Differential Geometry—MTG 6256—Fall 1999
Problem Set 2: Introduction to the Calculus of Variations

Later this semester or next semester we will be discussing geodesics on Riemannian manifolds. Geodesics are paths that minimize arclength among all paths between two specified points. In essence one analyzes this problem by doing calculus on the (infinite-dimensional) space of paths, arriving at a differential equation the geodesic must satisfy. This procedure is an example of the *calculus of variations*, which this problem set introduces.

In Lagrangian mechanics one studies "action functionals" of the form

$$S(f) = \int_a^b L(t, f(t), f'(t))\, dt,$$

where $f : [a, b] \to \mathbf{R}^n$ ($n = 1, 2,$ or 3) is meant to represent the motion of a particle between times a and b (usually with the positions $f(a)$ and $f(b)$ fixed), and where L, the *Lagrangian*, is a real-valued function which, after plugging in f and f', gives the difference between the kinetic and potential energies associated with the motion $t \mapsto f(t)$. Hamilton's *principle of least action* asserts that the physical trajectory of a particle is one that minimizes $S(f)$ over all f with the given endpoint constraints. Historically, inventors of this subject realized that at an extremum of such a problem (called a *variational problem*), the functional S must be stationary with respect to small changes in f (dubbed *variations*, whence the terms "variational problem" and "calculus of varations"). It was recognized early that this situation was formally similar to that of finding an extremum of a function of one variable (or finitely many), but it was not realized for some time that what was being done was not merely *similar* to critical-point calculus, but *identical* to it—once one realizes that the correct home for f is an appropriate Banach space, and has the right definition of derivative.

The problems below take you through the solution of this variational problem in the language of calculus on Banach spaces. To simplify the writing, I have assumed that the motion is one-dimensional and that L has no explicit t-dependence. The more general situation is no harder, just slightly more cumbersome to write down.

Setup. Let $[a, b]$ be a closed finite interval in $\mathbf{R}$. Recall that the *uniform norm* or *sup-norm* of a function $f : [a, b] \to \mathbf{R}$ is defined to be $\|f\|_\infty = \sup_{t\in[a,b]} |f(t)|$. For $k \geq 0$, the C^k*-norm* of a C^k function f is defined to be

$$\|f\|_{C^k} = \|f\|_\infty + \|f'\|_\infty + \ldots + \|f^{(k)}\|_\infty,$$

where $f^{(k)}$ denotes the k^{th} Calculus-1-derivative of f (using one-sided derivatives at the endpoints). Convergence in the sup-norm (or C^0-norm) is identical to uniform convergence. Recall from your undergraduate advanced calculus class that (1) the pointwise limit of a uniformly convergent sequence of continuous functions is continuous, and (2) if $\{f_n\}$ is a sequence of functions that converges at at least one point and for which $\{f_n'\}$ converges uniformly on $[a,b]$, then $\{f_n\}$ converges uniformly on $[a,b]$ and $(\lim f_n)'(t) = \lim(f_n'(t))$ for all $t \in [a,b]$. Let $C^k([a,b])$ denote the space of C^k real-valued functions on $[a,b]$. Facts (1) and (2) imply that for $k \geq 0$, $C^k([a,b])$ is complete in the C^k-norm, hence is a Banach space.

1. A *critical point* of a real-valued differentiable function F defined on some open subset of a Banach space is a point p for which $D_pF = 0$. Prove relative extrema of differentiable functions occur only at critical points.

For the remaining problems, an interval $[a,b]$ is fixed.

2. Let $L : \mathbf{R}^2 \to \mathbf{R}$ be C^∞. (Note: C^1 would actually suffice for this problem, and C^2 would suffice for problem 3.) Let $\partial_1 L$ and $\partial_2 L$ denote the partial derivatives of L with respect to its first and second variables (in terms of the notation used in class, at a point $(x,y) \in \mathbf{R}^2$ these would be $(D^{[1]}_{(x,y)}L)(1)$ and $(D^{[2]}_{(x,y)}L)(1)$ respectively).

(a) Define $\iota : C^1([a,b]) \to C^0([a,b] \times C^0([a,b])$ by $\iota(f) = (f, f')$. Show that ι is a continuous linear map, hence C^∞. What is $D_f\iota$?

(b) Define $\mathcal{L} : C^0([a,b]) \times C^0([a,b]) \to C^0([a,b])$ by $\mathcal{L}(f,g)(t) = L(f(t), g(t))$. Compute the directional derivatives of $\mathcal{L}$, and use this to show that $\mathcal{L}$ is continuously differentiable.

(c) Define $\mathcal{I} : C^0([a,b]) \to \mathbf{R}$ by $\mathcal{I}(f) = \int_a^b f(t)\, dt$. Show that $\mathcal{I}$ is a continuous linear map, hence C^∞. What is $D_f\mathcal{I}$?

(d) For $f \in C^1([a,b])$, define

$$S(f) = \int_a^b L(f(t), f'(t))\, dt.$$

Note that $S(f) = \mathcal{I} \circ \mathcal{L} \circ \iota(f)$. Using the Chain Rule and parts (a)-(c), show that $S : C^1([a,b]) \to \mathbf{R}$ is C^1, and compute the directional derivative $(D_fS)(h)$ for arbitrary $f, h \in C^1$.

3. Notation as in problem 2. (a) If f is C^2 rather than merely C^1, the Chain Rule implies that $\frac{d}{dt}(\partial_2 L(f(t), f'(t)))$ is defined and continuous. Show that for $f, h \in C^2([a,b])$,

$$(D_fS)(h) = h(t)\partial_2 L(f,f')(t)|_{t=a}^{t=b} + \int_a^b (\partial_1 L(f,f')(t) - \frac{d}{dt}\partial_2 L(f,f')(t))h(t)\, dt. \quad (1)$$

(b) Consider the restriction of S to the space of C^2 paths going from a point P on the real line to a point Q in time interval $[a,b]$ (i.e. we restrict the domain of S to the set $M = \{f \in C^2([a,b]) \mid f(a) = P,\ f(b) = Q\}$). Note that M is a translate of the vector subspace $V = \{f \in C^2([a,b]) \mid f(a) = 0 = f(b)\}$; given any $f_0 \in M$ we can write any $f \in M$ uniquely in the form $f_0 + g$ where $g \in V$. Fix such an f_0, and for $g \in V$ define

$$\tilde{S}(g) = S(f_0 + g).$$

Then the constrained minimization problem, minimizing S over the subset M—the set of paths with the given endpoints—is equivalent to minimizing $\tilde{S}$ over V. Since convergence in C^2 implies pointwise convergence, any the limit f of C^2-convergent sequence of functions in V still has $f(a) = 0 = f(b)$, so V is a closed subset of $C^2([a, b])$, hence complete. Therefore V, with the C^2-norm, is a Banach space. Show (quickly, using results above) that $\tilde{S}$ is C^1 on V, and that $(D_g\tilde{S})(h) = (D_{f_0+g})S(h)$. Since $h \in V$, to what formula does (1) simplify?

(c) Prove that $g \in V$ is a critical point of $\tilde{S}$ iff

$$\partial_1 L(f, f') - \frac{d}{dt}\partial_2 L(f, f') \equiv 0, \tag{2}$$

where $f = f_0 + g$. Equation (2) is known as the *Euler-Lagrange equation*; physicists usually write it as

$$\frac{\partial L}{\partial f} - \frac{d}{dt}\frac{\partial L}{\partial f'} = 0.$$

The preceding arguments prove that any minimum of the constrained minimization problem must be a solution of (2). (The converse, of course, is not true in general, since not every critical point is a minimum.)