Notes for Math 189 2/10/1990 R. Montgomery

Existence of minima in the calculus of variations.

We will use the methods of functional analysis to prove the existence and well-behavedness of minimizing paths for our "standard problem" in the calculus of variations. This is the problem whose Euler-Lagrange equations are Newton's equations (with a possibly variable mass matrix). The methods used are also used in modern nonlinear p.d.e., eg. nonlinear wave equations, fluid mechanics, and gauge theory, but our problem is of course significantly simpler than these.

We set

$$S(x) = \int_a^b L(x(t), \dot{x}(t), t)\, dt \quad ; \text{where} \quad L(x, \dot{x}, t) = \frac{1}{2} \langle \dot{x}, g(x,t)\dot{x} \rangle - V(x,t)$$

In the integrand $x(t) \in \mathbb{R}^n$, $a \le t \le b$, is taken to be a piecewise smooth function. The "metric" $g(x,t)$ is a symmetric postive definite nxn matrix which depends smoothly on x and t, and V, the "potential", is a smooth real-valued function of the same variables. All the essential difficulties are present in the case $n = 1$, so you may just want to assume this as you read along.

Question: Does S admit a minimum when restricted to the set of smooth paths $x: [a,b] \to \mathbb{R}^n$ which satisfy the boundary conditions

$$x(a) = x_0, \quad x(b) = x_1 \quad ?$$

If "yes", characterize the minima.

Answer: "Yes", provided that

(a) g is uniformly positive definite, and

(b) V is bounded above.

These minima are smooth, and satisfy the Euler-Lagrange equations. They may not be unique.

(a) means that there are positive constants $C > c$ such that for all $x, v \in \mathbb{R}^n$, and all $t \in [a,b]$ we have $c\|v\|^2 \leq \langle v, g(x,t)v\rangle \leq C\|v\|^2$. (b) means that there is a positive constant m such that $V(x,t) \leq m$, for all x, t. Conditions (a) and (b) are certainly not necessary. For example, (b) could be replaced by "$V(x)$ is bounded by const.$\|x\|^\alpha$ as $\|x\|$ goes to infinity, and where $\alpha < 2$". The only place where condition (b) is used is at step (1): showing that S is bounded below.

The rest of this manuscript is devoted to proving the validity of our answer. We begin by giving names to certain function spaces.

$$C_0 = \{ x: [a,b] \to \mathbb{R}^n \mid x \text{ is continuous } \}$$

$$C_1 = \{ x: [a,b] \to \mathbb{R}^n \mid x \text{ is continuously differentiable } \}$$

$$C_1[x_0, x_1] = \{ x \in C_1 \mid x(a) = x_0, x(b) = x_1 \},$$ an affine subspace of C_1.

We divide the proof into 7 steps. First we list these steps. Then we prove them.

(1). Show that S is bounded below, as a function on C_1.

(2). Take a sequence $\{x_i\}$ of paths in $C_1[x_0, x_1]$ such that $S(x_i) \to \inf S(x) = \gamma$, where the infimum is taken over $C_1[x_0, x_1]$. By step (1), $\gamma > -\infty$.

Summary of steps (3)-(5). The first goal, which will be achieved in step (5), is to show that we can find a subsequence of the sequence $\{x_i\}$ from step (2) which converges to some x_∞, and that $S(x_\infty) = \gamma$. To do this we will have to enlarge C_1 to some bigger space, denoted H_1, in order to incorporate the limit x_∞. H_1 will be the completion (in the sense of metric spaces) of C_1 with respect to the distance fuction defined by a certain norm $\|\cdot\|_1$. This norm is essentially defined by the kinetic energy term of S. (H_1 is the simplest example of a certain type of Hilbert space called a "Sobolev space".)

(3). Show that there is a subsequence of the $\{x_i\}$ which converges weakly to some x_∞. This means that it converges with respect to a certain topology on H_1 which is weaker (easier for things to converge) than that of the norm-topology. The only hypothesis needed to get this weak convergence is an inequality of the form $S(x) + \text{const.} \geq \|x\|_1^2$.

(4). Show that $S(x_\infty) \leq \inf S(x)$, where the infimum is taken over the set of x in H_1 satisfying the endpoint conditions. (This step is called proving the "weak * lower semi-continuity" of S.) The key here, and in a sense to the entire argument, is a certain inequality, the "Sobolev inequality" (simplest form) which has, as a corollary, that $H_1 \subset C_0$. Now we have our minimum, and it satisfies the boundary conditions.

(5). Show that a subsequence of the sequence in (3) converges strongly to x_∞ (i.e. in the original topology on H_1).

(6). Show that $dS(x_\infty) = 0$, as a linear functional on the subspace of H_1 corresponding to zero variations of the endpoint conditions. We say that such an x is a "weak (or distributional) solution to the Euler-Lagrange equations".

(7). Prove the fundamental lemma of the calculus of variations, in the form alluded to in class at the prompting of Mr. Lumley. This is essentially a justification of integration by parts. Now our x_∞ is a solution to the Euler-Lagrange equations, and in particular smooth. We are finally done!

(1) is immediate, since L is the sum of two terms, the first (kinetic energy) is positive, and the second is bounded below by m. Thus $S(x) \geq -m(b-a)$

(2). By definition of "infimum" we can find such a sequence.

(3). To simplify the consequent notation, we will denote the length of a vector $v \in \mathbb{R}^n$ as $|v|$. Define the norm

$$\|x\|_1^2 = |x(a)|^2 + \int_a^b |\dot{x}(t)|^2 \, dt$$

on C_1. Let H_1 denote the completion of C_1 with respect to the distance function which this norm defines. H_1 forms a Hilbert space (every Cauchy sequence converges) with the obvious inner product. It is the simplest example of a "Sobolev space".

(We recall this process of completion. For details, see any real analysis text, for example Rudin. A "Cauchy sequence" is a sequence $\{x_i\}$ such that $\|x_i - x_j\|_1 \to 0$ as $i, j \to \infty$. An element of H_1 is an equivalence class of such Cauchy sequences. This procedure of completion is the procedure by which the real numbers are created out of the rational numbers.)

Definitions. Suppose H is a real Hilbert space. A linear functional $\lambda : H \to \mathbb{R}$ is called "bounded" if there is a positive constant C such that $|\lambda(x)| \leq C \|x\|$. Example If $v \in H$, then $x \to \langle v,x \rangle$ is a bounded linear functional. (Proof: Cauchy-Schwartz inequality. The Riesz representation theorem says all bounded linear functionals have this form.)

Suppose that $\{x_i\}$ is a sequence in H. We say that "x_i converges weakly to $x_\infty \in H$" provided that for all bounded linear functionals λ we have $\lambda(x_i) \to \lambda(x_\infty)$.

Banach-Alaoglu Theorem. Every bounded sequence in a Hilbert space has a weakly convergentsubsequence.

(For a proof , see for eg. Royden's Real Analysis. p. 202. The theorem is a direct consequence of the Tychonoff product theorem.)

We will apply this theorem to our minimizing sequence $\{x_i\}$. By assumptions (a) and (b) of our "answer" we have

$$S(x) \geq c \int |\dot{x}|^2 - m(b-a) \qquad [3.1.1]].$$

Whichimpliesthat

$$\int |\dot{x}_i|^2 \leq M \text{, for some constant M.} \qquad [3.1.2],$$

where a plain $\int f$ always means $\int_a^b f(t)dt$. (Proof of [3.1.2]: since we only care about the tail of the sequence, The constant M can be taken to be $[(\gamma + \varepsilon) + m(b-a)]/c$, where from some point i_0 on, $S(x_i) < \gamma + \varepsilon$, and $\gamma = \inf S$.) Moreover $|x_i(a)|^2 = |x_0|^2$ for all i. It follows that $\|x_i\|_1^2$ is bounded. The Banach-Alaolglu theorem now implies that $\{x_i\}$ has a weakly convergent subsequence $\{x_{n_i}\}$. For simplicity, we relabel this subsequence $\{x_i\}$. This finishes step (3).

So far, we have done almost no work, except that of quoting theorems, and we know next to nothing about the weak limit, x_∞.

(4) & (5). Let $\{x_i\}$ be the just-constructed weakly convergent subsequence, and x_∞ be the vector that it weakly converges to. Most of our work is based on

Lemma 1. [Simplest case of a Sobolev inequality.] If x is continuously differentiable then

$$|x(t) - x(s)| \leq \sqrt{|t-s|}\sqrt{\int_s^t |\dot{x}(t)|^2 dt} \quad .$$

Proof. $|x(t) - x(s)| = |\int_s^t \dot{x}(t)\ dt\ |$

$$\leq \int_s^t |\ \dot{x}(t)\ |\ dt$$

Now apply the Cauchy-Schwartz inequality $\int |fg| \leq \sqrt{\int f^2}\ \sqrt{\int g^2}$ to $f = 1$, $g = |\dot{x}(t)|$.
■

Corollary of Lemma 1. (Sobolev embedding theorem, simplest case.)
(i) $H_1 \subset C_0$ and the inclusion is bounded and compact.
(ii) $H_1 \subset L_2$ = paths x with $\int |x|^2 < \infty$.

Explanation. The inclusion (i) means a number of things. First, every element of H_1 (an equivalence class of Cauchy sequences, literally speaking) is represented by a unique continuous function. If we denote this function by x(t) then $\sup_t |x(t)| \leq \|\ x\ \|_1$, for some constant C not depending of x. In particular any limit in H_1 also holds in C_o. Finally (compactness) every bounded sequence in H_1 has a C_0 - convergent subsequence.

Proof of Corollary. (i). $|\ x(t)| \leq |x(a)| + |x(t) - x(a)|$

$$\leq |x(a)| + \sqrt{|t - a|}\sqrt{\int_a^t |\ \dot{x}(t)\ |^2\ dt} \qquad \text{(lemma 1)}$$

$$\leq C\|x\|_1$$

where C is a constant depending only on b-a. Thus if $\{x_i\}$ is a $\|\cdot\|_1$ -Cauchy sequence in C_1 (or in H_1) it is uniformly bounded in C_o, i.e. the supremum of $|\ x_i(t)|$ is uniformly bounded. Also, lemma 1 implies that $\{x_i\}$ is a uniformly continuous family. We recall that this means that the δ in the definition of continuity (" for every ε there is a δ ... ") depends only on $\varepsilon = |t - s|$ and not on the "time" t at which we are checking continuity, nor on the index i of the sequence. (Take $\delta = \sqrt{\varepsilon / M}$ where M is a uniform bound on $\|x_i\|_1$.) The **Arzela-AscoliTheorem** from real analysis has exactly these hypothesis, and implies that $\{x_i\}$ has a C_0 - convergent subsequence.

(ii) Integrate the inequality $|x(t)| \leq C\|x\|_1$. ■

Consequences of the embedding theorem. The Sobolev embedding lemma give us a rough idea of what the Hilbert space H_1 look like: it is (canonically embedded as) a vector subspace of C_0. In particular, it makes sense to evaluate $x \in H_1$ at $t \in [a,b]$. Thus

$$H_1[x_0, x_1] = \{x \in H_1 \mid x(a) = x_0, x(b) = x_1\}.$$

makes sense, as a well-defined, closed subspace of H_1.

Also, by the compactness part of the embedding, a subsequence of our sequence $\{x_i\}$ converges in C_0. By uniqueness of limits, we have (after taking this subsequence, and again relabelling it $\{x_i\}$) that $x_i \to x_\infty$ in C_0. Since the sequence of vectors $\{x_i(a)\}$ and $\{x_i(b)\}$ are simply the constant vectors x_0 and x_1, and the condition "$x(t) = \text{const.}$" is a closed condition in H_1, x_∞ satisfies these same boundary values, so that $x_\infty \in H_1[x_0, x_1]$.

Now

$$\gamma = \liminf S(x_i) = \inf_{x \in C_1[x_0, x_1]} S(x) = \inf_{x \in H_1[x_0, x_1]} S(x).$$

This last equality holds because S is a continuous function on H_1 and because C_1 is dense in H_1 (and hence $C_1[x_0, x_1]$ is dense in $H_1[x_0, x_1]$).

Exercise: Check these statements!

Therefore

$$\lim S(x_i) \leq S(x_\infty).$$

However, we do not know that $S(x_\infty) = \lim S(x_i)$, because the x_i only converge weakly to x_∞, and it is much easier than for a sequence to converge weakly than in the H_1 norm. We will now do a series of calculations which simultaneously proves that, in fact, $S(x_\infty) = \lim S(x_i)$, and that $x_i \to x_\infty$ in the H_1 norm.

$$S(x_i) - S(x_\infty) = \int K_i - K_\infty - (V_i - V_\infty),$$

where the shorthand notation is self-evident. We note that

$$< \dot{x}_i - \dot{x}_\infty , (g_\infty)(\dot{x}_i - \dot{x}_\infty) > = < \dot{x}_i , (g_\infty)\dot{x}_i > - 2 < \dot{x}_i , (g_\infty)\dot{x}_\infty) > + < \dot{x}_\infty , (g_\infty)\dot{x}_\infty >,$$

and

$$2 < \dot{x}_i - \dot{x}_\infty , (g_\infty)\dot{x}_\infty > = 2 < \dot{x}_i \ (g_\infty)\dot{x}_\infty > - 2 < \dot{x}_\infty , (g_\infty)\dot{x}_\infty >,$$

where we suppressed the t-dependence of these functions. Thus

$$K_i - K_\infty = < \dot{x}_i - \dot{x}_\infty , (g_\infty)(\dot{x}_i - \dot{x}_\infty) > - 2 < \dot{x}_i - \dot{x}_\infty , (g_\infty)\dot{x}_\infty > + < \dot{x}_i , (g_i - g_\infty)\dot{x}_i > .$$

Claim. As $i \to \infty$ we have

(A). $\int < \dot{x}_i - \dot{x}_\infty , (g_\infty)\dot{x}_\infty > \to 0$

(B) $\int < \dot{x}_i , (g_i - g_\infty)\dot{x}_i > \to 0$

(C) $\liminf \int V_i \geq \int V_\infty$.

Suppose, for the moment, that we can prove these claims. Then

$$\liminf (S(x_i) - S(x_\infty)) = \liminf \int < \dot{x}_i - \dot{x}_\infty , (g_\infty)(\dot{x}_i - \dot{x}_\infty) > + \liminf \int (V_i - V_\infty)$$

[equation 4.1].

and the right hand side is the sum of two non-negative terms.

Corollary 2. $\lim S(x_i) \geq S(x_\infty)$.

Since have already shown that $\lim S(x_i) \le S(x_\infty)$, we now have

$$\lim S(x_i) = S(x_\infty) \quad .$$

This completes step (4), the proof of "the weak -* -lower semicontinuity of S at x_∞", modulo proving the claims (A) - (C).

Corollary 3. , and $x_i \to x_\infty$ in H_1 , and in C_0 .

Proof of Cor. 3. (Cor. 2 is immediate). "$x_i \to x_\infty$ in H_1" means that $\| x_i - x_\infty \|_1 \to 0$, which in turn means that

$$\text{(i) } |x_i(a) - x_\infty(a)| \to 0, \text{ and (ii) } \int |\dot{x}_i - \dot{x}_\infty|^2 \to 0.$$

We have already proved that (i) holds. (See **Consequences of the embedding lemma.**) To see that (ii) holds, use the fact that both terms on the right of equation 4.1 are non-negative, and the fact, just proven, that $\lim(S(x_i) - S(x_\infty)) = 0$, to conclude that

$$\lim \int < \dot{x}_i - \dot{x}_\infty, (g_\infty)(\dot{x}_i - \dot{x}_\infty) = 0$$

Finally use the bound, (assumption (a) on the metric), that

$$\int < \dot{x}_i - \dot{x}_\infty, (g_\infty)(\dot{x}_i - \dot{x}_\infty) > \ \ge c \int | \dot{x}_i - \dot{x}_\infty|^2 . \quad ■$$

It remains to prove the claims (A), (B), (C).

Proof of claim (A). $\lambda(v) = \int <v, (g_\infty)\dot{x}_\infty) >$ is a linear functional on H_1 . It is bounded by assumption (a) on the metric, and the Cauchy-Schwartz inequality. In fact $\lambda(v) \ge \sqrt{c} \| v \|_1 \| x_\infty \|$. Thus $\lambda(x_i) - \lambda(x_\infty) \to 0$, by the definition of weak convergence. This is the limit claimed in (A). ■

Proof of claim (B). The embedding lemma, part (i) implies that (after taking a subsequence of the $\{x_i\}$, and then relabelling it $\{x_i\}$) $x_i \to x_\infty$ uniformly in C_0 . This implies, by the continuity and uniform boundedness of g, that $g_i \to g_\infty$ in C_0 ([a,b]; matrices). We can use the "operator norm" $|A| = \sup_{v \ne 0} |Av|/|v|$, on n x n matrices, so that the last statement means $\sup_t |g(x_i(t), t) - g((x_\infty(t), t))| \to 0$. The operator norm satisfies $|Av| \leq |A|\,|v|$, for all vectors v. Set $\|g_i - g_\infty\|_b = \sup_t |g(x_i(t), t) - g((x_\infty(t), t))|$. Thus

$$< \dot{x}_i(t), (g_i - g_\infty)\dot{x}_i(t)> \quad \leq \|g_i - g_\infty\|_b\, |\dot{x}_i(t)|^2,$$

so that , upon consulting equation [3.1.2], we see that

$$\int < \dot{x}_i, (g_i - g_\infty)\dot{x}_i > \;\leq\; \|g_i - g_\infty\|_b \int |\dot{x}_i(t)|^2 \;\leq\; \|g_i - g_\infty\|_b M.$$

and this last sequence goes to 0 , as we just mentioned. ■

Alternately: show the closure of the point set $\{x_i(t)\}$ in $\mathbb{R}^n$ is compact. Hence get a uniform bound on $|V_i|$, so that $|V_i|$ is bounded below. on (for $a \leq t \leq b$)

Proof of claim (C). Since $x_i \to x_\infty$ pointwise, and V is continuous, we have that $V_i \to V_\infty$ pointwise. If we knew that V were bounded below also, the same proof as for claim B would have worked . Since we don't , we will have to use some measure theory. "Recall" **Fatou's lemma** (which is 90% of the Lebesque dominated convergence theorem) : If f_i are a sequence of non-negative integrable functions, and $f_i \to f$ pointwise, then $\liminf \int f_i \geq \int f$. Apply the lemma to $f_i = m - V_i$, where m is the upper bound on V. ■

Step 6. Proposition(step 6). For all $\delta x \in H_1$, with $\delta x(a) = \delta x(b) = 0$, we have $dS(x_\infty) \cdot \delta x = 0$.

We will begin by assuming that the derivative in fact exists, and will prove the proposition by contradiction. After that, we will prove that the derivative exists for all $x \in H_1$, and is given by our usual formula, before integration by parts.

Proof (by contradiction): Assume the contrary. Then there exists a δx, with $\delta x(a) = \delta x(b) = 0$ such that $dS(x_\infty) \cdot \delta x \neq 0$. We can assume , by multiplying δx by -1, if need be, that $dS(x_\infty) \cdot \delta x < 0$. Note that for each ε , the path $x + \varepsilon\delta x$ satisfies the boundary conditions, and so is in the affine space $H_1[x_0, x_1]$. Apply the mean value theorem to the differentiable function $S(x + \varepsilon\,\delta x) - S(x)$ to conclude that there is an ε_* with $S(x + \varepsilon_*\,\delta x) < 0$. This contradicts the fact , proved in corollary 2 above, that x_∞ minimizes S among all x in $H_1[x_0, x_1]$. ■

Existence of the derivative (a bit more carefully than in class).

$$S(x+\varepsilon\,\delta x) - S(x) = \int \Delta_1 K + \Delta_2 K + \Delta V$$

The integrands are functions on the interval [a,b] and have the form

$$2\Delta_1 K(t) = \langle \dot{x} + \varepsilon\,\delta\dot{x}\,,\, g(x)(\dot{x} + \varepsilon\delta\dot{x})\rangle - \langle \dot{x}\,,\, g(x)\,\dot{x}\rangle$$

$$2\Delta_2 K = \langle \dot{x} + \varepsilon\,\delta\dot{x}\,,\, [g(x+\varepsilon\,\delta x) - g(x)](\dot{x} + \varepsilon\delta\dot{x})\rangle$$

$$\Delta V = V(x+\varepsilon\,\delta x) - V(x)$$

Now $2\Delta_1 K = 2\varepsilon\langle \dot{x}\,,\, g(x)\,\delta\dot{x}\rangle + \varepsilon^2\langle \delta\dot{x}\,,\, g(x)\,\delta\dot{x}\rangle$

$$2\,\Delta_2 K = \varepsilon\langle \dot{x}\,,\, dg(x)\cdot\delta x\,)\,\dot{x}\rangle + o(\varepsilon)$$

$$\Delta V = dV(x)\cdot\delta x + o(\varepsilon)\,.$$

Here $o(\varepsilon)$ means any function $f(t,\varepsilon)$ satisfying the inequality $|f(t,\varepsilon)|/\varepsilon \to 0$ uniformly in t as $\varepsilon \to 0$. These $o(\varepsilon)$ terms come from Taylor's theorem with remainder, and the assumption that g and V are twice-continuously differentiable.

It follows that

$$dS(x)\cdot\delta x = \int \langle \dot{x}\,,\, g(x)\,\delta\dot{x}\rangle + \langle \dot{x}\,,\,(dg(x)\cdot\delta x\,)\,\dot{x}\rangle + dV(x)\cdot\delta x\,. \quad \blacksquare$$

Step 7. We begin by proving the following version of the

Fundamental lemma of the calculus of variations. Let p and f be absolutely integrable functions on the interval [a,b] with values in $\mathbb{R}^n$ (meaning that the length of the corresponding vectors in $\mathbb{R}^n$ are integrable functions). Suppose that

$$\int \langle p,\delta\dot{x}\rangle + \langle f,\delta x\rangle = 0 \text{ for all } \delta x \in H_1\,, \text{ with } \delta x(a) = \delta x(b) = 0. \qquad [*]$$

Then $dp/dt = f$ (almost everywhere).

Proof. It suffices to prove this for the case $n = 1$, since in the general case the integral is a sum of n independent such terms. It suffices to assume that $a = 0$, $b = 1$, by the translation invariance and scaling properties of integration. It suffices to prove the lemma for the smaller class of "test functions" δx which have the form $1-\cos(2\pi nt)$, and $\sin(2\pi nt)$, for $n = 1,2, \ldots$. That is, we will only assume that [*] holds for every δx in this small subclass of the functions in H_1 which vanish at the endpoints, but still arrive at the same conclusion. Finally it suffices to prove just the case $f = 0$. To see this, replace p by $P = p - F$, where $F(t) = \int_0^t f(x)\, dx$. P is still integrable, and we can integrate $- F\,\delta\dot{x}$ by parts, since both factors are differentiable. Then $\int P\,\delta\dot{x} = \int p\,\delta\dot{x} + \int f\,\delta x$ and $\frac{d}{dt}P = \frac{d}{dt}p - f$.

So, assume that $\int_0^1 P(t)\delta\dot{x}(t)\, dt = 0$

for all δx with the above trigonometric form. This is the statement that

$$\int_0^1 P(t)\sin(2\pi nt)\, dt = \int_0^1 P(t)\cos(2\pi nt)\, dt = 0, \text{ for } n = 1,2,3,\ldots .$$

Thus all the Fourier coefficients of P vanish except its mean, $A_0 = \int_0^1 P(t)dt$. Since an absolutely integrable function is uniquely determined by its Fourier series, we have $P(t) = A_0$, for all t. ■

According to the fundamental lemma, our weak equations: $dS(x)\cdot\delta x = 0$ for all $\delta x \in H_1$ vanishing at the endpoints imply the validity of the differential equations, which we write in the schematic form: $\frac{d}{dt}(g\,\dot{x}) = - dg\,\dot{x}^2 - dV$. The right hand side is only a measurable function. We do not even know that $\dot{x}$ exists everywhere. Integrate this equation to obtain

$$\dot{x}(t) = g^{-1}\int_a^t \{- dg\,\dot{x}^2 - dV\} + \text{const.}.$$

The right hand side is now a continuous function. This is because: (a) the indefinite integral of a measurable function is continuous, and (b) x(t) is continuous (Sobolev embedding), and thus so is $g(x(t),t)^{-1}$. Hence the left hand side, $\dot{x}$ is continuous everywhere! We apply the same argument again, using this new information about $\dot{x}$, (i.e. we notice that the indefinite integral of a continuous function is a differentiable function whose derivative is the integrand) to conclude that $\dot{x}$ is in fact twice continuously differentiable. Now the equation holds in a completely classical sense: all functions in it are continuous. Finally we have that $x = x_\infty$ is a classical solution to the Euler-Lagrange equations.

We can now keep differentiating the equation (or integrating if you prefer that argument) to obtain that x is smooth.

End Note.

The proofs of some of the steps required the following "big guns" from analysis: the Banach-Alaoglu theorem, the Arzela-Ascoli theorem, and Fatou's lemma. It also required the notion of completing a vector space with respect to a norm, and, at the end, some knowledge of Fourier series. A proof of the fundamental lemma of calculus of variations not relying on Fourier series can be found in L.C. Young's book "Optimal Control". He takes his "test functions" δx to be piecewise linear functions instead of trig functions.

If you want to understand the entire proof better, go back and reprove the theorem in the case $V(x,t) \leq A + M|x|$ for some constants A, M. Initially, you will fail at step 1, but you can use the inequalities from step 4 to show that S is in fact bounded below.