MORSE THEORY ON LOOP SPACES

ISHAN LEVY

1. Setup and Critical Points

A natural object of study in algebraic topology is the space of loops ΩX of a space, which is adjoint to the suspension Σ, so in particular we have nice properties such as $\pi_n\Omega X = \pi_{n+1}X$. Instead of studying the loop space, we may instead like to study the space of paths between two points on X, which is homotopy equivalent to the loop space if X is path connected. Our goal is to use Morse theory to learn about a CW-structure on the loop space ΩM of a smooth manifold M. To begin with, instead of considering the entire loop space, we will consider $\Omega^* M$, the space (to be topologized later) of piecewise smooth paths from p to q, two fixed points on M.

We should think about $\Omega^* M$ as an "infinite dimensional manifold", but to avoid doing Morse theory on it directly we will ultimately use instead a finite dimensional approximation of it. Our Morse function will be the energy function $E(\gamma) = \int \|\gamma'\|^2$. Analogously to the finite dimensional case, the tangent space T_γ of a path γ will be the space of vector fields on γ that vanish at the end points, as these correspond to the derivatives of variations.

The first thing we need to do is figure out what the critical points of E are. To do this, we need to compute the derivative of the E:

Lemma 1.1 (First Variation Formula)**.** *Let t_i be a subdivision on which the vector field $W \in T_\gamma$ is smooth, and let $\Delta_i\gamma'$ denote the jump of γ' at t_i. Then* $E'(W) = -2\sum_i \langle W, \Delta_i\gamma'\rangle - 2\int_0^1 \langle W, \gamma''(t)\rangle$

Proof. Let $\gamma_u(t)$ be a variation that corresponds to W. Then $E'(W) = \frac{d}{du}E(\gamma_u(t)) = \int_0^1 \frac{d}{du}\langle \frac{d}{dt}\gamma_u(t), \frac{d}{dt}\gamma_u(t)\rangle$

$$= 2\int_0^1 \langle \frac{D}{du}\frac{d}{dt}\gamma_u(t), \frac{d}{dt}\gamma_u(t)\rangle = 2\sum_i \int_{t_i}^{t_{i+1}} \langle \frac{D}{dt}\frac{d}{du}\gamma_u(t), \frac{d}{dt}\gamma_u(t)\rangle$$

$$= -2\sum_i \langle \frac{d}{du}\gamma_u(t), \Delta_i \frac{d}{dt}\gamma_u(t)\rangle - 2\int_0^1 \langle \frac{d}{du}\gamma_u(t), \frac{D}{dt}\frac{d}{dt}\gamma_u(t)\rangle$$

which yields the formula setting $u = 0$. □

Date: 12/8/17.

Corollary 1.2. *The critical points of E are the geodesics.*

Proof. By Lemma 1.1, geodesics are critical points. Conversely if γ is critical, choosing W in the direction of γ' but vanishing at the discontinuities, we γ is piecewise geodesic, but then by choosing W in the direction of discontinuities of γ', we see indeed there are no jumps in γ' so by uniqueness of solutions to ODEs, γ is a geodesic. $\square$

2. The Hessian and its null space

Next we would like to make a notion of the Hessian at a critical point, and study its null space and index. Then, analogously to the finite dimensional case, if $\gamma_{u,w}(t)$ is a 2-parameter variation, with $\frac{d}{du}\gamma_{u,w}(t) = W_2$, $\frac{d}{dw}\gamma_{u,w}(t) = W_1$ at $(u,w) = (0,0)$, define the Hessian as $H(\gamma_{u,w}(t)) = H(W_1, W_2) = \frac{d}{du}\frac{d}{dw}E(\gamma_{u,w}(t))$. This is clearly symmetric bilinear, and to see that it is well defined, we can compute a formula in terms of W_1 and W_1.

Lemma 2.1 (Second Variation Formula)**.** *Let $\gamma_{u,w}(t)$ be a 2-parameter variation as above, and let t_i be the discontinuities of $\frac{D}{dt}W_2$. Then $H(W_1, W_2) = -2\sum_i \langle W_1, \Delta_i \frac{D}{dt}W_2\rangle - \int_0^1 \langle W_1, \frac{D^2}{dt^2}W_2 + R(\gamma', W_2, \gamma')\rangle$.*

Proof. From Lemma 1.1 using the fact that γ is geodesic, at $w = 0$ we get:

$$H(W_1, W_2) = -2\sum_i \langle \frac{d}{dw}\gamma_{u,0}, \Delta_i \frac{D}{du}\frac{d}{dt}\gamma_{u,0}\rangle - 2\int_0^1 \langle \frac{d}{dw}\gamma_{u,0}, \frac{D}{du}\frac{D}{dt}\frac{d}{dt}\gamma_{u,0}(t)\rangle$$

which after rearranging, and commuting the $\frac{D}{du}$ and $\frac{D}{dt}$ we get the formula. $\square$

Now a **Jacobi field** on a geodesic γ' is a vector field satisfying the second order linear differential equation called the Jacobi equation, $\frac{D^2}{dt^2}J + R(\gamma', J, \gamma') = 0$. As it is second order linear, the space of solutions is $2n$ dimensional. The proof of Corollary 1.2 gives:

Corollary 2.2. *The Jacobi fields are exactly the null space of H.*

We say that p, q on γ are **conjugate** of multiplicity $\nu > 0$ if ν is the dimension of the space of Jacobi fields vanishing at p, q. Thus γ is nondegenerate if the endpoints are nonconjugate. We can characterize Jacobi fields in another way: namely they arise from variations through geodesics.

Proposition 2.3. *Jacobi fields are exactly those arising from variations through geodesics.*

Proof. Given a variation through geodesics $\gamma_u(t)$, $0 = \frac{D}{du}\frac{D}{dt}\frac{d}{dt}\gamma_u(t) = \frac{D^2}{dt^2}\frac{d}{du}\gamma_u(t) + R(\frac{d}{dt}\gamma_u(t), \frac{d}{du}\gamma_u(t), \frac{d}{dt}\gamma_u(t))$ which shows the Jacobi equation is satisfied. Now to show all Jacobi fields arise as such, it suffices to do this on an arbitrarily small piece of γ, and then note that the variations of geodesics can be locally uniquely extended by compactness of $[0,1]$ and that the resulting extension will yield the unique extension of the Jacobi equation by uniqueness of solutions to ODEs. Then in a small enough neighborhood of $\gamma(0)$ containing $\gamma(\epsilon)$, all geodesics are minimal, so in particular, we can take the variation that takes the unique geodesics between two paths touching $\gamma(0)$ and $\gamma(\epsilon)$ with prescribed derivatives at $\gamma(0)$ and $\gamma(\epsilon)$. This yields $2n$ linearly independent solutions of the Jacobi equation so it must yield all of them. □

We then easily get the next proposition that guarantees the existence of nonconjugate points:

Proposition 2.4. *The nullity of* $d\exp_p(v)$ *is the multiplicity of conjugacy of* p *and* $\exp_p(v)$ *along the geodesic* $\exp_p(v)$.

Proof. Given a curve $\beta(u)$ at v, we can consider the variation through geodesics $\exp_p(t\beta(u))$. By restricting to an interval where geodesics are minimal, we see that this yields all Jacobi fields that vanish at p by choosing different $\beta'(0)$. In particular, the Jacobi field vanishes at the ends iff $d\exp_p(\beta'(0)) = 0$, so we get the result. □

Corollary 2.5. *There are points nonconjugate along any geodesic.*

Proof. This follows from the previous proposition and Sard's Theorem. □

From Proposition 2.4, we can see that the multiplicity of conjugacy between two points on a geodesic can be at most $n-1$. Indeed by choosing $\beta'(0)$ pointing away from the origin, the resulting variation will have nonzero derivative at one end point. Alternatively one can directly see that the corresponding vector field $t\gamma'(t)$ indeed satisfies the Jacobi equation but doesn't vanish at the second endpoint.

Indeed we can see in the case of S^n, that two antipodal points are conjugate of multiplicity $n-1$, as the space of minimal geodesics is a copy of S^{n-1} inside the tangent space. In fact one can exploit this fact to prove the Freudenthal Suspension Theorem.

3. Index Theorem

Now given a critical point, we would also like to be able to compute its index. Here γ is a critical point from p to q. We have the following theorem:

Theorem 3.1 (Index Theorem)**.** *The index of* γ *is the number of points conjugate to* p *counted with multiplicity. Moreover this index is finite.*

Proof. First we will approximate the tangent space T_γ by a finite dimensional subspace. Namely, choose a subdivision of γ, t_i, such that γ restricted to each subinterval is in a neighborhood of minimal geodesics. Then we will split T_γ as $T^\gamma_{t_i} \oplus T'$, where

T_{t_i} consists of piecewise Jacobi fields on each interval, and T' consists of vector fields that vanish at the t_i. Indeed given tangent vectors at each of the t_i, there is a unique piecewise Jacobi field with those prescribed tangent vectors since the geodesic is minimal, so T_γ does split as such, and we also have $T^\gamma_{t_i} \cong \bigoplus_1^{n-1} TM_{\gamma(t_i)}$ as a result.

Since γ is a minimal geodesic on each subinterval, and variations corresponding to elements of T' can be chosen to fix the endpoints, γ is a minimum of each of E on each of these variations, so H is positive semidefinite on T'. Now if $H(W, W) = 0$ for $W \in T_i$, then W is in the null space of H by Lemma 2.1, so is a Jacobi field, but then it must be 0. Thus H is positive definite on T' so we may focus our attention on T_{t_i}. In particular, we already get that the index is finite.

Define $\lambda(t)$ to be the index of γ restricted to the subinterval $[0, t]$. It is clear that $\lambda(t)$ is monotonically increasing as negative definite subspaces can be extended as 0 to further along the curve. Moreover, near 0, since γ is minimal, $\lambda(t) = 0$ as well. Now fixing a point τ and choosing a subdivision with $t_i < \tau < t_{i+1}$, note that $T^{\gamma|_{[0,t]}}_{t_i} \cong \bigoplus_1^{i} TM_{\gamma(t_i)}$ is not dependant on t for $t_i < t < t_{i+1}$. Then H varies continuously on $T^{\gamma|_{[0,t]}}_{t_i}$ near τ. Then we get that $\lambda(\tau - \epsilon) \geq \lambda(\tau)$ for small ϵ, but by monotonicity, this must be an equality. Indeed since we also know the null space is dimension ν, we also get $\lambda(\tau + \epsilon) \leq \lambda(\tau) + \nu$, and to complete the proof it suffices to prove the reverse inequality.

Now let $Y_1, \dots, Y_{\lambda(t)}$ be orthonormal and linearly independent such that H is negative definite on them, and extend them to $t + \epsilon$ be setting them to 0. Let $W_1, \dots, W_\nu$ be linearly independent Jacobi fields that vanish at t, and extend them as well to $t + \epsilon$ as 0. Then choose X_i such that $\langle X_i(t), W_j(t)\rangle = \frac{1}{2}\delta_{ij}$. Looking at the Lemma 2.1 we get $H(W_i, X_j) = \delta_{ij}, H(W_i, Y_j) = 0$. We can write H then as a matrix on the span of $Y_1, \dots, Y_{\lambda(t)}, cX_1 - c^{-1}W_1, \dots, cX_\nu - c^{-1}W_\nu$ as

$$\begin{pmatrix} -I_{\lambda(t)} & cB \\ cB^t & -I_\nu + c^2 A \end{pmatrix}$$

which is certainly negative definite for small c. □

4. Approximating the loop space

Now we would like to use these results to learn about the topology of the loop space ΩM. First we would like to say that ΩM is homotopy equivalent to our space of piecewise smooth paths $\Omega^* M$. To do this, first topologize $\Omega^* M$ with the metric $d(\gamma_1, \gamma_2) = \sup_t \rho(\gamma_1(t), \gamma_2(t)) + \int_0^1 \|\gamma_1' - \gamma_2'\|^2$ where the second term is to ensure that E is continuous, and ρ denotes the metric from the Riemannian metric.

Proposition 4.1. *The inclusion $i : \Omega^* M \to \Omega M$ is a homotopy equivalence.*

Proof. We cover M by geodesically convex neighborhoods, i.e. neighborhood such that any two points are connected by a unique minimal geodesic that lies entirely inside the neighborhood. Now let ΩM_i be the collection of paths such that the $[\frac{k}{i}, \frac{k+1}{i}]$ interval lies entirely inside one of these neighborhoods, and let $\Omega^* M_i = i^{-1}\Omega M_i$. Then $\Omega^* M$ is the homotopy direct limit of the $\Omega^* M_i$, and likewise for ΩM, so it suffices to show that the inclusion $i : \Omega^* M_i \to \Omega M_i$ is a homotopy equivalence. To show this, our homotopy inverse j will take a path an return the unique geodesic that agrees on each $\frac{k}{i}$. $i \circ j$ is homotopic to the identity as given a path, we can continuously deform it to $i \circ j$ in a natural way: namely at the point $\frac{k}{i} < r < \frac{k+1}{i}$ let $H_r(\gamma)(t)$ be $\gamma(t)$ for $t > r$, on $[\frac{l}{i}, \frac{l+1}{i}]$ and on $[\frac{k}{i}, r]$, the unique minimal geodesic between those points. A similar construction shows that $j \circ i$ is homotopic to the identity. □

Thus we can use $\Omega^* M$ to study the loop space. However, we would like to approximate $\Omega^* M$ further, so that we can do Morse theory on it. Namely, let $\Omega^c = E^{-1}[0, c]$, and $\operatorname{Int}\Omega^c = E^{-1}[0, c)$.

Theorem 4.2. *Assume M complete.* $\operatorname{Int}\Omega^c$ *can be approximated by a finite dimensional manifold B in a natural way. E will be a smooth function on B, The tangent space of B will be the T_{t_i}, and the critical points and Hessian, will be as before. Moreover* $\operatorname{Int}\Omega^c$ *deformation retracts onto B.*

Proof. The metric ball of radius $\sqrt{c}$ is compact as M is complete, and paths in $\operatorname{Int}\Omega^c$ lie there by the Cauchy Schwarz inequality. Thus it has a positive injectivity radius ϵ, and we can make a subdivision t_i such that $t_{i+1} - t_i < \frac{\epsilon^2}{c}$. By the Cauchy Schwarz inequality again, we get that every element of $\operatorname{Int}\Omega^c$ is length $< \epsilon$ in each interval $[t_i, t_{i+1}]$. Then we can let B be the subspace of piecewise geodesics on this subdivision. It is naturally a manifold as each path in B is determined by its values on the t_i, so B embeds as an open subset of $M \times \cdots \times M$. Now we immediately see then that its critical points are unbroken geodesics, and that its tangent space at these critical points is exactly the T_{t_i} from the proof of the Index Theorem. A similar construction to Proposition 4.1 shows that B is a deformation retract of $\operatorname{Int}\Omega^c$. □

The previous theorem implies that we can do our usual Morse theory on the finite dimensional approximation to get results about the full loop space. As usual, we get:

Theorem 4.3 (Fundamental Theorem)**.** *Let p, q be nonconjugate along any geodesic, and M complete. Then ΩM has the homotopy type of a CW-complex with a λ-cell in each dimension for each geodesic with index λ.*

In the case of S^n, any two distinct non-antipodal points are nonconjugate, and we can see that there is a unique geodesic of index $k(n-1)$ for each k.

Corollary 4.4. ΩS^n *has the homotopy type of a CW-complex with one $k(n-1)$-cell for each $k \geq 0$.*