Chapter 5

Calculus of Variations

5.1 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. 5.1. Let the path pass through the point $(0, y)$ on the interface. Then the time T is a function of y:

$$T(y) = \frac{1}{v_1} \sqrt{x_1^2 + (y - y_1)^2} + \frac{1}{v_2} \sqrt{x_2^2 + (y_2 - y)^2} \ . \tag{5.1}$$

To find the minimum time, we set

$$\begin{aligned} \frac{dT}{dy} = 0 &= \frac{1}{v_1} \frac{y - y_1}{\sqrt{x_1^2 + (y - y_1)^2}} + \frac{1}{v_2} \frac{y_2 - y}{\sqrt{x_2^2 + (y_2 - y)^2}} \\ &= \frac{\sin \theta_1}{v_1} - \frac{\sin \theta_2}{v_2} \ . \end{aligned} \tag{5.2}$$

Thus, the optimal path satisfies

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{v_1}{v_2} \ , \tag{5.3}$$

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. In terms of n,

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \ . \tag{5.4}$$

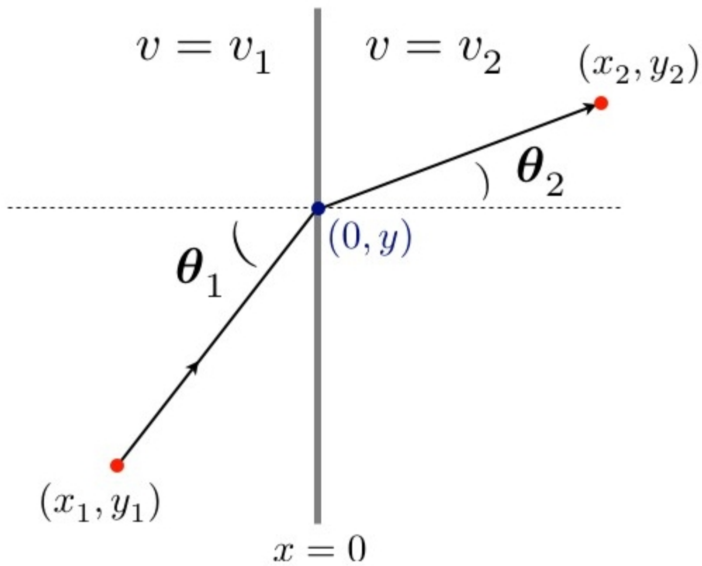

Figure 5.1: The shortest path between (x_1, y_1) and (x_2, y_2) is not a straight line, but rather two successive line segments of different slope.

If there are several interfaces, Snell's law holds at each one, so that

$$n_i \sin\theta_i = n_{i+1} \sin\theta_{i+1} \ , \tag{5.5}$$

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

Now let us imagine that there are many such interfaces between regions of very small thicknesses. We can then regard n and θ as continuous functions of the coordinate x. The differential form of Snell's law is

$$\begin{aligned} n(x)\,\sin\big(\theta(x)\big) &= n(x + dx)\,\sin\big(\theta(x + dx)\big) \\ &= (n + n'\,dx)\,\big(\sin\theta + \cos\theta\,\theta'\,dx\big) \\ &= n\,\sin\theta + \big(n'\,\sin\theta + n\,\cos\theta\,\theta'\big)\,dx \ . \end{aligned} \tag{5.6}$$

Thus,

$$\text{ctn}\,\theta\,\frac{d\theta}{dx} = -\frac{1}{n}\,\frac{dn}{dx} \ . \tag{5.7}$$

If we write the path as $y = y(x)$, then $\tan\theta = y'$, and

$$\theta' = \frac{d}{dx}\,\tan^{-1} y' = \frac{y''}{1 + y'^2} \ , \tag{5.8}$$

which yields

$$-\frac{1}{y'} \cdot \frac{y''}{1 + y'^2} = \frac{n'}{n} \ . \tag{5.9}$$

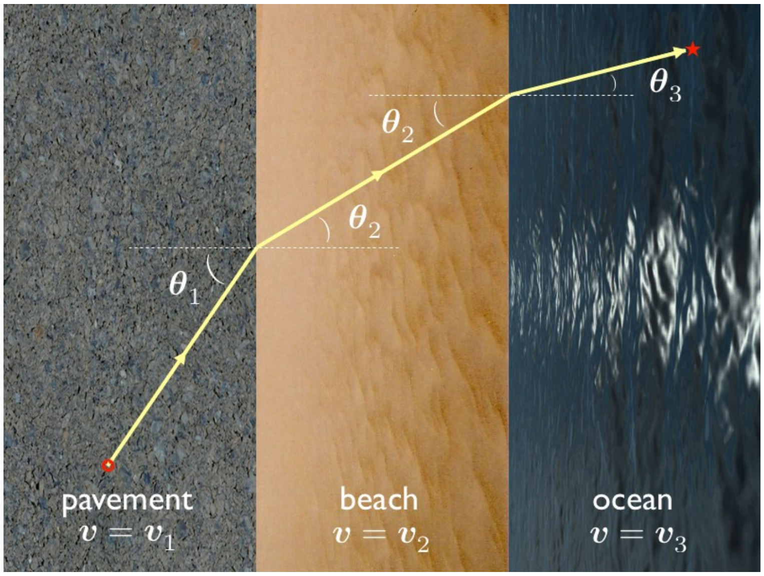

Figure 5.2: The path of shortest length is composed of three line segments. The relation between the angles at each interface is governed by Snell's Law.

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

$$T\big[y(x)\big] = \int \frac{ds}{v} = \frac{1}{c} \int\limits_{x_1}^{x_2} dx\, n(x)\, \sqrt{1 + y'^2} \tag{5.10}$$

is to be minimized.

5.2 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

$$T\big[y(x)\big] = \int\limits_{x_1}^{x_2} dx\, L(y, y', x) \ , \tag{5.11}$$

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

$$L(y, y', x) = c^{-1}\, n(x)\, \sqrt{1 + y'^2} \ . \tag{5.12}$$

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

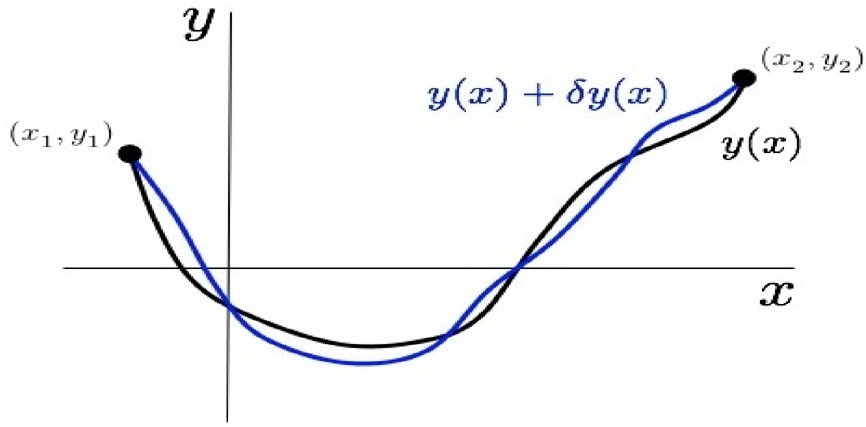

Figure 5.3: A path $y(x)$ and its variation $y(x) + \delta y(x)$.

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

$$f(x + dx) = f(x) + f'(x)\,dx + \tfrac{1}{2}\,f''(x)\,(dx)^2 + \ldots \; . \tag{5.13}$$

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

$$\begin{aligned} T\big[y(x) + \delta y(x)\big] &= \int_{x_1}^{x_2} dx\, L(y + \delta y, y' + \delta y', x) \\ &= \int_{x_1}^{x_2} dx \left\{ L + \frac{\partial L}{\partial y}\,\delta y + \frac{\partial L}{\partial y'}\,\delta y' + \mathcal{O}\big((\delta y)^2\big) \right\} \\ &= T\big[y(x)\big] + \int_{x_1}^{x_2} dx \left\{ \frac{\partial L}{\partial y}\,\delta y + \frac{\partial L}{\partial y'}\,\frac{d}{dx}\,\delta y \right\} \\ &= T\big[y(x)\big] + \int_{x_1}^{x_2} dx \left[\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) \right] \delta y + \frac{\partial L}{\partial y'}\,\delta y \,\Bigg|_{x_1}^{x_2} \; . \end{aligned} \tag{5.14}$$

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

$$\delta T = \int\limits_{x_1}^{x_2} dx \left[\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right)\right] \delta y \ . \tag{5.15}$$

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

$$\frac{\delta T}{\delta y(x)} = \left[\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right)\right]_x \ , \tag{5.16}$$

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 $y(x)$,

$$\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) = 0 \ , \tag{5.17}$$

known as the *Euler-Lagrange* equation. Since L is independent of y, we have

$$\begin{aligned} 0 = \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) &= \frac{1}{c}\frac{d}{dx}\left[\frac{n\,y'}{\sqrt{1+y'^2}}\right] \\ &= \frac{n'}{c}\,\frac{y'}{\sqrt{1+y'^2}} + \frac{n}{c}\,\frac{y''}{\big(1+y'^2\big)^{3/2}} \ . \end{aligned} \tag{5.18}$$

We thus recover the second order equation in 5.9. However, note that the above equation directly gives

$$n(x)\,\sin\theta(x) = \text{const.} \ , \tag{5.19}$$

which follows from the relation $y' = \tan\theta$. For $y(x)$ we obtain

$$\frac{n^2\,y'^2}{1+y'^2} \equiv \alpha^2 = \text{const.} \qquad \Rightarrow \qquad \frac{dy}{dx} = \frac{\alpha}{\sqrt{n^2(x)-\alpha^2}} \ . \tag{5.20}$$

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

$$\begin{aligned} F[y+\delta y] = F[y] &+ \int\! dx_1\, K_1(x_1)\,\delta y(x_1) + \tfrac{1}{2!}\int\! dx_1 \int\! dx_2\, K_2(x_1,x_2)\,\delta y(x_1)\,\delta y(x_2) \\ &+ \tfrac{1}{3!}\int\! dx_1 \int\! dx_2 \int\! dx_3\, K_3(x_1,x_2,x_3)\,\delta y(x_1)\,\delta y(x_2)\,\delta y(x_3) + \ldots \end{aligned} \tag{5.21}$$

and we write

$$K_n(x_1,\ldots,x_n) \equiv \frac{\delta^n F}{\delta y(x_1)\cdots\delta y(x_n)} \tag{5.22}$$

for the n^{th} functional derivative.

5.3 Examples from the Calculus of Variations

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

5.3.1 Example 1 : Minimal Surface of Revolution

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

$$A\big[y(x)\big] = \int\limits_{x_1}^{x_2} dx\, 2\pi y \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \ , \tag{5.23}$$

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 $y(x) \leftrightarrow q(t)$. We can then apply what we have derived for the mechanical action, with $L = L(q, \dot{q}, t)$, *mutatis mutandis.* Thus, the equation of motion is

$$\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) = \frac{\partial L}{\partial y} \ , \tag{5.24}$$

which is a second order ODE for $y(x)$. Rather than treat the second order equation, though, we can integrate once to obtain a first order equation, by noticing that

$$\begin{aligned} \frac{d}{dx}\left[y'\,\frac{\partial L}{\partial y'} - L\right] &= y''\,\frac{\partial L}{\partial y'} + y'\,\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) - \frac{\partial L}{\partial y'}\,y'' - \frac{\partial L}{\partial y}y' - \frac{\partial L}{\partial x} \\ &= y'\left[\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) - \frac{\partial L}{\partial y}\right] - \frac{\partial L}{\partial x} \ . \end{aligned} \tag{5.25}$$

In the second line above, the term in square brackets vanishes, thus

$$\mathcal{J} = y'\,\frac{\partial L}{\partial y'} - L \quad \Rightarrow \quad \frac{d\mathcal{J}}{dx} = -\frac{\partial L}{\partial x} \ , \tag{5.26}$$

and when L has no explicit x-dependence, $\mathcal{J}$ is conserved. One finds

$$\mathcal{J} = 2\pi y \cdot \frac{y'^2}{\sqrt{1 + y'^2}} - 2\pi y\,\sqrt{1 + y'^2} = -\frac{2\pi y}{\sqrt{1 + y'^2}} \ . \tag{5.27}$$

Solving for y',

$$\frac{dy}{dx} = \pm\sqrt{\left(\frac{2\pi y}{\mathcal{J}}\right)^2 - 1} \ , \tag{5.28}$$

which may be integrated with the substitution $y = \frac{\mathcal{J}}{2\pi} \cosh \chi$, yielding

$$y(x) = b \cosh\left(\frac{x-a}{b}\right) , \tag{5.29}$$

where a and $b = \frac{\mathcal{J}}{2\pi}$ 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

$$y_0 = b \cosh\left(\frac{x_0}{b}\right) \quad \Rightarrow \quad \gamma = \kappa^{-1} \cosh \kappa \ , \tag{5.30}$$

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 local minimum and one of which is a saddle point of $A[y(x)]$. The solution with the smaller value of κ (*i.e.* the larger value of $\operatorname{sech} \kappa$) yields the smaller value of A, as shown in Fig. 5.4. Note that

$$\frac{y}{y_0} = \frac{\cosh(x/b)}{\cosh(x_0/b)} \ , \tag{5.31}$$

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

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

$$y(x) = \begin{cases} y_1 & \text{if } x = x_1 \\ \\ 0 & \text{if } x_1 < x < x_2 \\ \\ y_2 & \text{if } x = x_2 \ . \end{cases} \tag{5.32}$$

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. 5.4, 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.

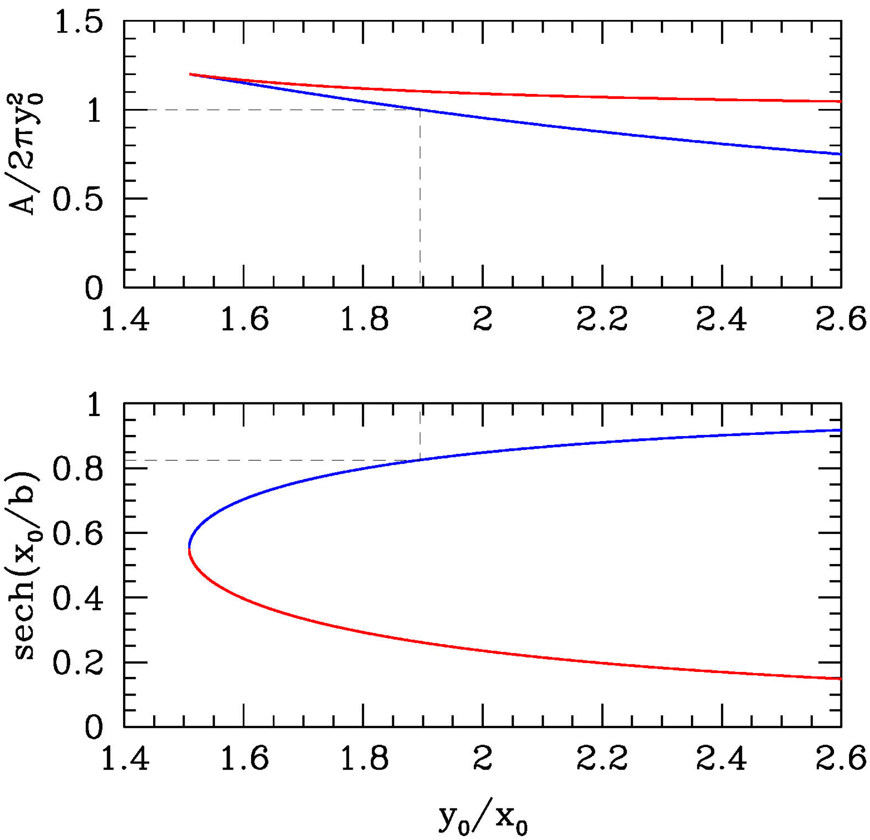

Figure 5.4: Minimal surface solution, with $y(x) = b\cosh(x/b)$ and $y(x_0) = y_0$. Top panel: $A/2\pi y_0^2$ *vs.* y_0/x_0. Bottom panel: $\mathrm{sech}(x_0/b)$ *vs.* y_0/x_0. The blue curve corresponds to a local minimum of $A[y(x)$, and the red curve to a saddle point.

Note that the functional derivative,

$$K_1(x) = \frac{\delta A}{\delta y(x)} = \left\{ \frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) \right\} = \frac{2\pi\big(1 + y'^2 - y y''\big)}{(1 + y'^2)^{3/2}} \ , \tag{5.33}$$

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. 5.4 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.

5.3.2 Example 2 : Geodesic on a Surface of Revolution

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

$$\begin{aligned} ds^2 &= d\rho^2 + \rho^2\, d\phi^2 + dx^2 \\ &= \Big\{ 1 + \big[z'(\rho)\big]^2 \Big\}\, d\rho + \rho^2\, d\phi^2 \ , \end{aligned} \tag{5.34}$$

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

$$D\big[\phi(\rho)\big] = \int\limits_{\rho_1}^{\rho_2} d\rho\, L(\phi, \phi', \rho) \ , \tag{5.35}$$

where

$$L(\phi, \phi', \rho) = \sqrt{1 + z'^2(\rho) + \rho^2\, \phi'^2(\rho)} \ . \tag{5.36}$$

The Euler-Lagrange equation is

$$\frac{\partial L}{\partial \phi} - \frac{d}{d\rho}\bigg(\frac{\partial L}{\partial \phi'}\bigg) = 0 \quad \Rightarrow \quad \frac{\partial L}{\partial \phi'} = \text{const.} \tag{5.37}$$

Thus,

$$\frac{\partial L}{\partial \phi'} = \frac{\rho^2\, \phi'}{\sqrt{1 + z'^2 + \rho^2\, \phi'^2}} = a \ , \tag{5.38}$$

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

$$d\phi = \frac{a\sqrt{1 + \big[z'(\rho)\big]^2}}{\rho\sqrt{\rho^2 - a^2}}\, d\rho \ , \tag{5.39}$$

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

$$d\phi = a\sqrt{1+\lambda^2}\, \frac{d\rho}{\rho\sqrt{\rho^2 - a^2}} = \sqrt{1+\lambda^2}\, d\tan^{-1}\sqrt{\frac{\rho^2}{a^2} - 1} \ , \tag{5.40}$$

which yields

$$\phi(\rho) = \beta + \sqrt{1+\lambda^2}\ \tan^{-1}\sqrt{\frac{\rho^2}{a^2} - 1} \ , \tag{5.41}$$

which is equivalent to

$$\rho\,\cos\left(\frac{\phi-\beta}{\sqrt{1+\lambda^2}}\right)=a\ . \tag{5.42}$$

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

5.3.3 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

$$\tfrac{1}{2}mv^2-mgy=mgy_1\ . \tag{5.43}$$

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

$$\begin{aligned}T\big[y(x)\big]&=\int\limits_{x_1}^{x_2}\frac{ds}{v}=\frac{1}{\sqrt{2g}}\int\limits_{x_1}^{x_2}\!dx\,\sqrt{\frac{1+y'^2}{y_1-y}}\\&\equiv\int\limits_{x_1}^{x_2}\!dx\,L(y,y',x)\ ,\end{aligned} \tag{5.44}$$

with

$$L(y,y',x)=\sqrt{\frac{1+y'^2}{2g(y_1-y)}}\ . \tag{5.45}$$

Since L is independent of x, eqn. 5.25, we have that

$$\mathcal{J}=y'\,\frac{\partial L}{\partial y'}-L=-\Big[2g\,(y_1-y)\,\big(1+y'^2\big)\Big]^{-1/2} \tag{5.46}$$

is conserved. This yields

$$dx=-\sqrt{\frac{y_1-y}{2a-y_1+y}}\,dy\ , \tag{5.47}$$

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

$$y_1-y=2a\sin^2(\tfrac{1}{2}\theta)\quad\Rightarrow\quad dx=2a\sin^2(\tfrac{1}{2}\theta)\,d\theta\ , \tag{5.48}$$

which results in the parametric equations

$$x-x_1=a\,\big(\theta-\sin\theta\big) \tag{5.49}$$

$$y-y_1=-a\,(1-\cos\theta)\ . \tag{5.50}$$

This curve is known as a *cycloid.*

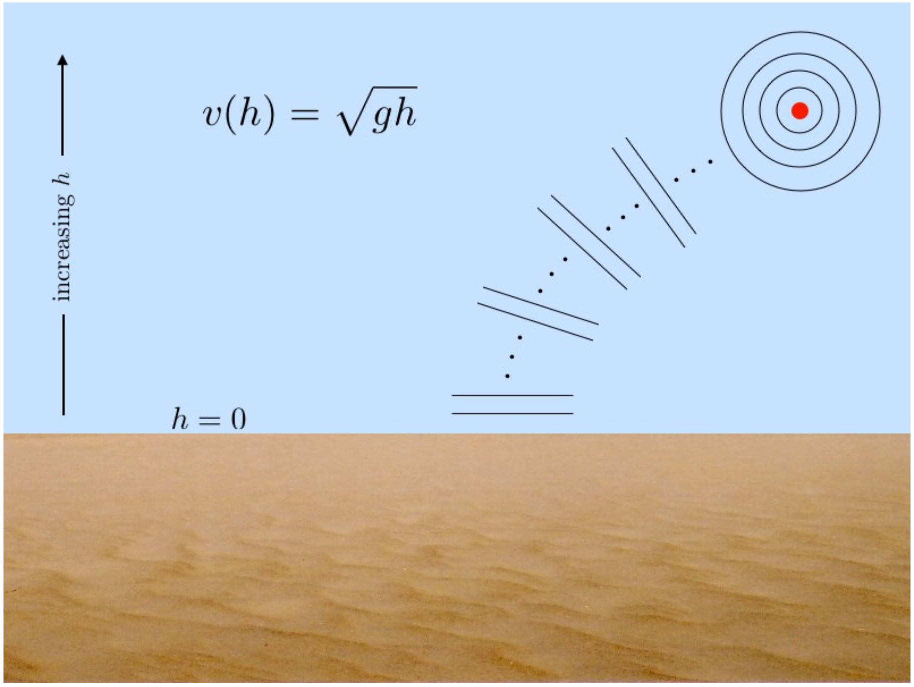

Figure 5.5: For shallow water waves, $v = \sqrt{gh}$. To minimize the propagation time from a source to the shore, the waves break parallel to the shoreline.

5.3.4 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

$$\omega(k) = \sqrt{gk\,\tanh kh} \approx \begin{cases} \sqrt{gh}\,k & \text{shallow } (kh \ll 1) \\ \\ \sqrt{gk} & \text{deep } (kh \gg 1)\ . \end{cases} \tag{5.51}$$

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

$$T\big[y(x)\big] = \int \frac{ds}{v} = \int_{x_1}^{x_2} dx \sqrt{\frac{1 + y'^2}{g\,h(y)}} \ . \tag{5.52}$$

We thus identify the integrand

$$L(y, y', x) = \sqrt{\frac{1 + y'^2}{g\,h(y)}} \ . \tag{5.53}$$

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

$$\mathcal{J} = y' \, \frac{\partial L}{\partial y'} - L = -\Big[g\,h(y)\,\big(1 + y'^2\big)\Big]^{-1/2} \tag{5.54}$$

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

$$\tan\theta = \frac{dy}{dx} = \sqrt{\frac{a}{h(y)} - 1} \ , \tag{5.55}$$

where $a = (g\mathcal{J})^{-1}$ 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

$$x(\theta) = b\,(\theta - \sin\theta) \tag{5.56}$$

$$y(\theta) = b\,(1 - \cos\theta) \ , \tag{5.57}$$

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

$$y(x) = \big(\tfrac{9}{2}\big)^{1/3} \, b^{1/3} \, x^{2/3} + \ldots \ . \tag{5.58}$$

A *tsunami* is a shallow water wave that manages 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 $h \sim 10\,\text{km}$. 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.