Variational Calculus
MANUSCRIPT

William F. Barnes
[1]

January 6, 2024

Contents

Chapter 1

Variational Calculus

1 Introduction

Any well-rounded student of natural philosophy or STEM field has at least one thorough encounter with the tenets of classical mechanics. A relatable picture of reality is constructed from Newton's laws of motion, along with conservation laws handling energy, momentum, and angular momentum. Toss in a few revelations from electromagnetism and thermodynamics and pre-1905 physics is complete, right?

Not quite. What if you learned that there is another tenant in the building living among the classical laws? That there is another principle of mechanics at play that not only contains, but runs a bit deeper than Newton's seventeenth-century picture? This is in fact the case, and it is called the *principle of least action.*

The tool kit used to grapple with 'least action' is called the *calculus of variations.* Variational calculus, as it's also called, is all about solving for *critical curves* rather than critical points as done in ordinary calculus.

As it pertains to classical mechanics, the principle of least action tells something very profound: *The true path on which a body moves is the one that minimizes the difference between kinetic and potential energy along that path.* Using symbols T for kinetic energy and U for potential, the above means

$$S = \int \left(T\left(v\left(t\right)\right) - U\left(x\left(t\right)\right)\right) dt$$

is always minimized if $x(t)$ and $v(t)$ represent the correct path of position and velocity. The quantity S is called the *action*. The action evaluates to a larger number if the wrong $x(t)$ or $v(t)$ are fed into the integral.

This is undoubtedly a strange way to think about mechanics, namely because there is no need for forces, momentum, or any vectors at all. As a matter of strictness, the idea of 'least' action is sometimes a misnomer, as sometimes the case of 'most' action is more applicable. In either case, we are safe saying 'principle of stationary action'.

2 Euler-Lagrange Equation

To begin we will work strictly on the xy plane without mentioning physics until the first solid result is gained. Consider two fixed points in the Cartesian plane

$$(x_0, y_0), (x_1, y_1)$$

that represent the initial and final position of any well-behaved curve $y(x)$. The curve $y(x)$ shall be considered the 'true' path that connects the initial and final fixed points:

$$y(x) = \text{true path from } x_0 \text{ to } x_1$$

2.1 Varied Path

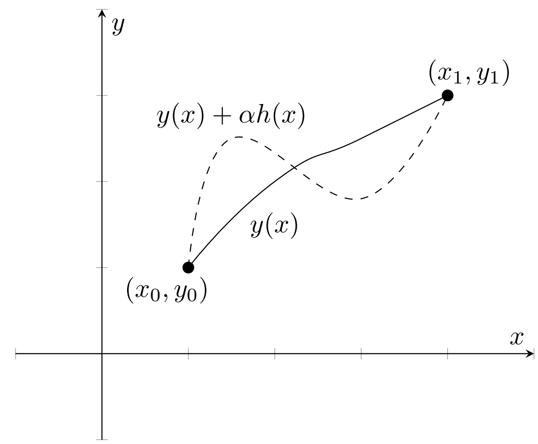
To accompany $y(x)$ in the plane, introduce a second curve that starts and finishes at the same endpoints, but is allowed to meander in the plane due to an additional term $\alpha h(x)$ such that

$$f(x) = y(x) + \alpha h(x) \ .$$

The variable α is a dimensionless *variational parameter*. The curve $h(x)$ is an 'incorrectness' that gives structure to the varied path, and this is decidedly zero at each endpoint:

$$h(x_0) = h(x_1) = 0$$

Summarizing this construction, we sketch the following figure in the Cartesian plane:



2.2 Writing the Action

Now, conceive of a new function Λ that depends of x, $f(x)$, and also the derivative $f'(x) = df(x)/dx$. Such a function $\Lambda(x, f(x), f'(x))$ can be arbitrary in most respects, whether it represent a physical quantity or a conceptual one is a mere formality as of now.

To construct an action from this, integrate the function Λ along the path from x_0 to x_1:

$$S = \int_{x_0}^{x_1} \Lambda(x, f(x), f'(x))\, dx$$

Keeping the notation from getting out of control, we hide the explicit mention of x from the f-related terms:

$$S = \int_{x_0}^{x_1} \Lambda(x, f, f')\, dx$$

2.3 Derivative of the Action

Buried in f and f' is the dependence on the variational parameter α, where we explicitly have

$$\begin{aligned} f &= y + \alpha h \\ f' &= y_x + \alpha h' \ , \end{aligned}$$

which means the action S is also a function of the parameter α. In traditional calculus-101 fashion, we then ask what happens by taking derivative of S with respect to α? Setting this up, the problem means to calculate

$$\frac{dS}{d\alpha} = \int_{x_0}^{x_1} \frac{d}{d\alpha}\left(\Lambda(x, f, f')\right) dx \ .$$

The parameter α does not depend on x itself (that's what h is for), which is why the α-derivative penetrates the integral without fuss.

To grapple with what $dA/d\alpha$ means, reach for partial derivatives to write

$$\frac{dA}{d\alpha} = \frac{\partial \Lambda}{\partial x}\cancel{\frac{dx}{d\alpha}} + \frac{\partial \Lambda}{\partial f}\frac{df}{d\alpha} + \frac{\partial \Lambda}{\partial f'}\frac{df'}{d\alpha} ,$$

where the first term is identically zero, and the derivative terms $df/d\alpha$ and $df'/d\alpha$ are nothing but h and h', respectively. So far then, we have:

$$\frac{dS}{d\alpha} = \int_{x_0}^{x_1} \left(\frac{\partial \Lambda}{\partial f} h + \frac{\partial \Lambda}{\partial f'} h' \right) dx \; .$$

2.4 Integrating by Parts

To proceed, we'll focus on the primed term in the above, and choose the following substitutions for integration by parts

$$\begin{aligned} u &= \partial \Lambda / \partial f' \\ dv &= h' \, dx \; , \end{aligned}$$

with

$$\begin{aligned} du &= \frac{d}{dx}\left(\frac{\partial \Lambda}{\partial f'} \right) dx \\ v &= h \; . \end{aligned}$$

Then, by the standard form

$$\int u \, dv = uv \Big| - \int v \, du \; ,$$

the derivative of the action looks like

$$\frac{dS}{d\alpha} = \int_{x_0}^{x_1} \left(\frac{\partial \Lambda}{\partial f} h - h \frac{d}{dx}\left(\frac{\partial \Lambda}{\partial f'} \right) \right) dx + \cancel{\frac{\partial \Lambda}{\partial f'} h \Big|_{x_0}^{x_1}} ,$$

where the boundary term wholly vanishes because $h(x)$ is zero at each endpoint.

2.5 Minimizing the Action

Summarizing our progress, the derivative of the action has taken the form

$$\frac{dS}{d\alpha} = \int_{x_0}^{x_1} h(x) \left(\frac{\partial \Lambda}{\partial f} - \frac{d}{dx}\left(\frac{\partial \Lambda}{\partial f'} \right) \right) dx \; .$$

Now comes the crucial observation regarding the variational parameter α. The derivative $dS/d\alpha$ goes to zero as α itself goes to zero:

$$\begin{aligned} dS/d\alpha &\rightarrow 0 \\ \alpha &\rightarrow 0 \end{aligned}$$

By making such a change, the varied path flattens down to the true path, which means $f(x)$ flattens down to $y(x)$:

$$f(x) \rightarrow y(x)$$

2.6 Euler-Lagrange Equation

In the zero-variation limit, the above becomes

$$0 = \int_{x_0}^{x_1} h(x) \underbrace{\left(\frac{\partial \Lambda}{\partial y} - \frac{d}{dx}\left(\frac{\partial \Lambda}{\partial y_x} \right) \right)}_{=0} dx \; ,$$

and the integral must clearly evaluate to zero. Since $h(x)$ is generally nonzero except for the endpoints, the parenthesized quantity must therefore be zero along the whole path. Plucking out this item from the integral, we arrive at the famed *Euler-Lagrange* equation:

$$0 = \frac{\partial \Lambda}{\partial y} - \frac{d}{dx}\left(\frac{\partial \Lambda}{\partial y_x} \right) \tag{1.1}$$

The Euler-Lagrange equation doesn't look like much at fist, perhaps a nice accident of the chain rule. It is interesting to fathom, though, that the above holds for whatever arbitrary function $\Lambda(x, y(x), y_x(x))$ is chosen.

2.7 Change of Domain

The whole derivation of Equation (1.1) can be repeated in the time domain, in which case t takes the place of x. Following this through, let us substitute

$$\begin{aligned} x &\rightarrow t \\ y &\rightarrow x(t) \\ y_x &\rightarrow v(t) \; , \end{aligned}$$

where $v(t)$ is the velocity $x'(t)$. While we're at it, relabel the arbitrary function Λ with the the letter L:

$$\Lambda \rightarrow L(t, x(t), v(t))$$

With all this, the Euler-Lagrange equation takes a form in the time domain:

$$0 = \frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial v} \right) \tag{1.2}$$

2.8 The Lagrangian

By choosing the proper function the proper function for L, something curious happens with the Euler-Lagrange equation (1.2). After some fiddling, one readily stumbles on the combination

$$L = T(v) - U(x) \; , \tag{1.3}$$

called the *Lagrangian*. The functions $T(v)$, $U(x)$ are the respective kinetic and potential energies of the body being considered.

By inserting Equation (1.3) into Equation (1.2), we write

$$0 = \frac{\partial \left(T - U\right)}{\partial x} - \frac{d}{dt}\left(\frac{\partial \left(T - U\right)}{\partial v}\right) ,$$

simplifying to

$$0 = -\frac{\partial U}{\partial x} - \frac{d}{dt}\left(\frac{\partial T}{\partial v}\right) .$$

From here, a deeply rich formulation of classical physics called *Lagrangian mechanics* can be developed. This topic won't be formally indulged here, but its foundations are explored here nonetheless.

2.9 Recovering Newton's Law

For the Newtonian regime, we have that the kinetic energy $T\left(v\right)$ is given by

$$T\left(v\right) = \frac{1}{2}mv^2 ,$$

where m is the mass of the body or particle. Then, the Euler-Lagrange equation with the choice $L = T - U$ becomes

$$0 = -\frac{\partial U}{\partial x} - \frac{d}{dt}\left(\frac{\partial}{\partial v}\frac{1}{2}mv^2\right) ,$$

readily simplifying to Newton's second law:

$$m\frac{d}{dt}v\left(t\right) = -\frac{\partial}{\partial x}U\left(x\right)$$

Astonishingly, the principle of least action seemingly *does* contain the classic laws of motion, and the true path followed by an object really is the one that minimizes the integral of $T - U$ along the path.

3 Formalism

While the above derivation of the Euler-Lagrange equation took place in two dimensions, the same reasoning apples when there are more variables at play. To this end, it helps to have an efficient symbol for certain derivative terms. If we have a function f, and we need the partial derivative with respect to x, we should be able to write

$$f_x = \partial f / \partial x$$
$$f_{xx} = \partial^2 f / \partial x^2$$

without ambiguity. To have the operator by itself with no mention of which function it's acting upon, we use the notation

$$\partial_x = \frac{\partial}{\partial x} = \partial / \partial x .$$

Of course, the same notation is useful for full derivatives, for instance

$$d_t f = \frac{d}{dt} f .$$

With such shortcuts, the Euler-Lagrange equation can be written in the most minimal way:

$$0 = L_x - d_t L_v$$

3.1 Formal Derivation

With the main ideas established, let us re-state the action calculation

$$S = \int_{x_0}^{x_1} \Lambda\left(x, f, f'\right) dx$$

in slightly different terms. The function $f\left(x\right)$ is still the varied path, except variations are represented by

$$f \rightarrow f + \Delta f$$
$$f_x \rightarrow f_x + \left(\Delta f\right)' ,$$

where Δf replaces the $\alpha h\left(x\right)$ construction, and Δf is zero at each endpoint.

The action S is no ordinary function, but is formally called a 'functional' depending on $f\left(x\right)$, denoted $S\left[f\right]$. So far then, we write

$$S\left[f\right] = \int_{x_0}^{x_1} \Lambda\left(x, f, x_x\right) dx .$$

Then a variation in S is written

$$\Delta S = \int_{x_0}^{x_1} \Lambda\left(x, f + \Delta f, f_x + \left(\Delta f\right)'\right) dx - S\left[f\right] .$$

By Taylor-expanding the inner quantity to first order, we further have

$$\Lambda\left(x, f + \Delta f, f_x + \left(\Delta f\right)'\right) = \Lambda\left(x, f, f_x\right) + \frac{\partial \Lambda}{\partial f}\Delta f + \frac{\partial \Lambda}{\partial f_x}\left(\Delta f\right)' ,$$

and ΔS simplifies to

$$\Delta S = \int_{x_0}^{x_1} \left(\frac{\partial \Lambda}{\partial f}\Delta f + \frac{\partial \Lambda}{\partial f_x}\left(\Delta f\right)'\right) dx .$$

From here, the derivation looks much like the 'informal' one we started with, and the steps to finish are the same. Letting ΔS and Δf go to zero simultaneously after integrating by parts, one finds the now-familiar Euler-Lagrange equation

$$0 = \frac{\partial \Lambda}{\partial f} - \frac{d}{dx}\left(\frac{\partial \Lambda}{\partial f_x}\right) .$$

3.2 Special Form

Starting with $\Lambda(x, f, f')$, let us calculate the total derivative in x, which is

$$\frac{d\Lambda}{dx} = \frac{\partial\Lambda}{\partial x} + \frac{\partial\Lambda}{\partial f}\frac{df}{dx} + \frac{\partial\Lambda}{\partial f_x}\frac{df_x}{dx} ,$$

and then replace $\partial\Lambda/\partial f$ using the Euler-Lagrange equation

$$\frac{d\Lambda}{dx} = \frac{\partial\Lambda}{\partial x} + \frac{d}{dx}\left(\frac{\partial\Lambda}{\partial f_x}\right) f_x + \frac{\partial\Lambda}{\partial f_x} f_{xx} .$$

Note that there is an equivalence between

$$\frac{df}{dx} \leftrightarrow \frac{\partial f}{\partial x}$$

because f is a function of only x. (This is certainly not true for Λ or any multivariate function.)

Off to the side, calculate the total derivative of $f_x \partial\Lambda/\partial f_x$, which looks like

$$\frac{d}{dx}\left(f_x \frac{\partial\Lambda}{\partial f_x}\right) = f_{xx}\frac{\partial\Lambda}{\partial f_x} + f_x \frac{d}{dx}\left(\frac{\partial\Lambda}{\partial f_x}\right) ,$$

containing the same f_{xx}-term as the previous result. Eliminating this between the two, we have, after simplifying:

$$\frac{dA}{dx} = \frac{\partial\Lambda}{\partial x} + \frac{d}{dx}\left(f_x \frac{\partial\Lambda}{\partial f_x}\right)$$

Putting everything on one side gives special form of the Euler-Lagrange equation:

$$0 = \frac{\partial\Lambda}{\partial x} - \frac{d}{dx}\left(\Lambda - f_x \frac{\partial\Lambda}{\partial f_x}\right) \tag{1.4}$$

The above is especially informative if the function Λ has no explicit dependence on x, for if this is the case then we can only have

$$\Lambda - f_x \frac{\partial\Lambda}{\partial f_x} = \text{constant} ,$$

where f_x is not zero. This result is sometimes called the Beltrami identity.

3.3 Constant of Motion

It's impossible to resist jumping back into time domain and make another connection to Newtonian mechanics. Momentarily make the same swap of variables

$$\begin{aligned} &x \to t \\ &f \to x(t) \\ &L \to T(v) - U(x) \ , \end{aligned}$$

and so on.

Since the Lagrangian has no explicit time dependence, may immediately use the Beltrami identity to write

$$L - v\frac{\partial L}{\partial v} = \text{constant} .$$

Knowing that the kinetic energy $T(v)$ takes the form $mv^2/2$, we find (as before) that

$$\frac{\partial L}{\partial v} = \frac{\partial T}{\partial v} = \frac{\partial}{\partial v}\left(\frac{1}{2}mv^2\right) = mv ,$$

meaning

$$v\frac{\partial L}{\partial v} = mv^2 = 2T(v) \ .$$

Putting it all together, we find

$$L - 2T = T - U - 2T = \text{constant} ,$$

or in other words,

$$T + U = -\text{constant} ,$$

therefore conservation of energy also emerges from the principle of least action.

3.4 More Dimensions

The formal derivation of the Euler-Lagrange equation that generalizes to N simultaneous functions

$$f \to \left\{ f^{(1)}, f^{(2)}, \ldots, f^{(N)} \right\} ,$$

each with a different variation

$$f^{(j)} \to f^{(j)} + \Delta f^{(j)} ,$$

with vanishing variation at the endpoints

$$\Delta f^{(j)}(x_0) = \Delta f^{(j)}(x_1) = 0$$

is straightforwardly written.

The function Λ becomes

$$\Lambda\left(x, f^{(1)}, f^{(2)}, \ldots, f^{(N)}, f_x^{(1)}, f_x^{(2)}, \ldots, f_x^{(N)}\right) ,$$

and gives rise to m Euler-Lagrange equations that all look the same up to index number:

$$0 = \frac{\partial\Lambda}{\partial f^{(j)}} - \frac{d}{dx}\left(\frac{\partial\Lambda}{\partial f_x^{(j)}}\right) \tag{1.5}$$

3.5 Conservation of Energy

When it comes to having more than one dimension, one wonders what the generalization of Equation (1.4) may be, and whether there is one new constant per added dimension. This is in fact *not* the case, but the real answer is more beautiful anyway. Anticipating the outcome of this analysis, consider a function E (for 'energy') that has all the same dependencies as Λ:

$$E = E\left(x, f^{(1)}, f^{(2)}, \ldots, f^{(N)}, f_x^{(1)}, f_x^{(2)}, \ldots, f_x^{(N)}\right)$$

Then, taking inspiration from the constant of motion found in the one-dimensional case, consider the relationship

$$E = \sum_{j=1}^{N} f_x^{(j)} \frac{\partial \Lambda}{\partial f_x^{(j)}} - \Lambda \, .$$

If E is to be constant, then the total derivative of E had better resolve to zero. Pursuing this, we write:

$$\frac{dE}{dx} = \sum_{j=1}^{N} f_{xx}^{(j)} \frac{\partial \Lambda}{\partial f_x^{(j)}} + \sum_{j=1}^{N} f_x^{(j)} \frac{d}{dx}\left(\frac{\partial \Lambda}{\partial f_x^{(j)}}\right) - \frac{d\Lambda}{dx}$$

From the chain rule we also have

$$\frac{d\Lambda}{dx} = \frac{\partial \Lambda}{\partial x} + \sum_{j=1}^{N} \frac{\partial \Lambda}{\partial f^{(j)}} f_x^{(j)} + \sum_{j=1}^{N} \frac{\partial \Lambda}{\partial f_x^{(j)}} f_{xx}^{(j)} \, ,$$

and note that the f_{xx}-sums occur in both equations, and so does $d\Lambda/dx$. Eliminating the common terms in each, we get the result

$$\frac{dE}{dx} = -\frac{\partial \Lambda}{\partial x} + \sum_{j=1}^{N} f_x^{(j)} \left(\frac{d}{dx}\left(\frac{\partial \Lambda}{\partial f_x^{(j)}}\right) - \frac{\partial \Lambda}{\partial f^{(j)}}\right) \, .$$

In the above, the parenthesized portion is none other than the Euler-Lagrange equation, and is identically zero. Evidently, the whole concern of energy conservation reduces to the statement

$$\frac{dE}{dx} = -\frac{\partial \Lambda}{\partial x} \, .$$

Note that we've worked with the unspecified function Λ depending fundamentally on x. To turn the above into a statement about physics, make the following replacement:

$$\begin{aligned} &x \to t \\ &\Lambda \to L = T(v) - U(x) \end{aligned}$$

3.6 Generalized Coordinates

As it pertains to physical systems, The Euler-Lagrange equation (1.5) occurs in the form

$$0 = \frac{\partial L}{\partial q^{(j)}} - \frac{d}{dt}\left(\frac{\partial L}{\partial q_t^{(j)}}\right) \, , \tag{1.6}$$

where the terms $q^{(j)}$ are *generalized coordinates*. The physical units of a given $q^{(j)}$ need not be spatial. The readiest example of this would be polar coordinates, in where the components of q are represented by

$$\vec{q} = \langle r, \theta \rangle \, .$$

Proof

To establish this firmly, let us pull no punches and do a proper proof. Supposing a system evolving by variable t depends on coordinates $x^{(j)}(t)$ and their derivatives $v^{(j)}(t)$, the Euler-Lagrange equation takes the form

$$0 = \frac{\partial \Lambda}{\partial x^{(j)}} - \frac{d}{dt}\left(\frac{\partial \Lambda}{\partial v^{(j)}}\right) \, .$$

Next, suppose the list of coordinates $\{x(t)\}$ can be defined in terms of another list $\{q(t)\}$ such that

$$x^{(j)} = x^{(j)}\left(t, q^{(1)}, q^{(2)}, \ldots, q^{(N)}\right)$$

if there are N members in $\{q(t)\}$. To first order, the two sets of coordinates further relate by

$$v^{(j)} = \frac{\partial x^{(j)}}{\partial t} + \sum_{k=1}^{N} \frac{\partial x^{(j)}}{\partial q^{(k)}} q_t^{(k)} \, .$$

By the chain rule, we further find

$$\frac{\partial x^{(j)}}{\partial q^{(k)}} = \frac{\partial x^{(j)}}{\partial q^{(k)}} \frac{dt}{dt} = \frac{\partial v^{(j)}}{\partial q_t^{(k)}} \, .$$

With the second set of coordinates, the proposed Euler-Lagrange equation reads

$$0 \stackrel{?}{=} \frac{\partial \Lambda}{\partial q^{(j)}} - \frac{d}{dt}\left(\frac{\partial \Lambda}{\partial q_t^{(j)}}\right) \, ,$$

which motivates picking on the inner term:

$$\frac{\partial \Lambda}{\partial q_t^{(j)}} = \sum_{k=1}^{N} \frac{\partial \Lambda}{\partial v^{(k)}} \frac{\partial v^{(k)}}{\partial q_t^{(j)}} = \sum_{k=1}^{N} \frac{\partial \Lambda}{\partial v^{(k)}} \frac{\partial x^{(k)}}{\partial q^{(j)}}$$

Take the time derivative of both sides to get

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial \Lambda}{\partial q_t^{(j)}}\right) = &\sum_{k=1}^{N} \frac{d}{dt}\left(\frac{\partial \Lambda}{\partial v^{(k)}}\right) \frac{\partial x^{(k)}}{\partial q^{(j)}} \\ &+ \sum_{k=1}^{N} \frac{\partial \Lambda}{\partial v^{(k)}} \frac{d}{dt}\left(\frac{\partial x^{(k)}}{\partial q^{(j)}}\right) \, , \end{aligned}$$

and simplify carefully to finish the proof:

$$\frac{d}{dt}\left(\frac{\partial\Lambda}{\partial q_t^{(j)}}\right) = \sum_{k=1}^{N}\left(\frac{\partial\Lambda}{\partial x^{(k)}}\frac{\partial x^{(k)}}{\partial q^{(j)}} + \frac{\partial\Lambda}{\partial v^{(k)}}\frac{\partial v^{(k)}}{\partial q^{(j)}}\right) = \frac{\partial\Lambda}{\partial q^{(j)}}$$

4 Motion on a Curve

In uniform gravity, consider a frictionless particle of mass m that sits on the curve

$$y\left(x\right) = \frac{k}{\alpha}\left(x-a\right)^{\alpha}$$

without departure. If the velocity of the particle is $\vec{v} = \langle \dot{x}, \dot{y}\rangle$, find the equations of motion of this system.

With the information provided, there is enough to write the kinetic and potential of this system all in terms of x-variables

$$T = \frac{1}{2}m\dot{x}^2\left(1 + k^2\left(x-a\right)^{2\alpha-2}\right)$$
$$U = \frac{mgk}{\alpha}\left(x-a\right)^{\alpha} \ ,$$

and the Lagrangian, of course, is $L = T - U$.

Applying the Euler-Lagrange equation, we must pursue

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = 0 \ .$$

Doing so and simplifying, arrive at a juicy differential equation:

$$0 = \ddot{x}\left(1 + k^2\left(x-a\right)^{2\alpha-3}\right) + \dot{x}\frac{k^2}{2}\left(2\alpha-2\right)\left(x-a\right)^{2\alpha-3} + gk\left(x-a\right)^{\alpha-1}$$

4.1 Particle in a Well

For the cases with $k > 0$ and $\alpha > 1$ is even, the particle is stuck in a 'well' centered at $x = a$. The simplest of these has $\alpha = 2$, where the above reduces to the differential equation

$$0 = \ddot{w}\left(1 + w^2\right) + \dot{w}^2 w + gkw \ ,$$

where $w = kz$ and $z = x - a$.

Despite the $\alpha = 2$ simplification, the above is still difficult to treat in the general case. We can do a quick reality check for small motions, which has $1 \gg w^2$ and $gk \gg \dot{w}^2$. For this we recover the setup for the simple harmonic oscillator, as expected:

$$\ddot{w} = -gkw$$

4.2 Particle on an Incline

An interesting modification to this setup has $\alpha = 1$, with $y\left(x\right)$ representing a straight line. Restarting the analysis from here and applying the Euler-Lagrange equation leads to

$$\ddot{x} = \frac{-gk}{1+k^2} \ ,$$

and similarly,

$$\ddot{y} = k\ddot{x} \ .$$

The total acceleration is the sum of square of each:

$$|a| = \sqrt{\ddot{x}^2 + \ddot{y}^2} = \frac{gk}{\sqrt{1+k^2}}$$

Letting θ denote the angle of incline, we further have

$$k = \tan\left(\theta\right)$$

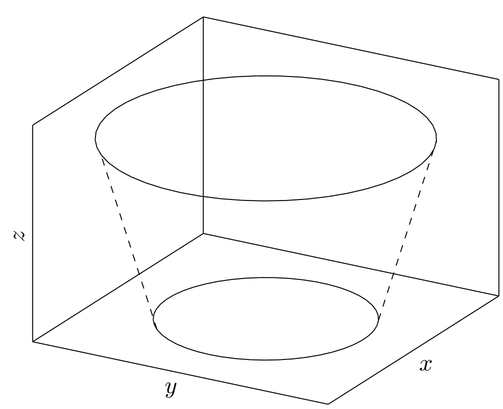
and

$$|a| = g\sin\left(\theta\right) \ .$$

5 Minimal Surface

5.1 Soap Film Problem

A stretchable membrane, such as a thin soap film, is trapped two hoops to form an axially-symmetric surface. One hoop is a circle laying on the xy-plane at $z = z_0$. the other hoop is parallel to the first and suspended at $z = z_1$. Ignoring gravity, determine the shape of the membrane that minimizes surface area. In other words, determine the proper shape of the dotted line profile indicated in the figure below.



The minimal surface problem is characterized by the surface area (conveniently named S) of the whole film. Using elementary methods, or by exploiting the axial symmetry of the system, the surface area reads

$$S = \int_{z_0}^{z_1} 2\pi r(z) \sqrt{1 + r_z^2} \, dz ,$$

where r_z is the derivative of r with respect to z.

Without needing to finish the integral, we pick out the working quantity to be

$$\Lambda = r\sqrt{1 + r_z^2} ,$$

which is a function of r and r_z, but not z itself. This warrants use of the special form of the Euler-Lagrange equation (1.4), which means, for this problem,

$$\Lambda - r_z \frac{\partial \Lambda}{\partial r_z} = C_0 ,$$

where C_0 is a constant.

Plugging Λ into the above and turning the crank leads to a separable differential equation

$$dz = \frac{dr}{\sqrt{r^2/C_0^2 - 1}} .$$

By the substitution

$$r = C_0 \cosh(\beta)$$

with

$$dr = C_0 \sinh(\beta) \, d\beta ,$$

the differential equation simplifies to

$$dz = C_0 \, d_\beta ,$$

revealing the simple relationship

$$z = C_0 \beta + C_1 ,$$

where C_1 is an integration constant. After eliminating β, suddenly we have an equation for $r(z)$:

$$r(z) = C_0 \cosh\left(\frac{z - C_1}{C_0}\right)$$

The constants C_0, C_1 are specified by the hoop radii, namely $r(z_0) = R_0$ and $r(z_1) = R_1$.

Fleshing out an easy example, the symmetric case with $R_1 = R_2$ and

$$-z_0 = z_1 = L$$

leads to

$$C_1 = \frac{z_0 + z_1}{2} = \frac{-L + L}{2} = 0 .$$

The $r(z)$ equation then says

$$R = C_0 \cosh\left(\frac{L}{C_0}\right) ,$$

containing one unknown.

5.2 Straight Line

A much easier minimal surface problem is to prove that the shortest path connecting two points is a straight line. Setting up the classic 'arc length' integral looks like

$$S = \int_{x_0}^{x_1} \sqrt{1 + y_x^2} \, dx ,$$

from which we pick out

$$\Lambda = \sqrt{1 + y_x^2} .$$

Unsurprisingly there is no explicit x-dependence in Λ, warranting the identity

$$\Lambda - y_x \frac{\partial \Lambda}{\partial y_x} = C ,$$

with C constant. Simplifying, we find

$$\frac{1}{\sqrt{1 + y_x^2}} = C ,$$

which can only mean y_x is a constant, or $y(x)$ is a straight line.

6 Motion on a Cycloid

6.1 Brachistochrone

The Ancient Greeks were interested in a curious problem that has widespread practical application. In uniform gravity, suppose a body at an initial height needs to slide down some kind of plank, ramp, or other curve so as to reach a lower height in the *shortest* time possible, ignoring friction. The curve that solves this problem is called the *brachistochrone*.

Starting from first principles, the time T taken to slide down such a curve is given by

$$T = \int_{y(t_0)}^{y(t_1)} dt ,$$

where T is the quantity to minimize. The differential arc length ds relates to dt by

$$ds = v(t) \, dt ,$$

where $v(t)$ is the velocity of the body at time t.

From geometry we have that

$$ds^2 = dx^2 + dy^2 \ ,$$

and meanwhile from energy conservation we know

$$v\left(x\right) = \sqrt{2g\left(y_0 - y\left(x\right)\right)} \ .$$

Updating the T-integral with this information gives

$$T = \int_{x_0}^{x_1} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \frac{dx}{\sqrt{2g\left(y_0 - y\left(x\right)\right)}} \ .$$

Now, we're trying to minimize T as a way of solving for y, which seems completely backwards until utilizing the calculus of variations. The above can be regarded as a functional

$$T = \int_{y_0}^{y_1} \Lambda\left(x, y, y_x\right) dx \ ,$$

with no explicit x-dependence. Reasoning from the special form of the Euler-Lagrange equation (1.4), we have

$$\Lambda - y_x \frac{\partial \Lambda}{\partial y_x} = C \ ,$$

where C is constant. Calculating this out, we find

$$\frac{1}{\sqrt{2g\left(y_0 - y\left(x\right)\right)}} \frac{1}{\sqrt{1 + \left(dy/dx\right)^2}} = C \ ,$$

which can be turned into an integral for $x\left(y\right)$:

$$\int dx = \int \sqrt{\frac{2gC^2\left(y_0 - y\right)}{1 - 2gC^2\left(y_0 - y\right)}} \, dy$$

Now introduce a peculiar trigonometric substitution in the variable θ such that

$$2gC^2\left(y_0 - y\right) = \frac{1 - \cos\left(\theta\right)}{2}$$

and

$$dy = -\frac{\sin\left(\theta\right)}{4gC^2} d\theta \ .$$

Running this substitution through the above integral, we have

$$\int dx = \frac{-1}{4gC^2} \int \left(1 - \cos\left(\theta\right)\right) d\theta \ ,$$

and evidently the combination $1/4gC^2$ has units of space, and this will be renamed to R, as in 'radius'.

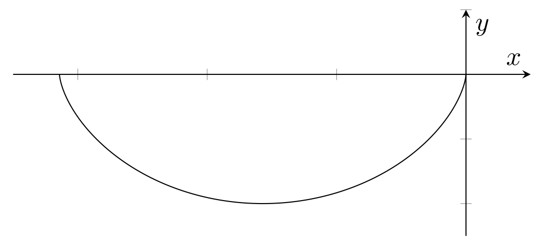
Finishing up the calculation for x, we finally have

$$x = x_0 - R\left(\theta - \sin\left(\theta\right)\right) \ , \qquad (1.7)$$

and solving similarly for y,

$$y = y_0 - R\left(1 - \cos\left(\theta\right)\right) \ . \qquad (1.8)$$

This pair of parametric equations is a surprising result, namely because a unified equation $y\left(x\right)$ is not straightforwardly attained. The shape described is known as a *cycloid*.



In the domain $0 \leq \theta \leq 2\pi$, we see the cycloid generates right-to-left. To rectify this, one can change $\theta \to -\theta$ to reverse the evolution of the x-equation while leaving the y-equation unchanged. To put the entire picture above the x-axis, one may flip the sign on R as if reversing the sign on g.

6.2 Tautochrone

A question similar to the brachistochrone is to seek a curve called the *tautochrone*, the shape with the property that a body can start sliding from rest anywhere on the curve and reach the lowest point in a fixed time.

It may be no surprise that the cycloid also solves this problem, so we start with this assumption and check that the claim is satisfied. To this end, let us take as a starting point

$$T = \int_{x_0}^{x_1} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \frac{dx}{\sqrt{2g\left(y_0 - y\left(x\right)\right)}} \ ,$$

where x_1 is located at the bottom of the cycloid (upturned as sketched above).

Knowing the solution to $y\left(\theta\right)$, we can write

$$y_0 - y = R\left(\cos\left(\theta_0\right) - \cos\left(\theta\right)\right) \ ,$$

where θ_0 characterizes the initial position of a body sliding from rest. Then, the equation for velocity as a function of θ reads

$$v\left(\theta\right) = \sqrt{2gR\left(\cos\left(\theta_0\right) - \cos\left(\theta\right)\right)} \ .$$

The remaining quantities in the integral must also be expressed in terms of θ. Doing so carefully, the T-integral becomes

$$T = \sqrt{\frac{R}{g}} \int_{\theta_0}^{\pi} \frac{\sqrt{1 - \cos\left(\theta\right)} \, d\theta}{\sqrt{\cos\left(\theta_0\right) - \cos\left(\theta\right)}} \ .$$

Note that the change of variables $\theta \to -\theta$ has been invoked, simply reversing the sign on the integral to make sure the result is positive.

The task is to solve the integral and ultimately show that θ_0 does not affect the result. For this we use a pair of trigonometric identities

$$\sin\left(\frac{\theta}{2}\right) = \sqrt{\frac{1}{2} - \frac{\cos(\theta)}{2}}$$
$$\cos(\theta) = 2\cos^2\left(\frac{\theta}{2}\right) - 1 \; ,$$

so now

$$T = \sqrt{\frac{R}{g}} \int_{\theta_0}^{\pi} \frac{\sin(\theta/2)\, d\theta}{\sqrt{\cos^2(\theta_0/2) - \cos^2(\theta/2)}} \; .$$

Proceed with the u-substitution

$$u = \cos(\theta/2) / \cos(\theta_0/2)$$
$$du = -d\theta \sin(\theta/2) / 2\cos(\theta_0/2) \; ,$$

and all θ_0-dependence vanishes from the integral, leaving

$$T = \sqrt{\frac{R}{g}} \int_1^0 \frac{-2\, du}{\sqrt{1-u^2}} \; .$$

To finish the integral one could proceed by elementary methods, however notice in our final u-substitution that the initial condition θ_0 has been divided out, and thus never mattered. Choosing a θ_0 that trivializes the integral, namely $\theta_0 = 0$, makes the hard part vanish. One way or the other, the answer boils down to

$$T = \pi \sqrt{\frac{R}{g}} \; ,$$

affirming the cycloid as the solution to the tautochrone problem.

7 Lagrange Multipliers

An item from the calculus tool chest, not particularly related to the Euler-Lagrange equation or Lagrange mechanics, is the idea of *Lagrange multiplier*. A Lagrange multiplier is used when solving an optimization problem subject to a particular constraint. For instance, if we have a two-variable function $f(x,y)$, there may be reason to extremize f subject to a different function $g(x,y) = c$, where c characterizes a level curve on g.

7.1 Parallel Gradients

Finding critical points in f subject to g entails noticing that the gradient of each function is the same at a critical point, up to a proportionality constant λ, the Lagrange multiplier:

$$\vec{\nabla} f(x,y) = \lambda \vec{\nabla} g(x,y)$$

Said a different way, a constrained optimization problem minimizes a functional $\vec{F}$

$$\vec{F}[x,y,\lambda] = \vec{\nabla} f(x,y) - \lambda \vec{\nabla} g(x,y)$$

at critical points (x_0, y_0).

When there are multiple constraint functions, each introduces a new and different λ. The gradient of the original function and all constraints remain proportional:

$$\vec{\nabla} f(x,y) = \lambda_1 \vec{\nabla} g_1(x,y) + \lambda_2 \vec{\nabla} g_2(x,y) + \cdots$$

Worked Example

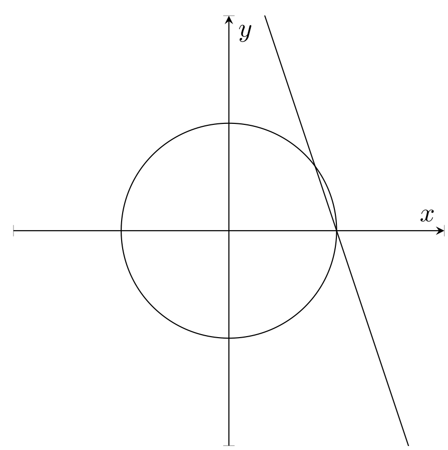
For an example, suppose we have a function

$$f(x,y) = x^2 + y^2 \; ,$$

and we are to find any critical points in f subject to the constraining line

$$y + 3x = 3 \; .$$

Sketched below is the constraining function along with a single level curve (f = constant) of $f(x,y)$.



Constructing the functional $\vec{F}[\lambda]$, we have

$$\vec{F}[\lambda] = \vec{\nabla} f(x,y) - \lambda \vec{\nabla} g(x,y)$$
$$= \langle 2x, 2y \rangle - \lambda \langle 3, 1 \rangle \; .$$

Setting the left side to zero, we gain two new equations

$$2x_0 = 3\lambda$$
$$2y_0 = \lambda \; ,$$

and a third equation is given by g:

$$y_0 + 3x_0 = 3$$

As a system of three equations and three unknowns, the results for λ, x_0, y_0 must be found simultaneously, resulting in

$$\begin{aligned} \lambda &= 3/5 \\ x_0 &= 9/10 \\ y_0 &= 3/10 \,, \end{aligned}$$

and the problem is solved.

Practice

Problem 1

Identify all critical points of the function $f(x, y) = x^2 + 2y^2$ that coincide with the unit circle $x^2 + y^2 = 1$.

Problem 2

Find the largest rectangle that fits inside the first quadrant of the ellipse $x^2/a^2 + y^2/b^2 = 1$. Answer:

$$\begin{aligned} x_0 &= a/\sqrt{2} \\ y_0 &= b/\sqrt{2} \end{aligned}$$

Problem 3

Find the largest square that fits inside the first quadrant of the ellipse $x^2/a^2 + y^2/b^2 = 1$. Hint:

$$\vec{\nabla}(xy) = \lambda_1 \vec{\nabla}\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1\right) + \lambda_2 \vec{\nabla}(x - y)$$

Problem 4

A cylinder of radius R and length L is capped on each end by a cone of height H. Maximize the volume for a given surface area. Hint:

$$\vec{\nabla}\left(\pi R^2 L + \frac{2}{3}\pi R^2 H\right) = \lambda \vec{\nabla}\left(2\pi R L + 2\pi R\sqrt{R^2 + H^2}\right)$$

Answer:

$$V = \frac{AR}{3} = A^{3/2}\left(2\pi\right)^{-1/2} \frac{5^{-1/4}}{3}$$

7.2 Constrained Systems

Let us return to Equation (1.5) representing a system of several variables, namely:

$$\frac{\partial \Lambda}{\partial f^{(j)}} - \frac{d}{dx}\left(\frac{\partial \Lambda}{\partial f_x^{(j)}}\right) = 0$$

With zero on the right side of the equation, one speaks of this as an 'unconstrained' case.

Lagrange multipliers are an elegant means for enforcing constraints of motion by modifying the right side:

$$\frac{\partial \Lambda}{\partial f^{(j)}} - \frac{d}{dx}\left(\frac{\partial \Lambda}{\partial f_x^{(j)}}\right) = \sum_{k=1}^{n} \lambda_k \frac{\partial g_k}{\partial f^{(j)}} \tag{1.9}$$

Replacing zero is any number n total constraint terms. Each involves a Lagrange multiplier λ and also a function g to specify the constraint itself. In particular, g_k must evaluate to zero when the constraint is satisfied.

7.3 Generalized Force

From a mechanical point of view, the right side of Equation (1.9) is called the *generalized force.* This is justified by taking a modified Lagrangian

$$L = T(v) - U(x) - \lambda g(x) \ ,$$

where $g(x)$ is a potential energy term. Then, the gradient factor

$$Q = -\lambda \frac{\partial}{\partial x} g(x)$$

when $g(x) = 0$ corresponds to the body obeying the constraint.

In terms of generalized coordinates, the mechanical analog to Equation (1.9) reads

$$\frac{\partial L}{\partial q^{(j)}} - \frac{d}{dt}\left(\frac{\partial L}{\partial q_t^{(j)}}\right) = Q_{(j)} = \sum_{k=1}^{n} \lambda_k \frac{\partial g_k}{\partial f^{(j)}} \ .$$

8 Sagging Cable

In gravity, a homogeneous cable of fixed length suspended between two points will sag downward to minimize the gravitational potential energy throughout the cable. The exact shape of the cable is straightforwardly attained using the calculus of variations with a Lagrange multiplier.

8.1 Setup

If the linear mass density of the cable is a constant ρ, then a functional representing the gravitational potential energy of the cable is written

$$U[y] = \rho g \int_{x_0}^{x_1} y(x) \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \ .$$

To instill the notion that the length of the cable is constant, a second functional is constructed:

$$V[y] = \int_{x_0}^{x_1} \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$$

The 'grand' functional with which we must work is one that relates U and V by a Lagrange multiplier such that

$$F[y] = U[y] - \lambda V[y] \ ,$$

and the working quantity becomes

$$F[y] = \int_{x_0}^{x_1} (\rho g y(x) - \lambda) \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \ .$$

8.2 Shape of the Cable

Introduce the substitution

$$u(x) = y(x) - \frac{\lambda}{\rho g}$$

such that the functional changes to

$$F[u] = \rho g \int_{x_0}^{x_1} u(x) \sqrt{1 + \left(\frac{du}{dx}\right)^2}\, dx \ ,$$

or in more symbolic notation,

$$F[u] = \int_{x_0}^{x_1} \Lambda(x, u, u')\, dx \ .$$

There is no explicit x-dependence in Λ, so the identity

$$\Lambda - u_x \frac{\partial \Lambda}{\partial u_x} = C_0$$

must hold, where C_0 is a constant. Substituting B into the above and simplifying results in a differential equation

$$\left(\frac{du}{dx}\right)^2 = \left(\frac{u}{C_0/\rho g}\right)^2 - 1 \ ,$$

which can be separated into x- and u-integrals:

$$\int dx = \int \frac{du}{\sqrt{u^2/(C_0/\rho g)^2 - 1}}$$

These are straightforwardly solved as

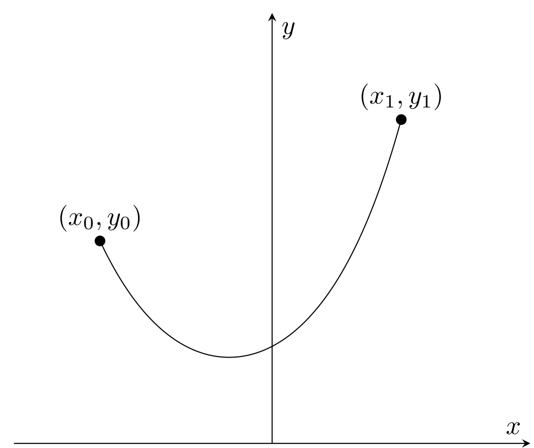
$$x = \frac{C_0}{\rho g} \cosh^{-1}\left(\frac{u}{C_0/\rho g}\right) + C_1 \ ,$$

where C_1 is an arbitrary constant.

Restoring the original y variable through the layers of substitutions, we end up with, for the final shape of the cable:

$$y(x) = \frac{\lambda}{\rho g} + \frac{C_0}{\rho g} \cosh\left(\frac{x - C_1}{C_0/\rho g}\right)$$

The answer you can walk away with is, 'the sagging cable makes the shape of a hyperbolic cosine'. There are three constants in the solution that grant all the flexibility for adjusting endpoints and cable length, but the shape is always governed by cosh as depicted:



8.3 Length of the Cable

None of λ, C_0, or C_1 alone specify the length of the cable. To determine this, we write

$$L = \int_{x_0}^{x_1} \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$$

and use the dy/dx derived from the shape $y(x)$. Carrying this out results in

$$L = \frac{C_0}{\rho g} \sinh\left(\frac{x - C_1}{C_0/\rho g}\right)\Bigg|_{x_0}^{x_1} .$$

8.4 Lowest Point

The location x^* at which the cable sags lowest is the point satisfying $dy/dx = 0$, a criteria easily written:

$$0 = \sinh\left(\frac{x^* - C_1}{C_0/\rho g}\right)$$

This is satisfied by $x^* = C_1$, thus C_1 is equal to the x with the lowest y. The lowest point reached by the cable is

$$y^* = \frac{\lambda + C_0}{\rho g} \ .$$

Having a complete description of the sagging cable doesn't always mean that solving problems is an easy chore. The constant C_0, for instance, is brutally tangled into the y-equation, and also makes an appearance in the formula for L. Once the problem is set up, the number crunching is best left to a computer.

Worked Example

To have an example, consider a cable hung between two endpoints of equal height and equal distance x_0 to the origin. If the cable length obeys $L = 4x_0$, determine the height difference H between the lowest point and the highest point.

The maximum height difference $H = y(\pm x_0) - y^*$ is straightforwardly written

$$H = \frac{C_0}{\rho g}\left(\cosh\left(\frac{x_0}{C_0/\rho g}\right) - 1\right) ,$$

and the length factors in via

$$L = 2\frac{C_0}{\rho g}\sinh\left(\frac{x_0}{C_0/\rho g}\right) ,$$

which is given as $4x_0$. The above is characterized by

$$2q = \sinh(q) ,$$

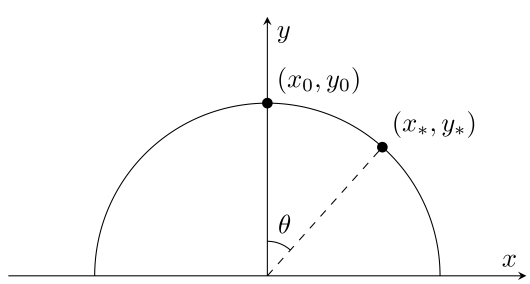
approximately solved by

$$\frac{x_0}{C_0/\rho g} = q \approx 2.177 .$$

With this, H comes out to

$$H \approx x_0\left(\frac{\cosh(2.177) - 1}{2.177}\right) \approx 1.59\, x_0 .$$

9 Sliding Down a Sphere



A particle of mass m sits at rest on a spherical surface of radius R. Receiving a very small nudge to the right, the particle slides down the sphere due to gravity. Measuring the particle's evolution using angle θ, determine the critical angle θ_* that corresponds to the particle leaving the sphere's surface and entering free-fall.

9.1 Newtonian Analysis

By standard Newtonian analysis, it's straightforward to write an equation in polar coordinates for the normal force N that keeps the particle on the surface

$$N = -\frac{mv^2}{R} + mg\cos(\theta) ,$$

where $-mv^2/R$ is the angular acceleration. The moment when the sliding particle enters free-fall, characterized by θ_*, v_*, occurs exactly when the normal force reaches zero:

$$0 = -\frac{mv_*^2}{R} + mg\cos(\theta_*)$$

Meanwhile, energy conservation gives us

$$E = \frac{1}{2}mv^2 + mgR\cos(\theta) ,$$

where E is constant and consists of a kinetic term and a potential term. This also holds until the condition θ_*, v_* is met:

$$mgR = \frac{1}{2}mv_*^2 + mgR\cos(\theta_*)$$

Eliminating v_* between the two equations and simplifying, we conclude easily that

$$\cos(\theta_*) = \frac{2}{3} ,$$

which means the particle leaves the sphere's surface at angle:

$$\theta_* = \arccos\left(\frac{2}{3}\right) \approx 0.841 \text{ rad} \approx 48.2^\circ$$

9.2 Constrained Motion Analysis

It it illustrative to solve the problem using constraints. To this end we will write the Lagrangian of the system with the assumption that r is allowed to vary in time:

$$L = \frac{1}{2}m\left((r_t)^2 + (r\theta_t)^2\right) - mgr\cos(\theta)$$

Then, the so-called constraint simply makes sure that r is a constant:

$$g(r) = r - R$$

With two variables in play, namely r and θ, the constrained Euler-Lagrange equation (1.9) yields two items:

$$\frac{\partial L}{\partial r} - \frac{d}{dt}\left(\frac{\partial L}{\partial r_t}\right) = \lambda\frac{\partial}{\partial r}(r - R)$$
$$\frac{\partial L}{\partial \theta} - \frac{d}{dt}\left(\frac{\partial L}{\partial \theta_t}\right) = \lambda\frac{\partial}{\partial \theta}(r - R)$$

Carrying these calculations through gives a pair of equations

$$\begin{aligned} mr^2\omega_t &= -2mrr_t\omega + mgr\sin(\theta) \\ mr_{tt} &= mr\omega^2 - mg\cos(\theta) - \lambda\,, \end{aligned}$$

where λ is the normal force that keeps the particle outside the sphere, and the angular speed θ_t is selectively replaced by another Greek letter ω:

$$\frac{\partial\theta}{\partial t} = \theta_t = \omega$$

The moment that λ goes to zero the the moment the particle leaves contact with the sphere. Motion is characterized by $r_{tt} = r_t = 0$ because $r = R$. At the critical point, the angular speed θ_t takes on a special value ω_*. Updating the above and simplifying, we have

$$\begin{aligned} \omega_t &= \frac{g}{R}\sin(\theta) \\ \omega_*^2 &= \frac{g}{R}\cos(\theta_*)\ . \end{aligned}$$

The top equation can be manipulated to write

$$R\int \omega\, d\omega = \int g\sin(\theta)\, d\theta\,,$$

which is indeed a statement of energy conservation:

$$E = mgR = \frac{1}{2}mR^2\omega^2 + mgR\cos(\theta)$$

The critical case $\theta = \theta_*$, $\omega = \omega_*$, produces the same answer as above.

10 Maximal Area

10.1 About Enough Length

Two points fixed on the x-axis are separated by Δx. The points are connected by a length L of string that is longer than Δx but shorter than $\pi\Delta x/2$:

$$\Delta x < L < \frac{\pi}{2}\Delta x$$

Maximize the area contained above the x-axis and under the string.

Supposing the path of the string is $y(x)$, the system is described by an area integral and an arc length integral:

$$\begin{aligned} A[y] &= \int_{x_0}^{x_1} y\, dx \\ L[y] &= \int_{x_0}^{x_1} \sqrt{1+y_x^2}\, dx \end{aligned}$$

Introduce a Lagrange multiplier λ to combine each quantity:

$$F[y] = \int_{x_0}^{x_1}\left(y - \lambda\sqrt{1+y_x^2}\right) dx$$

The working quantity

$$\Lambda = y - \lambda\sqrt{1+y_x^2}$$

has no explicit x-dependence, thus we reason from Equation (1.4) that

$$\Lambda - y_x\frac{\partial\Lambda}{\partial y_x} = C_1\,,$$

where C_1 is a constant. Running our function Λ through this results in

$$y - \frac{\lambda}{\sqrt{1+y_x^2}} = C_1\,,$$

which can be written as a separable differential equation:

$$\frac{dy}{dx} = \sqrt{\left(\frac{\lambda}{y-C_1}\right)^2 - 1}$$

Letting $w = y - C_1$ we quickly transform the above into an integral

$$\int dx = \int \frac{w\, dw}{\sqrt{\lambda^2 - w^2}}\,,$$

motivating another substitution

$$\begin{aligned} q &= \lambda^2 - w^2 \\ dq &= -2w\, dw\,. \end{aligned}$$

From here the integration results in

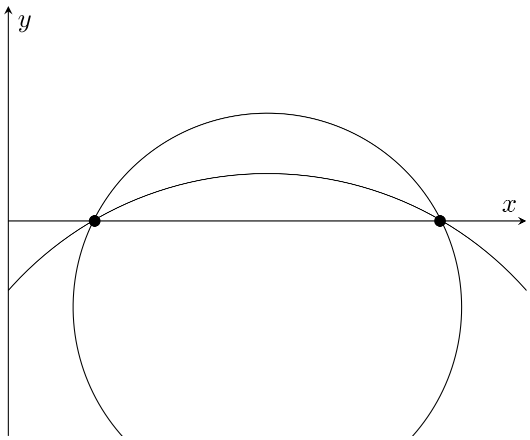
$$x = -\sqrt{q} + C_0\,,$$

where C_0 is a constant.

Restoring all of the substitutions back to y, we get a result that is undoubtedly a circle:

$$(x - C_0)^2 + (y - C_1)^2 = \lambda^2$$

Evidently, the shape that solves the problem is a semicircular arc with radius λ centered at (C_0, C_1) as shown:



10.2 Too Much Length

To solve the maximal area problem above, each integral is framed in terms of y and its derivative y_x, which is to assume that the curve never 'doubles back' in itself, i.e. $y(x)$ is a function. Clearly this won't work in the case where there is too much string, i.e.

$$L > \frac{\pi}{2}\Delta x ,$$

which is the case we ponder now.

By intuitive arguments, we could speculate (but not rely on) that the solution for the modified case is also semicircular. If so, it seems that the figure above already depicts the answer, as the quadrant under the x-axis happens to display circles that satisfy the criteria.

As it turns out, it's a bit easier to proceed by studying a closed loop which has no endpoints. To rationalize this we may assume the special case

$$\Delta x \ll L ,$$

so that the flat part of the resulting curve will be negligible. In effect, we're finding the otherwise-unconstrained finite closed loop that encloses the most area.

10.3 Cartesian Analysis

One way to frame the problem is to write integrals for A and L that do not depend on y being a function. To this end, we borrow from vector calculus and Green's theorem to write the following area formula for a closed curve:

$$A[x,y] = \oint \frac{1}{2}(x\,dy - y\,dx)$$
$$L[x,y] = \oint \sqrt{dx^2 + dy^2}$$

To proceed, frame each of x and y as parametric equations in the variable t. Then, using the chain rule, the above become

$$A[x,y] = \oint \frac{1}{2}(x\,y_t - y\,x_t)\,dt$$
$$L[x,y] = \oint \sqrt{x_t^2 + y_t^2}\,dt ,$$

and from here we combine these using a Lagrange multiplier:

$$F[x,y] = \oint \left(\frac{1}{2}(x\,y_t - y\,x_t) - \lambda\sqrt{x_t^2 + y_t^2}\right)dt$$

Picking out the working quantity

$$\Lambda = \left(\frac{1}{2}(x\,y_t - y\,x_t) - \lambda\sqrt{x_t^2 + y_t^2}\right)$$

and counting the variables, we can apply two instances of the Euler-Lagrange equation:

$$\frac{d}{dt}\left(\frac{\partial\Lambda}{\partial x_t}\right) - \frac{\partial\Lambda}{\partial x} = 0$$
$$\frac{d}{dt}\left(\frac{\partial\Lambda}{\partial y_t}\right) - \frac{\partial\Lambda}{\partial y} = 0$$

Substituting Λ into each gives

$$\frac{d}{dt}\left(\frac{-1}{2}y - \frac{\lambda x_t}{\sqrt{x_t^2 + y_t^2}}\right) = \frac{1}{2}y_t$$
$$\frac{d}{dt}\left(\frac{1}{2}x - \frac{\lambda y_t}{\sqrt{x_t^2 + y_t^2}}\right) = \frac{-1}{2}x_t ,$$

both of which are easily integrated. Introducing respective integrations constants C_0, C_1, we find, after simplifying:

$$x - C_0 = \frac{\lambda y_t}{\sqrt{x_t^2 + y_t^2}}$$
$$y - C_1 = \frac{-\lambda x_t}{\sqrt{x_t^2 + y_t^2}}$$

Square both sides and add to recover the formula for a circle:

$$(x - C_0)^2 + (y - C_1)^2 = \lambda^2$$

10.4 Polar Analysis

The same result can be framed more 'naturally' in polar coordinates, however the calculation that follows isn't any easier than the Cartesian analysis. The area and length functionals respectively take the form

$$A[r,\theta] = \oint \frac{1}{2}r^2\,d\theta$$
$$L[r,\theta] = \oint r\sqrt{1 + \left(\frac{1}{r}r_\theta\right)^2}\,d\theta ,$$

so the total functional using a Lagrange multiplier reads

$$F[r,\theta] = \oint \left(\frac{1}{2} r^2 - \lambda r \sqrt{1 + \left(\frac{1}{r} r_\theta \right)^2} \right) d\theta \,.$$

The working quantity

$$\Lambda = \frac{1}{2} r^2 - \lambda \sqrt{r^2 + r_\theta^2}$$

is a function of r and r_θ, but there is no explicit θ-dependence. By Equation (1.4), we additionally have

$$\Lambda - r_\theta \frac{\partial \Lambda}{\partial r_\theta} = C \,,$$

where C is constant. Substituting Λ into the above, after simplifying, gives

$$\frac{1}{2} r^2 - C = \frac{\lambda r^2}{\sqrt{r^2 + r_\theta^2}} \,.$$

Polar Frame

In a polar coordinate system, the differential arc length is given by

$$d\vec{s} = \langle dr, r\, d\theta \rangle \,,$$

having two components that form the sides of a right triangle with hypotenuse

$$ds = \sqrt{dr^2 + r^2 d\theta^2} \,.$$

The angle formed between the hypotenuse and $r\, d\theta$ shall be denoted ϕ and is given by:

$$\sin(\phi) = \frac{r\, d\theta}{\sqrt{dr^2 + r^2 d\theta^2}} = \frac{r}{\sqrt{r^2 + r_\theta^2}}$$

In terms of ϕ, the differential equation for this system now reads

$$\frac{1}{2} r^2 - C = \lambda r \sin(\phi) \,.$$

Now, if the path of integration is to make a closed loop somewhere in the plane, then the angle ϕ will (at least) hit all of its values in the domain $[0, 2\pi]$. From this we can write a pair of relations

$$\begin{aligned} \frac{1}{2} r_-^2 - C &= \lambda r_- \sin\left(\frac{-\pi}{2} \right) \\ \frac{1}{2} r_+^2 - C &= \lambda r_+ \sin\left(\frac{\pi}{2} \right) \,, \end{aligned}$$

readily implying

$$\lambda = \frac{1}{2} (r_+ + r_-) \,.$$

The distances r_+, r_- are interpreted as the respective furthest and nearest distances from the origin to the extremes of the path.

The case $r^2 = 2C$ corresponds to two points along the path where the $d\theta$-component of the arc length is zero, i.e. the displacement is momentarily parallel to the line made by r.

Cartesian Frame

It's possible to begin with the polar analysis of the problem, namely

$$\frac{1}{2} r^2 - C = \frac{\lambda r^2}{\sqrt{r^2 + r_\theta^2}} \,,$$

and end up with a Cartesian result.

Looking at the case $r^2 = 2C$, this corresponds to the scenario where the position vector is tangent to the curve. By specially tuning $C = 0$, we move the origin to somewhere on the curve itself, with the polar axis being along the tangent to the curve at that point. With this choice, we then have

$$r^2 = \frac{2\lambda r^2}{\sqrt{r^2 + r_\theta^2}} \,,$$

readily simplifying to a separable differential equation

$$\int d\theta = \int \frac{dr}{\sqrt{4\lambda^2 - r^2}} \,,$$

having solution

$$\theta = \theta_0 + \arcsin\left(\frac{r}{2\lambda} \right) \,,$$

where the integration constant $\theta_0 = 0$ by construction.

Rearranging the above, we ultimately find

$$r = 2\lambda \sin(\theta) \,.$$

Making use of the identities

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

we swiftly make out the fingerprint of a circle:

$$\begin{aligned} r^2 &= 2\lambda r \sin(\theta) \\ x^2 + y^2 &= 2\lambda y \\ x^2 + (y - \lambda)^2 &= \lambda^2 \end{aligned}$$