Lecture IIIA: h-cobordism Theorem, Cont'd

October 8, 2020

In this lecture we sketch the proof of the two cancellation theorems used in the proof of the h-cobordism theorem. These arguments follow closely those in Milnor's h-cobordism book.

1 The First Cancellation Theorem

Recall the statement of this result.

Theorem 1.1. *(First Cancellation Result)*
Suppose:

1. *M is compact (possibly with boundary) and $f\colon M \to \mathbb{R}$ is a Morse function.*
2. *p and q are critical points of indices $r-1$ and r respectively and with $f(p) < f(q)$; there are constants $a < f(p) < f(q) < b$ such that p and q are the only critical points with critical values in $[a,b]$.*
3. *χ is a gradient-like vector field for f.*
4. *Setting $N = f^{-1}(\ell)$ for some $\ell \in (f(p), f(q))$, the intersection of the descending disk for q, the ascending disk for p and N is a single point x of transverse intersection.*

Then there is a Morse function f' that agrees with f outside $f^{-1}([a,b])$ and has no critical points in $f^{-1}([a,b]) = (f')^{-1}([a,b])$ and a gradient-like vector field such that χ' for f' that agrees with χ outside $f^{-1}(a,b])$.

1.1 First Reduction

We temporarily replace M by $W = f^{-1}([a,b])$ and restrict f and χ to this submanifold. We have $\partial_- W = f^{-1}(a)$ and $\partial_+ W = f^{-1}(b)$. The points p

and q are the only critical points of $f|_W$. We consider the following localized version of the first cancellation theorem.

Theorem 1.2. *Under the hypotheses of the First Cancellation Theorem and with $W = f^{-1}([a,b])$, there is a nowhere zero vector field χ_1 on W agreeing with χ near ∂W and a Mose function $f_1\colon W \to \mathbb{R}$ without critical points for which χ_1 is a gradient-like vector field with the property that f_1 agrees with f near $\partial_- W$.*

The localized version easily implies the first cancellation theorem. For given such a f_1 and χ_1, we define

$$(f', \chi') = \begin{cases} (f, \chi) & \text{on } M \setminus f^{-1}(a,b) \\ (f_1, \chi_1) & \text{on } f^{-1}([a,b]) \end{cases}.$$

Then f' and χ' are the Morse function and gradient-like vector field whose existence is asserted in the first cancellation theorem.

It remains to prove Theorem 1.2.

Proof. For simplicity of the presentation we assume that $a = 0$ and $b = 1$.

The fourth hypothesis in the First Cancellation Theorem implies that there is a unique trajectory, say T, from p to q. The main step in the proof of this result is to show that there are local coordinates on a neighborhood U of T that make this neighborhood the result of a birth.

Proposition 1.3. *There is $\epsilon > 0$ such that for $Z \subset \mathbb{R}^n$ the open subset defined by $-(1+\epsilon) < x^1 < (1+\epsilon)$ and $\sum_{i=2}^n (x^i)^2 < \epsilon$, there is a neighborhood U of the unique trajectory T from p to q, a diffeomorphism $\varphi\colon U \to Z$ and a positive function k on U such that*

$$\varphi_*(k\chi) = \big((1-(x^1)^2, -2x^2, \ldots, -2x^r, 2x^{r+1}, \ldots, 2x^n\big).$$

Remark 1.4. We write $\mathbb{R}^n = \mathbb{R} \times \mathbb{R}^{r-1} \times \mathbb{R}^{n-r}$ We write (t, x, y) for the Euclidean coordinates on Z with the understanding that t is the usual coordinate on $\mathbb{R}$ and the x and y refer to the usual coordinates $(x^1, \ldots, x^{r-1})$ and $(y^1, \ldots, y^{n-r})$ on $\mathbb{R}^{r-1}$ and $\mathbb{R}^{n-r}$ respectively.

1.2 Proof of Theorem 1.2 assuming Proposition 1.3

We now assume this result and show how it implies the first cancellation result. We sketch the proof of this at the end of this section.

Claim 1.5. *Let U be a neighborhood of T There is a neighborhood $U' \subset U$ such that every partial flow line for χ that starts in U' and exits U never re-enters to U'*

Proof. If no such U' exits then there is a sequence of partial flow lines starting at r_k passing though a point s_k and then passing through t_k where r_k and t_k converge to T and $s_k \notin U$. By compactness we can assume that the s_k converge to $s \notin U$. Consider the flow line L through s. It cannot converge to p at $-\infty$ and to q at $+\infty$ since T is the only flow line with this property and $T \subset U$. Thus, either L begins in $\partial_- W$ or ends in $\partial_+ W$ (or both). For definiteness, let us assume that it begins in $\partial_- W$. Then for all s' in a compact neighborhood N of s the flow line through s' will also begin in $\partial_- W$. Hence, for all $s' \in N$ the partial flow lines from $\partial_- W$ to s' will remain outside some neighborhood of T. Since for all k sufficiently large, $s_k \in N$, this implies that for all k sufficiently large r_k is outside a fixed neighborhood of T. This contradicts the fact that the $r_k \mapsto T$ as $k \mapsto \infty$. $\square$

Remark 1.6. This is where we use the global statement that there are no other trajectories connecting p to q.

We use $\varphi \colon U \to Z$ to define local coordinates (t, x, y) on U, and let $U' \subset U$ be as in the conclusion of the previous claim. By shrinking U and Z we can make Z of the form $(-1-\epsilon, 1+\epsilon) \times B^k(\epsilon) \times B^{n-k-1}(\epsilon,$ the product on an interval in $\mathbb{R}$ with balls of radius $\epsilon > 0$ in Ar^k and $\mathbb{R}^{n-k-1}$, We can assume that for some $\delta > 0$ the neighborhood U' contains $(-1-\delta, 1+\delta) \times B^k(\delta) \times B^{n-k-1}(\delta)$. In these coordinates

$$\chi = \frac{1}{k(t,x,y)}(1-(t)^2, -2x, 2y),$$

for some positive function $k(t, x, y)$ on U.

Let $v_1(t)$ be a smooth function defined for all t that is everywhere negative and agrees with $1 - t^2$ for x outside of $(-1-\delta, 1+\delta)$. Let $\rho(r)$ be a smooth function that is identically 1 near 0 and identically 0 for $r \geq \delta$. Let $v(t,r) = \rho(r)v_1(t) + (1-\rho(r))(1-t^2)$. Now we define $V(t,x,y) = v(t, |x|^2 + |y|^2)$, and define

$$\chi' = \frac{1}{k(t,x,y)}(V(t,x,y), -2x, 2y)).$$

Notice that χ' agrees with χ off of a compact subset of U'. Thus, we can extend χ' to a vector field on all of W, one that agrees with χ outside U' and hence in a neighborhood of ∂W. Clearly χ' is nowhere zero.

Lemma 1.7. *Every flow line for χ' extends from $\partial_- W$ to $\partial_+ W$.*

Proof. If a flow of χ' never enters U', then it is a flow line of χ that never enters U'. Such flow cannot be asymptotic as $t \mapsto \pm\infty$ to either p or q, and hence flow from $\partial_- W$ to $\partial_+ W$.

Next, we show that no flow line of χ' can spend an infinite interval in U'. Let s be the flow parameter along the flow line. For suppose the flow line is in U' for all $s \geq s_0$. Then restricted to this half-infinite interval this flow line lines in a compact region in the coordinate space (t, x, y). Since y grows exponentially in forward time under χ', this implies that these coordinates are zero on this flow line. Since x decays exponentially in forward time under χ', eventually these are arbitrarily small. Thus, for all s sufficiently large the flow line at time s lies in the region where $\rho = 1$. In this region the t-coordinate of χ' is negative and bounded away from zero, so that the t-coordinate of the flow line is unbounded as $s \mapsto \infty$. This is a contradiction. The symmetric argument reversing the roles of x and y shows that the flow line for χ' cannot remain in U for an interval of time going to $-\infty$.

Thus, for any flow line L for χ' that is in U' at time s there are times $s_- < s < s_+$ such that L is outside U at times s_- and s_+. By the claim and since $\chi' = \chi$ outside of U', this implies that the flow line L is outside U' at all times $\geq s_+$ and at all times $\leq s_-$. In these time intervals the flow line L is a flow line for χ. There are no half-infinite flow lines for χ remaining outside U'. This implies that L goes from $\partial_- W$ to $\partial_+ W$. $\square$

This completes the proof of the lemma. $\square$

Now we are in position to prove the theorem. Since every flow line of χ' goes from $\partial_- W$ to $\partial_+ W$, we can define $\mu\colon \partial_- W \to \mathbb{R}$ by setting $\mu(w)$ equal to the time at which the flow line for χ' beginning at w arrives at $\partial_+ W$. This is a smooth positive function bounded away from zero. We define a new vector field $\chi_1(x)$ as follows. Let $a(x) \in \partial_- W$ be the initial point of the flow line for χ' though x and set $\chi_1(x) = \chi'(x)/\mu(a(x))$. The integral curves for χ_1 define a diffeomorphism

$$\varphi\colon (\partial_- W \times [0,1], \partial_- W \times \{0\}, \partial_- W \times \{1\}) \to (W, \partial_- W, \partial_+ W).$$

Using this diffeomorphism we write points of W as (w, t) with $w \in \partial_- W$ and $t \in [0,1]$. Then $\chi_1 = \partial/\partial t$, and t is a Morse function on W. It follows that, χ', which is a positive multiple of $\chi_1 = \partial/\partial t$, is also a gradient-like vector field for t. While the Morse function t takes the same value as f on ∂W, it does not necessarily agree with f near ∂W.

We must interpolate between f and t to produce the required Morse function. It is more convenient to interpolate their time derivatives. Any smooth positive function $\mu\colon W \to \mathbb{R}$ with the property that for all $w \in \partial_- W$ the integral

$$\int_0^1 \mu(w,t)dt = 1$$

can be used to define a Morse function

$$g(w,u) = \int_0^u \mu(w,t)dt$$

for W taking value 0 on $\partial_- W$ and value 1 on $\partial_+ W$ and for which $\partial/\partial t$, and hence χ', is a gradient-like vector field. If μ agrees with $\frac{\partial f}{\partial t}$ near ∂W, then g will agree with f near ∂W.

To create μ with these properties we shall interpolate between $\frac{\partial f}{\partial t}$ and a positive function $h(w,t) = h(w)$ depending only on $w \in \partial_- W$. Since $\chi_1(f) > 0$ on ∂W there is $\epsilon > 0$ such that $\chi_1(f) > 0$ on $\partial_- W \times ([0,\epsilon) \cup (1-\epsilon, 1])$. By taking $\epsilon > 0$ sufficiently small we can assume that $f(\partial_- W \times \{\epsilon\}) \leq 1/4$ and $f(\partial_- W \times \{1-\epsilon\}) \geq (3/4)$. Let $\lambda(t)$ be a C^∞ function $[0,1] \to [0,1]$ that is zero on $[\epsilon, 1-\epsilon]$ and identically 1 near 0 and 1.

Claim 1.8. *There is a smooth positive function $h(w)$ defined on $\partial_- W$ such that the function*

$$\mu(w,t) = \lambda(t)\frac{\partial f}{\partial t}(w,t) + (1-\lambda(t))h(w)$$

is everywhere positive and for every $w \in \partial_- W$ has

$$\int_0^1 \mu(w,t) = 1.$$

Proof. The integral condition translates into

$$h(w) = \Big(1 - \int_0^1 \lambda(t)\frac{\partial f}{\partial t}(w,t)dt\Big)\big/\Big(\int_0^1 (1-\lambda(t))dt\Big).$$

Since $\frac{\partial f}{\partial t}$ is positive where $\lambda \neq 0$ and $\lambda \geq 0$ we see that

$$\begin{aligned} 0 < \int_0^1 \lambda(t)\frac{\partial f}{\partial t}(w,t)dt &\leq \int_0^\epsilon \frac{\partial f}{\partial t}(w,t)dt + \int_{1-\epsilon}^1 \frac{\partial f}{\partial t}(w,t)dt \\ &= (f(w,\epsilon) - f(w,0)) + (f(w,1) - f(w,(1-\epsilon)) \leq 1/2. \end{aligned}$$

This implies that $h(w)$ is a positive function. Since $h(w)$ is positive and $\frac{\partial f}{\partial t}$ is positive where $\lambda(t) > 0$, it follows that $\mu(w,t)$ is positive. The resulting Morse function $g(w,u) = \int_0^u \mu(w,t)dt$ then agrees with f near ∂W and has $\partial/\partial t$, and hence χ', as a gradient-like vector field. $\square$

1.3 Proof of Proposition 1.3

First, to simplify the constants, we now compose the Morse function on $f\colon W \to \mathbb{R}$ with a diffeomorphism of $\mathbb{R}$ to arrange that the critical values are $f(p) = -(2/3)$ and $f(q) = (2/3)$

We begin with the standard model near the flow connection between p and q. Using the coordinate (t, x, y) on $\mathbb{R}^n = \mathbb{R} \times \mathbb{R}^{r-1} \times \mathbb{R}^{n-r}$ as before, we define

$$f_0(t, x, y) = t - \frac{1}{3}(t)^3 - |x|^2 + |y|^2.$$

The critical points of f_0 are $p_0 = (-1, 0, 0)$ and $q_0 = (1, 0, 0)$. They are non-degenerate and of indices $r-1$ and r and with critical values $f(p_0) = -(2/3)$ and $f(q_0) = (2/3)$. There is a single flow line T_0 in $\mathbb{R}^n$ connecting the p_0 to q_0. All this data match the corresponding data for f. Our goal is to find neighborhoods U of T and Z of T_0, a positive function $k(t, x, y)$ on U and a diffeomorphism $U \to Z$ sending $k(t, x, y)\chi|_U$ to the standard vector field ∇f_0. By a slight variant of the standard neighborhood result for non-degenerate critical points there are disjoint standard neighborhoods $N(p)$ of p and $N(q)$ of q and diffeomorphisms $g_p\colon N(p) \cong N(p_0) \subset \mathbb{R}^n$ and $g_q\colon N(q) \cong N(q_0) \subset \mathbb{R}^n$ carrying $\chi|_{N(p)}$ to $\nabla f_0|_{N(p_0}$ and similarly for q, and with $g_p^* f_0 = f|_{N(p)}$ and $g_q^* f_0 = f|_{N(q)}$. for $i = 0, 1$.

Fix $-(2/3) < a < b < (2/3)$ with $a + (2/3)$ and $b - (2/3)$ sufficiently close to zero that the level sets $f^{-1}(a)$ and $f^{-1}(b)$ meet $N(p)$ and $N(q)$ respectively. We truncate $N(p)$ from above at $f^{-1}(a)$ and $N(q)$ from below at $f^{-1}(b)$. We have the corresponding truncations of $N(p_0)$ at $f_0^{-1}(a)$ and $N(_0)$ at $f_0^{-1}(b)$. The flows for χ and ∇f_0 give diffeomorphisms $\widetilde{h}\colon f^{-1}(a) \to f^{-1}(b)$ and $\widetilde{h}_0\colon f_0^{-1}(a) \to f_0^{-1}(b)$, respectively. The pre-image under $\widetilde{h}$ of the descending sphere for q at level b with the ascending sphere for p at level a being the single point $T \cap f^{-1}(a)$ and the intersection of these submanifolds being transverse at this point. Similarly, for $\widetilde{h}_0$.

There is a small neighborhood $V_0(a)$ of $T_0 \cap f_0^{-1}(a)$ in $N(p_0) \cap f_0^{-1}(a)$ that flows under ∇f_0 diffeomorphically onto an open neighborhood $V_0(b)$ of $T_0 \cap f_0^{-1}(b)$ in $N(q_0) \cap f_0^{-1}(b)$. We fix a small neighborhood $V(a) \subset g_p^{-1}(V_0(a)) \subset N(p) \cap f^{-1}(a)$ of $T \cap f^{-1}(a)$ that flows under χ diffeomorphically onto a neighborhood $V(b) \subset g_q^{-1}(V_0(b))$ of $T \cap f^{-1}(b)$ in $N(q) \cap f^{-1}(b)$. Let $h\colon V(a) \to V(b)$ and $h_0\colon V_0(a) \to V_0(b)$ be the restrictions of $\widetilde{h}$ and $\widetilde{h}_0$. These are the diffeomorphisms that the flows for χ and ∇f_0 induce. Let $L = V(a) \times [a, b]$, resp. $L_0 = V_0(a) \times [a, b]$, be the product region produced by the flow for χ, resp., ∇f_0, and the Morse function f, resp f_0. Then L, resp. L_0, connects $N(p)$ to $N(q)$, resp. $N(p_0)$ to $N(q_0)$.

If $g_q \circ h = h_0 \circ g_p$ then we can use the flows to extend the diffeomorphsim

$$g_p \coprod g_q \colon N(p) \coprod N(q) \to N_0(p) \coprod N_0(q)$$

to an embedding

$$\varphi \colon N(p) \cup L \cup N(q) \to N_0(p) \cup L_0 \cup N_0(q)$$

that pulls f_0 back to f and sends the flow lines for χ to those for ∇f_0. The latter condition means that φ sendis a positive multiple $k(t,x,y)\chi$ of χ to ∇f_0 on these regions. Since these regions include neighborhoods of T and T' respectively, this would establish the result.

In fact, if $g_q \circ h = h_0 \circ g_p$ on some smaller neighborhood of $T \cap f^{-1}(a)$, then the result is established for the corresponding smaller region. The problem is that there is no reason to expect the flows for χ and ∇f_0 to produce diffeomorphisms h and h_0 satisfying $g_q \circ h \circ g_p^{-1} = h_0$ on any neighborhood of $T_0 \cap f^{-1}(a)$. We shall show:

Proposition 1.9. *There is an isotopy* $\widetilde{H}_t \colon V(a) \times I \to V(b)$ *satisfying:*

- $\widetilde{H}_0 = h$
- $g_q \circ \widetilde{H}_1(x) = h_0 \circ g_p(x)$ *for* x *in some neighborhood of* $T \cap V(a)$.
- $\widetilde{H}_t(T \cap V(a)) = T \cap V(b)$ *for all* t.
- *There is a compact subset* K *of* $T \cap V(a)$ *such that* $\widetilde{H}$ *is a constant isotopy outside of* K.
- *For each* $t \in [0,1]$ *the pre-image under* $\widetilde{H}_t$ *of the descending sphere of* q *at level* b *meets the ascending sphere for* p *at level* a *in a single point* $T \cap V(a)$ *and their intersection at this point is transverse.*

Given this proposition, we can extend the isotopy to an ambient isotopy $\widetilde{H}$ defined on all of $f^{-1}(a)$ that is constant outside of $V(a)$. For each t the image under $\widetilde{H}_t$ of the descending manifold for q intersects the ascending manifold for p only in $T \cap f^{-1}(b_0)$ and this intersection is transverse..

Given such an isotopy, then we can deform the vector field χ in $f^{-1}([a,b])$ by this isotopy to a new gradient-like vector field χ_1. The diffeomorphism $h' = \widetilde{H}_1$ is the result of flowing by χ_1. We have that $g_q \circ h' = h_0 \circ g_p$ near $T \cap f^{-1}(b_0)$. This shows that Proposition 1.9 implies Proposition 1.3.

We turn now to the proof of Proposition 1.9.

Proof. (Proposition 1.9) We use the given coordinates (t, x, y) restricted to Z. Then (x, y) are local coordinates on $V_0(a)$ and $V_0(b))$ centered at their intersections with T_0. The subspaces of $V_0(a)$ and $V_0(b)$ given by $\{x = 0\}$ and $\{y = 0\}$ are respectively the asscending manifold from p_0 and the descending manifold from q_0 for ∇f_0. Clearly, $h_0\colon V_0(a) \to V_0(b)$ preserves the subspaces $\{x = 0\}$ and $\{y = 0\}$. Since the flow χ and the flow for ∇f_0 agree under the identifications g_p and g_q, the identification $g_p\colon N(p) \to N(p_0)$ sends the intersection of $V(a)$ and the ascending sphere of p at level a defined by χ into the locus $\{x = 0\}$ in $V_0(a)$, and the identification $g_q\colon N(q) \to N(q_0)$ sends the intersection of $V(b)$ and the descending sphere of q at level b defined by χ into the locus $\{y = 0\}$ in $V_0(b)$. Let $P = h' g_0(V(a)) \subset V_0(b)$. We can take both P and $V_0(b)$ to be balls in the local coordinates (x, y) of $f_0^{-1}(b)$ and that g_p is defined on the closure of P. We introduce coordinates $u = (u^1, \ldots, u^{n-1})$ on $V_0(b)$ with $u = (x, y)$.

The statement that under the flow for χ the ascending manifold from p meets the descending manifold for q transversely in a single flow line means that the embedding $j = g_q h g_p^{-1}(h_0)^{-1}\colon P \to V_0(b)$ fixes the origin and that the image of subspace $\{x = 0\}$, is transverse to the subspace $\{y = 0\}$ at the origin, and otherwise the image of $\{x = 0\}$ is disjoint from $\{y = 0\}$.

Since j is an embedding fixing the origin we have

$$j(u^1, \ldots, u^{n-1}) = \sum_i u^i j_i(u^1, \ldots, u^{n-1}).$$

We define

$$j_t^0(u^1, \ldots, u^{n-1})) = \sum_i u^i j_i((1-t)u^1, \ldots, (1-t)u^{n-1})$$

this is the standard isotopy between j and the linear embedding Dj_0. (We may have to shrink P to a smaller ball to ensure that $j_t(P)$ lands in $V_0(b)$ for all $t \in [0, 1]$.) Of course, all of these maps fix the origin and are transverse to $\{y = 0\}$ there and are otherwise disjoint from $\{y = 0\}$. The linear map Dj_0 is an isomorphism transverse to $\{y = 0\}$. Next we deform Dj_0 through linear isomorphisms transverse to $\{y = 0\}$ to the identity. Putting these two isotopies together we have an isotopy defined on P, $J = j_t$, of j to the identity through embeddings fixing the origin and transverse to $\{y = 0\}$ at the origin and otherwise disjoint from $\{y = 0\}$.

Claim 1.10. *Fix a ball P_0 cetered at the origin with $\overline{P}_0 \subset P$. There are constants $a > 0$ and $A < \infty$ such that*

- *letting* $\pi\colon V_0(b) \to \{x=0\}$ *be the linear projection, for any* $u \in \{x = 0\} \cap \overline{P}_0$ *and any* $t \in [0,1]$ *we have* $|\pi j_t(u)| \geq a|u|x$*,*
- *and for any* $u \in \overline{P}_0$ *we have* $|\frac{\partial j(u,t}{\partial t}| \leq A|u|$*.*

Proof. The first is immediate from the compactness of the interval since for all t the map j_t is transverse to the origin. Since $j_t(0) = 0$ for all t we can write

$$j(u,t) = \sum_i u^i j_i(u,t),$$

and differentiating with respect to t gives the second inequality. □

Now fix a ball P_1 centered at the origin with $\overline{P}_1 \subset P_0$ sufficiently small so that $J(\overline{P}_1 \times I)$ is disjoint from $j(P \setminus P_0)$. We have an isotopy on $\overline{P}_1 \cup (P \setminus P_0)$ given by the union of $J|_{P_1 \times I}$ and the constant diffeomorphim j on $P \setminus P_0$. We can extend this to an isotopy of all of P. The way to this is to take restriction of the vector field ξ generating the isotopy J to a neighborhood W of $\overline{P}_1$ disjoint from $P \setminus \overline{P}_0$. Then damp $\xi|_W$ to zero by a function equal to 1 on $\overline{P}_1$ and with compact support in W. This new extended vector field then generates an isotopy isotopy on W that agrees with J on $\overline{P}_1$ and is the identity outside a compact set of W. We extend this isotopy by setting it equal to the identity on $P \setminus W$. It may happen that in the annular region $P_0 \setminus \overline{P}_1$ the image under the isotopy of the locus $\{x = 0\}$ meets $\{y = 0\}$. But because of the inequalities in the previous claim, the norm of the generating vector field at u is at most $A|u|$ where A is the constant in the preceding claim. Also, for any $u \in \{x = 0\}$ the projection $\pi(u)$ has norm at least $a|u|$. It follows that for $t \in [0,(a/A)]$ there are no points of $\{x = 0\}$ that are mapped by the isotopy into the locus $\{y = 0\}$ except the origin. This then produces the first step of the isotopy I^1. It is defined on the interval $[0,(a/A)]$; it agrees with $J|_{0,(a/A)]}$ on $\overline{P}_1$ and it is the identity outside a compact subset of P. Lastly, for each $t \in [0,(a/A)]$ the only point of intersection of $I^1_t(\{x = 0\})$ with $\{y = 0\}$ is the origin.

We use the isotopy I^1 as the beginning of the isotopy. Suppose by induction for some integer $s \geq 1$ with $s(a/A) < 1$ we have an isotopy I^s defined on $[0, s(a/A)]$ agreeing with $J_{[0,s(a/A)]}$ on a ball $\overline{P}_s \subset P_0$ centered at the origin equal to the identity outside a compact subset of P and with the image $I^s_t(\{x = 0\})$ meeting $\{y = 0\}$ only at the origin. We choose $\overline{P}_{s+1} \subset P_s$ such that $j_t(\overline{P}_{s+1})$ is disjoint from $j_{s(a/A)}(P \setminus P_s)$ for all $t \in [s(a/A),(s+1)(a/A)$. Extend the isotopy that is j_t for $t \in [s(a/A),(s+1)(a/A)]$ on $\overline{P}_{s+1}$ and the constant isotopy $I^s_{s(a/A)}$ on $P \setminus P_s$ as we extended before. The same argument using the inequalities in the preceding claim show that the image of $\{x = 0\}$ under this extended isotopy is disjoint from $\{y = 0\}$ except at the origin. This is the inductive step. (When $s(a/A) < 1 \leq (s+1)(a/A)$ we extend the isotopy in the same manner over the interval $[s(a/A), 1]$. Since this interval has length at most (a/A) the argument apples to it as well.)

This produces an isotopy I from $g_q h g_p^{-1}(h_0)^{-1}$ which at time 1 is the identity near the origin. The composition $I \circ h_0'$ is an isotopy from $g_q h g_p^{-1}$ to h_0 as required. This completes the proof of Claim 1.9. $\square$

This completes the proof of the proposition and hence of the first cancellation theorem.

2 The Second Cancellation Theorem

Recall the statement of this result.

Theorem 2.1. *(Second Cancellation Result) Suppose that W is a compact manifold of dimension $n \geq 6$ and that $f\colon W \to \mathbb{R}$ is a self-indexing Morse function without critical points of index $0, 1, n-1, n$. Suppose that $\partial_- W$ and $\partial_+ W$ are simply connected and let k be the smallest index of a critical point with p a critical point of index k and q a critical point of index $k+1$. Suppose that χ is a gradient-like vector field for f and that the ascending sphere of p at level $\ell = ((2k+1)/2)$ meets the descending sphere of q at level ℓ transversely. Fix orientations for the descending manifolds for p and q giving a sign to each intersection point of the descending sphere for q and the ascending sphere for p at level ℓ. If the sum of the local algebraic intersections is ± 1, then there is a new gradient-like vector field χ' for f, agreeing with χ outside $[k+(1/4), k+(3/4)]$ so that there is only one χ' flow connection between p to q, and this is a transverse flow line. Hence, we can cancel p and q using the First Cancellation Theorem applied to f and χ'.*

The key step in this proof is the *Whitney trick.*

Theorem 2.2. (Whitney Trick) *Let M be an oriented m-manifold with $m \geq 5$, and suppose that X^k and Y^{m-k} are smooth oriented submanifolds of M meeting transversely. Let p and q be two points of $X \cap Y$ with opposite intersection number, say intersection number -1 at q and $+1$ at p. Let*

$$N = \{(x,y) \big| x^2 + y^2 < (1+\epsilon)^2 \text{ and } y > -\epsilon\} \subset \mathbb{R}^2.$$

Suppose that there is a embedding of $N \subset M$ that meets X along $\gamma_X = N \cap \{y = 0\}$ and meets Y along $\gamma_Y = N \cap C$ where C is the unit circle and *that q, resp. p, is the image of* $(-1,0)$, *resp.* $(1,0)$. *Suppose that at any point of $a \in \gamma_X$ the normal vector to γ_X in N at b is not contained in T_bX and at any point $c \in \gamma_Y$ the normal vector to γ_Y in N at c is not contained in T_cY. Then there is an isotopy of X supported in a small neighborhood of γ_X that removes the points x and y from the intersection of X and Y and creates no new points of intersection.*

Proof. We endow N, the x-axis, and the unit circle C with their usual orientations. Let $\nu(N)$ denote the normal bundle of $N \subset M$. Let $\xi(X) \to \gamma_X$ be the normal bundle of $\gamma_X \subset X$ and $\xi(Y) \to \gamma_Y$ be the normal bundle of $\gamma_Y \subset Y$. The orientations of $N, M, \gamma_X, X, \gamma_Y$, and Y induce orientations on these normal bundles. We have decompositions of the fibers over q and p:

$$\nu(N)_q \cong \xi(X)_q \oplus \xi(Y)_q \tag{2.1}$$
$$\nu(N)_p \cong \xi(X)_p \oplus \xi(Y)_p. \tag{2.2}$$

These two isomorphisms of oriented bundles at p, q both act by $(-1)^{k-1}$ on the orientations.

We have sub-bundles

$$\xi(X) \subset \nu(N)|_{\gamma_X}$$
$$\xi(Y) \subset \nu(N)|_{\gamma_Y}.$$

Choose a metric on $\nu(N \subset M)$ such that over p and q the complementary subspaces $\xi(X)|_{\{p\}\cup\{q\}}$ and $\xi(Y)|_{\{p\}\cup\{q\}}$ are orthogonal. Let ζ_X be the sub-bundle over $\gamma_X \cup \gamma_Y$ that is equal to $\xi(X)$ over γ_X and is equal to the orthogonal complement to ξ_Y in $\nu(N)$ over γ_Y. Define ζ_Y symmetrically. Then ζ_X and ζ_Y are mutually orthogonal, complementary sub-bundles of $\nu(N)|_{\gamma_X\cup\gamma_Y}$. Because the signs of the maps induced on the orientations in Equations 2.1 and 2.2 are the same, each of these sub-bundles is orientable, and hence isomorphic to a trivial bundle. Lastly, we show that ζ_X and ζ_Y extend to complementary bundles $\widetilde{\zeta}_X$ and $\widetilde{\zeta}_Y$ of $\nu(N)$ defined over all of N. To do that we show that ζ_X and ζ_Y can be deformed to be constant with respect to the trivialization of $\nu(N)|_{\gamma_X\cup\gamma_Y}$ coming from the fact that the latter bundle extends on N which is contractible. Fixing this trivialization for $\nu(N)$ and a trivialization of ζ_X, the sub-bundle ζ_X is described as a family of oriented $(k-1)$-planes in $\mathbb{R}^{m-2}$ parameterized by a loop. The space of such oriented subspaces is $SO(m-2)/SO(k-1) \times SO(m-k-1)$. Since $m-2 \geq 3$, this space is simply connected and hence ζ_X deforms to a constant sub-bundle. Taking the orthogonal complements gives a deformation of ζ_Y through complementary sub-bundles to a oonstant sub-bundle.

Once we have deformed ζ_X and ζ_Y to constant sub-bundles of $\nu(N)$ in its trivialization, they extend to complementary constant sub-bundles $\widetilde{\zeta}_X$ and $\widetilde{\zeta}_Y$ of $\nu(N)$ Of course, these sub-bundles are trivial.

For any r, let B^r denote the unit ball in $\mathbb{R}^r$. The trivial complementary sub-bundles $\widetilde{\zeta}_X$ and $\widetilde{\zeta}_Y$ produce a product structure $N \times B^{k-1} \times B^{m-k-1}$ for a neighborhood of N in M meeting X in $U = \gamma_X \times B^{k-1} \times \{0\}$ and meeting Y is $\gamma_Y \times \{0\} \times B^{m-k-1}$.

Let $\rho\colon [0,1] \to [0,1]$ be a smooth function identically 1 near 0 and identically 0 near 1. Let $\varphi_t\colon \gamma_X \to N$ be an isotopy, constant on $(-1-\epsilon, -1-\epsilon/2) \cup (1+\epsilon/2, 1+\epsilon)$, from the identity to a map that sends γ_X to an arc missing the intersection of N with unit disk. Define an isotopy

$$H_t\colon U \to N \times B^{k-1} \times B^{m-k-1}$$

by $H_t(a, \overline{x}, 0) = (\varphi_{\rho(|\overline{x}|)t}(a), \overline{x}, 0)$. The isotopy H_t is an isotopy from the inclusion of U into M to a map $h_1\colon U \to M$ that is disjoint from Y and which agrees with the inclusion outside a compact subset of $\gamma_X \times B^{k-1}$. This deformation removes the points p, q from the intersection of U with Y without creating any addition points of intersection of U with Y. We extend this isotopy of U to an isotopy of X that is constant outside of U. This is the required isotopy of $X \subset M$ from the inclusion to a map $\widetilde{h}_1$ with the property that $\widetilde{h}_1(X) \cap Y$ is equal to $X \cap Y \setminus (\{p\} \cup \{q\})$ and such that $\widetilde{h}_1$ agree with the inclusion outside of U. $\square$

Proposition 2.3. *Suppose that W is a compact manifold of dimension $n \geq 6$ and $f\colon W \to \mathbb{R}$ is a self-indexing Morse function without critical points of index $0, 1, n-1, n$. Suppose that $\partial_- W$ and $\partial_+ W$ are simply connected. Suppose k is the smallest index of a critical point of f. Set $\ell = (2k+1)/2$ and suppose in the level set $f^{-1}(\ell)$ the intersection of the ascending sphere for a critical point p of index k and the descending manifold for a critical points q of index $k+1$ is transverse and contains two points x and y with opposite sign of intersection. Then there is an isotopy of the descending manifold for q that removes these two points of intersection without adding any additional points of intersection with any ascending sphere of any critical point of index k.*

Proof. Let q and p be the points of intersection of ascending sphere $S^{n-k-1}(p)$ and the descending sphere $S^k(q)$ at level $f^{-1}(\ell)$ with opposite sign of the local intersection numbers. Fix arcs $\gamma_X\colon (-1-\epsilon, 1+\epsilon) \to X$ and $\gamma_Y\colon (-1-\epsilon, 1+\epsilon) \to Y$ with $\gamma_X(-1) = \gamma_Y(-1) = q$ and $\gamma_X(1) = \gamma_Y(1) = p$ Let

$$N = \{(x,y) | x^2 + y^2 < (1+\epsilon)^2 \text{ and } Y > -\epsilon\}.$$

Let $A \subset N$ be the intersection of N with the x-axis and let $C \subset N$ be the intersection of N with the unit circle. There is an embedding of a regular neighborhood V of $C \cup A \subset N$ into M so that A maps to γ_X and C maps to γ_Y and the conditions required in the previous proposition for the normal vectors to γ_X and γ_Y are satisfied.

Claim 2.4. *The complement in $f^{-1}(\ell)$ of the union of the ascending spheres for critical points of index k and the descending spheres for critical points of index $k+1$ is simply connected.*

Suppose this claim for the moment. Let α be a loop in V contained in the interior of the intersection of N with the open unit disk that generates the fundamental group of V. Since α is homotopically trivial in the complement of the union of the ascending spheres for critical points of index k and the descending spheres for critical points of index $k+1$ and since the dimension of $f^{-1}(\ell)$ is at least 5t, it follows that α bounds an embedded disk disjoint from these two spheres. This allows us to extend the map $V \to M$ to a map $N \to M$ as required by the previous proposition with that the property that the intersection of N with the union of the ascending spheres for critical points of index k and descending spheres for critical points of index $k+1$ is exactly $\gamma_X \cup \gamma_Y$.. According to that proposition, this will allow us to cancel two points of intersection of $S^{n-k-1}(p)$ and $S^k(q)$. Examination of the proof shows that since N meets this intersection of spheres only in $\gamma_X \cup \gamma_Y$, the deformation adds no new points of intersection between these spheres.

Since the algebraic intersection of the spheres is ± 1 we can remove the points of intersection in pairs until only one is left. At this point we use the variation of critical values to lower the critical value of q slightly and raise the critical value of p slightly so that there is $\epsilon > 0$ such that setting $a = k + \epsilon$ and $b = k + 1 - \epsilon$, both a and b are regular values and $W' = f^{-1}([a, b])$ contains the critical points p and q in its interior. This puts us in a situation where the hypotheses of the First Cancellation Theorem are satisfied for $f|_{W'} \colon W' \to [a, b]$ with critical points q and p. Hence, we can cancel these, replacing $f|_{W'}$ by a Morse function g without critical points and agreeing with f near $\partial W'$. The union of g and $f|_{W \setminus \mathrm{int}(W')}$ has all the critical points of f except for p and q, which have been removed. This establishes the result modulo the claim.

Proof. (of claim) First, let us show that the level set $f^{-1}(\ell)$ is simply connected. The reason is that $f^{-1}((-\infty, \ell])$ is obtained from $\partial_- W$ by adding handles of dimension ≥ 2 and hence is simply connected, and also

$f^{-1}((-\infty, \ell])$ is obtained from $f^{-1}(\ell)$ by adding handles of dimension ≥ 3 implying that $f^{-1}(\ell)$ is also simply connected.

Now suppose that $2 < k < n-3$. In this case in $f^{-1}(\ell)$ the ascending spheres for the critical points of index k are of codimension $k \geq 3$ and the descending for the critical points of index $k+1$ are of codimension $n-k-1 \geq 3$. Hence, by general position, removing the union of all these spheres does not change the fundamental group, so that the complement of the union of these spheres in $f^{-1}((\ell)$ is simply connected, establishing the claim when $2 < k < n-3$.

Now suppose thar $k = 2$. Since there are no critical points of index 0 or 1, the complement X of the ascending spheres for the critical points of index 2 in $f^{-1}(\ell)$ is diffeomorphic to the complement in $\partial_- W$ of the union of the descending spheres (one-dimensional spheres). Since $\partial_- W$ is simply connected and of dimenson ≥ 5, it follows that X is simply connected. Now we must remove from X its intersection with the descending spheres for critical points of index 3. These spheres are of codimension $n-3$. Since $n \geq 6$, they are index ≥ 3. Hence, by general position removing them does not change the fundamental group. This establishes the claim when $k = 2$.

Lastly, we must consider the case $k = n-3$. In this case we are dealing with critical points of index $n-3$ and $n-2$. Turning the Morse function produces the situation $k = 2$, and the analysis in the case $k = 2$ applies to establish the claim in this case as well. $\square$

This completes the proof of the proposition and hence of the second cancellation theorem. $\square$