13 The Simplest Problem

Simple can be harder than complex: You have to work hard to get your thinking clean to make it simple. But it's worth it in the end because once you get there, you can move mountains.
—Steve Jobs

13.1 Derivation of Euler's Equation

Suppose we want to optimize the functional

$$J[y] = \int_a^b L(x, y, y')dx \tag{13.1}$$

where $y' = \frac{dy}{dx}$ and the endpoints are fixed, i.e. we are optimizing over $y(x)$ such that $y(a) = c$ and $y(b) = d$ where $c, d \in \mathbb{R}$. We will also suppose that $L(x, y, y')$ is continuous and differentiable with respect to y and y'. This defines what is usually referred to as the 'simplest problem'.

Definition 13.1.1. *A local extrema y_0 is called a 'weak extremum' of $J[y]$ if there exists $\varepsilon > 0$ such that $J[y] - J[y_0]$ has the same sign for all $y \in B(y_0, \varepsilon)$, where the norm defining the ball $B(y_0, \varepsilon)$ is that induced by $C^1[a, b]$.*

Remark 13.1.2. A 'strong extremum' uses the norm induced by $C[a, b]$, meaning that the functional evaluation of the extremum $J[y_0]$ is smaller (larger) than the functional evaluation $J[y]$ for all functions $y(x)$ that are 'close' to $y_0(x)$ in the $C[a, b]$ norm. This means that $y_0(x)$ is an extrema relative to continuous functions $y(x)$ whose values are close to $y_0(x)$.

On the other hand, weak extrema must be the minima/maxima not only in the function itself, but also in the first derivative of the function. Hence a strong minima/maxima is a stronger concept because it indicates that the extrema is optimal over a larger class of potential functions whereas weak extrema satisfy a much stricter criterion, i.e. weak extrema are extremal relative to all functions whose functional values *AND* their derivatives are 'close' in the $C^1[a, b]$ norm. See Figures 13.1 and 13.2 for an illustration of these differences.

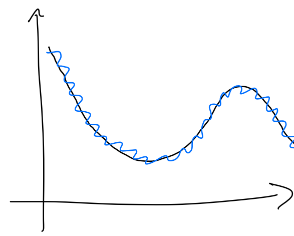

Figure 13.1: An illustration of 'closeness' of two functions in the strong sense. The blue and black curves illustrated here are within ε for all values of the independent variable even though their functional derivatives are vastly different. Hence the functions these two curves represent are close in the strong sense, but not in the weak sense.

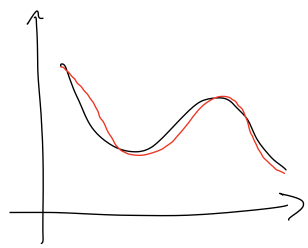

Figure 13.2: An illustration of 'closeness' of two functions in the weak sense. The red and black curves represent two functions that are close in the weak sense. Note that not only does the function need to nearly agree, but extreme differences in the derivative of the function are not allowed either.

To find a weak extremum of $J[y]$ we need to calculate its Gateaux differential (1st variation, ideally we would prefer calculating the Fréchet differential but that is not very practical). We consider $h \in C^1[a,b]$ such that $h(a) = h(b) = 0$ so that $y(x) + \varepsilon h(x)$ still satisfies the correct boundary conditions, i.e. $y(a) + \varepsilon h(a) = A$ and $y(b) + \varepsilon h(b) = B$. It follows that

$$\begin{aligned}\delta J[y;h] &= \frac{d}{d\varepsilon} J[y+\varepsilon h]\bigg|_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon}\int_a^b L(x, y+\varepsilon h, y'+\varepsilon h')dx\bigg|_{\varepsilon=0} \\ &= \int_a^b \left[L_y(x,y,y')h(x) + L_{y'}(x,y,y')h'(x)\right] dx\end{aligned}$$

For an extrema, we want $\delta J[y;h] = 0$ for all $h(x) \in C^1[a,b]$ with $h(a) = h(b) = 0$. Now, if we formally suppose that $L_{y'}(x,y,y')$ is continuously differentiable with respect to x, then we can integrate by parts to arrive at:

$$\delta J[y;h] = \int_a^b \left[L_y(x,y,y')h(x) - \frac{d}{dx}L_{y'}(x,y,y')h(x)\right] dx + L_{y'}(x,y,y')h(x)|_a^b.$$

Since $h(a) = h(b) = 0$ we see that $L_{y'}(x, y, y')h(x)|_a^b = 0$. Now, for $y(x)$ to be an extrema of $J[y]$, then we need to have $\delta J[y; h] = 0$ for all $h(x)$. Thus

$$\int_a^b \left[L_y(x, y, y') - \frac{d}{dx} L_{y'}(x, y, y') \right] h(x) dx = 0$$

for all $h(x) \in C^1[a, b]$, which at first glance isn't the most useful statement. This states that this integral must vanish which isn't all that helpful. The key is that this integral must vanish for all admissible $h(x)$ which does give us some information as shown below. We must make use of the following Lemma to come up with a computable equation for these extrema.

Lemma 13.1.3. *If $y \in C^1[a, b]$ and $\int_a^b \phi(x)h(x)dx = 0$ for every $h \in C^1[a, b]$ with $h(a) = h(b) = 0$. Then $\phi(x) = 0$ for all $x \in [a, b]$*

Proof. Assume by way of contradiction (and without loss of generality) that $\phi(\tilde{x}) > 0$ for some $\tilde{x} \in [a, b]$. It follows that because ϕ is continuous, there must exist a' and b' so that $\phi(x) > 0$ for $x \in [a', b'] \subset [a, b]$. Let

$$h(x) = \begin{cases} (x - a')^2(x - b')^2 & : x \in [a', b'] \\ 0 & : \text{otherwise} \end{cases}$$

Then $h \in C^1[a, b]$ and $h(a) = h(b) = 0$. However

$$\int_a^b \phi(x)h(x)dx > 0,$$

which is a contradiction, and thus $\phi(\tilde{x}) = 0$ for all $\tilde{x} \in [a, b]$. □

Remark 13.1.4. Note that we could have formalized this statement by considering the ramifications of a measure theoretic integral, i.e. the function $\phi(x)$ will vanish everywhere except on a set of measure zero. We avoid such complications here, and consider instead only the traditional Riemannian integral and consider continuously differentiable functions that are defined at all points in their domain, ignoring the effects of a possibly zero measurable set.

This indicates that extrema of (13.1) with fixed endpoints will satisfy the differential equation

$$L_y(x, y, y') - \frac{d}{dx} L_{y'}(x, y, y') = 0. \tag{13.2}$$

This is referred to as Euler's equation, and provides a method of solution to the simplest problem, i.e. the extrema is a solution to the inherently 2nd order differential equation (13.2).

The use of integration by parts in the derivation above is of course critically dependent on the differentiability of the functions themselves, i.e. we are assuming that we seek extrema in the weak sense. Extensions of this approach to strong solutions is feasible as illustrated below, but we do not overly focus on these extensions in the following Chapters of this text. In addition to making this formal derivation rigorous, consideration of strong extrema also leads to additional considerations near points where the function is allowed to be non-differentiable. Such points are often referred to as 'corner points' and specific conditions are needed to guarantee that the resultant function is still an appropriate extrema of the cost functional, just not in the weak sense.

13.1.1 *Rigorous justification without integration by parts

The calculations above relied on the single integration by parts that assumed the differentiability of the Lagrangian $L(x, y, y')$ with respect to x and y'. This is not necessary, as illustrated by the following Lemmata.

Lemma 13.1.5. *Let $y(x) \in C[a, b]$ and*

$$\int_a^b y(x)h'(x)dx = 0,$$

for all $h(x) \in C^1[a, b]$ such that $h(a) = h(b) = 0$. Then $y(x) =$ constant.

Proof. Let c be the constant defined by

$$c = \frac{1}{b-a}\int_a^b y(x)dx,$$

and

$$h(x) = \int_a^x [y(\xi) - c]d\xi,$$

implying that $h'(x) = y(x) - c$. Note that $h \in C^1[a, b]$ and $h(a) = h(b) = 0$. Now

$$\int_a^b [y(x) - c]h'(x)dx = \int_a^b y(x)h'(x)dx - c[h(b) - h(a)] = 0,$$

because

$$\int_a^b y(x)h'(x)dx = 0$$

and $c[h(b) - h(a)] = 0$, but

$$0 = \int_a^b [y(x) - c]h'(x)dx = \int_a^b [y(x) - c]^2 dx$$

and thus

$$y(x) - c = 0,$$

so that $y(x) = c$. □

This combined with the following Lemma (which is left as an exercise) provide an avenue for showing that (13.2) is satisfied by all extrema of (13.1) even without the supposition that the Lagrangian is differentiable with respect to y' and x.

Lemma 13.1.6. *If $y(x)$ and $u(x) \in C[a, b]$ and*

$$\int_a^b [y(x)h(x) + u(x)h'(x)]dx = 0$$

for all $h(x) \in C^1[a, b]$ with $h(a) = h(b) = 0$, then $u(x)$ is differentiable and $u'(x) = y(x)$.

Proof. Left as an exercise □

Remark 13.1.7. Note that Euler's equation (sometimes referred to as the Euler-Lagrange equation) is a necessary but not sufficient condition for an extrema of (13.1).

13.1.2 A 'simple example'

Example 13.1.8 (Shortest path on the plane). As shown in the previous chapter, the distance between two points given by (a, c) and (b, d) can be calculated along a curve defined by the function $y(x)$ connecting these two points as

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

where $L(x, y, y') = \sqrt{1 + (y')^2}$. Then $L_y - \frac{d}{dx}L_{y'} = 0$ implies that the path with the shortest distance will satisfy

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

where $y(a) = c$ and $y(b) = d$. This implies that $\frac{y'}{\sqrt{1+(y')^2}} = c_0$, a constant. Then

$$\frac{(y')^2}{1 + (y')^2} = c_0^2,$$

which can be solved to see that $(y')^2 = \frac{c_0^2}{1-c_0^2} = c_1^2$ so $y' = c_1$ and $y(x) = c_1 x + c_2$ where c_1 and c_2 can be determined by $y(a) = c$ and $y(b) = d$ i.e. the shortest path between two points in the plane is a line! Who would have thought of such a thing?

Remark 13.1.9. Note that because the Euler equation is only a necessary condition, we should re-evaluate whether this will really give us the minimal path connecting the two points in question (other than the obvious fact that we knew the answer before we started). Returning to the statement of the problem, we note that if we do find an extremizer it must necessarily be a minimizer, because it would be quite easy to construct a continuously differentiable curve $y(x)$ connecting the two points, that was as long as desirable, i.e. there is no maximum to the functional $J[y]$ in this case, so whatever we do find must be the minimizer.

Such is typically the case. Although there are conditions we can place on the extremizer (similar to the 2nd derivative test you learned in Multivariable Calculus), they are typically very difficult to use in practice, and usually if the variational problem is set up properly then there can be only one type of extrema, i.e. minima or maxima, and the physical considerations of the problem will indicate which one holds.

In particular, as you have already seen for optimization problems in finite dimensions, it is always a good idea to set up your cost functional to be globally convex. This will eliminate the need to consider whether you find a minimizer (global or not). Thus if there are options when setting up the problem, it is always best to choose a convex cost function.

13.2 More examples

In this section we will illustrate a few examples that indicate the true applicability and range of potential applications for the simplest problem. This is certainly not a comprehensive set of examples, but should imply the utility of these methods to a variety of different problems.

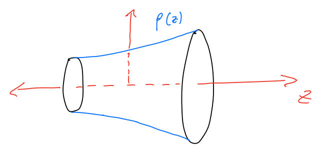

Figure 13.3: The minimal surface of Example 13.2.1. The z axis is taken along the axis of rotation of the two concentric rings with the origin at the central point between the two rings. The minimal surface is a surface (illustrated in blue here) that connects the two rings.

Example 13.2.1 (Minimal Surfaces). Consider two concentric rings with a thin soap film suspended between them, separated by a distance L. The smaller ring has radius ρ_A and the larger ρ_B. We choose cylindrical coordinates to describe the system, with the z-axis oriented along the line of symmetry and $z = -L/2$ at the center of the smaller ring and $z = L/2$ is at the center of the larger ring, i.e. the z-axis runs through the center of the two concentric rings. The shape of the soap film can be completely described by $\rho(z)$ for $z \in [-L/2, L/2]$. This is illustrated in Figure 13.3.

The dominant force that maintains the shape of the soap film is surface tension. Because the surface tension acts on the surface of the film, the soap film will naturally adjust to minimize its surface area. Because the soap film is radially symmetric its geometric description is independent of the angle ϕ, and hence the surface area for this soap film is given by

$$S = \int_0^{2\pi} \int \rho(z) dl d\phi,$$

where dl refers to the infinitesimal length of the soap film in a direction orthogonal to $d\phi$. Using cylindrical coordinates we find that $dl = (dz^2 + d\rho^2)^{\frac{1}{2}} = \sqrt{1 + \left(\frac{d\rho}{dz}\right)^2} dz$. Hence integrating over the angle ϕ, the soap film is minimizing

$$S[\rho] = 2\pi \int_{\frac{-L}{2}}^{\frac{L}{2}} \rho \sqrt{1 + \left(\frac{d\rho}{dz}\right)^2} dz,$$

where $\rho(\frac{-L}{2}) = \rho_A$ and $\rho(\frac{L}{2}) = \rho_B$. The Euler-Lagrange equations become something interesting that yield solutions of the form $\rho(z) = c_1 \cosh\left(\frac{z+c_2}{c_1}\right)$ where c_1 and c_2 are determined by the boundary conditions. This solution is called a ‘catenary’ and the surface it represents is called a ‘catenoid’

Remark 13.2.2. Minimal surfaces appear in a variety of different engineering applications and can be used as fascinating demonstrations of geometric structures (think of soap bubbles). They also appear in such common circumstances as the shape of a pringle chip, to the design of several canopies designed to minimize the wind resistance and hence potential turbulence generated under the canopy. Note that the generic description of a minimal surface is universal, i.e. the minimization of the surface area, but each particular situation introduces a unique set of constraints that modify the resulting surface. For instance this particular example included cylindrical symmetry and boundary conditions on $\rho(z)$. Other settings may have drastically different symmetry assumptions (or a lack thereof) and different boundary conditions that will modify the actual computed minimal surface for that setting.

Once again, there is clearly only a minimizer to this physically stated problem, it would be relatively straightforward to construct a surface connecting the two rings that folds in on itself frequently enough that its surface area is as large as we desire it to be.

Example 13.2.3 (Finance Problem (Retirement Planning)). Suppose at the onset of retirment, an individual decides that S amount of money is to be used over a period of time from $t = 0$ up to $t = t_f$. The enjoyment (utility) at a given time t is described by $u(r(t)) = 2\sqrt{r(t)}$ where $r(t)$ is the rate at which the money is spent. Noting that future utility is discounted (you always want to spend your money now more than you want to save it, at least if you have been raised in the 21st century) we want to maximize

$$J[r(t)] = \int_0^{t_f} e^{-\beta t} u[r(t)]dt,$$

where $e^{-\beta t}$ indicates that current utility is more desirable than at a later time. Starting with S money and ending with S_0 at time $t = t_f$. The only source of income is from investment (interest) so the total capital at time t, given by $x(t)$, is described by $x' = \alpha x(t) - r(t)$, where α is the interest rate. Thus we rewrite our cost functional as

$$J[x(t)] = \int_0^{t_f} e^{-\beta t} u(\alpha x(t) - x'(t))dx = 2\int_0^{t_f} e^{-\beta t}\sqrt{\alpha x(t) - x'(t)}dt,$$

where $x(0) = S$ and $x(t_f) = S_0$. We also must have $x(t) \geq 0$ for all t, but we see that this condition is implicitly maintained in the solution. The Euler-Lagrange equation is then given by noting that in this case the Lagrangian (term inside the integral) is $L = e^{-\beta t}\sqrt{\alpha x(t) - x'(t)}$ so that

$$\frac{\partial}{\partial x}\left(e^{-\beta t}\sqrt{\alpha x - x'}\right) - \frac{d}{dt}\frac{\partial}{\partial x'}\left(e^{-\beta t}\sqrt{\alpha x - x'}\right) = 0.$$

This can be simplified to:

$$\frac{\alpha e^{-\beta t}}{\sqrt{\alpha x(t) - x'(t)}} + \frac{d}{dt}\frac{e^{-\beta t}}{\sqrt{\alpha x(t) - x'(t)}} = 0.$$

Applying the chain rule to the second term this can be rewritten as:

$$\frac{(\alpha - \beta)e^{-\beta t}}{\sqrt{\alpha x(t) - x'(t)}} + e^{-\beta t}\frac{d}{dt}\frac{1}{\sqrt{\alpha x(t) - x'(t)}} = 0.$$

This implies that

$$\frac{d}{dt}\frac{1}{\sqrt{\alpha x(t) - x'(t)}} = \frac{(\beta-\alpha)}{\sqrt{\alpha x(t) - x'(t)}}. \tag{13.3}$$

Thus, letting

$$v(t) = (\alpha x(t) - x'(t))^{-\frac{1}{2}}$$

then we have shown that

$$\frac{d}{dt}v(t) = (\beta-\alpha)v(t),$$

so $v(t) = v(0)e^{(\beta-\alpha)t}$ and thus

$$\frac{1}{\sqrt{\alpha x(t) - x'(t)}} = \frac{e^{(\beta-\alpha)t}}{\sqrt{\alpha x(0) - x'(0)}}.$$

This can be solved to find that (recall that $x(0) = S$):

$$x'(t) = \alpha x(t) - (\alpha S - x'(0))e^{2(\alpha-\beta)t}. \tag{13.4}$$

This is a linear ODE that can be integrated (in the exercises, what lucky souls you are!) to find that

$$x(t) = \left(S - \frac{\alpha S - x'(0)}{2\beta-\alpha}\right)e^{\alpha t} + \frac{\alpha S - x'(0)}{2\beta-\alpha}e^{2(\alpha-\beta)t}.$$

The initial rate of change of the capital, $x'(0)$ can be found by imposing $x(t_f) = S_0$ For this solution, if $\alpha > \beta > \frac{\alpha}{2}$ then $x(t)$ will initially increase, followed by decreasing to $S_0 < S$. The final solution maintains $x(t) \geq 0$ which is rather convenient, although in today's economy of stacked mortgages and exorbitant credit card debt, it may not be completely necessary.

Remark 13.2.4. The previous example is most certainly overly simplified. In particular, human behavior is intrinsically best modeled with a stochastic or noisy influence, and the assumptions used above on the interest rates etc. are extremely simplistic. If world markets and retirement planning really functioned on such a level, then we wouldn't need financial advisors or degrees in economics. The surprising fact is that even these overly simplistic perspectives can lead to incredibly insightful observations that do correlate with qualitative observations in reality.

The next example is perhaps the most classical example (even more than walking in a straight line) and very possibly the very motivation for the creation of the Calculus of Variations in the first place.

Example 13.2.5 (Brachistochrone). There is a lot of history to this example going all the way back to Newton and the Bernoulli brothers which we will not review here (Google it and you will find a plethora of conflicting resources of course). Essentially the idea is that suppose you have a ball of a given mass that can frictionlessly roll down a path that you prescribe (and then build) from a prescribed height. Your goal is given an initial and final position (see Figure 13.4 to construct the path that gets to the final position as fast as possible where the only force acting on the ball is the acceleration due to gravity.

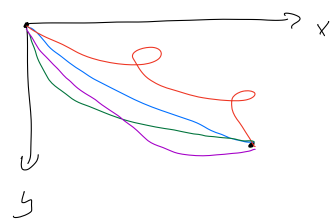

Figure 13.4: Setup for Example 13.2.5. The goal is to construct a path $y(x)$ that will get from the origin to the final point at the bottom right in the minimal amount of time under the action of gravity alone.

This simply reduces to minimizing the total time i.e. minimizing

$$\int dt.$$

We note that for a given velocity v then $dt = \frac{dl}{v}$ where dl represents the infinitesimal length of the curve at a given point in time (recall that $dl = \sqrt{1 + y'(x)^2}dy$ for a curve $y = y(x)$ in the plane). Conservation of energy (this isn't a Physics textbook, so we won't go into details here) enforces that

$$E = \frac{m}{2}v^2 - mgy,$$

is constant where m is the mass and g is the constant gravitational acceleration and we are taking y to be positive in the downward direction in Figure 13.4. Thus, we can replace v with

$$v = \sqrt{2E/m + 2gy}.$$

Further assuming that the ball starts from rest i.e. $E = 0$ then we see that we are hoping to minimize

$$J[y] = \int_0^a \frac{\sqrt{1 + (y')^2}}{\sqrt{2gy}} dx,$$

where $y(0) = 0$ and $y(a) = b$ is the final endpoint.

This rather quickly leads to the Euler-Lagrange equations

$$\frac{\sqrt{1 + (y')^2}}{2y^{3/2}} + \frac{d}{dx} \frac{y'}{2\sqrt{y}\sqrt{1 + (y')^2}} = 0.$$

The solution of this equation describes what is commonly referred to as the brachistochrone, and can best be described with x as a function of y i.e.

$$x(y) = -\sqrt{2cy - y^2} + c\cos^{-1}\left(\frac{c - y}{c}\right),$$

where c is a constant that can be specified based on the endpoint condition $x(b) = a$.

13.2.1 Special Cases

For a cost functional given by

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

the Euler-Lagrange (EL) equations are

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

If L does not depend on all three variables (x, y, y') explicitly there are some simplifications you can make

(i) $L = L(x, y)$ implies (EL) $L_y(x, y) = 0$

(ii) $L = L(x, y')$ implies (EL) $L_{y'}(x, y') = \text{constant}$

(iii) $L = L(y, y')$ implies (EL) $L(y, y') - y' L_{y'}(y, y') = \text{constant}$

These simplifications are called first integrals of the problem.

Remark 13.2.6. Finding geodesics on a given surface usually falls under the second case, i.e. typically the distance along the path is independent of the curve itself, and only depends on its derivative in space.

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.

13.1. *Let $y(x),\ u(x) \in C[a, b]$ and let

$$\int_a^b \left[y(x)h(x) + u(x)h'(x)\right] dx = 0,$$

for all $h \in C^1[a, b]$ such that $h(a) = h(b) = 0$. Prove that $u(x)$ is differentiable and $u'(x) = y(x)$.

For the following 4 problems, we are considering the flight path of an airplane traveling from point A to point B on a *flat earth*, both of which are at the same elevation (taken to be zero in this example) and separated by a distance d. Due to the density of the airplane relative to the atmosphere it is cheaper for the plane to fly at higher altitude than lower. We wish to minimize the cost of the specified flight if the cost of traveling a distance ds at an altitude h is constant and given by $e^{-h/H} ds$ where H is a reference altitude.

13.2. Choose a suitable coordinate system for this problem and sketch a picture of the situation.

13.3. Find an expression that describes the cost of the flight and express this as a variational problem.

13.4. Derive the Euler-Lagrange equations associated with the variational statement of the previous problem.

13.5. In reality the cost of ascending is more than descending. How might this be incorporated into the cost functional and hence the variational problem? Come up with a reasonable cost function for this situation (reasonable just means that it costs more as the plane ascends, and less as the plane descends).

13.6. Derive the Euler-Lagrange equation for the minimizer of the surface area of a soap film as discussed in class, i.e. for the functional

$$S[\rho] = 2\pi \int_{-L/2}^{L/2} \rho\sqrt{1+\left(\frac{d\rho}{dz}\right)^2}\,dz.$$

13.7. Show that $\rho(z) = C_1 \cosh\left(\frac{z+C_2}{C_1}\right)$ is a solution to the Euler-Lagrange equation derived in the previous problem.

13.8. Finishing up the example we started the first week, determine the geodesics on the sphere, i.e. quickest route between two points by finding the Euler-Lagrange equations to minimize the functional

$$S[\phi(\theta)] = \rho \int_{\theta_1}^{\theta_2} \left[\left(\frac{d\phi}{d\theta}\right)^2 + \sin^2\phi\right]^{1/2} d\theta.$$

13.9. *Solve the Euler-Lagrange equation for the previous problem to verify that great circles (feel free to google this if you feel the need) are the optimal paths on the sphere. This will take some serious trigonometric and spherical coordinate identities. In addition it may be useful to recall that $\frac{d\phi}{d\theta} = \left(\frac{d\theta}{d\phi}\right)^{-1}$. Feel free to ask Wolfram for help on this one.

13.10. *Solve the ODE given by (13.4) to obtain the solution described in the Example.

13.11. *Finish Example 13.2.3 by enforcing the final condition $x(T) = S_0$ (this is going to be rather messy, but it is fine if you just solve for $x'(0)$).

13.12. Verify that for the cost functional $J[y] = \int_a^b L(y, y')dx$ (there is no explicit dependence on the independent variable x) that the (EL) reduce to:

$$L(y, y') - y' L_{y'}(y, y') = \textit{constant}.$$

13.13. Fermat's Principle states that the path which light takes to move from one point to another is the one for which the time given by

$$T[y] = \int_a^b n\frac{ds}{c},$$

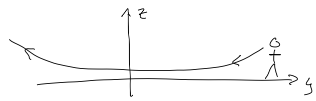

Figure 13.5: The path of light due to refraction on a hot summer day over an asphalt surface.

is a minimum. In this instance, ds is the differential path length, c is the speed of light in a vacuum, and n is the index of refraction. Consider an individual standing near a road on a very hot day. If the index of refraction increases linearly with vertical distance from the road, i.e. $n(z) = n_0 + n'z$ with $n' \ll n_0$, show that a light path that leaves the observer's eyes pointing initially downwards may bend. As a result, the observer may see the blue sky even if they are looking downwards. *HINT: Find an expression for $y'(z)$ and then show that $y''(z) \neq 0$.* See Figure 13.5 for an illustration of this problem complete with realistic depiction of the individual and road in question.

Notes