Lecture VI: Bott Periodicity

October 31, 2020

Our goal here is to give a proof of Bott periodicity for the unitary group. There is a similar proof for the orthogoal group but we will not discuss that case.

1 The Statement

Let $U(n) \subset GL_n(\mathbb{C})$ be the unitary group; that is to say the subgroup of matrices $A \in GL_n(\mathbb{C})$ satisfying $\overline{A}^{tr} A = \mathrm{Id}$. We have the closely related subgroup $SU(n) \subset U(n)$ of unitary matrices of determinant 1. Indeed there is a smooth, locally trivial fibration

$$SU(n) \to U(n) \xrightarrow{\det} S^1.$$

Since the natural inclusion $U(1) \subset U(n)$ gives a section of this fibration, we see

- $\det\colon U(n) \to S^1$ induces an isomorphism on π_1, so that $\pi_1(U(n)) \cong \mathbb{Z}$,
- the inclusion $SU(n) \subset U(n)$ induces an isomorphism on π_i for all $i > 1$.

There are natural inclusions $U(n) \subset U(n+1)$.

Claim 1.1. *There is a locally trivial smooth fibration*

$$U(n) \to U(n+1) \to S^{2n+1}.$$

In particular, the natural inclusion induces a map $\pi_i(U(n)) \to \pi_i(U(n+1))$ that is an isomorphism for all $i < 2n$ and a surjection on π_{2n}.

Proof. The columns of $A \in U(n+1)$ give a unitary basis for $\mathbb{C}^{n+1}$ with its usual hermitian inner product. The map $U(n+1) \to S^{2n+1}$ sends A to the unit vector in $\mathbb{C}^{n+1}$ given by the last column. The action of $U(n)$ by right multiplication on $U(n+1)$ has orbit space identified with S^{2n+1} by this map. □

This result tells us that the homotopy groups of $U(n)$ stabilize: For any i for all $n > (i/2)$, the group $\pi_i(U(n))$ is independent of n.

One way to view this is to take the direct limit U of the $U(n)$ under the natural inclusions. This is the group of complex matrices with rows and columns numbered $1, 2, \ldots$ which are the image of some element in $U(n)$ for some n under the inclusion induced by the map $C^n \subset \mathbb{C}^\infty$. Then $\pi_i(U) = \pi_i(U(n))$ for any $n > i/2$.

Bott periodicity is a computation of the homotopy groups of U, or equivalently of the homotopy groups of $\pi_i(U(n))$ for $i < 2n$.

Theorem 1.2. *(Bott Periodicity)*

$$\pi_i(U) \cong \begin{cases} 0 & \text{if } i \equiv 0 \pmod 2 \\ \mathbb{Z} & \text{if } i \equiv 1 \pmod 2. \end{cases}.$$

Furthermore, the second loop space $\Omega^2(SU)$ is homotopy equivalent to U.

2 Grassmannians

For any n, k consider the space of n-dimensional complex linear subspaces of $\mathbb{C}^{n+k}$. We denote this space by $Gr_{\mathbb{C}}(n, n+k)$. It is a manifold of dimension nk, and indeed it is the homogeneous space $U(n+k)/U(n) \times U(k)$ where $U(n) \times U(k)$ are block matrices

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

where $A \in (n)$ an $B \in U(k)$. Given $A \in U(n+k)$ we associate to A the linear subspace spanned (over $\mathbb{C}$) by its first n columns. It is elementary to see that the right action of $U(n) \times U(k)$ leaves this map invariant and identifies the quotient space with $Gr_{\mathbb{C}}(n, k)$.

The following is clear.

Let $Gr(n, \infty)$ be the limit of the Grassmannians of n-planes in $n+k$ space for fixed n as $k \mapsto \infty$. Let $V_n \subset U$ denote the subgroup of U consisting of matrices of the form

$$\begin{pmatrix} \mathrm{Id}_n & 0 \\ 0 & A \end{pmatrix}$$

where $A \in U$, and let $\mathcal{F}_{\mathbb{C}}(n, \infty)$ be the quotient of U/V. It is space of pairs consisting of a n-plane P in $\mathbb{C}^\infty$ together with a unitary frame (or basis) for P. The natural action of $U(n)$ is to act on the frame. This is a free action with quotient $Gr_{\mathbb{C}}(n, \infty)$.

Since the inclusion of $V \subset U$ is a homotopy equivalence, it follows that $\pi_i(Gr_{\mathbb{C}}(n,\infty)) \cong \pi_{i-1}(U)$. Even stronger, we have

Proposition 2.1. *There is a homotopy equivalence* $\Omega(Gr(n,\infty)) \to U$. *)*

We need a finite version of this result. Here is one that will suffice. Consider the frame bundle $Fr_{\mathbb{C}}(n,n) = U(2n)(\mathrm{Id}_n \times U(n))$ associated to the Grassmannian of n-planes in $\mathbb{C}^{2n}$. By Claim 1.1

$$\pi_i(Fr_{\mathbb{C}}(n,n)) = 0 \quad \text{for} \quad i \leq 2n.$$

From the homotopy exact sequence of the fibration

$$U(n) \to Fr_{\mathbb{C}}(n,n) \to Gr_{\mathbb{C}}(n,n)$$

we see that

Theorem 2.2. $\pi_i(U(n)) = \pi_{i+1}(Gr_{\mathbb{C}}(n,n))$ *for* $i < 2n$ *and the induced map* $\Omega(Gr_{\mathbb{C}}(n,n)) \to U(n)$ *induces an isomorphism on homotopy groups in dimensions* $< 2n$

3 Jacobi Fields in a symmetric space

Let us recall the definition of a symmetric space.

Definition 3.1. A Riemannian manifold M is said to be a *symmetric space* if for each $p \in M$ there is an isometry $T_p \colon M \to M$ that fixes p and whose differential at p is $-\mathrm{Id}$.

It is equivalent to say that T_p reserves any geodesic passing though p while reversing the direction of the geodesic..

Lemma 3.2. *Let* $\gamma \colon [0,a] \to M$ *be a geodesic parameterized at unit speed with endpoints* $\gamma(0) = p$ *and* $\gamma(a) = q$. *Let* $\mu(t) = T_q(\gamma(t))$ *and define*

$$\rho(t) = \begin{cases} \gamma(t) & \text{if } 0 \leq t \leq a \\ \mu(2a-t) & \text{if } a \leq t \leq 2a. \end{cases}$$

Then ρ *is a geodesic of length* $2a$ *containing* γ *as an initial segment.*

Proof. Since γ is a geodesic and T_q is an isometry μ is a geodesic. Thus, $\alpha = \rho|_{[0,a]}$ and $\beta = \rho|_{[a,2a]}$ are geodesics with $\alpha(a) = \beta(a)$. Also, $\alpha'(a) = \beta'(a)$, so that these geodesics fit together to form a geodesic. $\square$

As a consequence we see that the Riemannian manifold M is complete since geodesics extend forever.

A similar argument shows that if γ is a geodesic $\gamma(0) = p$ and $\gamma(a) = q$, then $T_q T_p(\gamma(t) = \gamma(t + 2a)$.

Lemma 3.3. *Let M be a symmetric space and let $R(u, v)$ be the Riemannian curvature operator:*

$$R(u, v) = \nabla_u \nabla_v - \nabla_v \nabla_u - \nabla_{[u,v]}.$$

Let γ be a geodesic with tangent vector V. Let W be a parallel vector field along γ. Then $R(W, V)V$ is also parallel along γ.

Proof. Let X be a vector field parallel along γ. Consider $\langle R(W, V)V, X\rangle$. Fix two points $p = \gamma(0)$ and $q = \gamma(a)$, and let $T = T_q T_p$. This is an isometry that stablizes γ and sends $\gamma(t)$ to $\gamma(t + 2a)$. Since T is an isometry,

$$\langle R(T(W), T(V))T(V), T(X)\rangle = \langle R(W, V)V, X\rangle.$$

Of course $T(V) = V$. On the other hand, $T_p(W(p)) = -W(p)$. Since W is parallel, it follows that $T_p(W) = -W$. By the same argument $T_q(W) = -W$ and hence $T(W) = W$. The same argument applies to X. It follows that $\langle R(W, V)V, X\rangle$ takes the same value at $\gamma(t)$ and at $\gamma(t + 2a)$. But a is arbitrary so $\langle R(V, W)V, X\rangle$ is constant.

Since X is parallel along γ it follows that $R(W, V)V$ is also. □

Notice that $W \mapsto R(W, V)V$ is a linear map $TM_{\gamma(t0)} \to TM_{\gamma(0)}$. We claim that it is self-adjoint with respect to the inner product on $TM_{\gamma(t)}$. This follows from the symmetry of the Riemannian curvature tensor and the skew-symkmetry of R under both the interchange of the first two variables and the interchange of last two variables that

$$\langle R(W_1, V)V, W_2\rangle = \langle R(W_2, V)V, W_1\rangle.$$

Thus, $R(\cdot, V)V : TM_{\gamma(0)} \to TM_{\gamma(0)}$ can be diagonalized in some orthonormal basis for $TM_{\gamma(0)}$. Let $\{e_1, \ldots, e_k\}$ be the eigenvalues of this quadratic form with respect to eigenvectors $\{w_1, \ldots, w_k\}$. Since the quadratic form is parallel along γ the w_i extend to parallel fields $W_i = w_i(t)$ along γ. Then $R(W_i, V)V = e_i W_i$. Writting $W = \sum_i a_i(t)W_i$ the Jacobi equation becomes the diagonal system of equations

$$\{\ddot{a}_i(t) + e_i a_i = 0\}_i$$

4 The case of $U(n)$

We use the metric on $U(n)$ which comes from the inner product on $u(n)$ given by $\langle A, B\rangle = -\text{trace}(AB)$

We consider geodesics in $U(n)$ from 1 to -1. First notice that any $A \in U(n)$ can be diagonalized with diagonal entries of norm 1. The corresponding statement for the Lie algebra is that any element of $u(n)$ is in the orbit under the adjoint action of $U(n)$ a diagonal matrix with purely imaginary entries down the diagronal. Given such a diagonal matrix in $u(n)$

$$\begin{pmatrix} ia_1 & 0 & \cdots & 0 \\ 0 & ia_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 \\ 0 & 0 & \cdots & ia_n \end{pmatrix}$$

with the $a_i \in \mathbb{R}$ the geodesic it generates is

$$\begin{pmatrix} e^{ita_1} & 0 & \cdots & 0 \\ 0 & e^{ita_2} & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 \\ 0 & 0 & \cdots & e^{ita_n} \end{pmatrix}$$

All such geodesicsγ begin at 1 in the sense that $\gamma(0) = 1$. The condition $\gamma(1) = -1$ is equivalent to a_i is an odd integer of π for all $1 \leq i \leq n$. Every geodesic with $\gamma(0) = 1$ and $\gamma(1) = -1$ is conjugate in $U(n)$ to one of this form.

Let us specialize to $SU(2n)$.

Proposition 4.1. *Every geodesic from* 1 *to* -1 *in* $SU(2n)$ *is conjugate in* $SU(2n)$ *to one of the form*

$$A = \begin{pmatrix} e^{ita_1} & 0 & \cdots & 0 \\ 0 & e^{ita_2} & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 \\ 0 & 0 & \cdots & e^{ita_{2n}} \end{pmatrix}$$

with the a_i *being odd integral multiples of* π *and with* $\sum_{i=1}^{2n} a_i = 0$. *The length of any such geodesic is* $\pi\sqrt{a_1^2 + \cdots + a_{2n}^2}$.

Proof. The first statement follows from the discussion above and the fact that an element in $su(2n)$ is an element of $u(2n)$ of trace zero. The last statement follows immediately from a computation of the trace of $(A')^2$. □

It follows directly that the minimal geodesics from 1 to -1 in $SU(2n)$ are those where each $a_i = \pm 1$ and n of them are $+1$ and n of them are -1. Associated to any minimal geodesic γ in $SU(2n)$ with $\gamma(0) = 1$ and $\gamma(1) = -1$ is an n-dimensional complex lineaer subspace of $\mathbb{C}^{2n}$, namely the $+i\pi$ eiigenspace of $\gamma'(0)$. Conversely, given any such subspace we let X be the matrix in $su(2n)$ that is $+i\pi$ on this subspace and $-i\pi$ on its orthogonal complement. This element X is tangent vector at $\gamma(0)$ of a unique geodesic as required. We have established:

Lemma 4.2. *The space of minimal geodesics from* 1 *to* -1 *in* $SU(2n)$ *is a Grassmannian* $Gr(n,n)$ *of* n*-dimensional complex linear subspaces in* $\mathbb{C}^{2n}$.

4.1 The Jacobi equation along a geodesic from 1 to -1 in $SU(2n)$

Let

$$\gamma(t) = \begin{pmatrix} e^{i\pi a_1 t} & 0 & 0 & \cdots & 0 \\ 0 & e^{i\pi i a_2 t} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & e^{i\pi a_{2n} t} \end{pmatrix}.$$

We compute the eigenvalues of the form $R(\cdot, V)V$ on the tangent bundle to $SU(2n)$ along γ.

Lemma 4.3. *The matrices in* $su(2n)$ *with only entries on the diagonal are in the* 0*-eigenspace of* $R(\cdot, V)V$. *For each* $i < j$ *the (real) two-dimensional subspace of* $su(2n)$ *with non-zero entries only in positions* (i,j) *and* (j,i) *is contained in the eigenspace of* $R(\cdot, V)V$ *with eigenvalue* $\frac{\pi^2}{4}((a_i - a_j)^2$.

Proof. (Following Milnor) Since the integral curves of any left-invariant vector field are geodesics, it follows $\nabla_X X = 0$ for X left-invariant. Let X and Y be left-invariant vector fields. Using this and computing $\nabla_{X+Y}(X+Y)$ yields $\nabla_X(Y) + \nabla_Y(X) = 0$. The torsion-free condition then implies

$$\nabla_X(Y) = \frac{1}{2}[X, Y].$$

It follows by direct computation that $R(W,V)V = \frac{1}{4}[V,[W,V]]$. Now we apply this with V a diagonal matrix with diagonal entries $ia_, \pi, \ldots, ia_{2n}\pi$ and W a matrix with $W_{ij} = w$ and $W_{ji} = -\overline{w}$ and all other entries zero. Clearly,

$$[V, W] = i\pi(a_j - a_i)W.$$

It follows that

$$R(W,V)V = \frac{1}{4}\pi^2(a_i - a_j)^2 W.$$

□

Corollary 4.4. *Let γ is a minimal geodesic from 1 to -1 in $SU(2n)$, then the dimension of the kernel of the Hessian of E at γ is of (real) dimension $2n^2$. The minimal critical points of E on $\Omega^{1,2}(SU(2n,1,-1)$ are Bott-Morse.*

Proof. In this case the tangent vector to γ at 1 has n eigenvectors with eigenvalue $i\pi$ and n eigenvectors with eigenvalue $-i\pi$. According to the previous computation there is a (real) $2n^2$ dimensional eigenspace for linear map $W \mapsto R(W,V)(V)$ with eigenvalue π^2 and the only other eigenvalue of this operator is 0. For any Jacobi field J along γ vanishing at 1 and with tangent vector W in this eigenspace satisfies

$$\nabla_{\dot\gamma}\nabla_{\dot\gamma}J + \pi^2 J = 0.$$

Letting W denote the extension of W to a parallel vector field along γ, we see that the Jacobi field vanishing at $\gamma(0) = 1$ is $\sin(\pi t)W(t)$. Hence, this Jacobi fields vanishes at $\gamma(1) = -1$. This gives a real $2n^2$ dimensional subspace in the kernel of the Hessian of E at γ. On the other hand, Jacobi fields vanishing at 1 and with tangent vector in the 0-eigenspace grow linearly and hence do not vanish at -1. Thus, the kernel of the Hessian of E at a minimal γ has real dimension $2n^2$. We have already seen that the space of minimal geodesics is $Gr(n,n)$ which has complex dimension n^2 and hence real dimension $2n^2$. It follows that the kernel of the Hessian at every minimal geodesic is the tangent space to the space of minimal geodesics at that point. □

4.2 The index at a non-minimal geodesic

Now we give a lower bound for the index of the Hessain at a non-minimal geodesic.

Proposition 4.5. *Let γ be a non-minimal geodesic. Then the index of the Hessian of E at γ is at least $2(n+1)$.*

This proposition relies on the following general result about the index of the energy functional at a geodesic.

Theorem 4.6. *Let $M \subset \mathbb{R}^N$ be a properly embedded C^{k+2} submanifold ($1 \leq k \leq \infty$) and let E be the energy functional on the path space of $W^{1,2}$-paths in M from p to q. Let γ be a critical point of E. Then γ is a C^{k+2}-curve that satisfies the geodesic equation $\nabla_{\dot{\gamma}}(\dot{\gamma}(t)) = 0$ for all $t \in [0.1]$ For each $t \in [0,1]$ let $J_\gamma(t)$ be lhe dimension of the space of Jacobi fields along $\gamma|_{[0,t]}$ vanishing at both end points. Then:*

1. *$J_\gamma(t) = \{0\}$ except for finitely many $t \in (0,1]$.*
2. *$J_\gamma(t)$ is finite dimensional for all $t \in [0,1]$.*
3. *The index of the $Hess_\gamma(E) = \sum_{t \in [0,1)} \dim J_\gamma(t)$.*

This theorem is proved in the last section of this lecture.

Proof. (of the proposition assuming the theorem.) The tangent vector to γ at $\gamma(0) = 1$ is conjugate to a diagonal matrix with diagaonal entries $i\pi a_1, \ldots, i\pi a_{2n}$ with the a_i being odd integers summing to zero. Also, there is at least one a_i with $|a_i| \geq 3$. If $a_i < a_j$, there is 2 dimensional eigenspace with eigenvalue $\pi^2(a_i - a_j)^2$ given by matrices with non-zero entries only in positions (i,j) land (j,i). For any parallel vector field W along γ with initial vector in this 2-dimensional eigenspace the Jacobi field vanishing at $\gamma(0) = 1$ with tangent vector at $\gamma(0) = 1$ equal to $W(0)$ is given by

$$\sin(\pi(a_j - a_i))tW(t).$$

This Jacobi field vanishes at $t = 2/(a_j - a_i), 4/(a_j - a_i), \cdots$. Thus, the number of zeros of such a Jacobi field in the interval $0 < t < 1$ is $(a_j - a_i - 2)/2$. Since this eigenspace is two-dimensional, it adds $(a_j - a_i - 2)$ to the dimension of the kernel of the Hessian of E at γ.

The following is an elementary exercise.

Claim 4.7. *Suppose that $a_1, \ldots, a_{2n}$ are odd integers summing to 0 and suppose that at least one of them has absolute value ≥ 3. Then*

$$\sum_{\{i,j | a_j > a_i\}} (a_j - a_i - 2) \geq 2(n+1).$$

This proves the proposition. □

5 Completion of the proof of Bott Periodicity

We have just established that the minimal geodesics in $\Omega^{1,2}(SU(2n), 1, -1)$ are identified with $Gr(n, n)$ and that the index on any non-minimal geodesic in this space is $2(n+1)$. Thus, the inclusion $Gr_{\mathbb{C}}(n, n) \subset \Omega^{1,2}(SU(2n), 1, -1)$ induces an isomorphism on the homotopy groups in diemsions $\leq 2n$. Hence, for $i \leq 2n$ we have

$$\pi_i(Gr(n, n)) = \pi_i(\Omega^{1,2}(SU(2n), 1, -1)) = \pi_i(\Omega(SU(2n), 1) = \pi_{i+1}SU(2n).$$

On the other hand, Theorem 2.2 says $\pi_{i-1}(U(n)) = \pi_i(Gr(n, n))$ for $i \leq 2n$. We conclude that $\pi_{i-1}(U(n)) = \pi_{i+1}(SU(2n))$ for all $i \leq 2n$. Since the inclusion $SU(k) \subset U(k)$ induces an isomorphism on all π_i for $i > 1$, we see that $\pi_{i-1}(U(n)) = \pi_{i+1}(U(2n))$ for all $1 < i \leq 2n$. Of course, by stability $\pi_i(U(n)) = \pi_i(U(2n))$ for $i \leq 2n + 1$. Thus, for all $1 < i \leq 2n$ we have

$$\pi_{i-1}(U(n)) = \pi_{i+1}(U(n)).$$

We can reformulate this as

$$\pi_{i-1}(U) = \pi_{i+1}(U) \quad \text{for all } i \geq 2.$$

This of gives that for all $i \geq 1$

$$\pi_{2i}(U) = 0$$

$$\pi_{2i-1}(U) \cong \mathbb{Z}.$$

Let Φ be the inclusion $Gr_{\mathbb{C}}(n, n) \to \Omega(SU(2n))$. It induces an isomorphism on homotopy groups in degrees $\leq 2n$. Let $F\colon U(n) \to \Omega(Gr_{\mathbb{C}}(n, n))$ be the map induced by the fibration

$$U(n) \to \mathcal{F}_{\mathbb{C}}(n, n) \to Gr_{\mathbb{C}}(n, n).$$

It induces an isonorphism on homotopy groups in degrees $< 2n$. Hence $\Omega(\Phi) \circ F\colon U(n) \to \Omega^2(SU(2n))$ induces an isomorphism on homotopy groups in degrees $< 2n$. Taking the limit as $n \mapsto \infty$ gives a homotopy equivalence

$$U \to \Omega^2(SU).$$

This completes the proof of Bott Periodicity.

6 The index of the Hessian

In this section fix a properly embedded smooth n-dimensional manifold $M \subset \mathbb{R}^N$ and we give a general description of the Hession of $E\colon \Omega^{1,2}(M,p,q) \to \mathbb{R}$ at a geodesic γ.

Theorem 6.1. *Let $M \subset \mathbb{R}^N$ be a properly embedded C^{k+2} submanifold ($1 \le k \le \infty$) and let E be the energy functional on the path space of $W^{1,2}$-paths in M from p to q. Let γ be a critical point of E. Then γ is a C^{k+2}-curve that satisfies the geodesic equation $\nabla_{\dot\gamma}(\dot\gamma(t)) = 0$ for all $t \in [0.1]$ For each $t \in [0,1]$ let $J_\gamma(t)$ be lhe dimension of the space of Jacobi fields along $\gamma|_{[0,t]}$ vanishing at both end points. Then:*

1. *$J_\gamma(t) = \{0\}$ except for finitely many $t \in (0,1]$.*
2. *$J_\gamma(t)$ is finite dimensional for all $t \in [0,1]$.*
3. *The index of the $Hess_\gamma(E) = \sum_{t\in[0,1)} \mathrm{dim} J_\gamma(t)$.*

Remark 6.2. What we are really saying is that the index of the Hessian of E at $\gamma|_{[0,1]}$ is a weakly increasing function of t with finitely many jumps. These occur at t for which $\gamma|_{[0,t]}$ is a degenerate critical point for the energy functional, corresponding to eigenvalues zero for the Hessian of E at these geodesics. Furthermore, all the zero eigenvalues at of the Hessian at $\gamma|_{[0,t]}$ are limits of positive eigenvalues of the Hessian at $\gamma|_{[0,t']}$ as t' approaches t from below. That is to say, it is never the case that as t increases a negative eigenvalue becomes zero. Also, the zero eigenvalues for the Hessian of the energy functional at $\gamma|_{[0,t]}$ cross zero and are positive eigenvalues for the Hessian of energy at $\gamma|_{[0,t']}$ for $t' < t$ and they are negative eigenvalues of the Hessian of the energy functional at $\gamma|_{[0,t']}$ for t' slightly greater than t. As t increases all the spectral flow is in the negative direction and this flow occurs anytime there is a zero eigenvalue and makes the full eigenspace of 0 at t cross from positive before t to negative after t.

6.1 Finiteness of the index

Suppose that the have a continuous $\gamma\colon (-\epsilon,\epsilon)\times[0,1] \to M$ with the property that there is a subdivision $0 = T_0 < T_1 < \ldots, T_N < T_{N+1} = 1$ such that the restriction of γ to each strip $(-\epsilon,\epsilon) \times [T_i, T_{i+1}]$ is C^{k+2}. Then for each $s \in (-\epsilon,\epsilon)$ the velocity $\dot\gamma(t,s)$ is a piecewise smooth, but possibly discontinuous at the T_i. Let $\lambda = ((\partial/\partial s)\gamma)(t,0)$ We denote $\delta_{T_i}(\dot\gamma)$ the difference of the limit of $\dot\gamma(t;)$ for $t' \mapsto t$ from above minus the limit from below. Then the first variation formula for the energy is

$$\frac{d}{ds}|_{s=0}E(s) = \int_0^1 \langle\dot\lambda(t), \dot\gamma(t)\rangle dt = -\int_0^1 \langle\lambda(t), \nabla_{\dot\gamma}\dot\gamma(t)\rangle dt - \sum_i \langle\lambda(T_i, \delta_{T_i}\dot\gamma\rangle.$$

The point is that shifting the t-derivative from the first factor to the second uses integration by parts of $\langle\lambda(t), \dot\gamma(t)\rangle$ and this function has discontinuities whose sum has to be included as an error term.

A similar argument shows that if $\gamma(t, s_1, s_2)$ is a two-parameter family of continuous curves with a division $0 = T_0 < T_1 \cdots < T_N < T_{N+1} = 1$ so that for every (s_1, s_2) the curve $\gamma(t, s_1, s_2)$ is smooth on each of the $T_i, T_{i+1}]$, then

$$\frac{\partial}{\partial s_1}\,|_{s_1=0}\,\frac{\partial}{\partial s_2}\,|_{s_2=0}\,E = -\int_0^1 \langle\lambda_2, \nabla_{\dot\gamma}\nabla_{\dot\gamma}\lambda_1 + R(\lambda_1, \dot\gamma)\dot\gamma(t)\rangle dt - \sum_{T_i}\langle\lambda_2(T_i), \delta_{T_i}(\lambda_1)\rangle.$$

We shall be working simultaneously with spaces associated with geodesics $\gamma|_{[a,b]}$ for $0 \le a < b \le 1$. We will keep the parametrization by t fixed, so that these curves are maps $[a, b] \to M$ from q_a to q_b where $q_t = \gamma(t)$. Such a geodesic sits in a space denoted $\Omega^{1,2}(M, q_a, q_b, [a, b]) \subset W^{1,2}([a, b], \mathbb{R}^N)$. The tangent space to this space at a curve γ defined on $[a, b]$ is a $W^{1,2}$-path $\lambda(t)$ defined for $a \le t \le b$ with $\lambda(t) \in T_{\gamma(t)}M$ for all t and $\lambda(a) = 0$ and $\lambda(b) = 0$. The energy functional is

$$E_{a,b} = \int_a^b |\dot\gamma(t)|^2 dt,$$

and critical points are curves defined on $[a, b]$ that are critical points. These are C^{k+2}-geodesics defined on $[a, b]$ parametrized at constant speed. By a Jacobi field along $\gamma|_{[a,b]}$ we mean a tangent vector λ satisfying the Jacobi equation

$$\nabla_{\dot\gamma}\nabla_{\dot\gamma}\lambda(t) + R(\lambda(t), \dot\gamma(t))\dot\gamma(t) = 0.$$

These are automatically C^{k+1}-paths.

Let γ be a geodesic in $\Omega^{1,2}(M, p, q)$ (defined on $[0, 1]$). Then there is an $\epsilon > 0$ such that for every $t \in [0, 1]$ the the restriction of the exponential map at $\gamma(t)$ to the ball of radius ϵ in $T_\gamma(t)M$ is a C^{k+2}-diffeomorphsim of this ball onto an open subset of M. We choose $0 = T_0 < T_1 \ldots < T_N < T_{N+1} = 1$ so that the length of $\gamma|_{[T_i,T_{i+1}]}$ is less than ϵ for all $1 \le i \le N$.

Claim 6.3. *A Jacobi field along $\gamma|_{[T_i,T_{i+1}]}$ is determined by its endpoints, and for each any values for $\lambda(T_i)$ and $\lambda(T_{i+1}$ there is a Jacobi field along $\gamma|_{[T_i,T_{i+1}]}$ with these values at the end points.*

Proof. The length of $\gamma|_{[T_i,T_{i+1}]}$ is $\dot\gamma(T_i)$ times $(T_{i+1}-T_i)$ so that $\exp_{\gamma(T_i)}(\dot\gamma)$ is a diffeomorphism on the ball of radius $T_{i+1}-T_i$. This means that any non-zero Jacobi field for $\gamma|_{[T_i,T_{i+1}]}$ vanishing at T_i is non-zero at T_{i+1}. Since the space of Jacobi fields for $\gamma|_{[T_i,T_{i+1}]}$ vanishing at T_i is of dimension n, the result follows. □

Definition 6.4. A *broken Jacobi field* along γ is a tangent $\lambda(t)$ that is a smooth Jacobi field for the $\gamma_{[T_i,T_{i+1}]}$ but it is not necessarily differentiable at the $\gamma(T_i)$. Let BJ_γ be the space of broken Jacobi fields that vanish at $\gamma(0)$ and $\gamma(1)$. The claim above shows that this space is the $2N$ dimensional manifold $\prod_{i=1}^N T_{\gamma(T_i)}M$, the diffeomorphism being given by associating to each broken Jacobi field its values at the T_i.

Let $\Omega_0^{1,2}(M,p,q)_\gamma$ be the subspace of tangent vectors that vanish on all the T_i.

Claim 6.5. *We have an orthogonal decomposition*

$$\Omega^{1,2}(M,p,q)_\gamma = \Omega_0^{1,2}(M,p,q)_\gamma \oplus BJ_\gamma.$$

Proof. The intersection of the two subspaces is $\{0\}$ since a broken Jacobi field that vanishes at the T_i is identically zero. On the other hand, given any tangent vector λ, let $J(\lambda)$ be the unique broken Jacobi field that takes the same values at the T_i as λ. Clearly, $\lambda - J(\lambda) \in \Omega_0^{1,2}(M,p,q)_\gamma$. This proves that we have a direct sum decomposition.

We show that the two summands are orthogonal; i.e., that

$$\int_0^1 \langle \lambda_2(t), \nabla_{\dot\gamma}\nabla_{\dot\gamma}\lambda_1(t) + R(\lambda_1,\dot\gamma)\dot\gamma(t)\rangle dt + \sum_i \langle \lambda_2(T_i), \delta_{T_i}\dot\lambda_1\rangle = 0$$

for $\lambda_1 \in BJ_\gamma$ and $\lambda_2 \in \Omega_0^{1,2}(M,p,q)_\gamma$. The point is that the integral vanishes since λ_1 is a Jacobi field on each of the subintervals $[T_i,T_{i+1}]$ and the finite sum vanishes since $\lambda_2(T_i)=0$ for all i. □

Claim 6.6. *The Hessian of E at γ is positive definite on $\Omega_0^{1,2}(M,p,q)_\gamma$.*

Proof. There is no non-zero Jacobi field for each of the $\gamma|_{[T_i,T_{i+1}]}$. This implies that the Hessian of E at $\gamma|_{[T_i,T_{i+1}]}$ is non-degenerate. The space $\Omega_0^{1,2}(M,p,q)_\gamma$ decomposes as a direct sum of the tangent spaces to the $\gamma|_{[T_i,T_{i+1}]}$. But each of these geodesics is the unique shortest geodesic between its endpoints. These two facts together imply that the Hessian for E at $\gamma|_{[T_i,T_{i+1}]}$ is positive definite. On the other hand, the Hessian for E at γ is the sum of the Hessians for E at the $\gamma|_{[T_i,T_{i+1}]}$, and consequently is also positive definite on their sum which is $\Omega_0^{1,2}(M,p,q)_\gamma$. □

Corollary 6.7. *The negative eigensace of the Hessian of E at γ is contained in BJ_γ and is the negative eigenspace of the restriction of the Hessian of E at γ to this subspace. Since* $\dim BJ_\gamma$ *is finite, the index of the Hessian for E at any geodesic γ is finite.*

6.2 Completion of the Proof of Theorem 6.1

Let $d(t)$ be the index of the Hessian of E at $\gamma|_{[0,t]}$. We complete the proof of the theorem by showing the following four things.
Step 1. $d(t) = 0$ for $t \in [0,1]$ sufficiently small.

Step 2. $d(t)$ is a weakly monotone increasing function; i.e., if $t' > t$ then $d(t') \geq d(t)$.

Step 3. If t_n is an increasing sequence converging to $t_\infty \in [0,1]$, then $d(t_\infty) = d(t_n)$ for all n sufficiently large.

Step 4. For any $t \in [0,1)$, there is $\epsilon > 0$ such that for any t' with $t < t' < t+\epsilon$ we have $d(t') = d(t) + \dim(J_{\gamma(t)})$.

Here are the proofs of the four steps.

Proof of Step 1. We have already seen that for $t > 0$ sufficiently small there are no non-zero Jacobi fields along $\gamma|_{[0,t]}$ vanishing at both end points. From this Step 1 is immediate.

Proof of Step 2. If $\lambda\colon [0,t] \to M$ is a tangent, then the extension $\widetilde{\lambda}$ of λ to be zero on $[t,t']$ is a tangent vector for $\gamma|_{[0,t']}$. This defines an embedding of $\varphi\colon \Omega_0^{1,2}(M,p,q_t,[0,t])_{\gamma|_{[0,t]}} \subset \Omega_0^{1,2}(M,p,q_{t'},[0,t'])|_{\gamma|_{[0,t']}}$ with

$$Hess_{\gamma|_{[0,t]}}(E)(\lambda_1,\lambda_2) = Hess_{\gamma|_{[0,t']}}(E)(\widetilde{\lambda}_1,\widetilde{\lambda}_2).$$

Thus, a maximal negative definite subspace of $Hess_{\gamma|_{[0,t]}}(E)$ maps to a negative definite subspace of $Hess_{\gamma|_{[0,t']}}(E)$. This proves Step 2.

Proof of Step 3. We let i_0 be the minimal index such that $t_\infty \leq T_{i_0}$. Then we can assume that $T_{i_0-1} < t_n \leq T_{i_0}$ for all n. Hence, for $1 \leq n \leq \infty$ the spaces $BJ_{\gamma|_{[0,t_n]}}$ are all identified with the space space, namely

$$\prod_{i=1}^{i_0-1} T_{\gamma(T_i)}M.$$

For $n \leq \infty$, let E_n, denote the energy functional on $\Omega^{1,2}(M,p,\gamma(t_n),[0,t_n])$ restricted to $BJ_{\gamma|_{[0,t_n]}}$.

By Corollary 6.7, we know that for $n \leq \infty$ the index of the E_n is the index of the energy functional E at $\gamma|_{[0,t_n]}$. We view these as smooth functions on $\prod_{i=1}^{i_0-1} T_{\gamma(T_i)}M$. Then E_n for $n < \infty$ converge smoothly to E_∞. Under these identifications the points $\gamma|_{[0,t_n]}$ are identified with the origin. Let $N(\infty)$ be the negative eigenspace of the Hessian for E_∞ at $\gamma|_{[0,t_\infty]}$. It follows that for all n sufficiently large, E_n is negative definite on $N(\infty)$, and hence that the index of E_n is at least as large as that of E_∞. Since $t_n < t_\infty$, we have already seen that the index of E_n is less than or equal to that of E_∞. Hence they are equal.

Proof of Step 4. By choosing a different set of division points we can assume that t is in the interior of one of the intervals; say $T_{i_0} < t < T_{i_0+1}$. Arguing exactly as before, for all t' sufficiently close to t we can identify all the $BJ_{\gamma|_{[0,t']}}$ with $\prod_{i=1}^{i_0} T_{\gamma(T_i)}M$ in such away that the $\gamma|_{[0,t']}$ are identified with the origin and the restrictions of the energy functionals for t' converge to that of t as $t' \mapsto t$. It follows that for any sequence $t'_n \mapsto t$, after replacing the sequence by a subsequence, the negative eigenspaces of the restriction to $BJ_{\gamma|_{[0,t'_n]}}$ of the Hessian of energy functional at $\gamma|_{[0,t'_n]}$ converge to a subspace of the non-positive eigenspace for the Hessian of the energy functional at $\gamma(t)$ restricted to $BJ_{\gamma|_{[0,t]}}$. Hence, the index of the Hessian for the energy functional at $\gamma|_{[0,t'_n]}$, $d(t'_n)$, is less than or equal to $d(t) + \dim(J_t(\gamma))$.

Let us establish the opposite inequality. Let $W_1, \ldots, W_k \in BJ_{\gamma|_{[0,t]}}$ be tangent vectors along $\gamma|_{[0,t]}$ that are a basis for the negative eigenspace of the Hessian of E at $\gamma|_{[0,t]}$, and let $J_1, \ldots, J_\ell$ be a basis for the Jacobi fields along $\gamma|_{[0,t]}$ vanishing at $\gamma(0)$ and $\gamma(t)$. We extend all these vector fields by zero along $\gamma|_{[t,t']}$ to become tangent vectors for $\gamma|_{[0,t']}$. Since the J_i are Jacobi fields on $[0,t]$ and identically zero on $[t,t']$,

$$Hess_{\gamma|_{[0,t']}}(E)(J_i, J_j) = 0$$

and

$$Hess_{|\gamma_{[0,t']}}(E)(J_i, W_\ell) = 0.$$

We know that the map $J_t(\gamma) \to T_{\gamma(t)}M$ given by $J_i \mapsto \nabla_{\dot\gamma} J(t)$ is an injection. Let $X_1, \ldots, X_\ell$ be smooth vector fields along $\gamma|_{[0,t']}$ such that

$$\big(A_{ij}\big) = \big(\langle X_i(t), \nabla_{\dot\gamma}(J_j)(t)\rangle\big)$$

is the identity matrix.

Let $N(c)$ be the subspace of the tangent space to $\gamma|_{[0,t']}$ generated by

$$\{W_1, \ldots, W_k, (c^{-1}J_1 - cX_1), \ldots, (c^{-1}J_s - cX_s)\}.$$

As $c \mapsto 0$

$$Hess_{\gamma|_{[0,t']}}(E)\big(W_1, \ldots, W_k, (c^{-1}J_1 - cX_1), \ldots, (c^{-1}J_s - cX_s)\big)$$

converges to a negative definite matrix. Thus, for all $c > 0$ sufficiently small the Hessian at $\gamma|_{[0,t']}$ is negative definite on the subspace $N(c)$. Thus, $d(t') \geq d(t) + \dim(J_t(\gamma))$.

This completes the proofs of the four steps and hence of the theorem.