{ \frac{\partial \mathcal{L}}{\partial \phi_a}\, \delta\phi_a + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\, \frac{\partial\, \delta\phi_a}{\partial x^\mu} \right\} \\ &= \oint\limits_{\partial\Omega} d\Sigma\, n^\mu\, \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\, \delta\phi_a + \int\limits_\Omega d^n\!x \left\{ \frac{\partial \mathcal{L}}{\partial \phi_a} - \frac{\partial}{\partial x^\mu}\bigg(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\bigg) \right\} \delta\phi_a \quad , \end{aligned} \tag{4.286}$$

where $\partial\Omega$ is the $(n-1)$-dimensional boundary of Ω, $d\Sigma$ is the differential surface area, and n^μ is the unit vector normal to $\partial\Omega$. If we demand $\partial\mathcal{L}/\partial(\partial_\mu\phi_a)\big|_{\partial\Omega} = 0$ or $\delta\phi_a\big|_{\partial\Omega} = 0$, the surface term vanishes, and we conclude

$$\frac{\delta S}{\delta \phi_a(\boldsymbol{x})} = \frac{\partial \mathcal{L}}{\partial \phi_a} - \frac{\partial}{\partial x^\mu}\bigg(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\bigg) \quad . \tag{4.287}$$

Next, consider the one-parameter family of *field* transformations

$$\phi_a(x) \to \widetilde{\phi}_a\big(\boldsymbol{\phi}(x), \zeta\big) \tag{4.288}$$

such that $\widetilde{\phi}_a\big(\boldsymbol{\phi}(x), \zeta = 0\big) = \phi_a(x)$. If the Lagrangian density $\mathcal{L}$ is independent of this transformation, then

$$\begin{aligned} \frac{d\mathcal{L}}{d\zeta}\bigg|_{\zeta=0} &= \frac{\partial \mathcal{L}}{\partial \phi_a}\, \frac{\partial \widetilde{\phi}_a}{\partial \zeta}\bigg|_{\zeta=0} + \sum_{\mu=1}^n \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\, \frac{\partial(\partial_\mu \widetilde{\phi}_a)}{\partial \zeta}\bigg|_{\zeta=0} \\ &= \sum_{\mu=1}^n \left\{ \frac{\partial}{\partial x^\mu}\bigg(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\bigg) \frac{\partial \widetilde{\phi}_a}{\partial \zeta}\bigg|_{\zeta=0} + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\, \frac{\partial}{\partial x^\mu}\bigg(\frac{\partial \widetilde{\phi}_a}{\partial \zeta}\bigg)_{\zeta=0} \right\} \\ &= \sum_{\mu=1}^n \frac{\partial}{\partial x^\mu}\bigg(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\, \frac{\partial \widetilde{\phi}_a}{\partial \zeta}\bigg)_{\zeta=0} \end{aligned} \tag{4.289}$$

We can write this as $\partial_\mu\, J^\mu = 0$, where

$$J^\mu \equiv \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\, \frac{\partial \widetilde{\phi}_a}{\partial \zeta}\bigg|_{\zeta=0} \quad . \tag{4.290}$$

We call $\Lambda = J^0/c$ the *total charge*. If we assume $\boldsymbol{J} = 0$ at the spatial boundaries of our system, then integrating the conservation law $\partial_\mu J^\mu$ (summation convention) over the spatial region Ω gives

$$\frac{d\Lambda}{dt} = \int\limits_\Omega d^3\!x\, \partial_0\, J^0 = -\int\limits_\Omega d^3\!x\, \boldsymbol{\nabla}\!\cdot\!\boldsymbol{J} = -\oint\limits_{\partial\Omega} d\Sigma\, \hat{\boldsymbol{n}} \cdot \boldsymbol{J} = 0 \quad , \tag{4.291}$$

assuming $\boldsymbol{J} = 0$ at the boundary $\partial\Omega$.

As an example, consider the case of a stretched string of linear mass density ρ and tension τ. The action is a functional of the height $y(x,t)$, where the coordinate along the string, x, and time, t, are the two independent variables. The Lagrangian density is

$$\mathcal{L} = \tfrac{1}{2}\,\rho \left(\frac{\partial y}{\partial t}\right)^{\!2} - \tfrac{1}{2}\,\tau \left(\frac{\partial y}{\partial x}\right)^{\!2} \quad . \tag{4.292}$$

The Euler-Lagrange equations are

$$\begin{aligned} 0 = \frac{\delta S}{\delta y(x,t)} &= -\frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial y'}\right) - \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{y}}\right) \\ &= \frac{\partial}{\partial x}\left(\tau\, \frac{\partial y}{\partial x}\right) - \rho\, \frac{\partial^2 y}{\partial t^2} \quad , \end{aligned} \tag{4.293}$$

where $y' = \partial y/\partial x$ and $\dot{y} = \partial y/\partial t$. We've assumed boundary conditions where $\delta y(x_a,t) = \delta y(x_b,t) = \delta y(x,t_a) = \delta y(x,t_b) = 0$. At this point, $\rho(x)$ and $\tau(x)$ may be position-dependent. For constant ρ and τ, we obtain the Helmholtz equation $\rho\ddot{y} = \tau y''$, where $c = (\tau/\rho)^{1/2}$ is the speed of wave propagation.

For practice with the Minkowski notation, we define $x^0 \equiv ct$ and $x^1 \equiv x$ and the two-dimensional space-time coordinate vector is then $x^\mu = (x^0, x^1) = (ct, x)$. The Lagrangian can then be written $\mathcal{L} = \frac{1}{2}\,\tau(\partial_\mu y)(\partial^\mu y)$, where $x_\mu = g_{\mu\nu}\, x^\nu = (ct, -x)$, in which case $\partial_\mu = \partial/\partial x^\mu$ and $\partial^\mu = \partial/\partial x_\mu$. Clearly $\mathcal{L}$ remains invariant under the one-parameter family of transformations $y \to y + \zeta$, and the conserved Noether current is

$$J^\mu = \tau\, \frac{\partial y}{\partial x_\mu} \quad , \tag{4.294}$$

and we have $\partial_\mu J^\mu = 0$, which is equivalent to $\partial^\mu J_\mu = 0$. (Upper indices are called *covariant* while lower ones are *contravariant*.) Current conservation in this system is simply a restatement of the Helmholtz equation.

Maxwell's equations

The Lagrangian density for an electromagnetic field with sources is

$$\mathcal{L} = -\frac{1}{16\pi}\, F_{\mu\nu}\, F^{\mu\nu} - \frac{1}{c}\, j_\mu\, A^\mu \quad . \tag{4.295}$$

The equations of motion are then

$$\frac{\partial \mathcal{L}}{\partial A^\mu} - \frac{\partial}{\partial x^\nu}\left(\frac{\partial \mathcal{L}}{\partial(\partial^\mu A^\nu)}\right) = 0 \quad \Rightarrow \quad \partial_\mu\, F^{\mu\nu} = \frac{4\pi}{c}\, j^\nu \quad , \tag{4.296}$$

which are Maxwell's equations.

Relativistic complex scalar field

As an example, consider the case of a complex scalar field, with Lagrangian density

$$\mathcal{L}(\psi, \psi^*, \partial_\mu \psi, \partial_\mu \psi^*) = \tfrac{1}{2} K \, (\partial_\mu \psi^*)(\partial^\mu \psi) - U\big(\psi^* \psi\big) \quad . \tag{4.297}$$

This is invariant under the transformation $\psi \to e^{i\zeta}\, \psi$, $\psi^* \to e^{-i\zeta}\, \psi^*$. Thus,

$$\frac{\partial \tilde{\psi}}{\partial \zeta} = i\, e^{i\zeta}\, \psi \qquad , \qquad \frac{\partial \tilde{\psi}^*}{\partial \zeta} = -i\, e^{-i\zeta}\, \psi^* \quad , \tag{4.298}$$

and, summing over both ψ and ψ^* fields, we have

$$\begin{aligned} J^\mu &= \frac{\partial \mathcal{L}}{\partial\,(\partial_\mu \psi)} \cdot (i\psi) + \frac{\partial \mathcal{L}}{\partial\,(\partial_\mu \psi^*)} \cdot (-i\psi^*) \\ &= \frac{K}{2i} \big(\psi^* \partial^\mu \psi - \psi\, \partial^\mu \psi^*\big) \quad . \end{aligned} \tag{4.299}$$

The potential, which depends on $|\psi|^2$, is independent of ζ. Hence, this form of conserved 4-current is valid for an entire class of potentials.

4.9.2 Gross-Pitaevskii model

As one final example of a field theory, consider the Gross-Pitaevskii model, with

$$\mathcal{L} = i\hbar\, \psi^* \, \frac{\partial \psi}{\partial t} - \frac{\hbar^2}{2m}\, \boldsymbol{\nabla}\psi^* \!\cdot\! \boldsymbol{\nabla}\psi - g\, \big(|\psi|^2 - n_0\big)^2 \quad . \tag{4.300}$$

This describes a nonrelativistic Bose fluid with repulsive short-ranged interactions. Here $\psi(\boldsymbol{x}, t)$ is again a complex scalar field, and ψ^* is its complex conjugate. Using the Leibniz rule, we have

$$\begin{aligned} \delta S[\psi^*, \psi] &= S[\psi^* + \delta\psi^*, \psi + \delta\psi] \\ &= \int\! dt \!\int\! d^d\!x \, \bigg\{ i\hbar\, \psi^*\, \frac{\partial \delta\psi}{\partial t} + i\hbar\, \delta\psi^*\, \frac{\partial \psi}{\partial t} - \frac{\hbar^2}{2m}\, \boldsymbol{\nabla}\psi^* \cdot \boldsymbol{\nabla}\delta\psi - \frac{\hbar^2}{2m}\, \boldsymbol{\nabla}\delta\psi^* \cdot \boldsymbol{\nabla}\psi \\ &\qquad\qquad\qquad - 2g\, \big(|\psi|^2 - n_0\big)\, (\psi^* \delta\psi + \psi \delta\psi^*) \bigg\} \\ &= \int\! dt \!\int\! d^d\!x \, \bigg\{ \bigg[-i\hbar\, \frac{\partial \psi^*}{\partial t} + \frac{\hbar^2}{2m}\, \nabla^2 \psi^* - 2g\, \big(|\psi|^2 - n_0\big)\, \psi^* \bigg]\, \delta\psi \\ &\qquad\qquad\qquad + \bigg[i\hbar\, \frac{\partial \psi}{\partial t} + \frac{\hbar^2}{2m}\, \nabla^2 \psi - 2g\, \big(|\psi|^2 - n_0\big)\, \psi \bigg]\, \delta\psi^* \bigg\} \quad , \end{aligned} \tag{4.301}$$

where we have integrated by parts where necessary and discarded the boundary terms. Extremizing $S[\psi^*, \psi]$ therefore results in the *nonlinear Schrödinger equation* (NLSE),

$$i\hbar\, \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m}\, \nabla^2 \psi + 2g\, \big(|\psi|^2 - n_0\big)\, \psi \tag{4.302}$$

as well as its complex conjugate,

$$-\,i\hbar\,\frac{\partial\psi^*}{\partial t} = -\frac{\hbar^2}{2m}\,\nabla^2\psi^* + 2g\,\big(|\psi|^2 - n_0\big)\,\psi^* \quad . \tag{4.303}$$

Note that these equations are indeed the Euler-Lagrange equations:

$$\begin{aligned}\frac{\delta S}{\delta\psi} &= \frac{\partial\mathcal{L}}{\partial\psi} - \frac{\partial}{\partial x^\mu}\bigg(\frac{\partial\mathcal{L}}{\partial\,\partial_\mu\psi}\bigg)\\ \frac{\delta S}{\delta\psi^*} &= \frac{\partial\mathcal{L}}{\partial\psi^*} - \frac{\partial}{\partial x^\mu}\bigg(\frac{\partial\mathcal{L}}{\partial\,\partial_\mu\psi^*}\bigg) \quad ,\end{aligned} \tag{4.304}$$

with $x^\mu = (t, \boldsymbol{x})$[7]. Plugging in

$$\frac{\partial\mathcal{L}}{\partial\psi} = -2g\,\big(|\psi|^2 - n_0\big)\,\psi^* \quad , \quad \frac{\partial\mathcal{L}}{\partial\,\partial_t\psi} = i\hbar\,\psi^* \quad , \quad \frac{\partial\mathcal{L}}{\partial\,\boldsymbol{\nabla}\psi} = -\frac{\hbar^2}{2m}\,\boldsymbol{\nabla}\psi^* \tag{4.305}$$

and

$$\frac{\partial\mathcal{L}}{\partial\psi^*} = i\hbar\,\psi - 2g\,\big(|\psi|^2 - n_0\big)\,\psi \quad , \quad \frac{\partial\mathcal{L}}{\partial\,\partial_t\psi^*} = 0 \quad , \quad \frac{\partial\mathcal{L}}{\partial\,\boldsymbol{\nabla}\psi^*} = -\frac{\hbar^2}{2m}\,\boldsymbol{\nabla}\psi \quad , \tag{4.306}$$

we recover the NLSE and its conjugate.

The Gross-Pitaevskii model also possesses a $\mathsf{U}(1)$ or $\mathsf{O}(2)$ invariance, *viz.*

$$\psi(\boldsymbol{x},t) \to \widetilde{\psi}(\boldsymbol{x},t) = e^{i\zeta}\,\psi(\boldsymbol{x},t) \quad , \quad \psi^*(\boldsymbol{x},t) \to \widetilde{\psi}^*(\boldsymbol{x},t) = e^{-i\zeta}\,\psi^*(\boldsymbol{x},t) \quad . \tag{4.307}$$

Thus, the conserved Noether current is then a $(d+1)$-dimensional vector with components

$$J^\mu = \frac{\partial\mathcal{L}}{\partial\,\partial_\mu\psi}\,\frac{\partial\widetilde{\psi}}{\partial\zeta}\bigg|_{\zeta=0} + \frac{\partial\mathcal{L}}{\partial\,\partial_\mu\psi^*}\,\frac{\partial\widetilde{\psi}^*}{\partial\zeta}\bigg|_{\zeta=0} \quad . \tag{4.308}$$

In terms of time ($\mu = 0$) and space ($\mu \in \{1,\ldots,d\}$) components, we have

$$\begin{aligned}J^0 &= -\hbar\,|\psi|^2\\ \boldsymbol{J} &= -\frac{\hbar^2}{2im}\,\big(\psi^*\boldsymbol{\nabla}\psi - \psi\boldsymbol{\nabla}\psi^*\big) \quad .\end{aligned} \tag{4.309}$$

Dividing out by $\hbar$, taking $J^0 \equiv -\hbar\rho$ and $\boldsymbol{J} \equiv -\hbar\boldsymbol{j}$, we obtain the continuity equation,

$$\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\!\cdot\!\boldsymbol{j} = 0 \quad , \tag{4.310}$$

where

$$\rho = |\psi|^2 \qquad , \qquad \boldsymbol{j} = \frac{\hbar}{2im}\,\big(\psi^*\boldsymbol{\nabla}\psi - \psi\boldsymbol{\nabla}\psi^*\big) \quad . \tag{4.311}$$

are the particle density and the particle current, respectively.

[7]In the nonrelativistic case, there is no utility in defining $x^0 = ct$, so we simply define $x^0 = t$.

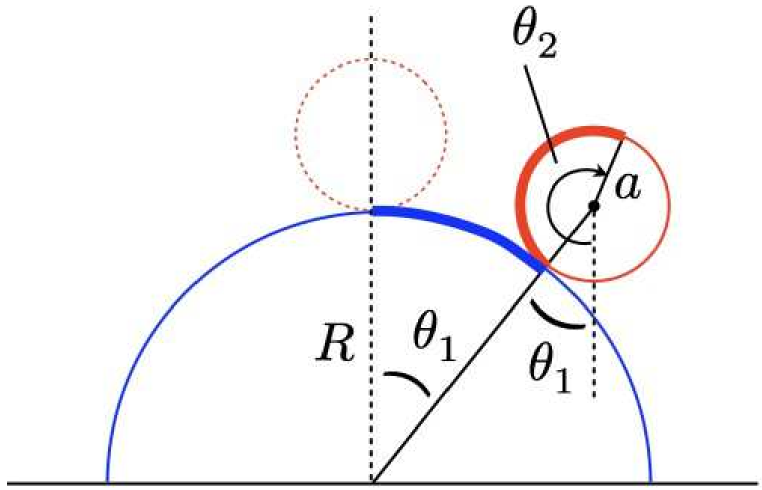

Figure 4.14: A cylinder of radius a rolls along a half-cylinder of radius R. When there is no slippage, the angles θ_1 and θ_2 obey the constraint equation $R\theta_1 = a(\theta_2 - \theta_1)$.

4.10 Constraints: General Theory

4.10.1 Introduction

A mechanical system of N point particles in d dimensions possesses $n = dN$ degrees of freedom[8]. To specify these degrees of freedom, we can choose any independent set of generalized coordinates $\{q_1, \ldots, q_n\}$. Oftentimes, however, not all n coordinates are independent.

Consider, for example, the situation in fig. 4.14, where a cylinder of radius a rolls over a half-cylinder of radius R. If there is no slippage, then the angles θ_1 and θ_2 are not independent, and they obey the *equation of constraint*,

$$R\,\theta_1 = a\,(\theta_2 - \theta_1) \quad . \tag{4.312}$$

In this case, we can easily solve the constraint equation and substitute $\theta_2 = \big(1 + \frac{R}{a}\big)\,\theta_1$. In other cases, though, the equation of constraint might not be so easily solved (*e.g.* it may be nonlinear). How then do we proceed?

4.10.2 Constrained extremization of functions: Lagrange multipliers

Given $F(x_1, \ldots, x_n)$ to be extremized subject to k constraints of the form $G_j(x_1, \ldots, x_n) = 0$ where $j = 1, \ldots, k$, construct

$$F^*\big(x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_k\big) \equiv F(x_1, \ldots, x_n) + \sum_{j=1}^{k} \lambda_j\, G_j(x_1, \ldots, x_n) \tag{4.313}$$

[8]For N rigid bodies, the number of degrees of freedom is $n' = \frac{1}{2}d(d+1)N$, corresponding to d center-of-mass coordinates and $\frac{1}{2}d(d-1)$ angles of orientation for each particle. The dimension of the group of rotations in d dimensions is $\frac{1}{2}d(d-1)$, corresponding to the number of parameters in a general rank-d orthogonal matrix (*i.e.* an element of the group $O(d)$).

which is a function of the $(n+k)$ variables $\{x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_k\}$, where the quantities $\{\lambda_1, \ldots, \lambda_k\}$ are *Lagrange undetermined multipliers*. We now *freely* extremize the extended function F^*:

$$\begin{aligned} dF^* &= \sum_{\sigma=1}^{n} \frac{\partial F^*}{\partial x_\sigma}\, dx_\sigma + \sum_{j=1}^{k} \frac{\partial F^*}{\partial \lambda_j}\, d\lambda_j \\ &= \sum_{\sigma=1}^{n} \left(\frac{\partial F}{\partial x_\sigma} + \sum_{j=1}^{k} \lambda_j\, \frac{\partial G_j}{\partial x_\sigma} \right) dx_\sigma + \sum_{j=1}^{k} G_j\, d\lambda_j = 0 \end{aligned} \tag{4.314}$$

This results in the $(n+k)$ equations

$$\begin{aligned} \frac{\partial F}{\partial x_\sigma} + \sum_{j=1}^{k} \lambda_j\, \frac{\partial G_j}{\partial x_\sigma} &= 0 \qquad (\sigma = 1, \ldots, n) \\ G_j &= 0 \qquad (j = 1, \ldots, k) \quad . \end{aligned} \tag{4.315}$$

The interpretation of all this is as follows. The first n equations in 4.315 can be written in vector form as

$$\boldsymbol{\nabla} F + \sum_{j=1}^{k} \lambda_j\, \boldsymbol{\nabla} G_j = 0 \quad . \tag{4.316}$$

This says that the (n-component) vector $\boldsymbol{\nabla} F$ is linearly dependent upon the k vectors $\boldsymbol{\nabla} G_j$. Thus, any movement in the direction of $\boldsymbol{\nabla} F$ must necessarily entail movement along one or more of the directions $\boldsymbol{\nabla} G_j$. This would require violating the constraints, since movement along $\boldsymbol{\nabla} G_j$ takes us off the level set $G_j = 0$. Were $\boldsymbol{\nabla} F$ linearly *independent* of the set $\{\boldsymbol{\nabla} G_j\}$, this would mean that we could find a differential displacement $d\boldsymbol{x}$ which has finite overlap with $\boldsymbol{\nabla} F$ but zero overlap with each $\boldsymbol{\nabla} G_j$. Thus $\boldsymbol{x} + d\boldsymbol{x}$ would still satisfy $G_j(\boldsymbol{x} + d\boldsymbol{x}) = 0$, but F would change by the finite amount $dF = \boldsymbol{\nabla} F(\boldsymbol{x}) \cdot d\boldsymbol{x}$.

Put another way, when we extremize $F(\boldsymbol{x})$ without constraints, we identify points $\boldsymbol{x} \in \mathbb{R}^n$ where the gradient $\boldsymbol{\nabla} F$ vanishes. However, when we have k constraints of the form $G_j(\boldsymbol{x}) = 0$, the subset

$$\Upsilon = \big\{ \boldsymbol{x} \in \mathbb{R}^n \,|\, G_j(\boldsymbol{x}) = 0 \; \forall \, j \in \{1, \ldots, k\} \big\} \tag{4.317}$$

is a hypersurface of dimension $n - k$. Generically we should not expect any of the solutions to $\boldsymbol{\nabla} F = 0$ to lie within the subspace Υ. Extremizing $F(\boldsymbol{x})$ subject to the k constraints $G_j(\boldsymbol{x}) = 0$ means that we must find the extrema of $F(\boldsymbol{x})$ for $\boldsymbol{x} \in \Upsilon \subset \mathbb{R}^n$. All such extrema satisfy that $\boldsymbol{\nabla} F(\boldsymbol{x})$ is *perpendicular* to the hypersurface Υ, *i.e.* $\boldsymbol{\nabla} F(\boldsymbol{x})$ must lie in the k-dimensional subspace spanned by the vectors $\boldsymbol{\nabla} G_j(\boldsymbol{x})$.

Example : volume of a cylinder

To see how this formalism works in practice, let's extremize the volume $V = \pi a^2 h$ of a cylinder of radius a and height h, subject to the constraint

$$G(a, h) = 2\pi a + \frac{h^2}{b} - \ell = 0 \quad . \tag{4.318}$$

We therefore define

$$V^*(a,h,\lambda) \equiv V(a,h) + \lambda\, G(a,h) \quad , \tag{4.319}$$

and set

$$\frac{\partial V^*}{\partial a} = 2\pi a h + 2\pi\lambda = 0 \tag{4.320}$$

$$\frac{\partial V^*}{\partial h} = \pi a^2 + 2\lambda\, \frac{h}{b} = 0 \tag{4.321}$$

$$\frac{\partial V^*}{\partial \lambda} = 2\pi a + \frac{h^2}{b} - \ell = 0 \quad . \tag{4.322}$$

Solving these three equations simultaneously gives

$$a = \frac{2\ell}{5\pi} \quad , \quad h = \sqrt{\frac{b\ell}{5}} \quad , \quad \lambda = -\frac{2}{5^{3/2}\pi}\, b^{1/2}\, \ell^{3/2} \quad , \quad V^* = \frac{4}{5^{5/2}\,\pi}\, \ell^{5/2}\, b^{1/2} \quad . \tag{4.323}$$

4.10.3 Constraints and variational calculus

Before addressing the subject of constrained dynamical systems, let's consider the issue of constraints in the broader context of variational calculus. Suppose we have a functional

$$F[\boldsymbol{y}(x)] = \int\limits_{x_a}^{x_b} dx\, L(\boldsymbol{y}, \boldsymbol{y}', x) \quad , \tag{4.324}$$

which we want to extremize subject to some constraints. Here $\boldsymbol{y}$ stands for an n-component vector of functions $\{y_\sigma(x)\}$. We assume that the endpoint values $y_\sigma(x_a)$ and $y_\sigma(x_b)$ are fixed for each σ. There are two classes of constraints we will consider:

1. *Integral constraints:* These are of the form

$$\int\limits_{x_a}^{x_b} dx\, N_j(\boldsymbol{y}, \boldsymbol{y}', x) = C_j \quad , \tag{4.325}$$

   where j labels the constraint.

2. *Holonomic constraints:* These are of the form

$$G_j(\boldsymbol{y}, x) = 0 \quad . \tag{4.326}$$

The cylinders system in fig. 4.14 provides an example of a holonomic constraint. There, $G(\theta,t) = R\,\theta_1 - a\,(\theta_2 - \theta_1) = 0$. As an example of a problem with an integral constraint, suppose we want to know the shape of a hanging rope of fixed length C. This means we minimize the rope's potential energy,

$$U[y(x)] = \rho g \int\limits_{x_a}^{x_b} ds\; y(x) = \rho g \int\limits_{x_a}^{x_b} dx\; y\sqrt{1 + y'^2} \quad , \tag{4.327}$$

where ρ is the linear mass density of the rope, subject to the fixed-length constraint

$$C = \int_{s_a}^{s_b} ds = \int_{x_a}^{x_b} dx\,\sqrt{1+y'^2} \quad . \tag{4.328}$$

Note $ds = \sqrt{dx^2+dy^2}$ is the differential element of arc length along the rope. To solve problems like these, we again use the method of Lagrange multipliers.

4.10.4 Extremization of functionals : integral constraints

Given a functional

$$F\big[\{y_\sigma(x)\}\big] = \int_{x_a}^{x_b} dx\, L\big(\{y_\sigma\},\{y'_\sigma\},x\big) \qquad (\sigma = 1,\ldots,n) \tag{4.329}$$

subject to boundary conditions $\delta y_\sigma(x_a) = \delta y_\sigma(x_b) = 0$ and k constraints of the form

$$\int_{x_a}^{x_b} dx\, N_l\big(\{y_\sigma\},\{y'_\sigma\},x\big) = C_l \qquad (l = 1,\ldots,k) \quad , \tag{4.330}$$

construct the extended functional

$$F^*\big[\{y_\sigma(x)\};\{\lambda_j\}\big] \equiv \int_{x_a}^{x_b} dx\, \bigg\{ L\big(\{y_\sigma\},\{y'_\sigma\},x\big) + \sum_{l=1}^k \lambda_l\, N_l\big(\{y_\sigma\},\{y'_\sigma\},x\big)\bigg\} - \sum_{l=1}^k \lambda_l\, C_l \tag{4.331}$$

and freely extremize over $\{y_1,\ldots,y_n;\lambda_1,\ldots,\lambda_k\}$. This results in $(n+k)$ equations

$$\begin{aligned} \frac{\partial L}{\partial y_\sigma} - \frac{d}{dx}\bigg(\frac{\partial L}{\partial y'_\sigma}\bigg) + \sum_{l=1}^k \lambda_l \bigg\{ \frac{\partial N_l}{\partial y_\sigma} - \frac{d}{dx}\bigg(\frac{\partial N_l}{\partial y'_\sigma}\bigg)\bigg\} &= 0 \qquad (\sigma = 1,\ldots,n) \\ \int_{x_a}^{x_b} dx\, N_l\big(\{y_\sigma\},\{y'_\sigma\},x\big) &= C_l \qquad (l = 1,\ldots,k) \quad . \end{aligned} \tag{4.332}$$

4.10.5 Extremization of functionals : holonomic constraints

Given a functional

$$F\big[\{y_\sigma(x)\}\big] = \int_{x_a}^{x_b} dx\, L\big(\{y_\sigma\},\{y'_\sigma\},x\big) \qquad (\sigma = 1,\ldots,n) \tag{4.333}$$

subject to boundary conditions $\delta y_\sigma(x_a) = \delta y_\sigma(x_b) = 0$ and k constraints of the form

$$G_j\big(\{y_\sigma(x)\},x\big) = 0 \qquad (j = 1,\ldots,k) \quad , \tag{4.334}$$

construct the extended functional

$$F^*\big[\{y_\sigma(x)\};\{\lambda_j(x)\}\big] \equiv \int\limits_{x_a}^{x_b} dx\, \bigg\{ L\big(\{y_\sigma\},\{y'_\sigma\},x\big) + \sum_{j=1}^{k} \lambda_j\, G_j\big(\{y_\sigma\},x\big)\bigg\} \tag{4.335}$$

and freely extremize over the $(n+k)$ *functions* $\{y_1(x),\ldots,y_n(x);\lambda_1(x),\ldots,\lambda_k(x)\}$:

$$\delta F^* = \int\limits_{x_a}^{x_b} dx\, \Bigg\{ \sum_{\sigma=1}^{n} \bigg(\frac{\partial L}{\partial y_\sigma} - \frac{d}{dx}\bigg(\frac{\partial L}{\partial y'_\sigma}\bigg) + \sum_{j=1}^{k} \lambda_j\, \frac{\partial G_j}{\partial y_\sigma}\bigg)\delta y_\sigma + \sum_{j=1}^{k} G_j\, \delta\lambda_j \Bigg\} = 0 \quad , \tag{4.336}$$

resulting in the $(n+k)$ equations

$$\begin{gathered} \frac{d}{dx}\bigg(\frac{\partial L}{\partial y'_\sigma}\bigg) - \frac{\partial L}{\partial y_\sigma} = \sum_{j=1}^{k} \lambda_j\, \frac{\partial G_j}{\partial y_\sigma} \qquad (\sigma = 1,\ldots,n) \\ G_j = 0 \qquad (j=1,\ldots,k) \quad . \end{gathered} \tag{4.337}$$

4.10.6 Examples of functional extremization with constraints

Hanging rope

We minimize the potential energy functional

$$U\big[y(x)\big] = \rho g \int\limits_{x_1}^{x_2} dx\; y\sqrt{1+y'^2} \quad , \tag{4.338}$$

where ρ is the linear mass density, subject to the constraint of fixed total length,

$$C\big[y(x)\big] = \int\limits_{x_1}^{x_2} dx\, \sqrt{1+y'^2} \quad . \tag{4.339}$$

Thus,

$$U^*\big[y(x),\lambda\big] = U\big[y(x)\big] + \lambda C\big[y(x)\big] = \int\limits_{x_1}^{x_2} dx\; L^*(y,y',x) \quad , \tag{4.340}$$

with

$$L^*(y,y',x) = (\rho g y + \lambda)\,\sqrt{1+y'^2} \quad . \tag{4.341}$$

Since $\partial L^*/\partial x = 0$ we have that

$$H = y'\,\frac{\partial L^*}{\partial y'} - L^* = -\frac{\rho g y + \lambda}{\sqrt{1+y'^2}} \tag{4.342}$$

is constant. Thus,

$$\frac{dy}{dx} = \pm H^{-1} \sqrt{(\rho g y + \lambda)^2 - H^2} \quad , \tag{4.343}$$

with solution

$$y(x) = -\frac{\lambda}{\rho g} + \frac{H}{\rho g} \cosh\left(\frac{\rho g}{H}\,(x-a)\right) \quad . \tag{4.344}$$

Here, H, a, and λ are constants to be determined by demanding $y(x_i) = y_i$ $(i = 1, 2)$, and that the total length of the rope is C.

Geodesic on a curved surface

Consider next the problem of a geodesic on a curved surface. Let the equation for the surface be

$$G(x, y, z) = 0 \quad . \tag{4.345}$$

We wish to extremize the distance,

$$D = \int\limits_a^b ds = \int\limits_a^b \sqrt{dx^2 + dy^2 + dz^2} \quad . \tag{4.346}$$

We introduce a parameter t defined on the unit interval: $t \in [0, 1]$, such that $x(0) = x_a$, $x(1) = x_b$, *etc.* Then D may be regarded as a functional, *viz.*

$$D\big[x(t), y(t), z(t)\big] = \int\limits_0^1 dt\, \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} \quad . \tag{4.347}$$

We impose the constraint by forming the extended functional, D^*:

$$D^*\big[x(t), y(t), z(t), \lambda(t)\big] \equiv \int\limits_0^1 dt \left\{ \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} + \lambda\, G(x, y, z) \right\} \quad , \tag{4.348}$$

and we demand that the first functional derivatives of D^* vanish:

$$\begin{aligned}
\frac{\delta D^*}{\delta x(t)} &= -\frac{d}{dt}\left(\frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}\right) + \lambda\,\frac{\partial G}{\partial x} = 0 \\
\frac{\delta D^*}{\delta y(t)} &= -\frac{d}{dt}\left(\frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}\right) + \lambda\,\frac{\partial G}{\partial y} = 0 \\
\frac{\delta D^*}{\delta z(t)} &= -\frac{d}{dt}\left(\frac{\dot{z}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}\right) + \lambda\,\frac{\partial G}{\partial z} = 0 \\
\frac{\delta D^*}{\delta \lambda(t)} &= G(x, y, z) = 0 \quad .
\end{aligned} \tag{4.349}$$

Thus,

$$\lambda(t) = \frac{v\ddot{x} - \dot{x}\dot{v}}{v^2\, \partial_x G} = \frac{v\ddot{y} - \dot{y}\dot{v}}{v^2\, \partial_y G} = \frac{v\ddot{z} - \dot{z}\dot{v}}{v^2\, \partial_z G} \quad , \tag{4.350}$$

with $v = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}$ and $\partial_x \equiv \frac{\partial}{\partial x}$, *etc.* These three equations are supplemented by $G(x, y, z) = 0$, which is the fourth.

4.10.7 Constraints in Lagrangian mechanics

Let us write our system of constraints in the differential form

$$\sum_{\sigma=1}^{n} g_{j\sigma}(q, t)\, dq_\sigma + h_j(q, t)\, dt = 0 \qquad (j = 1, \ldots, k) \quad . \tag{4.351}$$

If the partial derivatives satisfy

$$\frac{\partial g_{j\sigma}}{\partial q_{\sigma'}} = \frac{\partial g_{j\sigma'}}{\partial q_\sigma} \qquad , \qquad \frac{\partial g_{j\sigma}}{\partial t} = \frac{\partial h_j}{\partial q_\sigma} \quad , \tag{4.352}$$

then the k differentials can be integrated to give $dG_j(q, t) = 0$ for each $j \in \{1, \ldots, k\}$, where

$$g_{j\sigma} = \frac{\partial G_j}{\partial q_\sigma} \qquad , \qquad h_j = \frac{\partial G_j}{\partial t} \quad . \tag{4.353}$$

The action functional is

$$S\big[\{q_\sigma(t)\}\big] = \int\limits_{t_a}^{t_b} dt\, L\big(\{q_\sigma\}, \{\dot{q}_\sigma\}, t\big) \qquad (\sigma = 1, \ldots, n) \quad , \tag{4.354}$$

subject to boundary conditions $\delta q_\sigma(t_a) = \delta q_\sigma(t_b) = 0$. The first variation of S is given by

$$\delta S = \int\limits_{t_a}^{t_b} dt \sum_{\sigma=1}^{n} \left\{ \frac{\partial L}{\partial q_\sigma} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\sigma}\right) \right\} \delta q_\sigma \quad . \tag{4.355}$$

Since the $\{q_\sigma(t)\}$ are no longer independent, we cannot infer that the term in brackets vanishes for each index σ. What are the constraints on the variations $\delta q_\sigma(t)$? The constraints are expressed in terms of *virtual displacements* which take no time: $\delta t = 0$. Thus,

$$\sum_{\sigma=1}^{n} g_{j\sigma}(q, t)\, \delta q_\sigma(t) = 0 \quad , \tag{4.356}$$

where $j = 1, \ldots, k$ is the constraint index. We may now relax the constraint by introducing k undetermined functions $\lambda_j(t)$, by adding integrals of the above equations with undetermined coefficient functions to δS:

$$\sum_{\sigma=1}^{n} \left\{ \frac{\partial L}{\partial q_\sigma} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\sigma}\right) + \sum_{j=1}^{k} \lambda_j(t)\, g_{j\sigma}(q, t) \right\} \delta q_\sigma(t) = 0 \quad . \tag{4.357}$$

Now we can demand that the term in brackets vanish for all σ. Thus, we obtain a set of $(n+k)$ equations,

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\sigma}\right) - \frac{\partial L}{\partial q_\sigma} &= \sum_{j=1}^{k} \lambda_j(t)\, g_{j\sigma}(q,t) \equiv Q_\sigma \\ \sum_{\sigma=1}^{n} g_{j\sigma}(q,t)\, \dot{q}_\sigma + h_j(q,t) &= 0 \quad , \end{aligned} \tag{4.358}$$

in $(n+k)$ unknowns $\{q_1, \ldots, q_n, \lambda_1, \ldots, \lambda_k\}$. Here, Q_σ is the *force of constraint conjugate to the generalized coordinate* q_σ. Thus, with

$$p_\sigma = \frac{\partial L}{\partial \dot{q}_\sigma} \quad , \quad F_\sigma = \frac{\partial L}{\partial q_\sigma} \quad , \quad Q_\sigma = \sum_{j=1}^{k} \lambda_j\, g_{j\sigma} \quad , \tag{4.359}$$

we write Newton's second law as

$$\dot{p}_\sigma = F_\sigma + Q_\sigma \quad . \tag{4.360}$$

Note that we can write

$$\frac{\delta S}{\delta \boldsymbol{q}(t)} = \frac{\partial L}{\partial \boldsymbol{q}} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\boldsymbol{q}}}\right) \tag{4.361}$$

and that the *instantaneous* constraints may be written

$$\boldsymbol{g}_j \cdot \delta \boldsymbol{q} = 0 \qquad (j = 1, \ldots, k) \quad . \tag{4.362}$$

Thus, by demanding

$$\frac{\delta S}{\delta \boldsymbol{q}(t)} + \sum_{j=1}^{k} \lambda_j\, \boldsymbol{g}_j = 0 \tag{4.363}$$

we require that the functional derivative be linearly dependent on the k vectors $\boldsymbol{g}_j$.

4.10.8 Constraints and conservation laws

We have seen how invariance of the Lagrangian with respect to a one-parameter family of coordinate transformations results in an associated conserved quantity Λ, and how a lack of explicit time dependence in L results in the conservation of the Hamiltonian H. In deriving both these results, however, we used the equations of motion $\dot{p}_\sigma = F_\sigma$. What happens when we have constraints, in which case $\dot{p}_\sigma = F_\sigma + Q_\sigma$?

Let's begin with the Hamiltonian. We have $H = \dot{q}_\sigma\, p_\sigma - L$, hence

$$\begin{aligned} \frac{dH}{dt} &= \overbrace{\left(p_\sigma - \frac{\partial L}{\partial \dot{q}_\sigma}\right)}^{\text{this vanishes}} \ddot{q}_\sigma + \overbrace{\left(\dot{p}_\sigma - \frac{\partial L}{\partial q_\sigma}\right)}^{\text{this is } \dot{Q}_\sigma} \dot{q}_\sigma - \frac{\partial L}{\partial t} \\ &= Q_\sigma\, \dot{q}_\sigma - \frac{\partial L}{\partial t} \quad . \end{aligned} \tag{4.364}$$

We now use

$$Q_\sigma\, \dot{q}_\sigma = \lambda_j\, g_{j\sigma}\, \dot{q}_\sigma = -\lambda_j\, h_j \tag{4.365}$$

to obtain

$$\frac{dH}{dt} = -\lambda_j\, h_j - \frac{\partial L}{\partial t} \quad . \tag{4.366}$$

We therefore conclude that *in a system with constraints of the form* $g_{j\sigma}\, \dot{q}_\sigma + h_j = 0$, *the Hamiltonian is conserved if each* $h_j = 0$ *and if* L *is not explicitly dependent on time.* In the case of holonomic constraints, $h_j = \frac{\partial G_j}{\partial t}$, so H is conserved if neither L nor any of the constraints G_j is explicitly time-dependent.

Next, let us rederive Noether's theorem when constraints are present. We assume a one-parameter family of transformations $q_\sigma \to \tilde{q}_\sigma(\zeta)$ leaves L invariant. Then

$$\begin{aligned} 0 = \frac{dL}{d\zeta} &= \frac{\partial L}{\partial \tilde{q}_\sigma}\,\frac{\partial \tilde{q}_\sigma}{\partial \zeta} + \frac{\partial L}{\partial \dot{\tilde{q}}_\sigma}\,\frac{\partial \dot{\tilde{q}}_\sigma}{\partial \zeta} \\ &= \big(\dot{\tilde{p}}_\sigma - \tilde{Q}_\sigma\big)\,\frac{\partial \tilde{q}_\sigma}{\partial \zeta} + \tilde{p}_\sigma\,\frac{d}{dt}\bigg(\frac{\partial \tilde{q}_\sigma}{\partial \zeta}\bigg) \\ &= \frac{d}{dt}\bigg(\tilde{p}_\sigma\,\frac{\partial \tilde{q}_\sigma}{\partial \zeta}\bigg) - \lambda_j\,\tilde{g}_{j\sigma}\,\frac{\partial \tilde{q}_\sigma}{\partial \zeta} \quad . \end{aligned} \tag{4.367}$$

Now let us write the constraints in differential form as

$$\tilde{g}_{j\sigma}\, d\tilde{q}_\sigma + \tilde{h}_j\, dt + \tilde{k}_j\, d\zeta = 0 \quad . \tag{4.368}$$

We now have

$$\frac{d\Lambda}{dt} = \lambda_j\, \tilde{k}_j \quad , \tag{4.369}$$

which says that *if the constraints are independent of* ζ *then* Λ *is conserved.* For holonomic constraints, this means that

$$G_j\big(\tilde{q}(\zeta), t\big) = 0 \quad \Rightarrow \quad \tilde{k}_j = \frac{\partial G_j}{\partial \zeta} = 0 \quad , \tag{4.370}$$

i.e. $G_j(\tilde{q}, t)$ has no explicit ζ dependence.

4.11 Constraints: Worked Examples

Here we consider several example problems of constrained dynamics, and work each out in full detail.

4.11.1 One cylinder rolling off another

As an example of the constraint formalism, consider the system in fig. 4.14, where a cylinder of radius a rolls atop a cylinder of radius R. We have two constraints:

$$G_1(r, \theta_1, \theta_2) = r - R - a = 0 \qquad \text{(cylinders in contact)} \tag{4.371}$$

$$G_2(r, \theta_1, \theta_2) = R\,\theta_1 - a\,(\theta_2 - \theta_1) = 0 \qquad \text{(no slipping)} \quad , \tag{4.372}$$

from which we obtain the $g_{j\sigma}$:

$$g_{j\sigma} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & R+a & -a \end{pmatrix} \quad , \tag{4.373}$$

which is to say

$$\begin{aligned} \frac{\partial G_1}{\partial r} &= 1 \quad , \quad \frac{\partial G_1}{\partial \theta_1} = 0 \quad , \quad \frac{\partial G_1}{\partial \theta_2} = 0 \\ \frac{\partial G_2}{\partial r} &= 0 \quad , \quad \frac{\partial G_2}{\partial \theta_1} = R + a \quad , \quad \frac{\partial G_2}{\partial \theta_2} = -a \quad . \end{aligned} \tag{4.374}$$

The Lagrangian is

$$L = T - U = \tfrac{1}{2} M \big(\dot{r}^2 + r^2 \dot{\theta}_1^2\big) + \tfrac{1}{2} I \dot{\theta}_2^2 - M g r \, \cos\theta_1 \quad , \tag{4.375}$$

where M and I are the mass and rotational inertia of the rolling cylinder, respectively. Note that the kinetic energy is a sum of center-of-mass translation $T_{\rm tr} = \frac{1}{2} M \big(\dot{r}^2 + r^2 \, \dot{\theta}_1^2\big)$ and rotation about the center-of-mass, $T_{\rm rot} = \frac{1}{2} I \, \dot{\theta}_2^2$. The equations of motion are

$$\begin{aligned} \frac{d}{dt}\bigg(\frac{\partial L}{\partial \dot{r}}\bigg) - \frac{\partial L}{\partial r} &= M \ddot{r} - M r \, \dot{\theta}_1^2 + M g \cos\theta_1 = \lambda_1 \equiv Q_r \\ \frac{d}{dt}\bigg(\frac{\partial L}{\partial \dot{\theta}_1}\bigg) - \frac{\partial L}{\partial \theta_1} &= M r^2 \ddot{\theta}_1 + 2 M r \dot{r} \, \dot{\theta}_1 - M g r \sin\theta_1 = (R+a) \, \lambda_2 \equiv Q_{\theta_1} \\ \frac{d}{dt}\bigg(\frac{\partial L}{\partial \dot{\theta}_2}\bigg) - \frac{\partial L}{\partial \theta_2} &= I \ddot{\theta}_2 = -a \, \lambda_2 \equiv Q_{\theta_2} \quad . \end{aligned} \tag{4.376}$$

To these three equations we add the two constraints, resulting in five equations in the five unknowns $\big\{r, \theta_1, \theta_2, \lambda_1, \lambda_2\big\}$.

We solve by first implementing the constraints, which give $r = (R+a)$ a constant (*i.e.* $\dot{r} = 0$), and $\dot{\theta}_2 = \big(1 + \frac{R}{a}\big) \, \dot{\theta}_1$. Substituting these into the above equations gives

$$-M (R+a) \, \dot{\theta}_1^2 + M g \cos\theta_1 = \lambda_1 \tag{4.377}$$

$$M (R+a)^2 \ddot{\theta}_1 - M g (R+a) \sin\theta_1 = (R+a) \, \lambda_2 \tag{4.378}$$

$$I \bigg(\frac{R+a}{a}\bigg) \ddot{\theta}_1 = -a \lambda_2 \quad . \tag{4.379}$$

From eqn. 4.379 we obtain

$$\lambda_2 = -\frac{I}{a} \, \ddot{\theta}_2 = -\frac{R+a}{a^2} \, I \, \ddot{\theta}_1 \quad , \tag{4.380}$$

which we substitute into eqn. 4.378 to obtain

$$\bigg(M + \frac{I}{a^2}\bigg) (R+a)^2 \ddot{\theta}_1 - M g (R+a) \sin\theta_1 = 0 \quad . \tag{4.381}$$

Multiplying by $\dot{\theta}_1$, we obtain an exact differential, which may be integrated to yield

$$\tfrac{1}{2}M\left(1+\frac{I}{Ma^2}\right)\dot{\theta}_1^2 + \frac{Mg}{R+a}\,\cos\theta_1 = \frac{Mg}{R+a}\cos\theta_1^\circ \quad . \tag{4.382}$$

Here, we have assumed that $\dot{\theta}_1 = 0$ when $\theta_1 = \theta_1^\circ$, *i.e.* the rolling cylinder is released from rest at $\theta_1 = \theta_1^\circ$. Finally, inserting this result into eqn. 4.377, we obtain the radial force of constraint,

$$Q_r = \frac{Mg}{1+\alpha}\Big\{(3+\alpha)\,\cos\theta_1 - 2\cos\theta_1^\circ\Big\} \quad , \tag{4.383}$$

where $\alpha = I/Ma^2$ is a dimensionless parameter ($0 \le \alpha \le 1$). This is the radial component of the normal force between the two cylinders. When Q_r vanishes, the cylinders lose contact – the rolling cylinder flies off. Clearly this occurs at an angle $\theta_1 = \theta_1^*$, where

$$\theta_1^* = \cos^{-1}\left(\frac{2\cos\theta_1^\circ}{3+\alpha}\right) \quad . \tag{4.384}$$

The detachment angle θ_1^* is an increasing function of α, which means that larger I delays detachment. This makes good sense, since when I is larger the gain in kinetic energy is split between translational and rotational motion of the rolling cylinder. Note also that $Q_r(\theta_1^\circ) = Mg\cos\theta_1^\circ$ balances the initial radial component of the force of gravity.

Finally, note that the differential equation

$$dt = \left(\frac{R+a}{2g}\right)^{1/2} \frac{d\theta}{\sqrt{\cos\theta_1^\circ - \cos\theta_1}} \tag{4.385}$$

may be integrated to yield $\theta_1(t)$ for $t \in [0, t^*]$, where $\theta_1(t^*) = \theta_1^*$, *i.e.* t^* is the time to detachment.

4.11.2 Frictionless motion along a curve

Consider the situation in fig. 4.15 where a skier moves frictionlessly under the influence of gravity along a general curve $y = h(x)$. The Lagrangian for this problem is

$$L = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) - mgy \tag{4.386}$$

and the (holonomic) constraint is

$$G(x,y) = y - h(x) = 0 \quad . \tag{4.387}$$

Accordingly, the Euler-Lagrange equations are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\sigma}\right) - \frac{\partial L}{\partial q_\sigma} = \lambda\,\frac{\partial G}{\partial q_\sigma} \quad , \tag{4.388}$$

where $q_1 = x$ and $q_2 = y$. Thus, we obtain

$$\begin{aligned} m\ddot{x} &= -\lambda\, h'(x) = Q_x \\ m\ddot{y} + mg &= \lambda = Q_y \quad . \end{aligned} \tag{4.389}$$

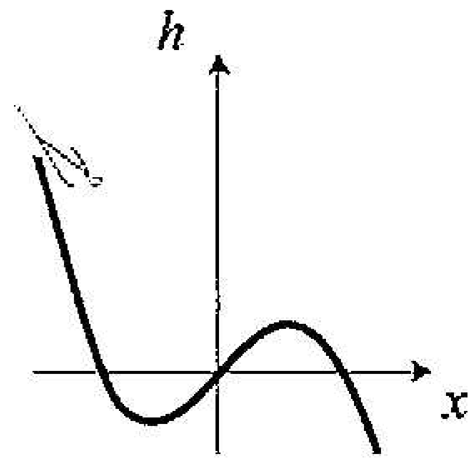

Figure 4.15: Frictionless motion under gravity along a curved surface. The skier flies off the surface when the normal force vanishes.

We eliminate y in favor of x by invoking the constraint. Since we need $\ddot{y}$, we must differentiate the constraint, which gives

$$\dot{y} = h'(x)\,\dot{x} \qquad , \qquad \ddot{y} = h'(x)\,\ddot{x} + h''(x)\,\dot{x}^2 \quad . \tag{4.390}$$

Using the second Euler-Lagrange equation, we then obtain

$$\frac{\lambda}{m} = g + h'(x)\,\ddot{x} + h''(x)\,\dot{x}^2 \quad . \tag{4.391}$$

Finally, we substitute this into the first E-L equation to obtain an equation for x alone:

$$\Big(1 + \big[h'(x)\big]^2\Big)\,\ddot{x} + h'(x)\,h''(x)\,\dot{x}^2 + g\,h'(x) = 0 \quad . \tag{4.392}$$

Had we started by eliminating $y = h(x)$ at the outset, writing

$$L(x,\dot{x}) = \tfrac{1}{2}m\Big(1 + \big[h'(x)\big]^2\Big)\,\dot{x}^2 - mg\,h(x) \quad , \tag{4.393}$$

we would also have obtained this equation of motion.

The skier flies off the curve when the vertical force of constraint $Q_y = \lambda$ starts to become negative, because the curve can only supply a positive normal force. Suppose the skier starts from rest at a height y_0. We may then determine the point x at which the skier detaches from the curve by setting $\lambda(x) = 0$. To do so, we must eliminate $\dot{x}$ and $\ddot{x}$ in terms of x. For $\ddot{x}$, we may use the equation of motion to write

$$\ddot{x} = -\bigg(\frac{gh' + h'\,h''\,\dot{x}^2}{1 + h'^2}\bigg) \quad , \tag{4.394}$$

which allows us to write

$$\lambda = m\bigg(\frac{g + h''\,\dot{x}^2}{1 + h'^2}\bigg) \quad . \tag{4.395}$$

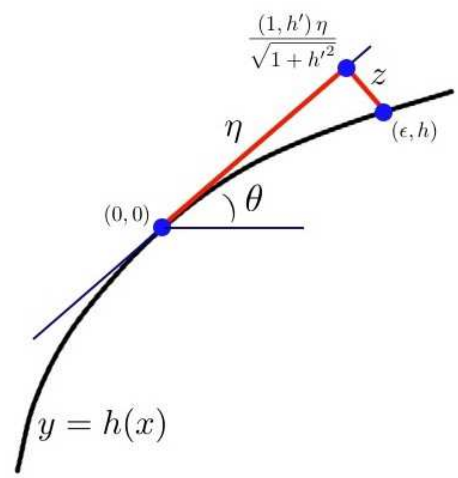

Figure 4.16: Finding the local radius of curvature: $z = \eta^2/2R$.

To eliminate $\dot{x}$, we use conservation of energy,

$$E = mgy_0 = \tfrac{1}{2}m\big(1 + h'^2\big)\,\dot{x}^2 + mgh \quad , \tag{4.396}$$

which fixes

$$\dot{x}^2 = 2g\bigg(\frac{y_0 - h}{1 + h'^2}\bigg) \quad . \tag{4.397}$$

Putting it all together, we have

$$\lambda(x) = \frac{mg}{\big(1 + h'^2\big)^2}\Big\{1 + h'^2 + 2(y_0 - h)\,h''\Big\} \quad . \tag{4.398}$$

The skier detaches from the curve when $\lambda(x) = 0$, *i.e.* when

$$1 + h'^2 + 2(y_0 - h)\,h'' = 0 \quad . \tag{4.399}$$

There is a somewhat easier way of arriving at the same answer. This is to note that the skier must fly off when the local centripetal force equals the gravitational force normal to the curve, *i.e.*

$$\frac{m\,v^2(x)}{R(x)} = mg\,\cos\theta(x) \quad , \tag{4.400}$$

where $R(x)$ is the local radius of curvature. Now $\tan\theta = h'$, so $\cos\theta = \big(1 + h'^2\big)^{-1/2}$. The square of the velocity is $v^2 = \dot{x}^2 + \dot{y}^2 = \big(1 + h'^2\big)\,\dot{x}^2$. What is the local radius of curvature $R(x)$? This can be determined from the following argument, and from the sketch in fig. 4.16. Writing $x = x^* + \epsilon$, we have

$$y = h(x^*) + h'(x^*)\,\epsilon + \tfrac{1}{2}h''(x^*)\,\epsilon^2 + \ldots \quad . \tag{4.401}$$

We now drop a perpendicular segment of length z from the point (x, y) to the line which is tangent to the curve at $\big(x^*, h(x^*)\big)$. According to fig. 4.16, this means

$$\begin{pmatrix} \epsilon \\ y \end{pmatrix} = \eta \cdot \tfrac{1}{\sqrt{1+h'^2}} \begin{pmatrix} 1 \\ h' \end{pmatrix} - z \cdot \tfrac{1}{\sqrt{1+h'^2}} \begin{pmatrix} -h' \\ 1 \end{pmatrix} \quad . \tag{4.402}$$

Thus, we have

$$\begin{aligned} y &= h'\,\epsilon + \tfrac{1}{2} h''\,\epsilon^2 \\ &= h' \left(\frac{\eta + z\,h'}{\sqrt{1+h'^2}} \right) + \tfrac{1}{2} h'' \left(\frac{\eta + z\,h'}{\sqrt{1+h'^2}} \right)^2 \\ &= \frac{\eta\,h' + z\,h'^2}{\sqrt{1+h'^2}} + \frac{h''\,\eta^2}{2\big(1+h'^2\big)} + \mathcal{O}(\eta z) = \frac{\eta\,h' - z}{\sqrt{1+h'^2}} \quad , \end{aligned} \tag{4.403}$$

from which we obtain

$$z = -\frac{h''\,\eta^2}{2\big(1+h'^2\big)^{3/2}} + \mathcal{O}(\eta^3) \tag{4.404}$$

and therefore

$$R(x) = -\frac{1}{h''(x)} \cdot \Big(1 + \big[h'(x)\big]^2\Big)^{3/2} \quad . \tag{4.405}$$

Thus, the detachment condition,

$$\frac{mv^2}{R} = -\frac{m\,h''\,\dot{x}^2}{\sqrt{1+h'^2}} = \frac{mg}{\sqrt{1+h'^2}} = mg\,\cos\theta \tag{4.406}$$

reproduces the result from eqn. 4.395.

4.11.3 Disk rolling down an inclined plane

A hoop of mass m and radius R rolls without slipping down an inclined plane. The inclined plane has opening angle α and mass M, and itself slides frictionlessly along a horizontal surface. Find the motion of the system.

Solution : Referring to the sketch in fig. 4.17, the center of the hoop is located at

$$\begin{aligned} x &= X + s\,\cos\alpha - a\,\sin\alpha \\ y &= s\,\sin\alpha + a\,\cos\alpha \quad , \end{aligned} \tag{4.407}$$

where X is the location of the lower left corner of the wedge, and s is the distance along the wedge to the bottom of the hoop. If the hoop rotates through an angle θ, the no-slip condition is $a\,\dot{\theta} + \dot{s} = 0$. Thus,

$$\begin{aligned} L &= \tfrac{1}{2} M \dot{X}^2 + \tfrac{1}{2} m \big(\dot{x}^2 + \dot{y}^2\big) + \tfrac{1}{2} I \dot{\theta}^2 - mgy \\ &= \tfrac{1}{2}\left(m + \frac{I}{a^2}\right)\dot{s}^2 + \tfrac{1}{2}(M+m)\dot{X}^2 + m\cos\alpha\,\dot{X}\,\dot{s} - mgs\,\sin\alpha - mga\,\cos\alpha \quad . \end{aligned} \tag{4.408}$$

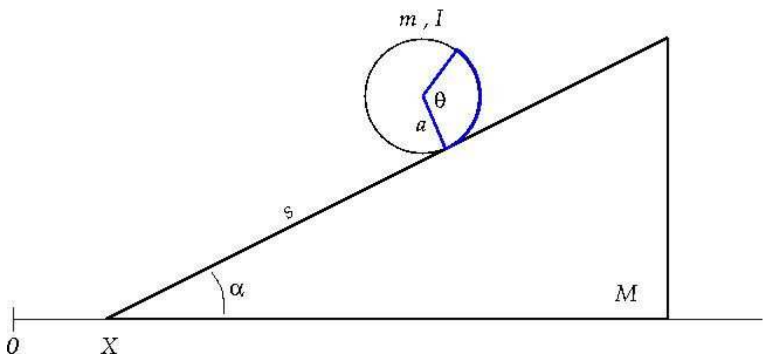

Figure 4.17: A hoop rolling down an inclined plane lying on a frictionless surface.

Since X is cyclic in L, the momentum

$$P_X = (M+m)\dot{X} + m\cos\alpha\,\dot{s} \quad , \tag{4.409}$$

is preserved: $\dot{P}_X = 0$. The second equation of motion, corresponding to the generalized coordinate s, is

$$\left(1+\frac{I}{ma^2}\right)\ddot{s} + \cos\alpha\,\ddot{X} = -g\,\sin\alpha \quad . \tag{4.410}$$

Using conservation of P_X, we eliminate $\ddot{s}$ in favor of $\ddot{X}$, and immediately obtain

$$\ddot{X} = \frac{g\sin\alpha\cos\alpha}{\left(1+\frac{M}{m}\right)\left(1+\frac{I}{ma^2}\right)-\cos^2\alpha} \equiv a_X \quad . \tag{4.411}$$

The result

$$\ddot{s} = -\frac{g\left(1+\frac{M}{m}\right)\sin\alpha}{\left(1+\frac{M}{m}\right)\left(1+\frac{I}{ma^2}\right)-\cos^2\alpha} \equiv a_s \tag{4.412}$$

follows immediately. Thus,

$$\begin{aligned} X(t) &= X(0) + \dot{X}(0)\,t + \tfrac{1}{2}a_X t^2 \\ s(t) &= s(0) + \dot{s}(0)\,t + \tfrac{1}{2}a_s t^2 \quad . \end{aligned} \tag{4.413}$$

Note that $a_s < 0$ while $a_X > 0$, *i.e.* the hoop rolls down and to the left as the wedge slides to the right. Note that $I = ma^2$ for a hoop; we've computed the answer here for general I.

4.11.4 Pendulum with nonrigid support

A particle of mass m is suspended from a flexible string of length ℓ in a uniform gravitational field. While hanging motionless in equilibrium, it is struck a horizontal blow resulting in an initial angular velocity ω_0. Treating the system as one with *two* degrees of freedom and a constraint, answer the following:

(a) Compute the Lagrangian, the equation of constraint, and the equations of motion.

Solution : The Lagrangian is

$$L = \tfrac{1}{2}m\left(\dot{r}^2 + r^2\,\dot{\theta}^2\right) + mgr\,\cos\theta \quad . \tag{4.414}$$

The constraint is $r = \ell$. The equations of motion are

$$\begin{aligned} m\ddot{r} - mr\,\dot{\theta}^2 - mg\,\cos\theta &= \lambda \\ mr^2\,\ddot{\theta} + 2mr\,\dot{r}\,\dot{\theta} - mg\sin\theta &= 0 \quad . \end{aligned} \tag{4.415}$$

(b) Compute the tension in the string as a function of angle θ.

Solution : Energy is conserved, hence

$$\tfrac{1}{2}m\ell^2\,\dot{\theta}^2 - mg\ell\cos\theta = \tfrac{1}{2}m\ell^2\,\dot{\theta}_0^2 - mg\ell\cos\theta_0 \quad . \tag{4.416}$$

We take $\theta_0 = 0$ and $\dot{\theta}_0 = \omega_0$. Thus,

$$\dot{\theta}^2 = \omega_0^2 - 2\,\Omega^2\left(1 - \cos\theta\right) \quad , \tag{4.417}$$

with $\Omega = \sqrt{g/\ell}$. Substituting this into the equation for λ, we obtain

$$\lambda = mg\left\{2 - 3\cos\theta - \frac{\omega_0^2}{\Omega^2}\right\} \quad . \tag{4.418}$$

(c) Show that if $\omega_0^2 < 2g/\ell$ then the particle's motion is confined below the horizontal and that the tension in the string is always positive (defined such that positive means exerting a pulling force and negative means exerting a pushing force). Note that the difference between a string and a rigid rod is that the string can only pull but the rod can pull or push. Thus, *the string tension must always be positive or else the string goes "slack"*.

Solution : Since $\dot{\theta}^2 \geq 0$, we must have

$$\frac{\omega_0^2}{2\Omega^2} \geq 1 - \cos\theta \quad . \tag{4.419}$$

The condition for slackness is $\lambda = 0$, or

$$\frac{\omega_0^2}{2\Omega^2} = 1 - \tfrac{3}{2}\cos\theta \quad . \tag{4.420}$$

Thus, if $\omega_0^2 < 2\Omega^2$, we have

$$1 > \frac{\omega_0^2}{2\Omega^2} > 1 - \cos\theta > 1 - \tfrac{3}{2}\cos\theta \quad , \tag{4.421}$$

and the string never goes slack. Note the last equality follows from $\cos\theta > 0$. The string rises to a maximum angle

$$\theta_{\text{max}} = \cos^{-1}\left(1 - \frac{\omega_0^2}{2\Omega^2}\right) \quad . \tag{4.422}$$

(d) Show that if $2g/\ell < \omega_0^2 < 5g/\ell$ the particle rises above the horizontal and the string becomes slack (the tension vanishes) at an angle θ^*. Compute θ^*.

Solution : When $\omega^2 > 2\Omega^2$, the string rises above the horizontal and goes slack at an angle

$$\theta^* = \cos^{-1}\left(\tfrac{2}{3} - \frac{\omega_0^2}{3\Omega^2}\right) \quad . \tag{4.423}$$

This solution craps out when the string is still taut at $\theta = \pi$, which means $\omega_0^2 = 5\Omega^2$.

(e) Show that if $\omega_0^2 > 5g/\ell$ the tension is always positive and the particle executes circular motion.

Solution : For $\omega_0^2 > 5\Omega^2$, the string never goes slack. Furthermore, $\dot\theta$ never vanishes. Therefore, the pendulum undergoes circular motion, albeit not with constant angular velocity.

4.11.5 Falling ladder

A uniform ladder of length ℓ and mass m has one end on a smooth horizontal floor and the other end against a smooth vertical wall. The ladder is initially at rest and makes an angle θ_0 with respect to the horizontal.

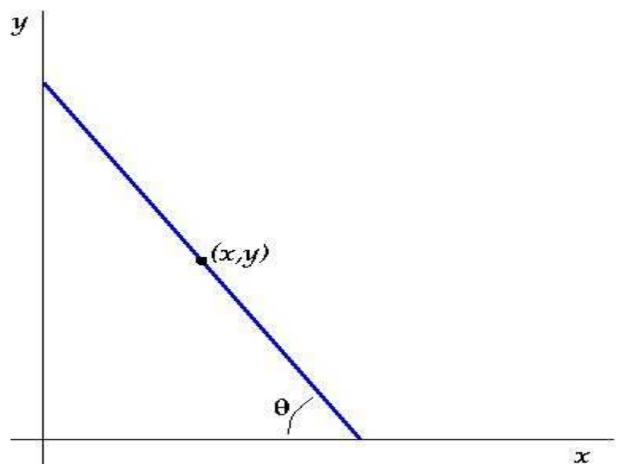

Figure 4.18: A ladder sliding down a wall and across a floor.

(a) Make a convenient choice of generalized coordinates and find the Lagrangian.

Solution : I choose as generalized coordinates the Cartesian coordinates (x, y) of the ladder's center of mass, and the angle θ it makes with respect to the floor. The Lagrangian is then

$$L = \tfrac{1}{2}\, m\, (\dot x^2 + \dot y^2) + \tfrac{1}{2}\, I\, \dot\theta^2 - mgy \quad . \tag{4.424}$$

There are two constraints: one enforcing contact along the wall, and the other enforcing contact along the floor. These are written

$$\begin{aligned} G_1(x, y, \theta) &= x - \tfrac{1}{2}\,\ell\,\cos\theta = 0 \\ G_2(x, y, \theta) &= y - \tfrac{1}{2}\,\ell\,\sin\theta = 0 \quad . \end{aligned} \tag{4.425}$$

(b) Prove that the ladder leaves the wall when its upper end has fallen to a height $\frac{2}{3} L \sin\theta_0$.

Solution : The equations of motion are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\sigma}\right) - \frac{\partial L}{\partial q_\sigma} = \sum_j \lambda_j \, \frac{\partial G_j}{\partial q_\sigma} \quad . \tag{4.426}$$

Thus, we have

$$\begin{aligned} m\ddot{x} &= \lambda_1 = Q_x \\ m\ddot{y} + mg &= \lambda_2 = Q_y \\ I\,\ddot{\theta} &= \tfrac{1}{2}\ell\big(\lambda_1 \sin\theta - \lambda_2 \cos\theta\big) = Q_\theta \quad . \end{aligned} \tag{4.427}$$

We now implement the constraints to eliminate x and y in terms of θ. We have

$$\begin{aligned} \dot{x} &= -\tfrac{1}{2}\,\ell\,\sin\theta\,\dot{\theta} \quad , \quad & \ddot{x} &= -\tfrac{1}{2}\,\ell\,\cos\theta\,\dot{\theta}^2 - \tfrac{1}{2}\,\ell\,\sin\theta\,\ddot{\theta} \\ \dot{y} &= \tfrac{1}{2}\,\ell\,\cos\theta\,\dot{\theta} \quad , \quad & \ddot{y} &= -\tfrac{1}{2}\,\ell\,\sin\theta\,\dot{\theta}^2 + \tfrac{1}{2}\,\ell\,\cos\theta\,\ddot{\theta} \quad . \end{aligned} \tag{4.428}$$

We can now obtain the forces of constraint in terms of the function $\theta(t)$:

$$\begin{aligned} \lambda_1 &= -\tfrac{1}{2}m\ell\,\big(\sin\theta\,\ddot{\theta} + \cos\theta\,\dot{\theta}^2\big) \\ \lambda_2 &= +\tfrac{1}{2}m\ell\,\big(\cos\theta\,\ddot{\theta} - \sin\theta\,\dot{\theta}^2\big) + mg \quad . \end{aligned} \tag{4.429}$$

We substitute these into the last equation of motion to obtain the result

$$I\,\ddot{\theta} = -I_0\,\ddot{\theta} - \tfrac{1}{2}mg\ell\,\cos\theta \quad , \tag{4.430}$$

which is to say $(1+\alpha)\,\ddot{\theta} = -2\omega_0^2\,\cos\theta$, with $I_0 = \frac{1}{4}m\ell^2$, $\alpha \equiv I/I_0$ and $\omega_0 = \sqrt{g/\ell}$. This may be integrated once (multiply by $\dot{\theta}$ to convert to a total derivative) to yield

$$\tfrac{1}{2}(1+\alpha)\,\dot{\theta}^2 + 2\,\omega_0^2\,\sin\theta = 2\,\omega_0^2\,\sin\theta_0 \quad , \tag{4.431}$$

which is of course a statement of energy conservation. This,

$$\dot{\theta}^2 = \frac{4\,\omega_0^2\,(\sin\theta_0 - \sin\theta)}{1+\alpha} \quad , \quad \ddot{\theta} = -\frac{2\,\omega_0^2\,\cos\theta}{1+\alpha} \quad . \tag{4.432}$$

We may now obtain $\lambda_1(\theta)$ and $\lambda_2(\theta)$:

$$\begin{aligned} \lambda_1(\theta) &= -\frac{mg}{1+\alpha}\,\big(3\sin\theta - 2\sin\theta_0\big)\,\cos\theta \\ \lambda_2(\theta) &= \frac{mg}{1+\alpha}\,\Big\{\big(3\sin\theta - 2\sin\theta_0\big)\,\sin\theta + \alpha\Big\} \quad . \end{aligned} \tag{4.433}$$

Demanding $\lambda_1(\theta) = 0$ gives the detachment angle $\theta = \theta_{\mathrm{d}}$, where

$$\sin\theta_{\mathrm{d}} = \tfrac{2}{3}\sin\theta_0 \quad . \tag{4.434}$$

Note that $\lambda_2(\theta_\mathrm{d}) = mg\alpha/(1+\alpha) > 0$, so the normal force from the floor is always positive for $\theta > \theta_\mathrm{d}$. The time to detachment is

$$T_1(\theta_0) = \int \frac{d\theta}{\dot\theta} = \frac{\sqrt{1+\alpha}}{2\,\omega_0} \int\limits_{\theta_\mathrm{d}}^{\theta_0} \frac{d\theta}{\sqrt{\sin\theta_0 - \sin\theta}} \quad . \tag{4.435}$$

(c) Show that the subsequent motion can be reduced to quadratures (*i.e.* explicit integrals).

Solution : After the detachment, there is no longer a constraint G_1. The equations of motion are

$$\begin{aligned} m\ddot x &= 0 \qquad \text{(conservation of } x\text{-momentum)} \\ m\ddot y + m\,g &= \lambda \\ I\,\ddot\theta &= -\tfrac{1}{2}\,\ell\,\lambda\,\cos\theta \quad , \end{aligned} \tag{4.436}$$

along with the constraint $y = \frac{1}{2}\,\ell\,\sin\theta$. Eliminating y in favor of θ using the constraint, the second equation yields

$$\lambda = mg - \tfrac{1}{2}m\ell\,\sin\theta\,\dot\theta^2 + \tfrac{1}{2}m\ell\,\cos\theta\,\ddot\theta \quad . \tag{4.437}$$

Plugging this into the third equation of motion, we find

$$I\,\ddot\theta = -2\,I_0\,\omega_0^2\,\cos\theta + I_0\,\sin\theta\,\cos\theta\,\dot\theta^2 - I_0\,\cos^2\theta\,\ddot\theta \quad . \tag{4.438}$$

Multiplying by $\dot\theta$ one again obtains a total time derivative, which is equivalent to rediscovering energy conservation:

$$E = \tfrac{1}{2}\big(I + I_0\,\cos^2\theta\big)\,\dot\theta^2 + 2\,I_0\,\omega_0^2\,\sin\theta \quad . \tag{4.439}$$

By continuity with the first phase of the motion, we obtain the initial conditions for this second phase:

$$\theta = \sin^{-1}\big(\tfrac{2}{3}\sin\theta_0\big) \qquad , \qquad \dot\theta = -2\,\omega_0\sqrt{\frac{\sin\theta_0}{3\,(1+\alpha)}} \quad . \tag{4.440}$$

Thus,

$$\begin{aligned} E &= \tfrac{1}{2}\big(I + I_0 - \tfrac{4}{9}\,I_0\,\sin^2\theta_0\big)\cdot\frac{4\,\omega_0^2\,\sin\theta_0}{3\,(1+\alpha)} + \tfrac{1}{3}\,mg\ell\,\sin\theta_0 \\ &= 2\,I_0\,\omega_0^2\cdot\bigg\{1 + \tfrac{4}{27}\,\frac{\sin^2\theta_0}{1+\alpha}\bigg\}\sin\theta_0 \quad . \end{aligned} \tag{4.441}$$

(d) Find an expression for the time $T(\theta_0)$ it takes the ladder to smack against the floor. Note that, expressed in units of the time scale $\sqrt{L/g}$, T is a dimensionless function of θ_0. Numerically integrate this expression and plot T *versus* θ_0.

Solution : The time from detachment to smack is

$$T_2(\theta_0) = \int \frac{d\theta}{\dot\theta} = \frac{1}{2\,\omega_0}\int\limits_0^{\theta_\mathrm{d}} d\theta\,\sqrt{\frac{1+\alpha\,\cos^2\theta}{\big(1-\frac{4}{27}\frac{\sin^2\theta_0}{1+\alpha}\big)\sin\theta_0 - \sin\theta}} \quad . \tag{4.442}$$

```
In[37]:= T[x_] := NIntegrate[√((4 / 3) / (x - Sin[y])), {y, ArcSin[2 x / 3], ArcSin[x] - 10^-9}] / 2

In[38]:= S[x_] := NIntegrate[
           √((1 + (4 / 3) (Cos[y]) ^2) / ((1 - (x / 3) ^2) x - Sin[y])), {y, 0, ArcSin[2 x / 3]}] / 2

In[39]:= Q[x_] := T[x] + S[x]

In[43]:= Plot[Q[x], {x, 0, 1}]
```

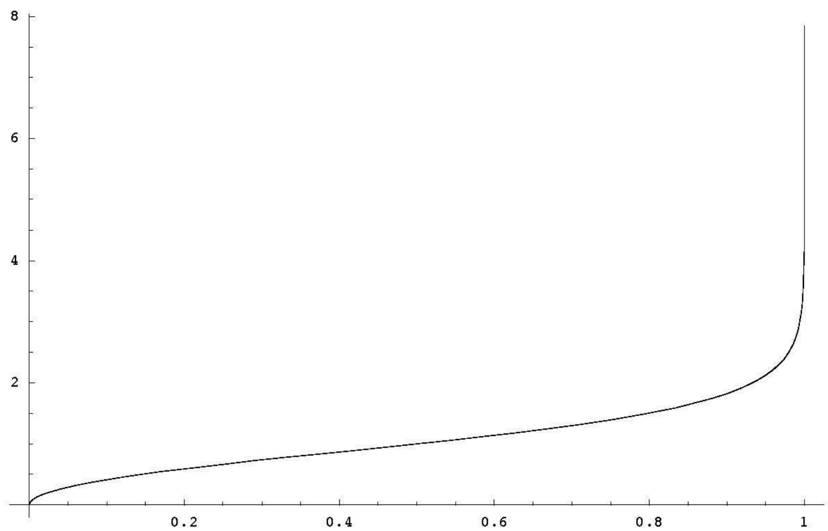

Figure 4.19: Plot of time to fall for the slipping ladder. Here $x = \sin\theta_0$.

The total time is then $T(\theta_0) = T_1(\theta_0) + T_2(\theta_0)$. For a uniformly dense ladder, $I = \frac{1}{12}\,m\ell^2 = \frac{1}{3}\,I_0$, and therefore $\alpha = \frac{1}{3}$.

(e) What is the horizontal velocity of the ladder at long times?

Solution : From the moment of detachment, and thereafter,

$$\dot{x} = -\tfrac{1}{2}\,\ell\,\sin\theta\,\dot{\theta} = \sqrt{\frac{4\,g\,\ell}{27\,(1+\alpha)}}\;\sin^{3/2}\!\theta_0 \quad . \tag{4.443}$$

(f) Describe in words the motion of the ladder subsequent to it slapping against the floor.

Solution : Only a fraction of the ladder's initial potential energy is converted into kinetic energy of horizontal motion. The rest is converted into kinetic energy of vertical motion and of rotation. The slapping of the ladder against the floor is an elastic collision. After the collision, the ladder must rise again, and continue to rise and fall *ad infinitum*, as it slides along with constant horizontal velocity.

4.11.6 Point mass inside rolling hoop

Consider the point mass m inside the hoop of radius R, depicted in fig. 4.20. We choose as generalized coordinates the Cartesian coordinates (X, Y) of the center of the hoop, the Cartesian coordinates (x, y) for the point mass, the angle ϕ through which the hoop turns, and the angle θ which the point mass makes with respect to the vertical. These six coordinates are not all independent. Indeed, there are only two independent coordinates for this system, which can be taken to be θ and ϕ. Thus, there are *four* constraints:

$$\begin{aligned} X - R\phi &\equiv G_1 = 0 \\ Y - R &\equiv G_2 = 0 \\ x - X - R\sin\theta &\equiv G_3 = 0 \\ y - Y + R\cos\theta &\equiv G_4 = 0 \quad . \end{aligned} \tag{4.444}$$

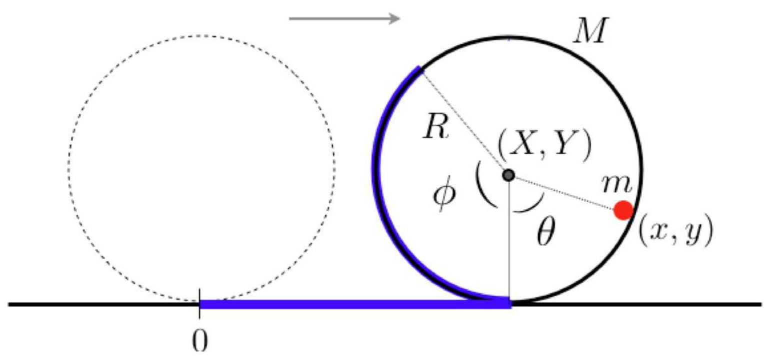

Figure 4.20: A point mass m inside a hoop of mass M, radius R, and moment of inertia I.

The kinetic and potential energies are easily expressed in terms of the Cartesian coordinates, aside from the energy of rotation of the hoop about its CM, which is expressed in terms of $\dot{\phi}$:

$$\begin{aligned} T &= \tfrac{1}{2}M(\dot{X}^2 + \dot{Y}^2) + \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \tfrac{1}{2}I\,\dot{\phi}^2 \\ U &= MgY + mgy \quad . \end{aligned} \tag{4.445}$$

The moment of inertia of the hoop about its CM is $I = MR^2$, but we could imagine a situation in which I were different. For example, we could instead place the point mass inside a very short cylinder with two solid end caps, in which case $I = \frac{1}{2}MR^2$. The Lagrangian is then

$$L = \tfrac{1}{2}M(\dot{X}^2 + \dot{Y}^2) + \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \tfrac{1}{2}I\,\dot{\phi}^2 - MgY - mgy \quad . \tag{4.446}$$

Note that L as written is completely independent of θ and $\dot{\theta}$!

Continuous symmetry

Note that there is a continuous symmetry to L which is satisfied by all the constraints, under

$$\tilde{X}(\zeta) = X+\zeta \quad , \quad \tilde{Y}(\zeta) = Y \quad , \quad \tilde{x}(\zeta) = x+\zeta \quad , \quad \tilde{y}(\zeta) = y \quad , \quad \tilde{\phi}(\zeta) = \phi+\frac{\zeta}{R} \quad , \quad \tilde{\theta}(\zeta) = \theta \quad . \tag{4.447}$$

Thus, according to Noether's theorem, there is a conserved quantity

$$\Lambda = \frac{\partial L}{\partial \dot{X}} + \frac{\partial L}{\partial \dot{x}} + \frac{1}{R}\,\frac{\partial L}{\partial \dot{\phi}} = M\dot{X} + m\dot{x} + \frac{I}{R}\,\dot{\phi} \quad . \tag{4.448}$$

This means $\dot{\Lambda} = 0$. This reflects the overall conservation of momentum in the x-direction.

Energy conservation

Since neither L nor any of the constraints are explicitly time-dependent, the Hamiltonian is conserved. And since T is homogeneous of degree two in the generalized velocities, we have $H = E = T + U$:

$$E = \tfrac{1}{2}M(\dot{X}^2 + \dot{Y}^2) + \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \tfrac{1}{2}I\,\dot{\phi}^2 + MgY + mgy \quad . \tag{4.449}$$

Equations of motion

We have $n = 6$ generalized coordinates and $k = 4$ constraints. Thus, there are four undetermined multipliers $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$ used to impose the constraints. This makes for ten unknowns: X, Y, x, y, ϕ, θ, λ_1, λ_2, λ_3, and λ_4. Accordingly, we have ten equations: six equations of motion plus the four equations of constraint. The equations of motion are obtained from

$$\frac{d}{dt}\bigg(\frac{\partial L}{\partial \dot{q}_\sigma}\bigg) = \frac{\partial L}{\partial q_\sigma} + \sum_{j=1}^{k} \lambda_j \, \frac{\partial G_j}{\partial q_\sigma} \quad . \tag{4.450}$$

Taking each generalized coordinate in turn, the equations of motion are thus

$$\begin{aligned}
M\ddot{X} &= \lambda_1 - \lambda_3 \\
M\ddot{Y} &= -Mg + \lambda_2 - \lambda_4 \\
m\ddot{x} &= \lambda_3 \\
m\ddot{y} &= -mg + \lambda_4 \\
I\,\ddot{\phi} &= -R\,\lambda_1 \\
0 &= -R\cos\theta\,\lambda_3 - R\sin\theta\,\lambda_4 \quad .
\end{aligned} \tag{4.451}$$

Along with the four constraint equations, these determine the motion of the system. Note that the last of the equations of motion, for the generalized coordinate $q_\sigma = \theta$, says that $Q_\theta = 0$, which means that the force of constraint on the point mass is radial. Were the point mass replaced by a rolling object, there would be an angular component to this constraint in order that there be no slippage.

Implementation of constraints

We now use the constraint equations to eliminate X, Y, x, and y in terms of θ and ϕ:

$$X = R\phi \quad , \quad Y = R \quad , \quad x = R\phi + R\sin\theta \quad , \quad y = R(1-\cos\theta) \quad . \tag{4.452}$$

We also need the derivatives:

$$\dot{x} = R\,\dot{\phi} + R\cos\theta\,\dot{\theta} \quad , \quad \ddot{x} = R\,\ddot{\phi} + R\cos\theta\,\ddot{\theta} - R\sin\theta\,\dot{\theta}^2 \quad , \tag{4.453}$$

and

$$\dot{y} = R\sin\theta\,\dot{\theta} \quad , \quad \ddot{y} = R\sin\theta\,\ddot{\theta} + R\cos\theta\,\dot{\theta}^2 \quad , \tag{4.454}$$

as well as

$$\dot{X} = R\,\dot{\phi} \quad , \quad \ddot{X} = R\,\ddot{\phi} \quad , \quad \dot{Y} = 0 \quad , \quad \ddot{Y} = 0 \quad . \tag{4.455}$$

We now may write the conserved charge as

$$\Lambda = \frac{1}{R}(I + MR^2 + mR^2)\,\dot{\phi} + mR\cos\theta\,\dot{\theta} \quad . \tag{4.456}$$

This, in turn, allows us to eliminate $\dot{\phi}$ in terms of $\dot{\theta}$ and the constant Λ:

$$\dot{\phi} = \frac{\gamma}{1+\gamma}\left(\frac{\Lambda}{mR} - \dot{\theta}\cos\theta\right) \quad , \tag{4.457}$$

where $\gamma = mR^2/(I + MR^2)$.

The energy is then

$$\begin{aligned} E &= \tfrac{1}{2}(I + MR^2)\,\dot{\phi}^2 + \tfrac{1}{2}m\big(R^2\,\dot{\phi}^2 + R^2\,\dot{\theta}^2 + 2R^2\cos\theta\,\dot{\phi}\,\dot{\theta}\big) + MgR + mgR(1-\cos\theta) \\ &= \tfrac{1}{2}mR^2\left\{\left(\frac{1+\gamma\sin^2\!\theta}{1+\gamma}\right)\dot{\theta}^2 + \frac{2g}{R}\,(1-\cos\theta) + \frac{\gamma}{1+\gamma}\left(\frac{\Lambda}{mR}\right)^2 + \frac{2Mg}{mR}\right\} \quad . \end{aligned} \tag{4.458}$$

The last two terms inside the big bracket are constant, so we can write this as

$$\left(\frac{1+\gamma\sin^2\!\theta}{1+\gamma}\right)\dot{\theta}^2 + \frac{2g}{R}\,(1-\cos\theta) = \frac{4gk}{R} \quad . \tag{4.459}$$

Here, k is a dimensionless measure of the energy of the system, after subtracting the aforementioned constants. If $k > 1$, then $\dot{\theta}^2 > 0$ for all θ, which would result in 'loop-the-loop' motion of the point mass inside the hoop – provided, that is, the normal force of the hoop doesn't vanish and the point mass doesn't detach from the hoop's surface.

Equation motion for $\theta(t)$

The equation of motion for θ obtained by eliminating all other variables from the original set of ten equations is the same as $\dot{E} = 0$, and may be written

$$\left(\frac{1+\gamma\sin^2\!\theta}{1+\gamma}\right)\ddot{\theta} + \left(\frac{\gamma\sin\theta\cos\theta}{1+\gamma}\right)\dot{\theta}^2 = -\frac{g}{R} \quad . \tag{4.460}$$

We can use this to write $\ddot{\theta}$ in terms of $\dot{\theta}^2$, and, after invoking eqn. 4.459, in terms of θ itself. We find

$$\begin{aligned} \dot{\theta}^2 &= \frac{4g}{R}\cdot\left(\frac{1+\gamma}{1+\gamma\sin^2\!\theta}\right)\big(k - \sin^2\tfrac{1}{2}\theta\big) \\ \ddot{\theta} &= -\frac{g}{R}\cdot\frac{(1+\gamma)\sin\theta}{\big(1+\gamma\sin^2\!\theta\big)^2}\Big[4\gamma\,\big(k - \sin^2\tfrac{1}{2}\theta\big)\cos\theta + 1 + \gamma\sin^2\!\theta\Big] \quad . \end{aligned} \tag{4.461}$$

Forces of constraint

We can solve for the λ_j, and thus obtain the forces of constraint

$$\begin{aligned}\lambda_3 = m\ddot{x} &= mR\,\ddot{\phi} + mR\cos\theta\,\ddot{\theta} - mR\sin\theta\,\dot{\theta}^2 \\ &= \frac{mR}{1+\gamma}\Big[\ddot{\theta}\,\cos\theta - \dot{\theta}^2\sin\theta\Big]\end{aligned} \tag{4.462}$$

and

$$\begin{aligned}\lambda_4 = m\ddot{y} + mg &= mg + mR\sin\theta\,\ddot{\theta} + mR\cos\theta\,\dot{\theta}^2 \\ &= mR\Big[\ddot{\theta}\,\sin\theta + \dot{\theta}^2\sin\theta + \frac{g}{R}\Big]\end{aligned} \tag{4.463}$$

and

$$\begin{aligned}\lambda_1 &= -\frac{I}{R}\,\ddot{\phi} = \frac{(1+\gamma)I}{mR^2}\,\lambda_3 \\ \lambda_2 &= (M+m)g + m\ddot{y} = \lambda_4 + Mg \quad .\end{aligned} \tag{4.464}$$

One can check that $\lambda_3\cos\theta + \lambda_4\sin\theta = 0$.

The condition that the normal force of the hoop on the point mass vanish is $\lambda_3 = 0$, which entails $\lambda_4 = 0$. This gives

$$-\,(1+\gamma\sin^2\!\theta)\,\cos\theta = 4(1+\gamma)\,\big(k - \sin^2\tfrac{1}{2}\theta\big) \quad . \tag{4.465}$$

Note that this requires $\cos\theta < 0$, *i.e.* the point of detachment lies above the horizontal diameter of the hoop. Clearly if k is sufficiently large, the equality cannot be satisfied, and the point mass executes a periodic 'loop-the-loop' motion. In particular, setting $\theta = \pi$, we find that

$$k_{\rm c} = 1 + \frac{1}{4(1+\gamma)} \quad . \tag{4.466}$$

If $k > k_{\rm c}$, then there is periodic 'loop-the-loop' motion. If $k < k_{\rm c}$, then the point mass may detach at a critical angle θ^*, but only if the motion allows for $\cos\theta < 0$. From the energy conservation equation, we have that the maximum value of θ achieved occurs when $\dot{\theta} = 0$, which means

$$\cos\theta_{\rm max} = 1 - 2k \quad . \tag{4.467}$$

If $\frac{1}{2} < k < k_{\rm c}$, then, we have the possibility of detachment. This means the energy must be large enough but not too large.

4.12 Appendix: Legendre Transformations

A *convex function* of a single variable $f(x)$ is one for which $f''(x) > 0$ everywhere. The *Legendre transform* of a convex function $f(x)$ is a function $g(p)$ defined as follows. Let p be a real number, and consider the

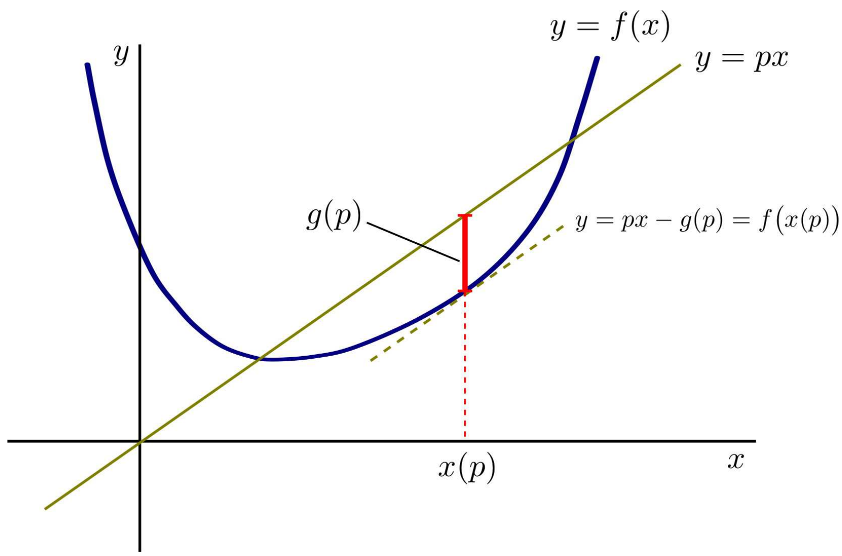

Figure 4.21: Construction for the Legendre transformation of a function $f(x)$.

line $y = px$, as shown in fig. 4.21. We define the point $x(p)$ as the value of x for which the difference $F(x,p) = px - f(x)$ is greatest. Then define $g(p) = F\big(x(p),p\big)$.[9] The value $x(p)$ is unique if $f(x)$ is convex, since $x(p)$ is determined by the equation

$$f'\big(x(p)\big) = p \quad . \tag{4.468}$$

Note that from $p = f'\big(x(p)\big)$ we have, according to the chain rule,

$$1 = \frac{d}{dp}\, f'\big(x(p)\big) = f''\big(x(p)\big)\, x'(p) \qquad \Longrightarrow \qquad x'(p) = \Big[f''\big(x(p)\big)\Big]^{-1} \quad . \tag{4.469}$$

From this, we can prove that $g(p)$ is itself convex:

$$\begin{aligned} g'(p) &= \frac{d}{dp}\Big[p\, x(p) - f\big(x(p)\big)\Big] \\ &= p\, x'(p) + x(p) - f'\big(x(p)\big)\, x'(p) = x(p) \quad , \end{aligned} \tag{4.470}$$

hence

$$g''(p) = x'(p) = \Big[f''\big(x(p)\big)\Big]^{-1} > 0 \quad . \tag{4.471}$$

In higher dimensions, the generalization of the definition $f''(x) > 0$ is that a function $F(x_1, \ldots, x_n)$ is convex if the matrix of second derivatives, called the *Hessian*,

$$H_{ij}(\boldsymbol{x}) = \frac{\partial^2 F}{\partial x_i\, \partial x_j} \tag{4.472}$$

[9] Note that $g(p)$ may be a negative number, if the line $y = px$ lies everywhere below $f(x)$.

is positive definite. That is, all the eigenvalues of $H_{ij}(\boldsymbol{x})$ must be positive for every $\boldsymbol{x}$. We then define the Legendre transform $\boldsymbol{G}(\boldsymbol{p})$ as

$$\boldsymbol{G}(\boldsymbol{p}) = \boldsymbol{p} \cdot \boldsymbol{x} - F(\boldsymbol{x}) \tag{4.473}$$

where $\boldsymbol{p} = \boldsymbol{\nabla} F$. Note that

$$dG = \boldsymbol{x} \cdot d\boldsymbol{p} + \boldsymbol{p} \cdot d\boldsymbol{x} - \boldsymbol{\nabla} F \cdot d\boldsymbol{x} = \boldsymbol{x} \cdot d\boldsymbol{p} \quad , \tag{4.474}$$

which establishes that G is a function of $\boldsymbol{p}$ and that

$$\frac{\partial G}{\partial p_j} = x_j \quad . \tag{4.475}$$

Note also that the Legendre transformation is *self dual*, which is to say that the Legendre transform of $G(\boldsymbol{p})$ is $F(\boldsymbol{x})$: $F \to G \to F$ under successive Legendre transformations.

We can also define a *partial Legendre transformation* as follows. Consider a function of q variables $F(\boldsymbol{x}, \boldsymbol{y})$, where $\boldsymbol{x} = \{x_1, \ldots, x_m\}$ and $\boldsymbol{y} = \{y_1, \ldots, y_n\}$, with $q = m + n$. Define $\boldsymbol{p} = \{p_1, \ldots, p_m\}$, and

$$G(\boldsymbol{p}, \boldsymbol{y}) = \boldsymbol{p} \cdot \boldsymbol{x} - F(\boldsymbol{x}, \boldsymbol{y}) \quad , \tag{4.476}$$

where

$$p_a = \frac{\partial F}{\partial x_a} \quad , \quad a \in \{1, \ldots, m\} \quad . \tag{4.477}$$

These equations are then to be inverted to yield

$$x_a = x_a(\boldsymbol{p}, \boldsymbol{y}) = \frac{\partial G}{\partial p_a} \quad . \tag{4.478}$$

Note that

$$p_a = \frac{\partial F}{\partial x_a} \big(\boldsymbol{x}(\boldsymbol{p}, \boldsymbol{y}), \boldsymbol{y}\big) \quad . \tag{4.479}$$

Thus, from the chain rule,

$$\delta_{ab} = \frac{\partial p_a}{\partial p_b} = \frac{\partial^2 F}{\partial x_a \, \partial x_c} \, \frac{\partial x_c}{\partial p_b} = \frac{\partial^2 F}{\partial x_a \, \partial x_c} \, \frac{\partial^2 G}{\partial p_c \, \partial p_b} \quad , \tag{4.480}$$

which says

$$\frac{\partial^2 G}{\partial p_a \, \partial p_b} = \frac{\partial x_a}{\partial p_b} = \mathsf{K}^{-1}_{ab} \quad , \tag{4.481}$$

where the $m \times m$ partial Hessian is

$$\frac{\partial^2 F}{\partial x_a \, \partial x_b} = \frac{\partial p_a}{\partial x_b} = \mathsf{K}_{ab} \quad . \tag{4.482}$$

Note that $\mathsf{K}_{ab} = \mathsf{K}_{ba}$ is symmetric. And with respect to the $\boldsymbol{y}$ coordinates,

$$\frac{\partial^2 G}{\partial y_\mu \, \partial y_\nu} = -\frac{\partial^2 F}{\partial y_\mu \, \partial y_\nu} = -\mathsf{L}_{\mu\nu} \quad , \tag{4.483}$$

where

$$\mathsf{L}_{\mu\nu} = \frac{\partial^2 F}{\partial y_\mu \, \partial y_\nu} \tag{4.484}$$

is the partial Hessian in the $\boldsymbol{y}$ coordinates. Now it is easy to see that if the full $q \times q$ Hessian matrix H_{ij} is positive definite, then any submatrix such as K_{ab} or $\mathsf{L}_{\mu\nu}$ must also be positive definite. In this case, the partial Legendre transform is convex in $\{p_1, \ldots, p_m\}$ and concave in $\{y_1, \ldots, y_n\}$.