15 Constraints

Problems are hidden opportunities, and constraints can actually boost creativity.
—Martin Villeneuve

There is ample evidence in all fields of science and engineering wherein the lack of constraints leads to intractable problems. For instance if we try to enclose the most area for a sheep farm, with no constraints on the cost of materials, there is no solution, nor any creativity. Without constraints on the systems, there are often no solutions, and occasionally the solution is far more relevant and even elegant in the presence of constraints.

When we consider an optimization problem with constraints, one perspective is to say that these constraints are changing the admissible set of potential optimizers. In other words, without constraints we may be interested in all functions that lie in $C^2(\mathbb{R}^2)$ for example, but then we introduce the constraint that the optimal path must also lie in a given region in $\mathbb{R}^2$, or perhaps the function has an upper bound. We are no longer optimizing over all of $C^2(\mathbb{R}^2)$ anymore, but must consider a more restricted class of potential optimizers. In what follows this is not the approach we take, but instead we will return to those wonderful objects referred to in multivariable calculus as Lagrange multipliers (although naturally we will make them worse in this context), but when optimization problems are solved numerically the typical approach to include constraints is to simply restrict the class over which the optimization routines search. Hence, although constraints often pose significant difficulties in the formal calculation of extremizers, the practical numerical implementation may actually be improved by adding constraints to the system.

15.1 Integral (Isoperimetric) Constraints, a rigorous formulation

Before proceeding we note that optimization with constraints is something that you have already seen in Calculus, but over finite dimensional spaces. In that setting you may have been asked to optimize a function $f(x, y, z)$ constrained to live on the generically two-dimensional surface prescribed by $G(x, y, z) = 0$. This led to the condition that $\nabla f = \lambda \nabla G$ (so long as $\nabla G \neq 0$) for some constant λ which also implies that the gradients of these two functions are parallel at the minima. We obtain a very similar result in the Calculus of Variations, but in this setting our gradient is the statement of the Euler-Lagrange equations.

Before we jump into how this works for the Calculus of Variations we will first consider one approach that justifies this in the finite dimensional setting. There are a few things we need to consider. First, note that $\nabla f = 0$ is equivalent to stating that

$$\nabla f \cdot \mathbf{v} = 0,$$

for all 'admissible' $\|\mathbf{v}\| = 1$ i.e. the directional derivative vanishes in all directions (this should sound a lot like our discussion of the Gateaux differential). Similarly to prove that

$$\nabla f = \lambda G,$$

we need only show that

$$\nabla f \cdot \mathbf{v} = \lambda \nabla G \cdot \mathbf{v},$$

for all $\|\mathbf{v}\| = 1$ (in the infinite dimensional setting we actually only work with the Gateuax differential which is the analog of the directional derivative and we never consider the Frechet derivative which is the analog of the gradient).

Now let $\boldsymbol{\eta}$ be an arbitrary vector (we can normalize it so that $\|\boldsymbol{\eta}\| = 1$ if we want, but it is relatively easy to see that won't matter here), and suppose that the point P is a local extrema (minimizer) of $f(x, y, z)$ constrained to the surface $G(x, y, z) = 0$. At this point, consider the vector $\mathbf{v}$ which is taken as an arbitrary vector that points along the surface in question and let $\mathbf{u} = \mathbf{v} - \boldsymbol{\eta}$. Because $\mathbf{v}$ points along the surface of constraint then $\nabla G \cdot \mathbf{v} = 0$. Otherwise, G would not be constant along this vector (which we already claimed pointed along the surface itself...recall that this is all evaluated at the point P which lies on the surface in question). Hence, we have the condition

$$\nabla G \cdot \boldsymbol{\eta} + \nabla G \cdot \mathbf{u} = 0.$$

Note that this condition is a condition on $\mathbf{u}$ and $\boldsymbol{\eta}$, not on G because we have already established the hypothesis that $\nabla G \neq 0$ at the point P.

Now if the point P is a local minima of f on the surface $G = 0$ then $\nabla f \cdot \mathbf{v} = 0$ or otherwise moving along the direction $\mathbf{v}$ (or $-\mathbf{v}$) will lead to a larger value of f. Thus we also have the condition

$$\nabla f \cdot \boldsymbol{\eta} + \nabla f \cdot \mathbf{u} = 0.$$

At this juncture it is important to recall that $\boldsymbol{\eta}$ is an arbitrarily chosen vector (direction). This leads to the following evaluated at the point P:

$$\begin{aligned} \nabla G \cdot \boldsymbol{\eta} &= -\nabla G \cdot \mathbf{u}, \\ \nabla f \cdot \boldsymbol{\eta} &= -\nabla f \cdot \mathbf{u}. \end{aligned}$$

Dividing both sides of the first equation by the same side of the second, we arrive at

$$\frac{\nabla G \cdot \boldsymbol{\eta}}{\nabla f \cdot \boldsymbol{\eta}} = \frac{\nabla G \cdot \mathbf{u}}{\nabla f \cdot \mathbf{u}} = -\lambda,$$

where λ is a constant which gives the desired condition on the arbitrarily chosen vector (direction) η.

15.1.1 The 'rigorous' justification

Suppose that we want to minimize the cost functional

$$J[y] = \int_a^b L(x, y, y')dx,$$

where $y(a) = A$ and $y(b) = B$, and subject to the integral constraint

$$I[y] = \int_a^b K(x, y, y')dx = c \text{ (constant)}.$$

This occurs in a variety of physically relevant circumstances including the classic problem of minimizing the perimeter of a particular shape for a fixed amount of enclosed area (think of the farmer trying to save on fencing material, but knowing exactly how much area he/she desired to enclose).

From multivariable calculus, we expect the following theorem to hold.

Theorem 15.1.1. *Let $\hat{y}(x)$ to be a minimizer of $J[y]$ satisfying $I[\hat{y}] = C$ (where $\hat{y}(x)$ is not an extrema of $I[y]$ so that the Euler-Lagrange equations of $I[y]$ are* NOT *satisfied for this particular $\hat{y}(x)$). Then it is necessary that $\hat{y}(x)$ be an extremal of*

$$\int_a^b [L(x, y, y') + \lambda K(x, y, y')]dx,$$

for some constant λ. (λ is most often referred to as the Lagrange multiplier for the problem).

Proof.

We begin by selecting two points x_1 and x_2 in the interval $[a, b]$ where x_1 is arbitrary, and x_2 satisfies

$$\left[\frac{\partial K}{\partial y} - \frac{d}{dx}\frac{\partial K}{\partial y'}\right]_{y=\hat{y}, x=x_2} \neq 0.$$

Such a point x_2 is guaranteed to exist because we assumed that $\hat{y}$ is NOT an extremal of the functional $I[y]$ (the utility of this condition is clearly not apparent yet, but will become clear later on in the proof).

Now, if $\hat{y}(x)$ is a minimizer of $J[y]$ then we will consider the variation $\hat{y}(x) \to \hat{y}(x) + \varepsilon h_1(x) + \varepsilon h_2(x)$ where $h_1(x)$ is nonzero only in a small interval around x_1 and $h_2(x)$ is similarly only nonzero in a small neighborhood around x_2. In the following we will basically treat $h_1(x)$ and $h_2(x)$ as smooth approximations of $\delta(x - x_1)$ and $\delta(x - x_2)$ respectively. Because $\hat{y}(x)$ is a minimizer of $J[y]$ it follows that

$$\delta J[\hat{y}; h_1 + h_2] = \int_a^b \left[L_y - \frac{d}{dx}L_{y'}\right][h_1 + h_2]\, dx = 0,$$

where the boundary terms vanished by assumption. For this particular choice of h_1 and h_2 (and ignoring a little bit of necessary rigorous justification for now, but assuming that these variations are indeed smooth approximations to the appropriate delta functions) this is equivalent to

$$\left[L_y - \frac{d}{dx}L_{y'}\right]_{x=x_1} + \left[L_y - \frac{d}{dx}L_{y'}\right]_{x=x_2} = 0.$$

Now we also need to guarantee that the variation $\hat{y}+\varepsilon h_1+\varepsilon h_2$ maintains the constraint $I[\hat{y} + \varepsilon h_1 + \varepsilon h_2] = I[\hat{y}]$ which up to $O(\varepsilon)$ leads to the condition that

$$\delta I[\hat{y}; h_1 + h_2] = 0,$$

which, as before we can take in the limit to become:

$$\left[K_y - \frac{d}{dx}K_{y'}\right]_{x=x_1} = -\left[K_y - \frac{d}{dx}K_{y'}\right]_{x=x_2} \neq 0,$$

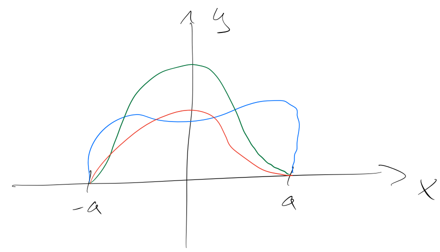

Figure 15.1: An illustration of potential shapes of the optimal Rathtar pen. The horizontal axis is the boundary with the lake (which Rathtars will not cross evidently).

because as our assumption on the point x_2 indicated, the right hand side of this equation is nonzero. Now we will note that because this quantity is nonzero then

$$\left.\frac{L_y - \frac{d}{dx}L_{y'}}{K_y - \frac{d}{dx}K_{y'}}\right|_{x=x_1} = -\left.\frac{L_y - \frac{d}{dx}L_{y'}}{K_y - \frac{d}{dx}K_{y'}}\right|_{x=x_2} = -\lambda,$$

which is a constant. Thus we see that at the point $x = x_1$ (which was chosen arbitrarily)

$$L_y - \frac{d}{dx}L_{y'} + \lambda\left(K_y - \frac{d}{dx}K_{y'}\right) = 0.$$

Because $x = x_1$ is an arbitrary point in the interval $[a, b]$ then we see that this is satisfied for every point $x \in [a, b]$ which satisfies the conditions of the theorem. □

15.1.2 Examples of integral constraints

After all the fun we had in the previous section, we now get to see how to use these Lagrange multipliers in this infinite dimensional setting.

Example 15.1.2 (Isoperimetric Problem). Suppose that you have been hired by the King of Prana to build an enclosure for his recently acquired pair of Rathtars. Rathtars evidently detest water, so one side of their enclosure will be the (straight) edge of a lake. The rest of the enclosure will be constructed from some very expensive, indestructible (almost) Mandalorian iron (this is what Beskar armor is made from). Unfortunately to make the enclosure high enough, you only have enough Mandalorian iron to construct an enclosure with perimeter L.

To solve this problem, denote the curve $y(x)$ to be the shape of the enclosure with prescribed length L, and bounded by the x-axis on the bottom. We are interested in this curve enclosing the largest area (why do we not need to worry about a minimum happening here?). We can assume that the endpoints (intersection points with the x-axis) are at $(-a, 0)$ and $(a, 0)$ (see Figure 15.1 for some sample shapes of the pen). The exact value of a is determined by the solution. The area we want to optimize is given by

$$J[y] = \int_{-a}^{a} y(x)dx$$

with the constraints $y(-a) = y(a) = 0$, and

$$I[y] = \int_{-a}^{a} dl = \int_{-a}^{a} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx = L.$$

Using the Lagrange multiplier formulation of the last section, this means that we want to optimize

$$\int_{-a}^{a} (y + \lambda\sqrt{1 + (y')^2}) dx.$$

The (EL) become

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

These can be simplified to

$$\begin{aligned}
\frac{d}{dx} \frac{y'}{\sqrt{1 + (y')^2}} &= \frac{1}{\lambda} \\
\Rightarrow \frac{y'}{\sqrt{1 + (y')^2}} &= \frac{1}{\lambda}(x + c_1) \\
\Rightarrow \frac{\lambda y'}{\sqrt{1 + (y')^2}} &= (x + c_1) \\
\Rightarrow \frac{\lambda^2 (y')^2}{1 + (y')^2} &= (x + c_1)^2 \\
\Rightarrow y' = \frac{dy}{dx} = \frac{\pm(x + c_1)}{\sqrt{\lambda^2 - (x + c_1)^2}} & \\
\Rightarrow y(x) = \mp\sqrt{\lambda^2 - (x + c_1)^2} + c_2. &
\end{aligned}$$

It follows that $(x + c_1)^2 + (y - c_2)^2 = \lambda^2$ which is the equation for a circle of radius λ, assuming the symmetry about the origin. Thus the extremizer is a semi-circle of raduis λ, and since it is centered at the origin then $c_1 = c_2 = 0$ and thus $\lambda = a$ the perimeter of the semi-circle is $L = \pi a$. Thus $a = \frac{L}{\pi}$.

This is of course one of the simplest examples of something like this, but there are of course many more examples of similar problems that are actually of interest to individuals that live outside of Prana. For instance, a higher dimensional isoperimetric problem is to determine the shape of a surface that will enclose the most area (given some boundary constraints) for a specified amount of available material.

Example 15.1.3. Consider the less motivated but more computable maximization of

$$J[y] = \int_0^1 \left[x^2 - (y')^2\right] dx,$$

where $y(0) = y(1) = 0$ and the integral constraint

$$\int_0^1 y^2 dx = 2,$$

must be satisfied.

In this case, the (EL) with Lagrange multiplier λ become

$$\begin{aligned} 2\lambda y + 2\frac{d}{dx}y' &= 0, \\ \Rightarrow \lambda y &= -y'', \end{aligned}$$

which we can readily recognize has solution given by

$$y(x) = c_0 \sin(\sqrt{\lambda}x) + c_1 \cos(\sqrt{\lambda}x).$$

Satisfying the boundary conditions $y(0) = y(1) = 0$ leads to (recall this from earlier in the book when we were priveleged to work with eigenvalue problems that were very similar)

$$y_k(x) = c_0 \sin(k\pi x),$$

with $\lambda_k = k^2\pi^2$. Note that $k \neq 0$ since $y_0 \equiv 0$ does not satisfy the integral constraint. The constant c_0 is found by satisfying the constraint

$$\int_0^1 y_k^2 dx = 2,$$

which becomes

$$\begin{aligned} 2 &= \int_0^1 c_0^2 \sin^2(k\pi x) dx, \\ &= \frac{c_0^2}{2}, \end{aligned}$$

which leads to $c_0 = \pm 2$, so that the critical solutions are given by

$$y_k^{\pm}(x) = \pm 2 \sin(k\pi x).$$

Evaluating these functions in the functional, we find

$$J[y_k^{\pm}] = \int_0^1 [x^2 - 4\cos^2(k\pi x)] dx = \frac{1}{3} - 2k^2\pi^2.$$

Thus, if $y_1^{\pm} = \pm 2\sin(\pi x)$ are local maxima of J, they are global maxima because the functional value of $y_k^{\pm}$ for $k > 1$ is less than the value for $k = 1$.

Remark 15.1.4. Note how similar the previous problem was to identifying eigenfunctions of a linear operator as discussed for solutions of PDEs. This is not a coincidence, in fact the existence and uniqueness of eigenfunctions and eigenvalues is typically stated as a variational problem.

Another example where this type of a constraint may appear is given below:

Example 15.1.5. We desire to minimize

$$J[u] = \left(\int_0^1 (u')^2 dx\right)\left(\int_0^1 (u+1) dx\right),$$

where $u(0) = 0$ and $u(1) = a$ are fixed. To approach this problem, we will first designate

$$I_1[u] = \int_0^1 (u')^2 dx, \quad \text{and} \quad I_2 = \int_0^1 (u+1)dx,$$

so that (Gateaux differentials obey the product rule...this should be apparent as the Gateaux differential is defined via a one-dimensional derivative)

$$\begin{aligned}\delta J[u;h] &= \delta I_1[u;h] I_2[u] + I_1[u]\delta I_2[u;h] \\ &= I_2[u]\left\{\delta I_1[u;h] + \frac{I_1[u]}{I_2[u]}\delta I_2[u;h]\right\} = 0.\end{aligned}$$

Now we note that $\frac{I_1[u]}{I_2[u]} = \mu$ is just a constant value for a prescribed $u(x)$ so that really what we need to consider is

$$\delta I_1[u;h] + \mu\delta I_2[u;h] = 0,$$

to find the optimizer of $J[u]$. This leads to a setup that is very similar to an isoperimetric constraint.

The previous example demonstrates a very important concept that allows us to optimize functionals that are not necessarily linear in the input function $y(x)$. Just as with finite dimensional Calculus, we can make use of the product, quotient, and chain rules when computing the Gateaux differential, and thus reduce the calculation of a nonlinear functional to one that is easier to work with.

15.2 Rigorous derivation of Non-Integral (Finite) Constraints

Now if we return to computing geodesics on a given surface we note that this is equivalent to specifying a pointwise constraint, i.e. minimize

$$J[\mathbf{u}] = \int_\Omega L(x;\mathbf{u},\mathbf{u}')dx$$

given that $G(x,\mathbf{u}) = c$ where $\mathbf{u}(x)$ is a vector valued function. In this case, there is a very similar result, but now the Lagrange multiplier is actually a function itself.

Theorem 15.2.1. *For the functional $J[\mathbf{u}] = \int_a^b L(x,\mathbf{u},\mathbf{u}')dx$ with the constraint $G(x,\mathbf{u}) = 0$ being satisfied and $\mathbf{u}(x)$ fixed at the boundaries $x = a$ and $x = b$. Let $\hat{\mathbf{u}}(x)$ be an extremal of this problem. If $\nabla_\mathbf{u} G \neq 0$ for $\mathbf{u}(x) = \hat{\mathbf{u}}(\mathbf{x})$ then there exists a function $\lambda(x)$ such that $\tilde{\mathbf{u}}(x)$ is an extremal of*

$$J^*[\mathbf{u},\lambda] = \int_a^b [L(x,\mathbf{u},\mathbf{u}') + \lambda(x)G(x,\mathbf{u})]dx$$

i.e. the Euler-Lagrange equations become:

$$L_{u_k} + \lambda G_{u_k} - \frac{d}{dx}L_{u_k'} = 0, \text{ for all } k, \quad G(x,\mathbf{u}) = 0.$$

Remark 15.2.2. There are several different ways to prove this Theorem. We have chosen to go the route here that is the most illustrative and (hopefully) informative.

Proof. Just as we did for the proof of the integral constraint we will start by considering variations that ensure that the constraint is satisfied, and then we connect these to general variations. Consider variations $\mathbf{u}(x) = \hat{\mathbf{u}}(x) + \varepsilon \mathbf{h}(x)$ where the constraint remains satisfied i.e.

$$G(x, \mathbf{u} + \varepsilon \mathbf{h}) = 0, \tag{15.1}$$

where $\mathbf{h}(a) = \mathbf{h}(b) = 0$. Differentiating this constraint with respect to ε and setting $\varepsilon = 0$ we obtain the relation

$$\nabla_{\mathbf{u}} G(x, \hat{\mathbf{u}}) \cdot \mathbf{h}(x) = 0. \tag{15.2}$$

The non-degeneracy condition $\nabla_{\mathbf{u}} G(x, \hat{\mathbf{u}}) \neq 0$ guarantees via the Implicit Function Theorem that the variations $\hat{\mathbf{u}}(x) + \varepsilon \mathbf{h}(x)$ exist.TODO: maybe clarify this a bit more?

We can set the Gateaux differential to zero for these variations $\mathbf{h}(x)$ but they are not generic i.e. we can not then apply the Euler-Lagrange equations because the corresponding integral only vanishes for very special $\mathbf{h}(x)$ that ensure the constraint remains satisfied. We need to rework this so that a different set of equations are satisfied for all generic variations $\boldsymbol{\eta}(x)$. With this in mind, we define $\mathbf{h}(x)$ as follows for any admissible $\boldsymbol{\eta}(x)$ (regardless of whether the variation $\boldsymbol{\eta}(x)$ maintains the constraint or not),

$$\mathbf{h}(x) = \boldsymbol{\eta}(x) - \frac{\boldsymbol{\eta}(x) \cdot \nabla_{\mathbf{u}} G(x, \hat{\mathbf{u}})}{\|\nabla_{\mathbf{u}} G(x, \hat{\mathbf{u}})\|^2} \nabla_{\mathbf{u}} G(x, \hat{\mathbf{u}}). \tag{15.3}$$

[29]

Note that

$$\begin{aligned}
\nabla_{\mathbf{u}} G(x, \hat{\mathbf{u}}) \cdot \mathbf{h}(x) &= \nabla_{\mathbf{u}} G(x, \hat{\mathbf{u}}) \cdot \boldsymbol{\eta}(x) - \frac{\boldsymbol{\eta}(x) \cdot \nabla_{\mathbf{u}} G(x, \hat{\mathbf{u}})}{\|\nabla_{\mathbf{u}} G(x, \hat{\mathbf{u}})\|^2} \|\nabla_{\mathbf{u}} G(x, \hat{\mathbf{u}})\|^2 \\
&= \nabla_{\mathbf{u}} G(x, \hat{\mathbf{u}}) \cdot \boldsymbol{\eta}(x) - \boldsymbol{\eta}(x) \cdot \nabla_{\mathbf{u}} G(x, \hat{\mathbf{u}}) = 0,
\end{aligned}$$

and so this $\mathbf{h}(x)$ ensures the constraint is satisfied. Thus we consider the variation

$$\begin{aligned}
0 = \delta J[\hat{\mathbf{u}}; \mathbf{h}] &= \int_a^b \left(\nabla_{\mathbf{u}} L - \frac{d}{dx} \nabla_{\mathbf{u}'} L \right) \cdot \mathbf{h}(x) dx \\
&= \int_a^b \left(\nabla_{\mathbf{u}} L - \frac{d}{dx} \nabla_{\mathbf{u}'} L - \lambda(x) \nabla_{\mathbf{u}} G(x, \hat{\mathbf{u}}) \right) \cdot \boldsymbol{\eta}(x) dx,
\end{aligned}$$

where

$$\lambda(x) = \frac{\left(\nabla_{\mathbf{u}} L - \frac{d}{dx} \nabla_{\mathbf{u}'} L \right) \cdot \nabla_{\mathbf{u}} G(x, \hat{\mathbf{u}})}{\|\nabla_{\mathbf{u}} G(x, \hat{\mathbf{u}})\|^2}. \tag{15.4}$$

As this is satisfied for an arbitrary admissible $\boldsymbol{\eta}(x)$ then the result follows. □

Remark 15.2.3. We note that this is satisfied at the minimizer $\hat{\mathbf{u}}(x)$ i.e. the Lagrangian L and it's gradient(s) are also evaluated at this point.

[29] Sometimes in Mathematical proofs there is a moment when there is a sudden assignment/assumption that miraculously makes the proof work out. This is just such a moment. Just like the magician that pulls the rabbit out of a hat, we have defined $\mathbf{h}(x)$ without giving any context, other than the inherent faith that you (the reader) have that everything will work out. In reality this choice of $\mathbf{h}(x)$ only comes about after carefully working through this proof line-by-line and concluding that this is how the constraint is placed on the variations themselves.

Remark 15.2.4. For the sake of including another remark, we comment on the method of proof for both this (finite constraints) and the previous case from the last two sections. The primary idea is that the Euler-Lagrange equations are derived relying on the vanishing of some integral quantity whose argument is multiplied by an arbitrary function (the only constraints on said function can be how differentiable it is). When we have a specific constraint such as those discussed so far, we can't rely on this same argument because the arbitrary function is no longer arbitrary but is constrained to a subset of the entire functional space.

In other words, we can't derive a version of the Euler-Lagrange equations when we simply restrict our variations (usually called the $h(\mathbf{x})$ in our formal derivations). We instead need to modify the cost functional so that the variations can be defined arbitrarily and yet the constraint is still satisfied.

15.2.1 Examples of non-integral constraints

Example 15.2.5. Consider a disk rolling without slipping (BB-8 is good example) on an inclined plane $y = R\theta$ where R is the radius of the disk; see Figure 15.2. We can specify the equation of constraint as

$$G(y, \theta) = y - R\theta = 0.$$

If l is the length of the inclined side of the plane and we suppose that there is zero potential energy when the disk reaches the bottom, Hamilton's principle (see the next Chapter) tells us that the disk's motion will minimize

$$J[y, \theta] = \int_{t_0}^{t_1} L(t, y, y', \theta, \theta')dt$$

where

$$L(t, y, y', \theta, \theta') = \frac{1}{2}m(y')^2 + \frac{1}{4}mR^2(\theta')^2 + mg(y - l)\sin\phi$$

where m is the mass of the disk and g is the gravitational acceleration constant. Thus the (EL) become

$$L_y + \lambda G_y - \frac{d}{dt}L_{y'} = 0$$
$$L_\theta + \lambda G_\theta - \frac{d}{dt}L_{\theta'} = 0,$$

and $G = 0$ which for our problem becomes

$$mg\sin\phi + \lambda - \frac{d}{dt}(my') = 0 \tag{15.5}$$

$$-R\lambda - \frac{d}{dt}\left(\frac{1}{2}mR^2\theta'\right) = 0 \tag{15.6}$$

$$y = R\theta. \tag{15.7}$$

From the final equation we also have $y'' = R\theta''$. From the second equation above, we have

$$-R\lambda = \frac{1}{2}mR^2\theta''.$$

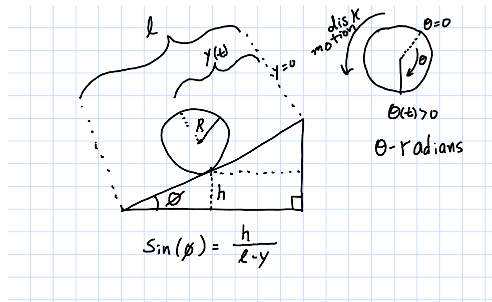

Figure 15.2: Figure of disk rolling without slipping.

With the constraint, this implies that $\lambda = \frac{-1}{2}my''$. Now with the first of the (EL),

$$y'' = g\sin\phi - \frac{1}{2}y''$$

implying that $y'' = \frac{2g\sin\phi}{3}$. Using the previous solutions, we have $\lambda = \frac{-mg\sin\phi}{3}$ and $\theta'' = \frac{2g\sin\phi}{3R}$. Note that without friction (the constraint) then $y'' = g\sin\phi$ so the constraint effectively reduces the gravitational force by $\frac{1}{3}$, i.e. the frictional force in this case is exactly $\lambda = \frac{-mg\sin\phi}{3}$.

The general solution to $y'' = \frac{2g\sin\phi}{3}$ is given by $y(t) = \frac{g\sin\phi}{3}(t-t_0)^2 + c_1 t + c_2$. Recalling that $y(0) = 0$ and assuming that the disk has velocity zero at the starting time t_0, we see that $y(t) = \frac{g\sin\phi}{3}(t-t_0)^2$, and $\theta(t) = \frac{g\sin\phi}{3R}(t-t_0)^2$.

Remark 15.2.6. Lagrange multipliers can be used as well when $G(x, y, y') = 0$, i.e. the constraint depends on derivatives of the extremal. This is referred to as a 'non-holonomic' constraint. Holonomic constraints are not functions of any of the derivatives of the extremals. Some holonomic constraints appear to depend on derivatives, but they can be integrated (cleverly) to resemble $G(\mathbf{x}, \mathbf{y}) = 0$ where $\mathbf{x}$ and $\mathbf{y}$ can be vectors.

Example 15.2.7. As an illustration of the use of Lagrange multipliers, we will derive the Euler-Lagrange equations for a cost function that is dependent on the curve $y(x)$, and its first two derivatives. Consider

$$J[y] = \int_a^b L(x, y, y', y'')dx$$

rewritten as

$$J[y] = \int_a^b L(x, y, y', z')dx,$$

where $z = y'$ yields the constraint $G(y', z) = z - y' = 0$ This leads to (EL)

$$L_y - \frac{d}{dx}(L_{y'} - \lambda) = 0$$
$$\lambda - \frac{d}{dx}L_{z'} = 0.$$

Since $z' = y''$, the latter equation gives

$$\lambda = \frac{d}{dx} L_{z'} = \frac{d}{dx} L_{y''}.$$

Putting everything together then yields

$$L_y - \frac{d}{dx} L_{y'} + \frac{d^2}{dx^2} L_{y''} = 0,$$

which are the (EL) for this particular cost functional.

Example 15.2.8 (Stokes equations). Suppose that $\mathbf{u}(\mathbf{x})$ is the velocity of a divergence free fluid, meaning that $\nabla \cdot \mathbf{u} = 0$. If we let Ω be a nice smooth region in $\mathbb{R}^3$ whereon $\mathbf{u} = 0$ on $\partial\Omega$ the boundary of Ω, then we can find the evolution of the fluid from a minimization principle. It turns out that the fluid in a steady state (meaning the flow is not changing in time) will minimize

$$J[\mathbf{u}] = \int_\Omega \left[\frac{1}{2} |\nabla \mathbf{u}|^2 - \mathbf{f} \cdot \mathbf{u} \right] d\mathbf{x},$$

where $\mathbf{f}(\mathbf{x})$ is a 'body force' acting on the fluid such as a wave-maker or perhaps a turbine. We also need to enforce the constraint that $\nabla \cdot \mathbf{u} = 0$, i.e. we want to find the (EL) for the cost functional

$$\int_\Omega \left[\frac{1}{2} |\nabla \mathbf{u}|^2 - \mathbf{f} \cdot \mathbf{u} - p(\mathbf{x}) \left(\nabla \cdot \mathbf{u} \right) \right] d\mathbf{x}.$$

After a bit of work and effort (review some of the previous chapter to see how this works), one can convince oneself that these are exactly the steady-state Stokes equations for an incompressible fluid, i.e.

$$\begin{aligned} -\Delta \mathbf{u} &= \mathbf{f} - \nabla p \\ \nabla \cdot \mathbf{u} &= 0, \end{aligned}$$

in the interior of Ω and $\mathbf{u}$ vanishes on the boundary of Ω.

In this particular case, the Lagrange multiplier $p(\mathbf{x})$ has a distinct physical interpretation (this is not guaranteed). $p(\mathbf{x})$ in this case is the pressure at the point $\mathbf{x}$.

Remark 15.2.9. A word of caution is necessary after the last example. Interpretation of Lagrange multipliers as a physically interesting quantity is a dangerous objective and one that we plan to avoid if possible. Occasionally we are lucky and the multiplier has a clear interpretation (such as the previous example, or often in economic examples as the enforced cost per unit). The reader should be aware that such instances are exceptions, and should be careful when trying to interpret the meaning of Lagrange multipliers.

Remark 15.2.10. One can further generalize the use of Lagrange multipliers to the case of multiple constraints. For example if we desired to optimize the cost functional

$$J[y] = \int_a^b L(x; y, y') dx,$$

subject to the constraints

$$G(x, y, y') = 0, \quad \text{and} \quad H(x, y, y') = 0,$$

then this would be equivalent to optimizing the cost functional

$$J^*[y] = \int_a^b \left[L(x; y, y') + \lambda(x)G(x, y, y') + \mu(x)H(x, y, y')\right] dx,$$

where $\mu(x)$ and $\lambda(x)$ are two scalar valued functions. Inclusion of an additional integral constraint would follow immediately as well.

15.3 Inequality constraints, setting up the problem

Now, consider the Underminer (see if you can find that reference) who is not really restricted to living on the surface of the earth, but must live in some shell between the surface and the edge of the mantle. In this case, if the Underminer is trying to determine the shortest path from one point to another (for instance he is aiming for the bank before the Incredibles get there), then he is no longer restricted to the standard constraint that $x^2 + y^2 + z^2 = R^2$ (assuming that the earth is a sphere), but instead he is restricted to $r^2 \leq x^2 + y^2 + z^2 \leq R^2$ where r is the radius that defines the edge of the mantle-crust boundary.

In this setting, the Underminer wants to find the minimum of

$$J[x, y, z] = \int \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} dt,$$

where

$$r^2 \leq x^2 + y^2 + z^2 \leq R^2.$$

In other words, we are looking at a finite inequality constraint. The general setting for this would be:

$$\text{minimize } J[\mathbf{y}] = \int_\Omega L(\mathbf{x}; \mathbf{y}, \nabla \mathbf{y}) d\mathbf{x}, \tag{15.8}$$

$$\mathbf{H}(\mathbf{x}) \leq 0, \tag{15.9}$$

where the last inequality is defined pointwise, that is every entry of the vector valued function $\mathbf{H}(\mathbf{x})$ is negative definite. In the case of the Underminer, there are two such inequality statements, one for r and one for R.

Rigorous results justifying inequality constraints are rather ornerous to come by, so we instead motivate everything here by examples. Making this approach rigorous requires resorting to a finite discretization, using the Karush-Kuhn-Tucker (KKT) conditions in finite dimensional optimization, and taking the continuum limit of the discretization. This isn't a terribly painful process, but would require more time than we would like to spend on the topic, particularly because the details of such a proof are actually not very instructive, and don't provide any insight into the methodology.

Hence, we will delve into the use of inequality constraints by returning to a tale of some other creatures that prefer to dwell in the ground, as explained by none other than J. R. R. Tolkien.

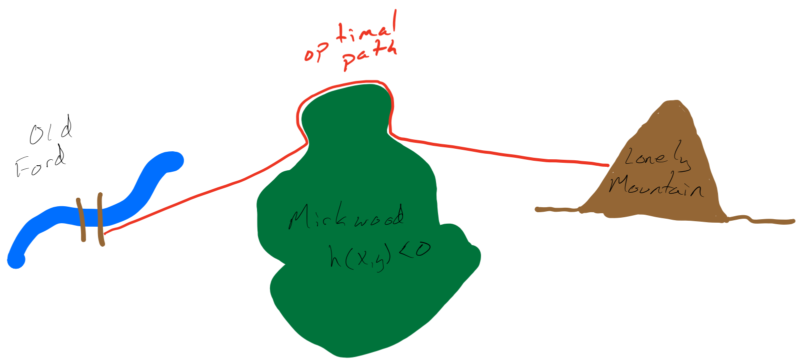

Figure 15.3: Gimli's optimal path for traveling from the Old Ford to the Lonely Mountain while stealthily avoiding any travel through Mirkwood.

Example 15.3.1. Consider the problem of determining the shortest distance between two points on the plane with an impassable object of some determined shape lying between the two points. For example, suppose that in his younger years Gimli the dwarf is trying to get from the Old Ford to the Lonely Mountain, and will avoid passing through Mirkwood at all costs. What would be his optimal path to get from the Old Ford to the Lonely Mountain while avoiding the Elven forest altogether? The answer is quite intuitive, Gimli would make an angled path that intersects the edge of Mirkwood, then he would skirt around the forest to the other side, making an identical angled, but straight line to Erebor on the other side.

Now that we know the answer in this case, we will formulate the problem as a Calculus of Variations problem. We desire to minimize the distance traveled, i.e.

$$J[y] = \int_c^d \sqrt{1+(y')^2}dx, \tag{15.10}$$

subject to the inequality constraint

$$H(x, y(x)) \leq 0, \tag{15.11}$$

where equality $H(x, y(x)) = 0$ defines the boundary of Mirkwood, and $H(x, y(x)) < 0$ indicates that the path lies outside the forest. This is very similar to the Underminer staying in the crust of the earth when he travels, but in this case we are only going to have to worry about a single inequality constraint.

The requirement we are interested in is to not enter Mirkwood, i.e. $H(x, y(x)) \leq 0$, which we have split up into 2 parts, when $H(x, y(x)) = 0$ (Gimli is skirting the edge of the forest, carefully watching for strange happenings within) and when $H(x, y(x)) < 0$ (Gimli is headed to or away from the forest). In essence we want to incorporate this as two distinct options. To do so, we will use a Lagrange multiplier $\mu(x)$ for the case when $H(x, y(x)) = 0$, i.e. when Gimli is skirting the edge of the forest we want to minimize the modified cost functional

$$J^*[y] = \int_c^d \left[\sqrt{1+(y')^2} + \mu(x)H(x, y(x))\right] dx, \tag{15.12}$$

however when Gimli is not right on the edge of the forest we no longer want to enforce this restriction. Thus, we will add the additional condition that $\mu(x) = 0$ whenever $H(x, y(x)) < 0$ or $H(x, y(x)) \neq 0$. Thus, whenever Gimli is not on the edge of Mirkwood the constraint vanishes identically. This can be restated as

$$\text{minimize } J^*[y] = \int_c^d \left[\sqrt{1+(y')^2} + \mu(x)H(x, y(x))\right] dx, \tag{15.13}$$

$$\mu(x)H(x, y(x)) = 0, \tag{15.14}$$

where the final condition is referred to as the complementary slackness condition.

This is not quite enough however, because it also turns out that $\mu(x) \geq 0$ for all x (as is shown below). With these conditions in mind, Gimli is prepared to adequately avoid the Elven forest, and still arrive at the Lonely Mountain in a somewhat reasonable time.

To understand where the condition $\mu(x) \geq 0$ arises from, we will consider the derivation of the complementary slackness conditions in a rather formal setting. Notice that another way of enforcing the constraint $H(x, y(x)) \leq 0$ is to introduce the variable $v(x)$ so that we enforce the exact constraint

$$H(x, y(x)) + v(x)^2 = 0. \tag{15.15}$$

We can see immediately that when the inequality constraint is satisfied exactly, i.e. $H = 0$ then this occurs when $v(x) = 0$, and otherwise $v(x) \neq 0$ yielding the strict inequality constraint. If we let $\mu(x)$ be the Lagrange multiplier to enforce (15.15) then we seek to minimize the cost function given by:

$$\tilde{J}[y, \mu, v] = \int_a^b \left[\sqrt{1+(y')^2} + \mu(x)\left(H(x, y(x)) + v(x)^2\right)\right] dx. \tag{15.16}$$

The Euler-Lagrange equations for $y(x)$ are no different, and those for $\mu(x)$ recover the constraint exactly, but now we also have an (EL) for $v(x)$:

$$2\mu(x)v(x) = 0. \tag{15.17}$$

Recalling that $v(x) = 0$ implies that $H(x, y(x)) = 0$ and vice versa, we recognize (acknowledging that this is only possible because the right hand side of the previous equation is zero) that this condition is equivalent to

$$\mu(x)H(x, y(x)) = 0. \tag{15.18}$$

Now, we have still not said anything about the sign of $\mu(x)$ other than the complementary slackness condition. This comes formally by recognizing that if we are minimizing $\tilde{J}$ in $v(x)$ then the second variation must be positive definite (if this seems frightening don't worry, this is discussed in more detail in a later Chapter), and this leads to (varying $v(x) \to v(x) + \varepsilon\beta(x)$):

$$\delta^2 J[v; \beta] = \int_a^b \mu(x)\beta(x)^2 dx > 0 \Rightarrow \mu(x) \geq 0. \tag{15.19}$$

Remark 15.3.2. The previous discussion is not completely rigorous, partially because to make these arguments rigorous would require far more expertise in non-smooth analysis, and that is just not 'cool'. In addition, as we will see when considering 2nd order sufficient conditions for minima of an infinite dimensional problem (something to look forward to!), exciting things are highly likely to happen.

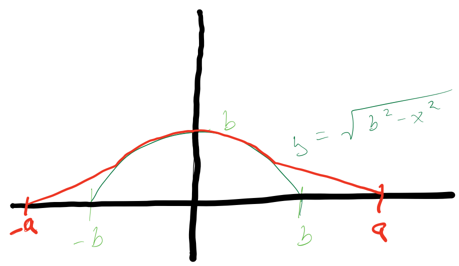

Figure 15.4: Optimal path for the simplified example of Gimli skirting the forest.

Summary of conditions for inequality constraints

In summary then, for the minimization of

$$J[y] = \int_a^b L(x, y, y')dx, \tag{15.20}$$

subject to both the equality and inequality constraints:

$$g(x, y) = 0 \qquad H(x, y) \leq 0, \tag{15.21}$$

the extremal will satisfy the Euler-Lagrange equations of the modified cost functional with modified Lagrangian given by

$$\mathscr{L} = L + \lambda g + \mu H, \tag{15.22}$$

i.e. the extremal $\tilde{y}(x)$ will satisfy

$$L_y + \lambda g_y + \mu H_y - \frac{d}{dx} L_{y'} = 0, \tag{15.23}$$

$$g(x, y) = 0, \quad H(x, y) \leq 0, \quad \mu(x)H(x, y) = 0, \quad \mu(x) \geq 0. \tag{15.24}$$

15.4 Actual computation of a solution using the KKT conditions

The problem here is that we have not quantified the point at which Gimli will reach the edge of Mirkwood, i.e. for what $\mathbf{x}^*$ does $\mu(\mathbf{x}^*) = 0$. To see how this might work we will consider a simplified version of the boundary of Mirkwood.

Example 15.4.1 (Continuation of Gimli's Mirkwood excursion). Suppose that the boundary of Mirkwood is given by $H(x, y)$ once again, but it is symmetric about $x = 0$ (the simplest example might be $H(x, y) = \sqrt{b^2 - x^2} - y$, i.e. a semi-circle centered at the origin...we will use this explicit example throughout when possible). Then we can simplify the problem to one of finding the shortest path from the initial point (which we will take to be $x = -a,\ y = 0$) to the edge of Mirkwood, and then to its peak at $(0, y(0))$ where $H(0, y(0) = 0$.

For the case where Mirkwood is modeled as a perfect semi-circle, this means we want to find the path from the initial point $(-a, 0)$ to the edge of the semi-circle which is located at the point $(x^*, y(x^*))$, where $(x^*)^2 + y(x^*)^2 = b^2$. Then we know that Gimli will follow the semi-circle exactly until he gets to the 'peak' at $(0, b)$. Positive values of x will be exactly symmetric, i.e. he will follow the semi-circle down to $-x^*$ and follow a symmetric curve to the Lonely Mountain located at $(a, 0)$.

In the general setting, this becomes the following minimization problem (employing the inequality constraint):

$$\min J[y] = \int_{-a}^{0} \left[\sqrt{1 + (y')^2} + \mu(x)\left(\sqrt{b^2 - x^2} - y\right)\right] dx,$$

subject to:

$$H(x, y) = \sqrt{b^2 - x^2} - y \leq 0, \quad \mu(x)H(x, y) = 0, \quad \mu(x) \geq 0.$$

We can rewrite this minimization problem as the following (using $L(x, y, y') = \sqrt{1 + (y')^2}$):

$$J[y] = \int_{-a}^{x^*} L(x, y, y')dx + \int_{x^*}^{0} \left[L(x, y, y') + \mu(x)H(x, y)\right] dx.$$

In other words, $[-a, x^*]$ is the interval on which $H(x, y(x)) < 0$ so that $\mu(x) = 0$ and on $[x^*, 0]$ then $H(x, y(x)) = 0$ so that we want $\mu(x) > 0$ there. Now with the assumption of no singular points on the edge of $H(x, y(x)) = 0$ (meaning there are no points where $\frac{\partial H}{\partial y} = 0$) then we can say that this is satisfied by the sufficiently smooth (C^1) curve $g(x)$, that is $H(x, g(x)) = 0$ defines the curve that we are aiming for. In a full general setting, there is no guarantee that such a function exists, but it certainly makes the notation of the following derivation much simpler to write out. For the semi-circular renormalization of Mirkwood then $g(x) = \sqrt{b^2 - x^2}$.

Now with this in mind we consider the total cost of Gimli's brave foray as

$$J[y] = \int_{-a}^{x^*} L(x, y, y')dx + \int_{x^*}^{0} \left[L(x, y, y') + \mu(x)H(x, y(x)\right] dx$$

Before proceeding however we note that variations of the second integral are already known, i.e. we are forced to follow the curve $y = g(x)$ and hence we can omit the multipler and rewrite the cost functional as

$$J[u] = \int_{-a}^{x^*} L(x, y, y')dx + \int_{x^*}^{0} L(x, g(x), g'(x))dx,$$

where now the second integral only varies by changes in x^*, i.e. we know that the equality constraint $y(x) = g(x)$ is enforced exactly on this interval. With this in mind, this cost functional has first variation (once again describing variations in x^* as $x^* \to x^* + f(\varepsilon)$ where $f(\varepsilon = 0) = 0$):

$$\begin{aligned}\delta J &= \frac{\partial}{\partial \varepsilon}\left\{\int_{-a}^{x^*+f(\varepsilon)} L(x, y+\varepsilon h, y'+\varepsilon h')dx + \int_{x^*+f(\varepsilon)}^{0} L(x, g(x), g'(x))dx\right\}\Bigg|_{\varepsilon=0} \\ &= f'(\varepsilon)\left[L(x^*, y(x^*), y'(x^*)) - L(x^*, g(x^*), g'(x^*))\right] + h(x^*)L_{y'}(x^*, y(x^*), y'(x^*)) \\ &\quad + \int_{-a}^{x^*}\left[L_y - \frac{d}{dx}L_{y'}\right]dx = 0.\end{aligned}$$

Because this must be satisfied for all potential functions $h(x)$ and variations in x^* prescribed by $f(\varepsilon)$ including those for which $h(x^*) = f'(\varepsilon) = 0$ then the Euler-Lagrange equations must be satisfied on the interval $[-a, x^*]$. Hence, as we already expected the optimal path for this segment is going to be a straight line to the edge of Mirkwood...the question we must answer though is what the value of x^* is.

To treat the boundary terms at $x = x^*$ we need to revisit our previous derivation of the transversality condition. To do this we note that continuity of Gimli's path guarantees that as the optimal path is varied, Gimli still much reach the edge of Mirkwood, i.e.

$$y(x^* + f(\varepsilon)) + \varepsilon h(x^* + f(\varepsilon)) = g(x^* + f(\varepsilon)).$$

Applying a Taylor series about $\varepsilon = 0$ to this relation and recalling that $y(x^*) = g(x^*)$, we find that

$$\begin{aligned}&y(x^*) + \varepsilon f'(\varepsilon)y'(x^*) + \varepsilon h(x^*) = g(x^*) + \varepsilon f'(\varepsilon)g'(x^*) + O(\varepsilon^2),\\ &\Rightarrow f'(\varepsilon)y'(x^*) + h(x^*) = f'(\varepsilon)g'(x^*)\\ &\Rightarrow h(x^*) = -f'(\varepsilon)(y'(x^*) - g'(x^*)).\end{aligned}$$

Using this final line in the boundary terms generated above, we arrive at our final transversality condition for this problem:

$$\begin{aligned}&f'(\varepsilon)\left[L(x^*, y(x^*), y'(x^*)) - L(x^*, g(x^*), g'(x^*)) - (y'(x^*) - g'(x^*)L_{y'}(x^*, y(x^*), y'(x^*))\right] = 0\\ &\Rightarrow L(x^*, y(x^*), y'(x^*)) - L(x^*, g(x^*), g'(x^*)) - (y'(x^*) - g'(x^*)L_{y'}(x^*, y(x^*), y'(x^*)) = 0.\end{aligned}$$

This is not a very pleasant exercise to enforce however, even noting that $g(x^*) = y(x^*)$ is enough to simplify things. We can make life a bit easier however by observing that the difference in the first two terms can be rewritten as

$$L(x^*, g(x^*), y'(x^*)) - L(x^*, g(x^*), g'(x^*)) = (y'(x^*) - g'(x^*))L_{y'}(x^*, y(x^*), c),$$

where c lies between $g'(x^*)$ and $y'(x^*)$ (Intermediate Value Theorem). Hence the full transversality condition can be rewritten as

$$(y'(x^*) - g'(x^*))\left(L_{y'}(x^*, y(x^*), y'(x^*)) - L_{y'}(x^*, y(x^*), c)\right) = 0,$$

implying that either $L_{y'}$ is constant in y' or $y'(x^*) = g'(x^*)$. Gimli's skirting around Mirkwood is the latter case, as demonstrated in the homework exercises.

Remark 15.4.2. This section only contains a single example for inequality constraints, but the reader should be able to easily guess why. Such problems are notoriously difficult to solve by hand, and most often result in solutions that should not be restricted to the pages of a textbook. Thus, unfortunately we relegate ourselves to this single example for the implementation of inequality constraints in infinite dimensions.

Should this sufficiently bother the reader so they are compelled to create a more accessible example, at least one of the authors of these notes would be delighted to include such useful information in later editions.

Exercises

Note to the student: Each section of this chapter has several corresponding exercises, all collected here at the end of the chapter. The exercises between the first and second line are for Section 1, the exercises between the second and third lines are for Section 2, and so forth.

You should **work every exercise** (your instructor may choose to let you skip some of the advanced exercises marked with *). We have carefully selected them, and each is important for your ability to understand subsequent material. Many of the examples and results proved in the exercises are used again later in the text. Exercises marked with ⚠ are especially important and are likely to be used later in this book and beyond. Those marked with † are harder than average, but should still be done.

Although they are gathered together at the end of the chapter, we strongly recommend you do the exercises for each section as soon as you have completed the section, rather than saving them until you have finished the entire chapter.

15.1. Write down the Euler-Lagrange equations that give a necessary condition for a minimizer of

$$J[y] = \int_a^b \left(p(x)y'^2 + q(x)y^2\right) dx$$

subject to the integral constraint

$$I[y] = \int_a^b r(x)y^2 dx = 1,$$

where $p, q, r \in C^2[a, b]$ and $y(a) = y(b) = 0$.

15.2. Find an extremal of

$$J[y] = \int_0^1 (y'^2 + y^2)dx, \quad y(0) = 0,\ y(1) = 0,$$

subject to the constraint

$$I[y] = \int_0^1 (y + y')dx = 1.$$

The following 2 questions deal with this problem:

Consider a cable suspended between two points (power lines suspended between two poles are one example). From experience we know that if the cable is longer than the horizontal distance between the two points, it will sag. If we specify the length of the cable as L, and note that the system will naturally minimize the potential energy which for a length ds of cable is given by $\rho g y ds$ where $y(x)$ is the height of the cable and ρ is the (constant) linear density of the cable and g is the acceleration due to gravity.

15.3. Set up the constrained variational problem and find the corresponding Euler-Lagrange equation.

15.4. Show that

$$y(x) = \frac{c_0}{\rho g} \cosh\left(\frac{\rho g}{c_0}(x - c_1)\right) - \frac{\lambda}{\rho g}$$

is a solution of the resulting PDE where λ is the Lagrange multiplier and c_0 c_1 are constants to be specified by the boundary conditions.

15.5. Complete Example 15.1.5 by deriving the corresponding (EL), and finding the optimal solution $u(x)$.

15.6. The critical load for a vertical beam is given by

$$\min_u \frac{EI \int_0^l (u''(z))^2 dx}{\int_0^l (u'(z))^2 dz},$$

where E is the modulus of elasticity and I is the surface moment of inertia (EI is called the bending stiffness and is a constant in this example), and $u(z)$ represents the shape of the beam. If the beam is fixed at the two endpoints $u(0) = u(l) = 0$ then what is the critical load of the beam (you will need natural boundary conditions on the second derivative)? *(HINT: This is very similar to Example 15.1.5).*

15.7. Repeat the previous problem, but now when the upper end point is not fixed, i.e. $u(0) = u'(l) = 0$.

15.8. Set up the problem to determine the shortest path from two different points on a volcano, i.e. find the shortest path between two (x, y, z) points on the conical surface $x^2 + y^2 = (1 - z)^2$. Use cylindrical coordinates (r, θ, z).

15.9. *Find the shortest path from the points $(0, -1, 0)$ to $(0, 1, 0)$ on the conical surface $x^2 + y^2 = (1 - z)^2$ as described in the previous problem. Note that in this setting $dz = -dr$ in cylindrical coordinates, and use the change of variables for the angle $\beta = \frac{\theta}{\sqrt{2}}$.

15.10. Set up the cost function to find the shortest path from one point to another on a sphere using a finite constraint that forces the Arctic tern to stay on the sphere (you are neglecting vertical changes in the flight path). Determine the Euler-Lagrange equations for this problem as well.

15.11. Although there are many other examples of using non-integral constraints, we will return to the shortest path on a weird surface. Suppose that Dr. Strange is stuck on the hyperboloid given by $z = \frac{y^2}{9} - \frac{x^2}{4}$. He is at the point $(4, 9, 5)$, and needs to reach the origin $(0, 0, 0)$ before Kaecilius. What is the fastest route he can take to get there? Set up the variational problem with relevant constraints for this problem.

15.12. For the previous problem, determine the Euler-Lagrange equations (solve if you dare).

15.13. Now for the same problem as the previous two, suppose that even though reality has been bent to a hyperboloid, Dr. Strange gets disoriented every time he changes z values, i.e. he wants to avoid changing z any more than he must. Incorporate this into the cost, and re-derive the corresponding Euler-Lagrange equations. Is this going to change the optimal path, and if so will it change it by a lot?

15.14. Set up a final version of the problem for the Underminer to travel within the Earth's crust from one point to another. This should include the cost function, with appropriate inequality constraints. Then find the Euler-Lagrange equations and other conditions that must be satisfied by the optimal path.

15.15. Modify the previous problem to account for the fact that the Underminer can get very close to the mantle $x^2 + y^2 + z^2 \approx r^2$, but would prefer not to get to close. How does this modify the resultant (EL) and other conditions?

15.16. Now suppose that the Underminer is trying to get from one point on a perfectly conal volcano to another. Using some of the other exercises above, set up a problem that would allow the Underminer to tunnel through the volcano to get from one point on its surface to another. Set up the cost functional for this problem and find the corresponding Euler-Lagrange equations as well as all other pertinent conditions.

15.17. Complete Example 15.4.1 by first showing that for this setting, $L_{y'}$ is not constant in its third argument (y').

15.18. Now, if the Old Ford is indeed located at $x = -a,\ y = 0$ and Mirkwood is defined as the interior of the perfect circle $x^2+y^2 = b^2$, and the Lonely Mountain is located at $x = a, y = 0$, then find the optimal point (x^*, y^*) where Gimli will meet the edge of the forest. Please do this for generic values of $a > 0$ and $b > 0$. *HINT: You will likely get a rational value of x^* if you use $a = 2$ and $b = \sqrt{\frac{3}{2}}$. (Don't despair if this isn't the case...these hints aren't guaranteed after all).*

15.19. Now Suppose that unbeknown to Gimli, the Ents have moved the boundaries of Mirkwood so that it is now best described by everything lying below the parabola $y = -x^2 + 5$. Find the optimal path for Gimli to travel from $(-3, 0)$ to $(3, 0)$ while avoiding Mirkwood. In this case you should get a nice integer value for the point x^* where Gimli intersects the boundary of Mirkwood. You will get 2 possible answers for x^* but only one of them actually makes sense.

Notes