17 Necessary and Sufficient Conditions for Weak Extrema and conditions for strong extrema

Learning and innovation go hand in hand. The arrogance of success is to think that what you did yesterday will be sufficient for tomorrow.
—William Pollard

In any good movie there is a villain (sometimes multiple villains and often the villains are actually just forces of nature, but stay with me for a bit here) that at some point or another is given to some lengthy diatribe, unless of course the movie was very well scripted and the audience actually knew the plot and storyline without it being told to them line by line. Either way, most movies do have this lengthy venting session where the villain explains their motives etc. etc.. Several movies also have periods of time where the hero does the same, now explaining his motives and choices (as if the audience couldn't have guessed at it already). Anyway, if you are one of those movie-goers who watches attentively and can often guess the plot lines before such moments then these soliloquoys will be very dull and monotonous, no matter how necessary they are to the development of the plot. Often that is what we encounter in applied mathematics. We get all the excitement of solving big problems using very fun tools and theorems, but eventually we do need to prove these theorems and justify our existence (you can pick whether you think that is the job of the villain or the hero). This Chapter is just such a time. Motivation here is difficult because it is not likely that the theorems presented will be used very frequently, but they are essential to the development of the Calculus of Variations nonetheless.

One distinction that will become important as we proceed is the difference between necessary and sufficient conditions. To discuss the differences between necessary and sufficient conditions we consider the abstract event A that we are interested in predicting. A necessary condition for A to happen is one that is required for A to occur, but if the necessary condition is satisfied it does not immediately mean that A will occur. That is what a sufficient condition is for. A sufficient condition will cause A to happen. In the case of minimizers, the first variation vanishing is a necessary condition for a minimizer to be present, but it is not sufficient. We saw this even in multivariable Calculus where saddle points or other types of critical points could be present. In the infinite dimensional setting that we are in now such geometric objects are even more prevalent and even more difficult to distinguish.

In the following, we first consider necessary conditions for the presence of a minima/-maxima of a cost functional, and then we also consider the sufficient conditions for the same. We emphasize here that the results described in this Chapter are restricted to a simple class of problems and generalizations of these results are not straightforward. Establishing the existence and potentially the uniqueness of a minimizer to a cost function in the Calculus of Variations requires a substantial amount of functional analysis that we are not equipped to handle.

17.1 Necessary Conditions to distinguish between maxima/minima

In Multivariable Calculus it was necessary for an extrema of $f(\mathbf{x})$ to satisfy $\nabla f(\tilde{\mathbf{x}}) = 0$ and the relevant sufficient condition for a local minima is that $\nabla f(\tilde{\mathbf{x}}) = 0$ and $D^2 f(\tilde{\mathbf{x}})$ is positive definite, meaning that $D^2 f(\tilde{\mathbf{x}})_{\mathbf{u},\mathbf{v}} = \mathbf{u}^T H \mathbf{v} > 0 \ \forall \mathbf{u}, \mathbf{v}$ and H is the Hessian matrix of the second derivatives of the function $f(\mathbf{x})$. This comes from Taylor's theorem expanded about the minima at $\mathbf{x} = \tilde{\mathbf{x}}$, i.e.

$$f(\tilde{\mathbf{x}} + \varepsilon \mathbf{h}) = f(\tilde{\mathbf{x}}) + \varepsilon \nabla f(\tilde{\mathbf{x}}) \cdot \mathbf{h} + \frac{\varepsilon^2}{2} \mathbf{h}^T H \mathbf{h} + O\left(\varepsilon^2\right).$$

If ε is sufficiently small (we are in a local neighborhood) and the necessary and sufficient conditions are satisfied at $\tilde{\mathbf{x}}$ then $f(\tilde{\mathbf{x}} + \varepsilon \mathbf{h}) \geq f(\tilde{\mathbf{x}})$ (assuming that $f(\mathbf{x})$ is smooth enough for Taylor's Theorem to apply). In a single variable these conditions become $f'(x) = 0$ and $f''(x) > 0$.

The goal of this Chapter is to extend these ideas from Multivariable Calculus to minimizers/maximizers of functionals, i.e. the infinite dimensional setting. One would hope that all we need to do is compute the second Frechét derivative and determine whether or not it is positive or negative definite. Unfortunately not only is the second order Frechét derivative extremely difficult to compute in practice, but this actually doesn't work out as one would hope. This is partially because Frechét derivatives are not exact analogues of the gradient/derivative in finite dimensions, but also because the concept of locally convex is fundamentally different in the infinite dimensional setting.

We will instead focus on the simpler case of sufficient conditions based on the Gateaux differential. An analogy in this setting is that we are considering local convexity in certain directions only, i.e. it is much more difficult to estimate convexity in all directions in this setting (infinite dimensional), but we can consider convexity in some subset of all possible directions. To clarify this we will consider a minimum $y(x)$ of the cost functional $J[y]$ (we will focus on the simplest problem below). If $y(x)$ is truly a minimum then $j(\varepsilon) = J[y + \varepsilon h]$ will have a minimum at $\varepsilon = 0$. This is equivalent to

$$\left.\frac{dj}{d\varepsilon}\right|_{\varepsilon=0} = 0, \quad \text{and} \quad \left.\frac{d^2 j}{d\varepsilon^2}\right|_{\varepsilon=0} > 0. \tag{17.1}$$

If we can establish this for all possible directions $h(x)$ then this leads to a natural connection with Gateaux differentials.

Definition 17.1.1. *Let $J[y]$ be a functional acting on $y(x) \in C^2[a,b]$. Define the 2nd order Gateaux differential (2nd variation) to be*

$$\delta^2 J[y; h] = \frac{d^2}{d\varepsilon^2} \left. J[y + \varepsilon h] \right|_{\varepsilon=0},$$

provided that the relevant limits exist

In general, $\delta^2 J[y;h]$ is quadratic in $h(x)$ and $h'(x)$ (kind of like when we refer to the Hessian being positive definite in the finite dimensional case where we look at the quadratic term $\mathbf{u}^T H\mathbf{v}$). For a functional of the form

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

then

$$\begin{aligned}
\delta^2 J[y;h] &= \frac{d^2}{d\varepsilon^2}\int_a^b [L(x,y+\varepsilon h, y'+\varepsilon h')]dx\Big|_{\varepsilon=0}\\
&= \int_a^b \frac{d}{d\varepsilon}\left[L_y(x,y+\varepsilon h,y'+\varepsilon h')h(x) + L_{y'}(x,y+\varepsilon h,y'+\varepsilon h')h'(x)\right]dx\Big|_{\varepsilon=0}\\
&= \int_a^b \left[L_{yy}(x,y,y')h^2(x) + 2L_{yy'}(x,y,y')h(x)h'(x) + L_{y'y'}(x,y,y')(h'(x))^2\right]dx.
\end{aligned}$$

If we are considering $h(x)\in C^1[a,b]$ with $h(a)=h(b)=0$ (the endpoints of $y(x)$ are fixed) then we can integrate by parts. Noting that $2hh' = \frac{d}{dx}(h^2)$,

$$\int_a^b 2L_{yy'}hh'dx = -\int_a^b \frac{d}{dx}\left(L_{yy'}\right)h^2dx + L_{yy'}h^2\big|_a^b,$$

where the boundary term will vanish so that

$$\begin{aligned}
\delta^2 J[y;h] &= \int_a^b \left\{\left[L_{yy} - \frac{d}{dx}L_{yy'}\right]h^2 + \int_a^b L_{y'y'}(h')^2\right\}dx\\
&= \int_a^b \left\{Qh^2 + P(h')^2\right\}dx,
\end{aligned}$$

where $Q = L_{yy} - \frac{d}{dx}L_{yy'}$ and $P = L_{y'y'}$. This defines a quadratic form, similar to those that occur in Linear Algebra optimization problems. In Linear Algebra we would be trying to optimize $\mathbf{x}^T A\mathbf{x}$. As we will see later for optimal control, quadratic forms are a useful modeling tool. Here things are a bit more complicated than what we saw in Linear Algebra, but in some sense also nicer because the functions $h(x)$ and $h'(x)$ (which are the analogues of entries in the vector $\mathbf{x}$) are not unrelated. In fact, as one may suppose it turns out that P is the dominating term here (in general the derivative of a given function will dominate the function itself).

Remark 17.1.2. At issue with all of this discussion of course is that we are trying to derive a necessary condition for the existence of a minimizer based on the Gateaux differential only, which as we have already addressed, does not comprehensively represent all directions of ascent/descent as the Frechét derivative would. Hence, while we are truly establishing a necessary condition for the existence of a local minimizer, we must keep in mind that these same arguments fall apart when we try to extend them to sufficient conditions for the same.

Lemma 17.1.3. *A necessary condition for the quadratic functional*

$$\int_a^b (Qh^2 + P(h')^2)dx,$$

(where Q and P are sufficiently smooth) to be nonnegative for all $h(x)\in C^1[a,b]$ with $h(a)=h(b)=0$, is that $P(x)\geq 0$ for all $x\in[a,b]$.

Proof. Suppose to the contrary that there is some $x_0 \in [a,b]$ so that $P(x_0) < 0$ but where the quadratic functional is nonnegative. Let $P(x_0) = -2\gamma$ for some $\gamma > 0$. Since P is continuous then there is some $\delta > 0$ so that $[x_0-\delta, x_0+\delta] \subseteq [a,b]$ and $P(x) < -\gamma$ for all $x \in [x_0-\delta, x_0+\delta]$. Now let

$$h(x) = \begin{cases} \sin^2\left(\frac{\pi(x-x_0)}{\delta}\right) & : x \in [x_0-\delta, x_0+\delta] \\ 0 & : \text{otherwise} \end{cases}.$$

Inserting this back into the quadratic functional, we arrive at

$$\begin{aligned} \int_a^b (Qh^2 + P(h')^2)dx &= \int_{x_0-\delta}^{x_0+\delta} Q\sin^4\left(\frac{\pi(x-x_0)}{\delta}\right)dx \\ &+ \int_{x_0-\delta}^{x_0+\delta} P\left[2\frac{\pi}{\delta}\sin\left(\frac{\pi(x-x_0)}{\delta}\right)\cos\left(\frac{\pi(x-x_0)}{\delta}\right)\right]^2 dx \\ &= \int_{x_0-\delta}^{x_0+\delta} Q\sin^4\left(\frac{\pi(x-x_0)}{\delta}\right)dx + \frac{\pi^2}{\delta^2}\int_{x_0-\delta}^{x_0+\delta} P\sin^2\left(\frac{2\pi(x-x_0)}{\delta}\right)dx \\ &< 2M\delta - \frac{2\pi^2\gamma}{\delta}, \end{aligned}$$

where $M = \max_{x\in[a,b]} Q(x)$ and we used the trigonometric identity

$$\frac{1}{2}\sin\theta = \sin\frac{\theta}{2}\cos\frac{\theta}{2}.$$

Now for $\delta < \sqrt{\frac{\gamma\pi^2}{2M}}$ then this implies that $\int_a^b (Qh^2+P(h')^2)dx < 0$ which proves the lemma. ☐

With this in mind we are prepared to establish the conditions that prove the Theorem which yields necessary conditions for a minima (maxima) of $J[y]$.

Theorem 17.1.4 (Legendre's Condition). *Consider the minimization of*

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

for $y(x) \in C^1[a,b]$ with fixed endpoints and $L(x,y,y')$ being 2 times continuously differentiable in both y and y'. Suppose that $\tilde{y}(x)$ gives a local minimum of J, then $\tilde{y}(x)$ satisfies (EL) $L_y - \frac{d}{dx}L_{y'} = 0$, and the Legendre Condition: $L_{y'y'} \geq 0$.

Proof. As in the proof for the (EL) let $f(\varepsilon) = J[\tilde{y} + \varepsilon h]$. Then if $J[y]$ has a minimum at $y = \tilde{y}$ then $f(\varepsilon)$ will have a minimum at $\varepsilon = 0$. From earlier courses, then $f'(0) = 0$ (EL) and $f''(0) = \delta^2 J[\tilde{y}; h] \geq 0$. From the lemma, and the previous computations we get the Legendre condition. ☐

Example 17.1.5. Recall the problem of finding the shortest path between two points on the plane with Lagrangian given by

$$L = \sqrt{1+(y')^2}.$$

Then $L_{y'} = \frac{y'}{\sqrt{1+(y')^2}}$ and

$$\begin{aligned} L_{y'y'} &= \frac{1}{\sqrt{1+(y')^2}} - \frac{(y')^2}{(1+(y')^2)^{\frac{3}{2}}} \\ &= \frac{1}{(1+(y')^2)^{\frac{3}{2}}} \geq 0. \end{aligned}$$

Thus, gratefully the straight line is indeed the shortest (and not the longest) path between two fixed points.

Example 17.1.6 (Brachistochrone again). As yet another example where we can walk through the calculation in reasonable time, recall that the optimal path taken for a ball rolling downhill under the influence of gravity alone is given by minimizing

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

This leads to:

$$\begin{aligned} L_{y'} &= \frac{y'}{\sqrt{2gy(1+(y')^2)}}, \\ \Rightarrow L_{y'y'} &= \frac{1}{\sqrt{2gy}} \frac{1}{(1+(y')^2)^{3/2}} \geq 0, \end{aligned}$$

for all reasonable profiles $y(x)$ and hence the Brachistochrone really does optimize how quickly the ball can get to the bottom of the hill.

17.2 Sufficient Conditions for a maximum/minimum

In order to consider generic sufficient conditions we will need a few additional preliminaries. The sufficiency condition in the finite dimensional setting was the same as the necessary condition which in essence was a local convexity condition, i.e. a the second derivative is positive definite. In this case we need to be a bit more careful. The strengthened Legendre condition which is given by:

$$L_{y'y'} > 0,$$

is not quite enough here. We will consider such sufficient conditions in two different classes. First we will consider the more traditional (at least as compared to multivariable calculus) route of calculating the second variation and determining conditions that guarantee the positive definiteness of the resultant quadratic form. Finally we will consider a convexity condition that can yield a global sufficiency condition.

17.2.1 Local sufficiency

The real concern with local sufficiency is that the primary tool we have relied on is the Gateaux differential which while providing a necessary condition, is not sufficiently general to provide a sufficient one. Instead we need to work with the full Frechét derivative, and consider convergence in terms of the relevant functional norm (usually $\|\cdot\|_{C^1}$). In the spirit of communicating the key ideas without getting bogged down in details, we present these results below without dwelling on the specifics, i.e. we always assume that a Taylor series is available to us (in the Frechét sense of course), and we will proceed as though convergence of the higher order terms is guaranteed. In reality this is a dangerous assumption, particularly when we note that we are optimizing over functionals that integrate over the domain in question so that there is no guarantee that the optimizer need be continuous let alone differentiable. Hence, even though we restrict ourselves to C^1 here, this is not the natural space to consider optimization in the Calculus of Variations. We should instead be broaching the subject via Sobolev spaces which show up in functional analysis. Alas, this is the last time you will hear us mention such beautiful spaces. If you desire to learn more about the theory supporting this majestic field of mathematics we encourage you to investigate advanced graduate textbooks in functional analysis and/or PDEs.

To foreshadow what we are trying to do in this section:

- We want to derive a condition that if fulfilled by the curve $y(x)$, will guarantee (give a sufficient condition) that $y(x)$ is a minimizer of the cost functional $J[y]$.
- To identify this sufficiency condition, we define a concept called conjugate points in the interval $[a, b]$ of interest. If the strengthened Legendre condition is satisfied, and there are no conjugate points in $[a, b]$ then the curve $y(x)$ is indeed a minimizer of $J[y]$, i.e. 'critical' points of the cost functional are indeed minimizers.

To further illustrate the point we are trying to make, consider the following unexample

Unexample 17.2.1. Consider the optimization problem

$$\min_{y(x)} J[y] = \int_0^4 \left[(y')^2 - y^2\right] dx.$$

Note that

$$L_y = -2y, \quad L_{y'}. = 2y', \quad L_{y'y'} = 2,$$

so that $y_0(x) = 0$ satisfies both the (EL) and the Legendre condition. We will show that despite this, $y_0(x)$ is *NOT* a local minimizer of $J[y]$. Define $y_\varepsilon(t) = \varepsilon t(4-t)$. Then it follows that

$$\begin{aligned} J[y_\varepsilon] &= \int_0^4 \left[(y_\varepsilon')^2 - y_\varepsilon^2\right] dt \\ &= \varepsilon^2 \int_0^4 \left[(4-2t)^2 - t^2(4-t)^2\right] dt \\ &= -\varepsilon^2 \frac{64}{5}. \end{aligned}$$

If we take $\varepsilon \to 0$ then clearly y_ε is close to y_0, but $J[y_\varepsilon] < J[y_0] = 0$ so y_0 can not be a local minimizer of $J[y]$.

To see what could go wrong, we consider the second variation (using the fact that $\frac{d}{dx}\tan x = \tan^2 x + 1$):

$$\begin{aligned}
\delta^2 J[y;h] &= 2\int_0^4 \left[(h')^2 - h^2\right] dx \\
&= 2\int_0^4 \left[(h')^2 - (1+\tan^2 x - \tan^2 x)h^2\right] dx \\
&= 2\int_0^4 \left[(h')^2 - \frac{d}{dx}(\tan x)h^2 + \tan^2 x h^2\right] dx \\
&= 2\int_0^4 \left[(h')^2 + 2\tan x h h' + \tan^2 x h^2\right] dx \\
&= 2\int_0^4 \left[h' + \tan x h\right]^2 dx,
\end{aligned}$$

where the second to last line was achieved via integration by parts. The ODE $h'(x) = -\tan x h(x),\ h'(0) = 0$ has a unique solution ($h(x) = 0$) over $[0,1]$, but does it over $[0,4]$? It turns out this matters, as we will see below.

In order to consider sufficient conditions for a minimum, we need a few preliminaries. Consider the simplest variational problem, i.e. optimizing

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

with an extremal at $\tilde{y}(x)$. Suppose that $y^*(x) = \tilde{y}(x) + h(x)$ is also an extremal. In other words, $y^*(x)$ and $\tilde{y}(x)$ are two 'nearby' extrema and arbitrarily close. The (EL) for these two extrema are:

$$\begin{aligned}
L_y(x,\tilde{y},\tilde{y}') - \frac{d}{dx}L_{y'}(x,\tilde{y},\tilde{y}') &= 0 \\
L_y(x,\tilde{y}+h,\tilde{y}'+h') - \frac{d}{dx}L_{y'}(x,\tilde{y}+h,\tilde{y}'+h') &= 0.
\end{aligned}$$

If we take the Taylor expansion of the last equation about $\tilde{y}(x)$, assuming that $h(x)$ is small in some norm then we arrive at

$$L_{yy}h + L_{yy'}h' - \frac{d}{dx}\left(L_{y'y}h + L_{y'y'}h'\right) = O(\|h\|).$$

Neglecting $\mathscr{O}(||h||)$ terms, and recognizing that

$$\frac{d}{dx}(L_{yy'}h) = \left(\frac{d}{dx}L_{yy'}\right)h + L_{yy'}h',$$

then we see that

$$\left(L_{yy} - \frac{d}{dx}L_{yy'}\right)h - \frac{d}{dx}\left(L_{y'y'}h'\right) = 0, \tag{17.2}$$

which is called Jacobi's equation. This is referred to as the variational equation of the (EL). Returning to the 2nd variation, note that Jacobi's equation is equivalent to $Qh - \frac{d}{dx}(Ph') = 0$ where the second variation is given by

$$\delta^2 J[y;h] = \int_a^b (Qh^2 + P(h')^2)dx,$$

and

$$Q = L_{yy} - \frac{d}{dx}L_{yy'} \quad \text{and} \quad P = L_{y'y'},$$

i.e. Jacobi's equation is the (EL) applied to the second variation.

Note that (17.2) always has $h \equiv 0$ as a solution, which also satisfies the boundary conditions $h(a) = h(b) = 0$. There are other possible, non-trivial solutions to 17.2. These are the solutions that we wish to consider in greater detail.

Definition 17.2.2. *The point $\tilde{a} \in (a, b]$ is 'conjugate' to a if Jacobi's equation has a nontrivial solution $h(x)$ for which $h(a) = 0$, $h'(a) = 1$ and $h(\tilde{a}) = 0$.*

Remark 17.2.3. Note that Unexample 17.2.1 indicates that there are no conjugate points for the cost functional in question on the interval $[0, 1]$, however there are potentially conjugate points on the interval $[0, 4]$. Why is that the case?

In reality we are only concerned with the condition $h(a) = 0$, but if $h(x)$ is such a solution then so is $ch(x)$ for any $c \in \mathbb{R}$. Hence the condition $h'(a) = 1$ selects a specific constant for c. This condition appears when we are trying to determine what form of P (and Q) enforce $\delta^2 J[y; h] > 0$. The basic approach (although this is likely overly simplified) is to complete the square (or add zero cleverly disguised).

Theorem 17.2.4. *If $P(x) > 0$ for $x \in [a, b]$ and if $[a, b]$ contains no points conjugate to a, then*

$$\int_a^b (Qh^2 + P(h')^2)dx$$

is positive definite for all non-trivial $h \in C^1[a, b]$ with $h(a) = h(b) = 0$.

Proof. Note that for any nice (continuously differentiable is likely sufficient) $w(x)$,

$$0 = wh^2\big|_a^b = \int_a^b \frac{d}{dx}\left(wh^2\right) dx,$$

so that adding this term onto the quadratic form does not change anything. This is how we will complete the square, by adding zero. The goal is to select $w(x)$ so that

$$Qh^2 + P(h')^2 + \frac{d}{dx}(wh^2) = (Q + w')h^2 + 2whh' + P(h')^2.$$

This is a perfect square if

$$(Q + w')h^2 + 2whh' + P(h')^2 = (qh + ph')^2,$$

where we still need to determine q and p. Factoring things out, we see that

$$(qh + ph')^2 = q^2h^2 + 2qphh' + p^2(h')^2,$$

so that $p^2 = P$, $q^2 = Q + w'$, and $pq = w$, indicating that

$$p = \sqrt{P}, \quad q = \sqrt{Q + w'}, \quad \text{and} \quad w = \sqrt{P(Q + w')}.$$

Finally this leads to

$$P(Q + w') = w^2. \tag{17.3}$$

If we can find such a $w(x)$ then

$$\int_a^b Qh^2 + P(h')^2 dx = \int_a^b P\left(\frac{w}{P}h + h'\right)^2 dx.$$

Thus if there is a nontrivial solution to (17.3) then $\delta^2 J[y;h]$ is manifestly non-negative. In fact, because $P(x) > 0$, if $\delta^2 J[y;h] = 0$ then

$$h'(x) + \frac{w}{P}h(x) = 0$$

for all $x \in [a,b]$. This is a 1st-order equation in h, and we have the condition that $h(a) = 0$ so $h(x) = 0$ is a solution and as long as $\frac{w}{P}$ is at least continuous, then uniqueness of solutions from ODE's guarantees that $h(x) \equiv 0$ is the only solution. Thus $\delta^2 J[y;h]$ is positive definite, i.e. vanishes only for $h(x) \equiv 0$.

It follows that positivity of $\delta^2 J[y;h]$ reduces to finding continuous (non-trivial) solutions of (17.3). It turns out that (17.3) is a Riccati equation, and can be reduced to a 2nd order, linear equation. To see this, let $w = -\frac{u'}{u}P$, then (17.3) becomes

$$P\left(Q - \frac{(Pu')'}{u} + \frac{P(u')^2}{u^2}\right) = P^2\frac{(u')^2}{u^2}$$

which implies that

$$P\left(Q - \frac{(Pu')'}{u}\right) = 0$$

so that

$$Q - \frac{(Pu')'}{u} = 0,$$

assuming $u \neq 0$ in the region of interest (meaning there are no conjugate points). This leads to

$$Qu - \frac{d}{dx}(Pu') = 0,$$

which is Jacobi's equation. Thus if there are no conjugate points to a, then w is well defined on all of $[a,b]$. $\square$

Remark 17.2.5. The last theorem tells us that the quadratic functional

$$\int_a^b \left(Qh^2 + P(h')^2\right) dx > 0$$

for all admissible $h(x) \in C^1[a,b]$ with $h(a) = h(b) = 0$ so long as $P > 0$ and there are no points conjugate to a in $[a,b]$. It turns out that this statement goes both ways, i.e. if the quadratic form is positive definite for all such $h(x)$, then $P > 0$ and there are no points conjugate to a. We do not prove the other direction here, but the interested reader is encouraged to seek out a simple proof.

We are finally prepared to make a definitive statement regarding the local existence of a weak minimum for the simplest problem (just imagine what we need to do for the hardest one).

Theorem 17.2.6. *For the functional*

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

with $y(a)$ and $y(b)$ fixed, suppose that an admissible curve $\tilde{y}(x)$ satisfies the following 3 conditions

(*i*) *$\tilde{y}(x)$ is an extremizer, i.e. it satisfies the (EL) $L_y - \frac{d}{dx}L_{y'} = 0$.*

(*ii*) *Evaluated along $y = \tilde{y}(x)$, $P = L_{y'y'} > 0$ (Strengthend Legrendre Condition).*

(iii) The interval $[a,b]$ has no points conjugate to a (Strengthend Jacobi Condition). Then $\tilde{y}(x)$ is a weak minimizer of $J[y]$.

Proof. If there are no conjugate points to a in $[a,b]$ and $P(x) > 0$ then continuity (of both $P(x)$ and solutions to Jacobi's equation) guarantee that there is an $\varepsilon > 0$ so there are also no conjugate points in $[a, b+\varepsilon]$ and $P(x) > 0$ for $x \in [a, b+\varepsilon]$.

Now consider

$$\int_a^b (Qh^2 + P(h')^2)dx - \alpha^2 \int_a^b (h')^2 dx$$

with corresponding (EL) (Jacobi's equation)

$$Qh - \frac{d}{dx}[(P-\alpha^2)h'] = 0. \tag{17.4}$$

Since $P(x) > 0$ and continuous, we can pick α^2 sufficiently small so that

(i) $P - \alpha^2 > 0$ for $x \in [a,b]$.

(ii) Solutions of (17.4) satisfying $h(a) = 0$, $h'(a) = 1$ do not vanish for $x \in [a,b]$.

From the previous theorem, if α^2 is sufficiently small then solutions are continuously dependent on α, and

$$\int_a^b (Qh^2 + P(h')^2)dx - \alpha^2 \int_a^b (h')^2 dx > 0. \tag{17.5}$$

Now consider an admissible curve $y(x) = \tilde{y}(x) + h(x)$ where $\|h\|_{C^1}$ is small ($\|h\|_{C^1} = \max_{x\in[a,b]} |h(x)| + \max_{x\in[a,b]} |h'(x)|$) Then using Taylor's theorem:

$$J[\tilde{y}+h] - J[\tilde{y}] = \int_a^b (Qh^2 + P(h')^2)dx + \int_a^b (\eta_1 h^2 + \eta_2 (h')^2)dx,$$

where $|\eta_1(x)||\eta_2(x)| = O\left(\|h\|_{C^1}\right)$ for all $x \in [a,b]$. Suppose that $\|h\|_{C^1}$ is sufficiently small so that $|\eta_1(x)|, |\eta_2(x)| \le \delta$ for all $x \in [a,b]$, Now using the Cauchy-Schwartz inequality ($\int fg dx \le \|f\|_{L^2}\|g\|_{L^2}$), for some $\delta > 0$,

$$\begin{aligned}
h^2(x) &= \left(\int_a^x h'(\xi)d\xi\right)^2 \\
&\le \left[\left(\int_a^x d\xi\right)^{\frac{1}{2}} \left(\int_a^x (h'(\xi))^2 d\xi\right)^{\frac{1}{2}}\right]^2 \\
&= (x-a)\int_a^x (h'(\xi))^2 d\xi \\
&\le (x-a)\int_a^b (h'(\xi))^2 d\xi.
\end{aligned}$$

Thus

$$\begin{aligned}
\int_a^b h^2(x)dx &\le \left(\int_a^b (x-a)dx\right)\left(\int_a^b (h'(\xi))^2 d\xi\right) \\
\int_a^b h^2(x)dx &\le \frac{(b-a)^2}{2}\left(\int_a^b (h'(\xi))^2 d\xi\right) \\
\Rightarrow \left|\int_a^b (\eta_1 h^2 + \eta_2 (h')^2)dx\right| &\le \delta\left(1 + \frac{(b-a)^2}{2}\right)\int_a^b (h'(x))^2 dx \\
\Rightarrow J[\tilde{y}+h] - J[\tilde{y}] &> \alpha^2 \int_a^b (h')^2 dx - \delta\left(1 + \frac{(b-a)^2}{2}\right)\int_a^b (h'(x))^2 dx.
\end{aligned}$$

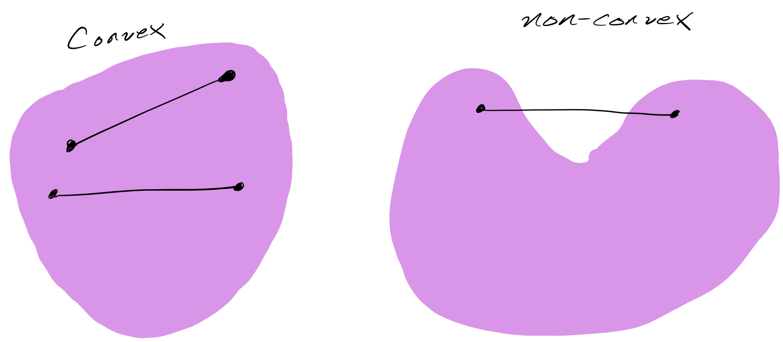

Figure 17.1: Demonstration of the difference between a convex and non-convex set. Note that in higher dimensions the same intuitive picture holds, that is lines that connect two points in the set must remain in the set for convexity to hold.

Hence, as with any good proof, we now (naturally) select

$$\delta < \frac{\alpha^2}{2 + (b-a)^2}$$

so that

$$J[\tilde{y} + h] - J[\tilde{y}] > c \int_a^b (h')^2 dx > 0,$$

for $\|h\|_{C^1}$ being sufficiently small. Thus $J[\tilde{y}+h] > J[\tilde{y}]$ for all admissible 'small' increments $h(x)$. $\square$

17.2.2 Convexity

The condition that $\delta^2 J[y; h]$ is positive definite is really a local convexity condition. To see this, recall the following definition from Volume I TODO: get reference

Definition 17.2.7. *A subset C of some linear vector space is convex if $\mathbf{u}t + (1-t)\mathbf{v} \in C$ for all $\mathbf{u}, \mathbf{v} \in C$ and $t \in [0, 1]$.*

Remark 17.2.8. A convex subset is exactly what you would think, any line segment connecting two points in the set must also lie entirely within the set itself. Thus in $\mathbb{R}^2$ the circle is a convex set, as is a rectangle, but a bean shaped object would not be (a torus is not a convex subset in $\mathbb{R}^3$).

Definition 17.2.9. *A functional $J[y]$ defined on a convex subset C is convex if $J[ty_1 + (1-t)y_2] \le tJ[y_1] + (1-t)J[y_2]$ for all $y_1, y_2 \in C$ and $t \in [0, 1]$.*

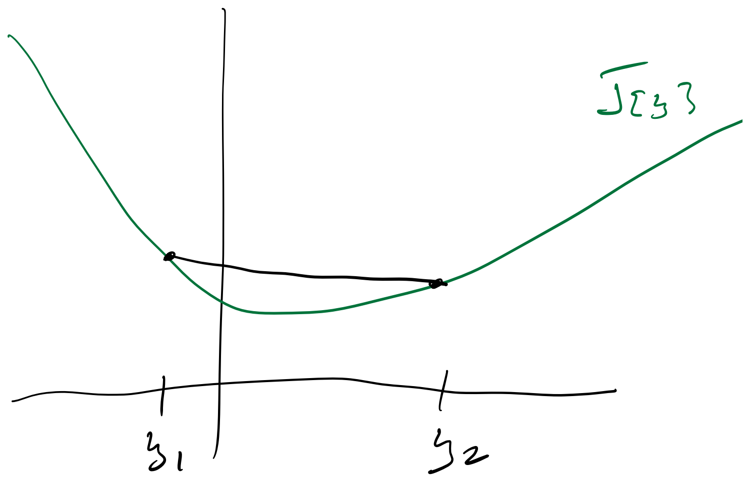

Figure 17.2: Example of a convex function. Note that the line connecting the two points y_1 and y_2 lies above the function $J[y]$ itself (and that this would be true for any values of y_1 and y_2).

Convex functionals are also easy to understand, for instance $y = x^2$ is one of the simplest globally convex functionals there is. You should feel comfortable recognizing that this function has only one unique minimum, and there are no other extrema possible. Although $y = x^2$ may seem to simple to be of any use, it is in fact illustrative of everything that goes into the construction and pursuit of minima for more complicated functionals. Generically if you have a choice in the development and production of some variational statement it is best to try and make it convex, particularly globally convex because then the minimization is not only guaranteed but any extrema that is found will be the unique one, i.e. there is no need to resort to the trickery/ detail of the previous section.

Theorem 17.2.10. *Let $J[y]$ be a convex functional defined on a convex set C Let $\alpha = \inf_{y \in C} J[y]$ then the subset $\Omega \subseteq C$ defined by*

$$\Omega = \{y \in C : J[y] = \alpha\},$$

is convex. In addition if $\tilde{y}$ is a local minimum of $J[y]$ then $J[\tilde{y}] = \alpha$, and it is the global minimum.

Proof. Let $y_1, y_2 \in \Omega$ and $y = ty_1 + (1-t)y_2$ for some $t \in (0,1)$. Then it follows that

$$J[y] \leq tJ[y_1] + (1-t)J[y_2] = \alpha,$$

but $J[y] \geq \alpha$ for all $y \in C$ so that $J[y] = \alpha$ indicating that $y \in \Omega$ and hence Ω is convex.

Consider a neighborhood of $\tilde{y}$, $B(\tilde{y}, \varepsilon)$ in which $\tilde{y}$ minimizes J. Now for any $y_1 \in C$, there exists $y_2 \in B(\tilde{y}, \varepsilon)$ so that $y_2 = t\tilde{y} + (1-t)y_1$ for some $t \in (0,1]$. Thus $J[\tilde{y}] \leq J[y_2] \leq tJ[\tilde{y}] + (1-t)J[y_1]$ which implies $J[\tilde{y}] \leq J[y_1]$ so $\tilde{y}$ is a global minimizer. $\square$

Thus, while we may go through much pain and suffering to get sufficient conditions for local extrema using the second variation, it is much easier if we can simply construct a cost functional that is automatically convex (or concave in the case of a maximum) without needing to evaluate near each extrema.

17.3 *Strong Extrema

17.3.1 Strong versus Weak

This is not a discussion on the virtue of hard work, nor on the domination of the strong over the weak (be that physically or mentally), but is restricted to a discussion on the merits of differentiability. It is rather ironic that if you had a child look at the plot of two extrema, one weak and the other strong, they would likely refer to the strong extrema as 'broken' and the weak extrema as 'whole'. Hence, the irony of naming anything in mathematics, for what is a clear and obvious naming convention to one is obtuse and convoluted to another.

So far we have lived in $C^1[a,b]$. This seemed like a good idea at the time, and why not since it was a rather safe place to be? There are occasions however where it is necessary to consider extremals that are not differentiable. In such cases we need to formalize what happens when we have a strong extremum that is not weak, i.e. the extremum lives in $C[a,b]$ but not $C^1[a,b]$.

17.3.2 Weirstrass-Erdmann corner conditions

First we must clarify what points are non-differentiable for a given function.

Definition 17.3.1. *For a function $y(t) \in C[a,b]$ where $y'(c^+) = \lim_{t\to c^+} y'(t) \neq \lim_{t\to c^-} y'(t) = y'(c^-)$, the point $t = c$ is referred to as a corner point.*

Consider the functional

$$J[y] = \int_a^b L(t, y, y')dt,$$

with an extremal at $\tilde{y}$ with only one corner point $c \in [a,b]$. We break $\tilde{y}$ into two pieces as $y_1(t) : [a,c] \to \mathbb{R}$ and $y_2(t) : [c,b] \to \mathbb{R}$ where $y_1(t)$ and $y_2(t)$ are in C^1.

Now we will vary $\tilde{y}(t)$ by varying $y_1(t)$ and $y_2(t)$ separately, i.e.

$$y(t) + \varepsilon h(t) = \begin{cases} y_1(t) + h_1(t) \\ y_1(t) + h_2(t) \end{cases},$$

where $h_1(a) = h_2(b) = 0$. This is not quite enough however, because we really have no reason to suppose that we know c ahead of time. In other words, c may vary itself, i.e. the corner point will now look like $c + \varepsilon(\delta t)$, but we still want continuity at the corner point, i.e.

$$y_1(c + \varepsilon(\delta t)) + \varepsilon h_1(c + \varepsilon(\delta t)) = y_2(c + \varepsilon(\delta t)) + \varepsilon h_2(c + \varepsilon(\delta t)).$$

Without loss of generality, consider $\varepsilon, \delta t > 0$. For $y_1 + \varepsilon h_1$ to be defined on $[a, c + \varepsilon(\delta t)]$ we consider the linear continuation of $y_1(t)$ onto this interval: for $t > c$,

$$y_1(t) = y(c) + y'(c^-)(t - c).$$

The continuation of $y_2(t)$ to the left can be defined similarly if ε or δt were negative.

Remark 17.3.2. The continuation need not be linear, but could be quadratic which would then allow $y_1(t) \in C^1[a, c+\varepsilon(\delta t)]$, but this is more regularity than we are interested in for this discussion. We are sufficiently happy with a linear continuation that guarantees that $y_1(c) = y(c)$ and $y_1'(c) = y'(c^-)$.

Now we rewrite the 'cost functional' $J[y]$ as two pieces over the two intervals $[a, c]$ and $[c, b]$.

$$J[y+\varepsilon h] = \int_a^{c+\varepsilon(\delta t)} L(t, y_1+\varepsilon h, y_1'+\varepsilon h')dt + \int_{c+\varepsilon(\delta t)}^b L(t, y_2+\varepsilon h_2, y_2'+\varepsilon h_2')dt.$$

Let

$$J_1[y_1+\varepsilon h_1] = \int_a^{c+\varepsilon(\delta t)} L(t, y_1+\varepsilon h_1, y_1'+\varepsilon h_1)dt,$$

and

$$J_2[y_2+\varepsilon h_2] = \int_{c+\varepsilon(\delta t)}^b L(t, y_2+\varepsilon h_2, y_2'+\varepsilon h_2)dt.$$

Before we get to the calculation of the first variation in this case, we must recall Leibniz's Rule (which you have probably needed on the homework by now) which states that

$$\frac{d}{dx}\int_{a(x)}^{b(x)} f(x,y)dy = \int_{a(x)}^{b(x)} \frac{\partial f}{\partial x}(x,y)dy + f(x, b(x))b'(x) + f(x, a(x))a'(x). \tag{17.6}$$

Thus

$$\delta J[y;h] = \delta J_1[y_1, h_1] + \delta J_2[y_2, h_2],$$

where

$$\delta J_1[y_1, h_1] = \int_a^c [L_y(t, y_1, y_1')h_1 + L_{y'}(t, y_1, y_1')h_1']dt + L(c, y_1(c), y_1'(c))(\delta t).$$

Integrating by parts we arrive at

$$\begin{aligned}\int_a^c L_{y'}(t, y_1, y_1')h_1'dt &= -\int_a^c \frac{d}{dt}L_{y'}(t, y_1, y_1')h_1 dt + L_{y'}(t, y_1(t), y_1'(t))h_1(t)\Big|_{t=a}^{t=c} \\ &= -\int_a^c \frac{d}{dt}(L_y(t, y, y'))h_1 dt + L_{y'}(c, y_1(c), y_1'(c))h_1(c).\end{aligned}$$

Recalling that $y_1(c) = y(c)$ and $y_1'(c) = y'(c^-)$,

$$\delta J_1[y;h] = \int_a^c [L_y - \frac{d}{dt}L_{y'}]h_1(t)dt + L(c, y(c), y'(c^-))(\delta t) + L_{y'}(c, y(c), y'(c^-))h_1(c).$$

Similarly we can find that

$$\delta J_2[y;h] = \int_c^b [L_y - \frac{d}{dt}L_{y'}]h_2(t)dt - L(c, y(c), y'(c^+))(\delta t) - L_{y'}(c, y(c), y'(c^+))h_2(c).$$

We are interested in finding $y(x)$ so that $\delta J = 0$ for all variations that maintain $h_1(a) = h_2(b) = 0$. Then a subset of these variations are when $(\delta t) = 0$ and $h_1(c) = h_2(c) = 0$, so that the extremal must satisfy the (EL) on intervals with no corner points on the interior, i.e.

$$L_y - \frac{d}{dt}L_{y'} = 0$$

for $y = y_1$ on $[a, c]$ and $y = y_2$ on $[c, b]$. We also expect δJ to vanish for the cases when $\delta t \neq 0$ and $h_1(c) \neq 0,\ h_2(c) \neq 0$ implying that

$$\left[L(c, y(c), y'(c^-)) + L(c, y(c), y'(c^+))\right](\delta t) + L_{y'}(c, y(c), y'(c^-))h_1(c) + L_{y'}(c, y(c), y'(c^+))h_2(c) = 0. \tag{17.7}$$

Continuity of the variations at $t = c$ requires that

$$y_1(c+\varepsilon(\delta t)) + \varepsilon h_1(c+\varepsilon(\delta t)) = y_2(c+\varepsilon(\delta t)) + \varepsilon h_2(c+\varepsilon(\delta t)) = y(c) + \varepsilon(\delta y).$$

Expanding this continuity condition in ε leads to

$$y_1(c) + \varepsilon(\delta t)y_1'(c) + \varepsilon h_1(t) = y_2(c) + \varepsilon(\delta t)y_2'(c) + \varepsilon h_2(c) = y(c) + \varepsilon(\delta y) + O\left(\varepsilon^2\right).$$

Because $y(c) = y_1(c) = y_2(c)$ we see that this leads to the $O(\varepsilon)$ condition:

$$\delta y = y'(c^-)(\delta t) + h_1(c) = y'(c^+)(\delta t) + h_2(c),$$

which in turn indicates that

$$h_1(c) = \delta y - (\delta t)y'(c^-) \qquad h_2(c) = \delta y - (\delta t)y'(c^+).$$

Inserting these back into the boundary terms arising from the requirement that the first variation will vanish (17.7), leads to

$$(\delta t)\left[L(c^-) - L(c^+) - L_{y'}(c^-)y'(c^-) + L_{y'}(c^+)y'(c^+)\right] + (\delta y)\left[L_{y'}(c^-) - L_{y'}(c^+)\right] = 0$$

Since (δt) and (δy) are independent then this implies that

- $L_{y'}$ is continuous at $t = c$
- $L - y'L_{y'}$ is continuous at $t = c$

These are called the *Weirstrass-Erdmann corner conditions*.

For a strong extrema $\tilde{y}$ of

$$J[y] = \int_a^b L(t, y, \dot{y})dt,$$

then at every corner point the Weirstrass-Erdmann conditions must be satisfied, i.e. the Hamiltonian $H = L - y'L_{y'}$ and the canonical momentum $p = L_{y'}$ must be continuous as a necessary condition for an extrema $J[y]$.

17.3.3 Weirstrass Function and Necessary Conditions for Extrema

Now we would like to have something similar to a Legendre condition for strong extrema, and in fact that is a reasonable place to start, but suppose that since we are dealing with strong extrema where differentiability is in question, that differentiability of the Lagrangian is also in question. If this is the case, then we can not resort to Legendre's condition where the second derivative of the Lagrangian was used to verify the local convexity of the cost functional. Instead we will consider a more subtle, yet also far more general approach.

Definition 17.3.3. *For the cost function*

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

we will define the E-function or Weirstrass function as

$$E(t, y, y', p) = L(t, y, y') - L(t, y, p) - (y' - p)L_{y'}(t, y, p).$$

Remark 17.3.4. At first glance there is very little that makes sense about this function, unless we think of the Lagrangian L as a function of y' alone. For example, consider a function $f(z)$ whose linear approximation about the point $z = p$ is given by the tangent line $f(p) + (z-p)f'(p)$. Then the difference between the function evaluated at $z = y$, and the linear approximation found at $z = p$ will be $f(y) - f(p) - (y-p)f'(p)$. Substituting the Lagrangian as a function of y' only for this function $f(z)$ then we arrive at the Weirestrass E-function. Hence, $E(t, y, y', p)$ can be thought of as the difference between the actual Lagrangian at one point, and it's linear approximation at a nearby point. In this simplified version of things, we can recognize from a simple sketch that $f(y) - f(p) - (y-p)f'(p) > 0$ for all y, p in some neighborhood, then the function $f(z)$ will be convex in that neighborhood, this being determined without resorting to trying to evaluate the second derivative of $f(z)$.

Remark 17.3.5. Always wondered why these things were called remarks. Why not refer to them as 'comments'? It seems rather arrogant to say that we are 'remarking' on the progress of mathematical thought, and not commenting on it (not as if speaking as 'we' is arrogant after all).

All of this indicates the following Theorem, which is just like the Theorem in the previous chapter where we established Legendre's condition will dictate that $\delta^2 J$ is positive semi-definite. This again indicates the significance of bounding the behavior of a function by dealing with its derivative. Just as when we dealt with Legendre's condition, the local convexity of the entire cost functional can be determined by the local convexity of the Lagrangian with respect to y'.

Theorem 17.3.6. *If $\tilde{y}(t)$ is a strong minimum of*

$$J[y] = \int_a^b L(t, y, y')dt,$$

then $\tilde{y}(t)$ satisfies the modified (EL) with Weirstrass-Erdmann corner conditions satisfied at the corner points, and $E(t, y, y', p) \geq 0$ for y and p in some neighborhood of $\tilde{y}$.

Proof. [Sketch of Proof] This really works for specific values of p, which requires more background. Recall that for an extremal to exist, the (EL) are satisfied, which generically yields a 2nd-order differential equation. Without imposing boundary conditions, the (EL) give a 'Field of extremals' $y_\alpha(t)$ parameterized by α. The slope of one of those solutions is given by $p(t, y_\alpha)$. Since $y_\alpha(t)$ is an extremal then

$$\frac{\partial L}{\partial y} - \frac{d}{dt}\frac{\partial L}{\partial p} = 0,$$

where $L = L(t, y, p)$.

For $y(t)$ near the strong minimum $\tilde{y}(t)$, define the functional (Hilbert's invariant integral)

$$K[y] = \int_a^b \left[L(t, y, p) + (y' - p)\frac{\partial L}{\partial p}(t, y, p) \right] dt. \tag{17.8}$$

Note that if $y' = p$, then $K[y] = J[y]$. This can be rewritten as a path integral,

$$K[y] = \int_C [udt + vdy],$$

where

$$u(t, y) = L(t, y, p) - p\frac{\partial L}{\partial p}(t, y, p), \qquad v(t, y) = \frac{\partial L}{\partial p}(t, y, p).$$

As we find in the homework, the path integral $K[y]$ is invariant (only depends on the endpoints and not the curve C itself.

Since $K[y]$ is independent of the path $\mathscr{C}$ (if the endpoints remain the same) then $K[y] = K[\tilde{y}]$. Thus

$$\begin{aligned} J[y] - J[\tilde{y}] &= J[y] - K[\tilde{y}] \\ &= J[y] - K[y] \\ &= \int_a^b [L(t, y, y') - L(t, y, p) - (y' - p)\frac{\partial L}{\partial p}(t, y, p)]dt \\ &= \int_a^b E(t, y, y', p)dt \geq 0. \end{aligned}$$

□

17.3.4 Sufficient conditions for strong extrema

Sufficient conditions for strong minima necessarily include the restriction of conjugate points to a set of measure zero, the Weirestrass-Erdmann corner conditions, Jacobi's condition, a strengthened Legendre condition, and a lot of functional analysis. We do not delve into the gory details here, but suffice it to say that the same ideas and concepts that were put into practice in the previous chapter come into play here, but with some slight nuances to avoid taking or considering to many derivatives of anything.

17.3.5 Final Discussion

One final topic to discuss is that strong extrema are typically rather difficult to find/compute. The basic method is to derive the Euler-Lagrange equations and then find the 'weak' solutions to them (recall that this is related to the distributional derivative). Once the weak solutions to the (EL) are found then you have to check all of their corner points to ensure that the Weirstrass-Erdmann corner conditions are satisfied. Unfortunately we did not have time to carefully discuss the weak solutions to a boundary value problem, so we will not be able to work through very many interesting examples.

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.

17.1. Calculate the second variation (2nd order Gateaux differential) of

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

where both $y(x)$ and $y'(x)$ are fixed at the endpoints. Make certain to state what class of variations you are considering.

17.2. Calculate the second variation of a $e^{J[y]}$ where $J[y]$ is a twice differentiable functional.

17.3. Recall that a surface of revolution minimizes the action integral with Lagrangian $L = \rho\sqrt{1+(\rho')^2}$ and has a solution given by $\rho(z) = C_1 \cosh\left(\frac{z+C_2}{C_1}\right)$. Using Legendre's condition show that this is indeed a minimum in this case. You can assume that $C_1 \geq 0$.

17.4. Find an extremizer of

$$J[y] = \int_0^1 (x(y')^4 - 2y(y')^3)dx$$

where $y(0) = a, y(1) = b$ are fixed.

17.5. Use Legendre's condition to derive necessary conditions on a and b for the solution of the previous problem to be a maximum or a minimum.

17.6. (Necessary versus sufficient for the Legendre condition) For the problem considered in the previous two questions, consider the case of $a = b = 0$. What is the value of the functional evaluated at the extremal found previously in this case? Also, supposing that the derivatives could be calculated nicely (just assume that the derivative is point wise defined), calculate the value of the functional for

$$y(x) = \begin{cases} x & \text{for } 0 \leq x \leq h \\ \frac{h}{h-1}(x-1) & \text{for } h \leq x \leq 1. \end{cases}$$

Evaluate this (probably numerically) for some small h, such as $h = 1/10$ and discuss what this means.

17.7. *Prove that if a function $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ is convex and in C^3 then the Hessian $D^2 f$ is positive definite.

17.8. Show that the path integral form of Hilbert's invariant integral

$$\int_C [udt + vdy],$$

is invariant (dependent only on the endpoints of the path) for $u(t,y) = L(t,y,p) - p\frac{\partial L}{\partial p}(t,y,p)$ and $v(t,y) = \frac{\partial L}{\partial p}(t,y,p)$. Recall that $L(t,y,p)$ satisfies the Euler-Lagrange equations.

17.9. Consider the problem of minimizing $J[y] = \int_{-1}^{1}(y'(x))^3 dx$ subject to $y(-1) = y(1) = 0$. Characterize all piecewise continuously differentiable externals of J.

17.10. For the previous problem, check and see which solutions satisfy the Weirstrass-Erdmann corner conditions. For each one that does, check and see if it is a minimum (using either the Weirstrass function or Legendre's condition if the extremal is weak).

17.11. Show that

$$x(t) = \begin{cases} 0, & -1 \leq t \leq 0 \\ t, & 0 \leq t \leq 1 \end{cases},$$

is a strong minimum for $J[x] = \int_{-1}^{1} x^2(1-\dot{x})^2 dt$.

Notes