THE BRACHISTOCHRONE PROBLEM.

Imagine a metal bead with a wire threaded through a hole in it, so that the bead can slide with no friction along the wire.[1] How can one choose the shape of the wire so that the time of descent under gravity (from rest) is smallest possible? (One can also phrase this in terms of designing the least-time roller coaster track between two given points.)

A little more precisely, suppose the shape of the track is given by the graph of a function $y(x)$, with $y(0) = 0, y(L) = L$ (we orient the y axis downwards). For which function $y(x)$ is the time T of descent minimized? Galileo experimented with objects rolling down tracks of different shapes, and comments made in a book published in 1638 suggest he may have believed the optimal track would have the shape of a quarter-circle from (0,0) to (L,L), which is not the right answer. The problem of the determining the *brachistochrone* (shortest-time curve) was formally posed by Johann Bernouilli in 1696 as a challenge to the mathematicians of his day. Bernouilli himself found the solution (using a physical argument partly suggested by Fermat's 'least-time' derivation of Snell's law of refraction in geometrical optics), and Newton had a different proof that it was the correct one. This problem is widely regarded as the founding problem of the 'calculus of variations' (finding the curve, or surface, minimizing a given integral), and the solution described below is in the spirit of the approach developed by L.Euler (in 1736) and J-L. Lagrange (in 1755) to deal with general problems of this kind.

The first thing to do is to express the time of descent as an integral involving $y(x)$. Let $v = \frac{ds}{dt}$ be the speed of the bead along the track, where ds is arc length along the graph of $y(x)$. From conservation of energy, at height y below $y = 0$, we have:

$$\frac{1}{2}v^2 = gy, \text{ or } v = \sqrt{2gy}.$$

The time of descent T is therefore given by:

$$T = \int dt = \int \frac{ds}{v} = \frac{1}{\sqrt{2g}} \int \frac{\sqrt{dx^2 + dy^2}}{\sqrt{y}},$$

[1]The contents of this handout are adapted with minor changes from Chapter 6 of 'When Least is Best', by Paul J.Nahin, Princeton U.P. (2004), a beautifully written account of the history of variational problems, at a level accessible to smart undergraduates. Anyone with an interest in mathematics should go out and get this book immediately. In fact, advertising this book (and this class of problems, see also Math 534) is the main purpose of this handout.

that is, by the definite integral:

$$T = \frac{1}{\sqrt{2g}} \int_0^L \frac{\sqrt{1+(y'(x))^2}}{\sqrt{y(x)}} dx.$$

The problem then is to find, among all functions $y(x)$ satisfying the boundary conditions $y(0) = 0, y(L) = L$, the one for each T is smallest possible. This sounds vaguely like the minimization problems of calculus; the difference is that, instead of minimizing a function over an interval on the real line, we're trying to minimize the value of a *definite integral* over a *family of functions.*

The method of Euler and Lagrange applies to 'variational problems' of the following kind. Given a function of three variables $f(x, y, p)$, find the function $y(x)$ (satisfying given 'boundary conditions' $y(a) = y_a, y(b) = y_b$) for which the integral:

$$\mathcal{F}[y] = \int_a^b f(x, y(x), y'(x))dx$$

has the smallest possible value. Clearly the brachistochrone problem is of this form. Euler (and later, independently Lagrange) discovered that the function $y(x)$ achieving the minimum (if one exists) must satisfy a second-order differential equation, the *Euler-Lagrange equation:*

$$\frac{\partial F}{\partial y}_{|(x,y(x),y'(x))} = \frac{d}{dx}[\frac{\partial F}{\partial p}_{|(x,y(x),y'(x))}].$$

The right-hand side is understood as follows: take partial derivatives of F with respect to the argument p, evaluate the resulting function of (x, y, p) at the point $(x, y(x), y'(x))$ to obtain a function of x only, then take its (one-variable) derivative with respect to x.

For example, in the brachistochrone problem we have (ignoring the constant $\sqrt{2g}$):

$$F(x, y, p) = \sqrt{\frac{1+p^2}{y}}, \quad \frac{\partial F}{\partial p} = \frac{p}{\sqrt{y(1+p^2)}},$$

and upon calculation (check!):

$$\frac{d}{dx}[\frac{\partial F}{\partial p}_{|(x,y(x),y'(x))}] = \frac{1}{\sqrt{y(1+(y')^2)}}(\frac{y''}{1+(y')^2} - \frac{1}{2}\frac{(y')^2}{y}).$$

The left-hand side is:

$$\frac{\partial F}{\partial y}_{|(x,y(x),y'(x))} = -\frac{\sqrt{1+(y')^2}}{2y^{3/2}}.$$

Thus the Euler-Lagrange equation says, in this case (after simplification):

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

a non-linear second order equation. The solution of this equation will give the least-time function $y(x)$.

Before we solve this equation, consider the Euler-Lagrange equation for a simpler problem. Suppose we want to connect $(0,0)$ to (a,b) by a curve of least possible *length* (we know the answer to this, right?) That is, we want to minimize the integral:

$$L = \int_0^a \sqrt{1+(y')^2}dx,$$

among all functions $y(x)$ satisfying $y(0) = 0, y(a) = b$. We have:

$$F(x,y,p) = \sqrt{1+p^2}, \quad \frac{\partial F}{\partial p} = \frac{p}{\sqrt{1+p^2}}, \quad \frac{\partial F}{\partial y} = 0,$$

so the Euler-Lagrange equation reads:

$$\frac{d}{dx}\frac{y'(x)}{\sqrt{1+(y')^2}} = 0,$$

or:

$$\frac{y''}{(1+(y')^2)^{3/2}} = 0,$$

which is simply $y'' = 0$. So (as expected) $y(x)$ is linear, $y(x) = (b/a)x$.

The first step in the solution of the Euler-Lagrange equation for the brachistochrone problem:

$$2yy'' + 1 + (y')^2 = 0$$

is to reduce it to a first-order equation. Multiplying by y':

$$y' + 2yy'y'' + (y')^3 = 0,$$

where the left-hand side equals $[y + y(y')^2]'$. We conclude:

$$y(1 + (y')^2) = C,$$

or:

$$\frac{dy}{dx} = \sqrt{\frac{C-y}{y}}, \quad \frac{dx}{dy} = \sqrt{\frac{y}{C-y}}.$$

This is separable, but it is easier to change variables from x to φ, the angle the tangent to the curve makes with the *vertical.* We have $dx/dy = \tan\varphi$, so:

$$\frac{y}{C-y} = \frac{\sin^2\varphi}{\cos^2\varphi},$$

or:

$$y = y\cos^2\varphi + y\sin^2\varphi = C\sin^2\varphi.$$

Differentiating in φ, we find:

$$\frac{dy}{d\varphi} = 2C\sin\varphi\cos\varphi,$$

and consequently:

$$\frac{dx}{d\varphi} = \sqrt{\frac{y}{C-y}}\frac{dy}{d\varphi} = 2C\sin^2\varphi = C(1-\cos 2\varphi)$$

(the reader should verify this result.) This can be integrated to give:

$$x = C\int(1-\cos 2\varphi)d\varphi = \frac{C}{2}(2\varphi - \sin 2\varphi).$$

(Here to set the constant of integration equal to zero we used the fact that $x = 0$ corresponds to $y = 0$, and hence to $\varphi = 0$, incidentally showing that the tangent to the curve is vertical at the starting point.) To summarize the result, we have derived the equations:

$$x = \frac{C}{2}(2\varphi - \sin 2\varphi), \quad y = \frac{C}{2}(1-\cos 2\varphi),$$

which can be thought of as a set of parametric equations for the solution curve $y(x)$ (with φ as parameter, of course.)

The curve described by these parametric equations was familiar to Bernouilli, and is just as familiar to calculus students: it is the *cycloid*, an evolute of the circle. The cycloid is the path described by a fixed point on a circle of radius a, as the circle rolls on a fixed line. Taking the line to be the x-axis, and choosing as parameter t the angle (in radians, measured from the center of the circle) formed by the point with the (downwards oriented) vertical, we have the usual parametric equations:

$$x = at - a\sin t, \quad y = a - a\cos t,$$

the same as above (after an obvious identification of the parameters.) It is easy to see one can adjust the parameter C so as to achieve the other boundary condition $y(L) = L$.

Conclusion: the brachistochrone is the cycloid.

One cannot end this discussion without mentioning the following remarkable property of this 'least-time curve': if you release your metal bead (or roller coaster car, or Matchbox car, take your pick) *from rest* at *any* point on the track, the time of descent to the lowest point will be *the same*, regardless of where on the track you release it.

Does this sound impossible? That's what proofs are for. If the object starts at (x_0, y_0), the time of descent will be:

$$T = \int_{x_0}^{\pi a} \sqrt{\frac{1 + (y')^2}{2g(y - y_0)}} dx = \sqrt{\frac{a}{g}} \int_{t_0}^{\pi} \sqrt{\frac{1 - \cos t}{\cos t_0 - \cos t}} dt,$$

using the parametric equation in t.

Exercise 1. Use the half-angle formulas:

$$\sin(t/2) = (1 - \cos t)^{1/2}/\sqrt{2}, \quad \cos t = 2\cos^2(t/2) - 1$$

in the numerator and denominator (resp.), followed by the substitution $u = \cos(t/2)/\cos(t_0/2)$ to show the value of this integral is π, for any t_0. (Or look it up in Nahin's book, p.226.) Thus the time of descent (from any point) is $\pi\sqrt{a/g}$. (Note that $\pi/\sqrt{10} \sim 0.99$, so this is practically $\sqrt{a}$ seconds, if a is in meters.)

See also: *Moby Dick*, by Herman Melville, ch. 96 'The Try-Works': *'It was in the left-hand try-pot of the Pequod...that I was first indirectly struck by the remarkable fact that, in geometry, all bodies gliding along a cycloid, my soapstone for example, will descend from any point in precisely the same time.'* (This remarkable reference is pointed out by Nahin on p.228.)

Exercise 2. The catenary revisited. Consider the *constrained* variational problem: given a cable of constant linear mass density ρ and given length L

connecting two points $(x_1, y_1), (x_2, y_2)$, find the shape $y(x)$ of the cable that minimizes total potential energy:

$$U = \int_{x_1}^{x_2} \rho g y ds = \rho g \int_{x_1}^{x_2} y\sqrt{1 + (y')^2} dx.$$

The total length constraint is:

$$L = \int_{x_1}^{x_2} ds = \int_{x_1}^{x_2} \sqrt{1 + (y')^2} dx.$$

One way to deal with the constraint is to use a 'Lagrange multiplier' λ, just as in ordinary calculus. That is, consider the problem of minimizing:

$$U + \lambda L = \int_{x_1}^{x_2} (\rho g y + \lambda)\sqrt{1 + (y')^2} dx,$$

now without constraints (except for the boundary conditions $y(x_i) = y_i$). λ is a real constant, and finding its value is part of the problem. Show that the Euler-Lagrange equation, for the integrand:

$$F(x, y, p) = (\rho g y + \lambda)\sqrt{1 + p^2}$$

takes the form:

$$y'' = (\rho g/C)\sqrt{1 + (y')^2},$$

for some constant $C > 0$ (The solution is on pp. 248-9 of Nahin's book.) This is exactly the equation derived in this course by a different method [Tenenbaum-Pollard p.506]. Recall it is solved by introducing $v = y'$ (leading to a first-order equation of 'homogeneous type'), and the solution is hyperbolic cosine (the 'catenary').

Conclusion: the catenary minimizes total potential energy, among curves of fixed length joining two given points.

Non-technical derivation of the Euler-Lagrange equation.

Suppose $y(x)$ minimizes the integral:

$$\mathcal{F}[y] = \int_a^b F(x, y(x), y'(x)) dx,$$

among all functions $y(x)$ satisfying $y(a) = y_a, y(b) = y_b$. Let $w(x)$ be another function, completely arbitrary except for the requirement $w(a) = w(b) = 0$. Then for any $\epsilon \in \mathbb{R}$ the function:

$$y_\epsilon(x) := y(x) + \epsilon w(x)$$

satisfies the same boundary conditions as $y(x)$; and if $|\epsilon|$ is small enough, $F(x, y_\epsilon, y'_\epsilon(x))$ is also defined. Now consider the function of one variable:

$$g(\epsilon) = \mathcal{F}[y_\epsilon] = \int_a^b F(x, y_\epsilon(x), y'_\epsilon(x))dx.$$

The key observation is that since $y(x)$ minimizes $\mathcal{F}$ and $y_\epsilon(x)$ is a competeting function, necessarily $\epsilon = 0$ is a minimum point for $g(\epsilon)$. By one-variable calculus, this implies $g'(0) = 0$. So we just have to compute this derivative:

$$\frac{dg}{d\epsilon}_{|\epsilon=0} = \int_a^b [\frac{\partial F}{\partial y}w(x) + \frac{\partial F}{\partial p}w'(x)]dx,$$

where the partial derivatives of F are evaluated at the point $(x, y(x), y'(x))$. Now we integrate by parts the term involving $w'(x)$. Noting that there is no 'endpoint term' (since $w(a) = w(b) = 0$), we find:

$$0 = \int_a^b w(x)[\frac{\partial F}{\partial y} - \frac{d}{dx}(\frac{\partial F}{\partial p}_{|(x,y(x),y'(x))})]dx.$$

Now use the following fact: if

$$\int_a^b w(x)u(x)dx = 0$$

for any function $w(x)$ vanishing at a and b, it follows necessarily that $u \equiv 0$ in $[a, b]$ (easy calculus-well, maybe 'analysis'- exercise). We conclude:

$$\frac{\partial F}{\partial y} - \frac{d}{dx}(\frac{\partial F}{\partial p}_{|(x,y(x),y'(x))}) = 0,$$

the Euler-Lagrange equation.

Remark 1: I haven't stated any technical conditions on the functions involved, so this is really a 'physicists's proof'. Would have been fine for Euler, but probably not Weierstrass.

Remark 2 (for the mathematically-minded) A serious weakness in the foregoing discussion is the assumption that, just because the integrals are (defined and) bounded below in a given class of functions, the minimum exists in that class. For example, consider the variational problem [Morrey, p.5]:

$$\text{minimize } \mathcal{F}[y] = \int_0^1 [1 + (y'(x))^2]^{1/4}dx \text{ over } y \in C^1[0, 1], y(0) = 0, y(1) = 1.$$

Clearly $\mathcal{F}[y] > 1$ for any such y, but for any $r < 1$ consider:

$$y_r(x) = \begin{cases} 0, \quad 0 \le x \le r, \\ -1 + [1 + 3(x-r)^2/(1-r)^2]^{1/2}, \quad r \le x \le 1 \end{cases}$$

Each $y_r(x)$ is in the competing class of functions, and $\mathcal{F}[y_r] \to 1$ as $r \to 1^-$ (*exercise*). Thus $\mathcal{F}$ can be made arbitrarily close to 1 in this class of functions, but the *infimum* (namely, 1) of $\mathcal{F}$ is not attained in the class, since clearly $\mathcal{F}[y] = 1$ forces y to be constant, which is not compatible with the boundary conditions.

Note that the functional $\mathcal{F}$ differs from arc length by a mere innocent-looking power.