

## 2. The Calculus of Variations

Michael Fowler

### Introduction

We've seen how Whewell solved the problem of the equilibrium shape of chain hanging between two places, by finding how the forces on a length of chain, the tension at the two ends and its weight, balanced. We're now going to look at a completely different approach: the equilibrium configuration is an energy minimum, so small deviations from it can only make second-order changes in the gravitational potential energy. Here we'll find how analyzing that leads to a differential equation for the curve, and how the technique developed can be successfully applied of a vast array of problems.

### The Catenary and the Soap Film

The catenary is the curved configuration  $y = y(x)$  of a uniform inextensible rope with two fixed endpoints at rest in a constant gravitational field. That is to say, it is the curve that minimizes the gravitational potential energy

$$J[y(x)] = 2\pi \int_{x_1}^{x_2} y ds = 2\pi \int_{x_1}^{x_2} y \sqrt{1 + y'^2} dx, \quad (y' = dy/dx)$$

where we have taken the rope density and  $g$  both equal to unity for mathematical convenience.

Usually in calculus we minimize a function with respect to a single variable, or several variables. Here the potential energy is a function of a function, equivalent to an infinite number of variables, and our problem is to minimize it with respect to arbitrary small variations of that function. In other words, if we nudge the chain somewhere, and its motion is damped by air or internal friction, it will settle down again in the catenary configuration.

Formally speaking, there will be no change in that potential energy to leading order if we make an infinitesimal change in the curve,  $y(x) \rightarrow y(x) + \delta y(x)$  (subject of course to keeping the length the same, that is  $\delta \int ds = 0$ .)

This method of solving the problem is called the *calculus of variations*: in ordinary calculus, we make an infinitesimal change in a variable, and compute the corresponding change in a function, and if it's zero to leading order in the small change, we're at an extreme value.

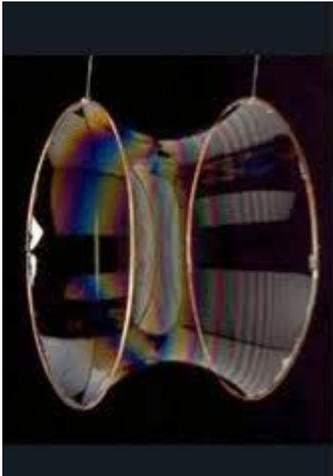
(*Nitpicking footnote*: Actually this assumes the second order term is nonzero—what about  $x^3$  near the origin? But such situations are infrequent in the problems we're likely to encounter.)

The difference here is that the potential energy of the hanging chain isn't just a function of a variable, or even of a number of variables—it's a *function of a function*, it depends on the position of every point on the chain (in the limit of infinitely small links, that is, or equivalently a continuous rope).

So, we're looking for the configuration where the potential energy doesn't change to first order for *any* infinitesimal change in the curve of its position, subject to fixed endpoints, and a fixed chain length.

As a warm up, we'll consider a simpler—but closely related—problem.

### A Soap Film Between Two Horizontal Rings: the Euler-Lagrange Equation



This problem is very similar to the catenary: surface tension will pull the soap film to the minimum possible total area compatible with the fixed boundaries (and neglecting gravity, which is a small effect).

(Interestingly, this problem is also closely related to *string theory*: as a closed string propagates, its path traces out as “world sheet” and the string dynamics is determined by that sheet having minimal area.)

Taking the axis of rotational symmetry to be the  $x$ -axis, and the radius  $y(x)$ , we need to find the function

$y(x)$  that minimizes the total area ( $ds$  is measured along the curve of the surface). Think of

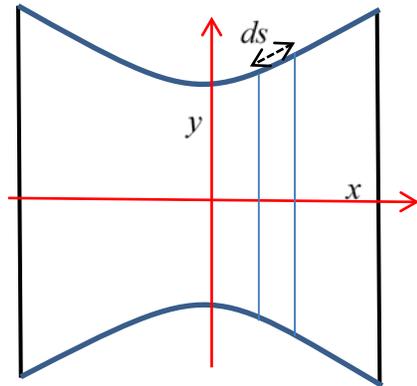
the soap film as a sequence of rings or collars, of radius  $y$ , and therefore area  $2\pi y ds$ . The total area is given by integrating, adding all these incremental collars,

$$J[y(x)] = 2\pi \int_{x_1}^{x_2} y ds = 2\pi \int_{x_1}^{x_2} y \sqrt{1 + y'^2} dx$$

given that  $y(x_1) = y_1$ ,  $y(x_2) = y_2$ .

(You might be thinking at this point: isn't this *identical* to the catenary equation? The answer is yes, *but* the chain has an additional requirement—it *has a fixed length*. The soap film is not constrained in that way, it can stretch or contract to minimize the total area, so this is a different problem!)

That is, we want  $\delta J = 0$  to first order, if we make a change  $y(x) \rightarrow y(x) + \delta y(x)$ . Of course, this also means  $y'(x) \rightarrow y'(x) + \delta y'(x)$  where  $\delta y' = \delta(dy/dx) = (d/dx)\delta y$ .



## General Method for the Minimization Problem

To emphasize the generality of the method, we'll just write

$$J[y] = \int_{x_1}^{x_2} f(y, y') dx \quad (y' = dy/dx).$$

Then under any infinitesimal variation  $\delta y(x)$  (equal to zero at the fixed endpoints)

$$\delta J[y] = \int_{x_1}^{x_2} \left[ \frac{\partial f(y, y')}{\partial y} \delta y(x) + \frac{\partial f(y, y')}{\partial y'} \delta y'(x) \right] dx = 0.$$

To make further progress, we write  $\delta y' = \delta(dy/dx) = (d/dx)\delta y$ , then integrate the second term by parts, remembering  $\delta y = 0$  at the endpoints, to get

$$\delta J[y] = \int_{x_1}^{x_2} \left[ \frac{\partial f(y, y')}{\partial y} - \frac{d}{dx} \left( \frac{\partial f(y, y')}{\partial y'} \right) \right] \delta y(x) dx = 0.$$

Since this is true for any infinitesimal variation, we can choose a variation which is only nonzero near one point in the interval, and deduce that

$$\frac{\partial f(y, y')}{\partial y} - \frac{d}{dx} \left( \frac{\partial f(y, y')}{\partial y'} \right) = 0.$$

This general result is called the **Euler-Lagrange equation**. It's very important—you'll be seeing it again.

## An Important First Integral of the Euler-Lagrange Equation

It turns out that, *since the function  $f$  does not contain  $x$  explicitly*, there is a simple first integral of this equation. Multiplying throughout by  $y' = dy/dx$ ,

$$\frac{\partial f(y, y')}{\partial y} \frac{dy}{dx} - \frac{d}{dx} \left( \frac{\partial f(y, y')}{\partial y'} \right) y' = 0.$$

Since  $f$  doesn't depend explicitly on  $x$ , we have

$$\frac{df}{dx} = \frac{\partial f}{\partial y} \frac{dy}{dx} + \frac{\partial f}{\partial y'} \frac{dy'}{dx}$$

and using this to replace  $\frac{\partial f(y, y')}{\partial y} \frac{dy}{dx}$  in the preceding equation gives

$$\frac{df}{dx} - \frac{\partial f}{\partial y'} \frac{dy'}{dx} - \frac{d}{dx} \left( \frac{\partial f}{\partial y'} \right) y' = 0,$$

then multiplying by  $-1$  (to match the equation as usually written) we have

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

giving a **first integral**

$$y' \frac{\partial f}{\partial y'} - f = \text{constant.}$$

For the soap film between two rings problem,

$$f(y, y') = y\sqrt{1 + y'^2},$$

so the Euler-Lagrange equation is

$$\sqrt{1 + y'^2} - \frac{d}{dx} \frac{yy'}{\sqrt{1 + y'^2}} = 0.$$

and has first integral

$$y' \frac{\partial f}{\partial y'} - f = \frac{yy'^2}{\sqrt{1 + y'^2}} - y\sqrt{1 + y'^2} = -\frac{y}{\sqrt{1 + y'^2}} = \text{constant.}$$

We'll write

$$\frac{y}{\sqrt{1 + y'^2}} = a,$$

with  $a$  the constant of integration, which will depend on the endpoints.

This is a first-order differential equation, and can be solved.

Rearranging,

$$\frac{dy}{dx} = \sqrt{\left(\frac{y}{a}\right)^2 - 1},$$

or

$$dx = \frac{ady}{\sqrt{y^2 - a^2}}.$$

The standard substitution here is  $y = a \cosh \xi$ , from which

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

Here  $b$  is the second constant of integration, the fixed endpoints determine  $a, b$ .

### The Soap Film and the Chain

We see that the soap film profile function and the hanging chain have identical analytic form. This is not too surprising, because the potential energy of the hanging chain in simplified units is just

$$\int y ds = \int y (1 + y'^2)^{\frac{1}{2}} dx,$$

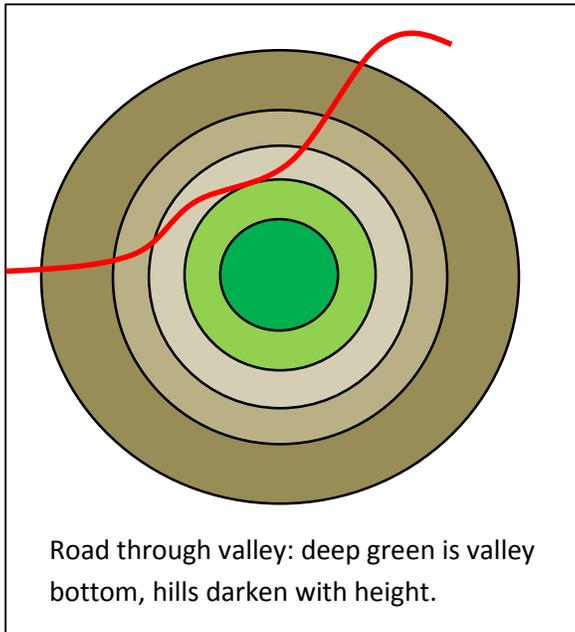
the same as the area function for the soap film. But there's an important *physical* difference: the chain has a fixed length. The soap film is free to adjust its "length" to minimize the total area. The chain isn't—it's *constrained*. How do we deal with that?

### Lagrange Multipliers

The problem of finding minima (or maxima) of a function subject to constraints was first solved by Lagrange. A simple example will suffice to show the method.

Imagine we have some smooth curve in the  $(x, y)$  plane that does not pass through the origin, and we want to find the point on the curve that is its closest approach to the origin. A standard illustration is to picture a winding road through a bowl shaped valley, and ask for the low point on the road. (We'll also assume that  $x$  determines  $y$  uniquely, the road doesn't double back, etc. If it does, the method below would give a series of *locally* closest points to the origin, we would need to go through them one by one to find the *globally* closest point.)

Let's write the curve, the road,  $g(x, y) = 0$  (the wiggly red line in the figure below).



To find the closest approach point algebraically, we need to minimize  $f(x, y) = x^2 + y^2$  (square of distance to origin) subject to the constraint  $g(x, y) = 0$ .

In the figure, we've drawn curves  $f(x, y) = x^2 + y^2 = a^2$  for a range of values of  $a$  (the circles centered at the origin).

We need to find the point of intersection of  $g(x, y) = 0$  with the smallest circle it intersects—and it's clear from the figure that it must **touch** that circle (if it crosses, it will necessarily get closer to the origin).

Therefore, at that point, the curves  $g(x, y) = 0$  and  $f(x, y) = a_{\min}^2$  are parallel.

Therefore the normals to the curves are *also* parallel:  $(\partial f / \partial x, \partial f / \partial y) = \lambda (\partial g / \partial x, \partial g / \partial y)$ .

**(Note:** yes, those *are* the directions of the normals—for an infinitesimal displacement along the curve  $f(x, y) = \text{constant}$ ,  $0 = df = (\partial f / \partial x) dx + (\partial f / \partial y) dy$ , so the vector  $(\partial f / \partial x, \partial f / \partial y)$  is perpendicular to  $(dx, dy)$ . This is also analogous to the electric field  $\vec{E} = -\vec{\nabla} \phi$  being perpendicular to the equipotential  $\phi(x, y) = \text{constant}$ .)

The constant  $\lambda$  introduced here is called a **Lagrange multiplier**. It's just the ratio of the lengths of the two normal vectors (of course, "normal" here means the vectors are perpendicular to the curves, they are *not* normalized to unit length!) We can find  $\lambda$  in terms of  $x, y$  but at this point we don't know their values.

The equations determining the closest approach to the origin can now be written:

$$\begin{aligned} \frac{\partial}{\partial x}(f - \lambda g) &= 0, \\ \frac{\partial}{\partial y}(f - \lambda g) &= 0, \\ \frac{\partial}{\partial \lambda}(f - \lambda g) &= 0. \end{aligned}$$

(The third equation is just  $g(x_{\min}, y_{\min}) = 0$ , meaning we're on the road.)

We have transformed a *constrained* minimization problem in two dimensions to an *unconstrained* minimization problem in three dimensions!

The first two equations can be solved to find  $\lambda$  and the ratio  $x/y$ , the third equation then gives  $x, y$  separately.

*Exercise for the reader:* Work through this for  $g(x, y) = x^2 - 2xy - y^2 - 1$ . (There are two solutions because the curve  $g = 0$  is a hyperbola with two branches.)

Lagrange multipliers are widely used in economics, and other useful subjects such as traffic optimization.

### Lagrange Multiplier for the Chain

The catenary is generated by minimizing the potential energy of the hanging chain given above,

$$J[y(x)] = \int y ds = \int y (1 + y'^2)^{\frac{1}{2}} dx,$$

but now subject to the constraint of fixed chain length,  $L[y(x)] = \int ds = \ell$ .

The Lagrange multiplier method generalizes in a straightforward way from variables to variable functions. In the curve example above, we minimized  $f(x, y) = x^2 + y^2$  subject to the constraint  $g(x, y) = 0$ . What we need to do now is minimize  $J[y(x)]$  subject to the constraint  $L[y(x)] - \ell = 0$ .

For the minimum curve  $y(x)$  and the correct (so far unknown) value of  $\lambda$ , an arbitrary infinitesimal variation of the curve will give zero first-order change in  $J - \lambda L$ , we write this as

$$\delta \left\{ J[y(x)] - \lambda L[y(x)] \right\} = \delta \left\{ \int_{x_1}^{x_2} (y - \lambda) ds \right\} = \delta \left\{ \int_{x_1}^{x_2} (y - \lambda) \sqrt{1 + y'^2} dx \right\} = 0.$$

Remarkably, the effect of the constraint is to give a simple adjustable parameter, the origin in the  $y$  direction, so that we can satisfy the endpoint and length requirements.

The solution to the equation follows exactly the route followed for the soap film, leading to the first integral

$$\frac{y - \lambda}{(1 + y'^2)^{\frac{1}{2}}} = a,$$

with  $a$  a constant of integration, which will depend on the endpoints.

Rearranging,

$$\frac{dy}{dx} = \sqrt{\left(\frac{y-\lambda}{a}\right)^2 - 1},$$

or

$$dx = \frac{ady}{\sqrt{(y-\lambda)^2 - a^2}}.$$

The standard substitution here is  $y - \lambda = c \cosh \xi$ , we find

$$y = \lambda + a \cosh\left(\frac{x-b}{a}\right).$$

Here  $b$  is the second constant of integration, the fixed endpoints and length give  $\lambda, a, b$ . In general, the equations must be solved numerically. To get some feel for why this will always work, note that changing  $a$  varies how rapidly the cosh curve climbs from its low point of  $(x, y) = (b, \lambda + a)$ , increasing  $a$  "fattens" the curve, then by varying  $b, \lambda$  we can move that lowest point to the lowest point of the chain (or rather of the catenary, since it might be outside the range covered by the physical chain).

Algebraically, we know the curve can be written as  $y = a \cosh(x/a)$ , although at this stage we don't know the constant  $a$  or where the origin is. What we do know is the length of the chain, and the horizontal and vertical distances  $(x_2 - x_1)$  and  $(y_2 - y_1)$  between the fixed endpoints. It's straightforward to calculate that the length of the chain is  $\ell = a \sinh(x_2/a) - a \sinh(x_1/a)$ , and the vertical distance  $v$  between the endpoints is  $v = a \cosh(x_2/a) - a \cosh(x_1/a)$  from which  $\ell^2 - v^2 = 4a^2 \sinh^2[(x_2 - x_1)/2a]$ . All terms in this equation are known except  $a$ , which can therefore be found numerically. (This is in Wikipedia, among other places.)

**Exercise:** try applying this reasoning to finding  $a$  for the soap film minimization problem. In that case, we know  $(x_1, y_1)$  and  $(x_2, y_2)$ , there is no length conservation requirement, to find  $a$  we must eliminate the unknown  $b$  from the equations  $y_1 = a \cosh((x_1 - b)/a)$ ,  $y_2 = a \cosh((x_2 - b)/a)$ . This is not difficult, but, in contrast to the chain, does *not* give  $a$  in terms of  $y_1 - y_2$ , instead,  $y_1, y_2$  appear separately. Explain, in terms of the physics of the two systems, why this is so different from the chain.

## The Brachistochrone

Suppose you have two points, A and B, B is below A, but not directly below. You have some smooth, let's say frictionless, wire, and a bead that slides on the wire. The problem is to curve the wire from A down to B in such a way that the bead makes the trip as quickly as possible.

This optimal curve is called the "brachistochrone", which is just the Greek for "shortest time".

But what, exactly, is this curve, that is, what is  $y(x)$ , in the obvious notation?

This was the challenge problem posed by Johann Bernoulli to the mathematicians of Europe in a Journal run by Leibniz in June 1696. Isaac Newton was working fulltime running the Royal Mint, recoinning England, and hanging counterfeiters. Nevertheless, ending a full day's work at 4 pm, and finding the problem delivered to him, he solved it by 4am the next morning, and sent the solution anonymously to Bernoulli. Bernoulli remarked of the anonymous solution "I recognize the lion by his clawmark".

*This was the beginning of the Calculus of Variations.*

Here's how to solve the problem: we'll take the starting point A to be the origin, and for convenience measure the  $y$ -axis positive *downwards*. This means the velocity at any point on the path is given by

$$\frac{1}{2}mv^2 = mgy, \quad v = \sqrt{2gy},$$

So measuring length along the path as  $ds$  as usual, the time is given by

$$T = \int_A^B \frac{ds}{v} = \int_A^B \frac{ds}{\sqrt{2gy}} = \int_0^x \frac{\sqrt{1+y'^2} dx}{\sqrt{2gy}}.$$

Notice that this has the same form as the catenary equation, the only difference being that  $y$  is replaced by  $1/\sqrt{2gy}$ , the integrand does not depend on  $x$ , so we have the first integral:

$$y' \frac{\partial f}{\partial y'} - f = \text{constant}, \quad f = \sqrt{\frac{1+y'^2}{2gy}}.$$

That is,

$$\frac{y'^2}{\sqrt{(1+y'^2)2gy}} - \sqrt{\frac{1+y'^2}{2gy}} = -\frac{1}{\sqrt{(1+y'^2)2gy}} = \text{constant},$$

so

$$\left(\frac{dy}{dx}\right)^2 + 1 = \frac{2a}{y}$$

$2a$  being a constant of integration (the 2 proves convenient).

Recalling that the curve starts at the origin A, it must begin by going vertically downward, since  $y = 0$ . For small enough  $y$ , we can approximate by ignoring the 1, so  $\sqrt{2ax} \cong \sqrt{y}dy$ ,  $\sqrt{2ax} \cong \frac{2}{3}y^{3/2}$ . The curve must however become horizontal if it gets as far down as  $y = 2a$ , and it cannot go below that level.

Rearranging in order to integrate,

$$dx = \frac{dy}{\sqrt{\frac{2a}{y}-1}} = \sqrt{\frac{y}{2a-y}} dy.$$

This is not a very appealing integrand. It looks a little nicer on writing  $y = a - az$ ,

$$dx = -a\sqrt{\frac{1-z}{1+z}} dz.$$

Now what? We'd prefer for the expression inside the square root to be a perfect square, of course. You may remember from high school trig that  $1 + \cos \theta = 2 \cos^2(\theta/2)$ ,  $1 - \cos \theta = 2 \sin^2(\theta/2)$ . This gives immediately that

$$\frac{1 - \cos \theta}{1 + \cos \theta} = \tan^2 \frac{\theta}{2},$$

so the substitution  $z = \cos \theta$  is what we need. Then  $dz = -\sin \theta d\theta = -2 \sin(\theta/2) \cos(\theta/2) d\theta$ ,

$$dx = -a \tan \frac{\theta}{2} dz = 2a \tan \frac{\theta}{2} \sin \frac{\theta}{2} \cos \frac{\theta}{2} d\theta = 2a \sin^2 \frac{\theta}{2} d\theta = a(1 - \cos \theta) d\theta.$$

This integrates to give

$$\begin{aligned} x &= a(\theta - \sin \theta) \\ y &= a(1 - \cos \theta), \end{aligned}$$

where we've fixed the constant of integration so that the curve goes through the origin (at  $\theta = 0$ ).

To see what this curve looks like, first ignore the  $\theta$  term in  $x$ , leaving  $x = -a \sin \theta$ ,  $y = -a \cos \theta$ . Evidently as  $\theta$  increases from zero, the point  $(x, y)$  goes anticlockwise around a circle of radius  $a$  centered at  $(0, -a)$ , that is, touching the  $x$ -axis at the origin.

Now adding the  $\theta$  back in, this circular motion move steadily to the right, in such a way that the initial direction of the path is vertically down. (For very small  $\theta$ ,  $y \sim \theta^2 \gg x \sim \theta^3$ ).

Visualizing the total motion as  $\theta$  steadily increases, The center moves from its original position at  $(0, -a)$  to the right at a speed  $a\theta$ . Meanwhile, the point is moving round the circle anticlockwise at this same speed. Putting together the center's linear velocity with the corresponding angular velocity, we see the motion  $(x(\theta), y(\theta))$  is the path of a point on the rim of a wheel rolling without sliding along a road (upside down in our case, of course). This is a *cycloid*.

### Fastest Curve for Given Horizontal Distance

Suppose we want to find the curve a bead slides down to minimize the time from the origin to some specified horizontal displacement  $X$ , but we don't care what vertical drop that entails.

Recall how we derived the equation for the curve:

At the minimum, under any infinitesimal variation  $\delta y(x)$ ,

$$\delta J[y] = \int_{x_1}^{x_2} \left[ \frac{\partial f(y, y')}{\partial y} \delta y(x) + \frac{\partial f(y, y')}{\partial y'} \delta y'(x) \right] dx = 0.$$

Writing  $\delta y' = \delta(dy/dx) = (d/dx)\delta y$ , and integrating the second term by parts,

$$\delta J[y] = \int_{x_1}^{x_2} \left[ \frac{\partial f(y, y')}{\partial y} - \frac{d}{dx} \left( \frac{\partial f(y, y')}{\partial y'} \right) \right] \delta y(x) dx + \left[ \frac{\partial f(y, y')}{\partial y'} \delta y(x) \right]_0^x = 0.$$

In the earlier treatment, both endpoints were fixed,  $\delta y(0) = \delta y(X) = 0$ , so we dropped that final term.

However, we are now trying to find the fastest time for a given horizontal distance, so the final vertical distance is an adjustable parameter:  $\delta y(X) \neq 0$ .

As before, since  $\delta J[y] = 0$  for arbitrary  $\delta y$ , we can still choose a  $\delta y(x)$  which is only nonzero near some point not at the end, so we must still have

$$\frac{\partial f(y, y')}{\partial y} - \frac{d}{dx} \left( \frac{\partial f(y, y')}{\partial y'} \right) = 0.$$

However, we must also have  $\frac{\partial f(y(X), y'(X))}{\partial y'} \delta y(X) = 0$ , to first order for arbitrary infinitesimal

$\delta y(X)$ , (imagine a variation  $\delta y$  only nonzero near the endpoint), this can only be true if  $\frac{\partial f(y, y')}{\partial y'} = 0$  at  $x = X$ .

For the brachistochrone,  $f = \sqrt{\frac{1+y'^2}{2gy}}$ ,  $\frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{2gy(1+y'^2)}}$  so  $\frac{\partial f(y, y')}{\partial y'} = 0$  at  $x = X$  means that

$f' = 0$ , the curve is horizontal at the end  $x = X$ . So the curve that delivers the bead a given horizontal distance the fastest is the half-cycloid (inverted) flat at the end. It's easy to see this fixes the curve uniquely: think of the curve as generated by a rolling wheel, one half-turn of the wheel takes the top point to the bottom in distance  $X$ .

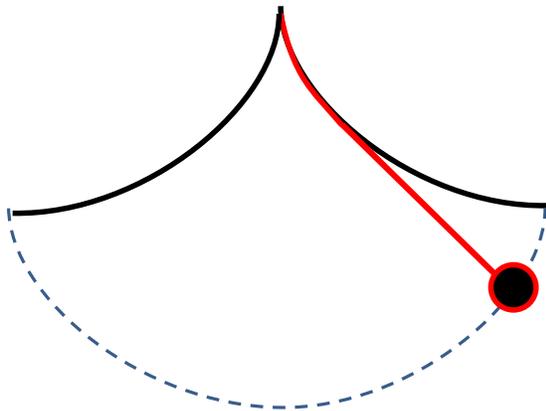
*Exercise:* how low does it go?

### The Perfect Pendulum

Around the time of Newton, the best timekeepers were pendulum clocks—but the time of oscillation of a simple pendulum depends on its amplitude, although of course the correction is small for small amplitude. The pendulum takes longer for larger amplitude. This can be corrected for by having the string constrained between enclosing surfaces to steepen the pendulum's path for larger amplitudes, and thereby speed it up.

It turns out (and was proved geometrically by Newton) that the ideal pendulum path is a cycloid. Thinking in terms of the equivalent bead on a wire problem, with a symmetric cycloid replacing the circular arc of an ordinary pendulum, if the bead is let go from rest at any point on the wire, it will reach the center in the same time as from any other point. So a clock with a pendulum constrained to such a path will keep very good time, and not be sensitive to the amplitude of swing.

The proof involves similar integrals and tricks to those used above:



$$T(y_0) = \int \frac{ds}{\sqrt{2g(y-y_0)}} = \int \frac{\sqrt{1+y'^2}}{\sqrt{2g(y-y_0)}} dx$$

And with the parameterization above,

$y' = \sin \theta / (1 - \cos \theta)$ , the integral becomes

$$\sqrt{\frac{a}{g}} \int_{\theta_0}^{\pi} \sqrt{\frac{1 - \cos \theta}{\cos \theta_0 - \cos \theta}} d\theta.$$

As before, we can now write  $1 - \cos \theta = 2 \sin^2(\theta/2)$ , etc., to find that  $T(y_0)$  is in fact independent of  $y_0$ .

This is left as an *exercise* for the reader. (*Hint*: you may find  $\int_a^b dx / \sqrt{(x-a)(b-x)} = \pi$  to be useful. Can you prove this integral is correct? Why doesn't it depend on  $a, b$ ?)

*Exercise*: As you well know, a simple harmonic oscillator, a mass on a linear spring with restoring force  $-kx$ , has a period independent of amplitude. Does this mean that a particle sliding on a cycloid is equivalent to a simple harmonic oscillator? Find out by expressing the motion as an equation  $F = ma$  where the distance variable from the origin is  $s$  measured along the curve.

Check out sliding along a cycloid [here!](#)

## Calculus of Variations with Many Variables

We've found the equations defining the curve  $y(x)$  along which the integral

$$J[y] = \int_{x_1}^{x_2} f(y, y') dx$$

has a stationary value, and we've seen how it works in some two-dimensional curve examples.

But most dynamical systems are parameterized by more than one variable, so we need to know how to go from a curve in  $(x, y)$  to one in a space  $(x, y_1, y_2, \dots, y_n)$ , and we need to minimize (say)

$$J[y_1, y_2, \dots, y_n] = \int_{x_1}^{x_2} f(y_1, y_2, \dots, y_n, y_1', y_2', \dots, y_n') dx.$$

In fact, the generalization is straightforward: the path deviation simply becomes a vector,

$$\delta \bar{y}(x) = (\delta y_1(x), \delta y_2(x), \dots, \delta y_n(x))$$

Then under any infinitesimal variation  $\delta \mathbf{y}(x)$  (writing also  $\mathbf{y} = (y_1, \dots, y_n)$ )

$$\delta J[\bar{y}] = \int_{x_1}^{x_2} \sum_{i=1}^n \left[ \frac{\partial f(\bar{y}, \bar{y}')}{\partial y_i} \delta y_i(x) + \frac{\partial f(\bar{y}, \bar{y}')}{\partial y_i'} \delta y_i'(x) \right] dx = 0.$$

Just as before, we take the variation zero at the endpoints, and integrate by parts to get now  $n$  separate equations for the stationary path:

$$\frac{\partial f(\bar{y}, \bar{y}')}{\partial y_i} - \frac{d}{dx} \left( \frac{\partial f(\bar{y}, \bar{y}')}{\partial y'_i} \right) = 0, \quad i = 1, \dots, n.$$

### Multivariable First Integral

Following and generalizing the one-variable derivation, multiplying the above equations one by one by the corresponding  $y'_i = dy_i / dx$  we have the  $n$  equations

$$\frac{\partial f(\bar{y}, \bar{y}')}{\partial y_i} \frac{dy_i}{dx} - \frac{d}{dx} \left( \frac{\partial f(\bar{y}, \bar{y}')}{\partial y'_i} \right) y'_i = 0.$$

Since  $f$  doesn't depend explicitly on  $x$ , we have

$$\frac{df}{dx} = \sum_{i=1}^n \left( \frac{\partial f}{\partial y_i} \frac{dy_i}{dx} + \frac{\partial f}{\partial y'_i} \frac{dy'_i}{dx} \right),$$

and just as for the one-variable case, these equations give

$$\frac{d}{dx} \left( \sum_{i=1}^n y'_i \frac{\partial f}{\partial y'_i} - f \right) = 0,$$

and the (important!) **first integral**

$$\sum_{i=1}^n y'_i \frac{\partial f}{\partial y'_i} - f = \text{constant.}$$