On the Diffeomorphism Group of $S^1 \times S^2$

Allen Hatcher

This is a revision, written in December 2003, of a paper of the same title that appeared in the Proceedings of the AMS 83 (1981), 427-430.

The main result of this paper is that the group $\mathrm{Diff}(S^1 \times S^2)$ of diffeomorphisms $S^1 \times S^2 \to S^1 \times S^2$ has the homotopy type one would expect, namely the homotopy type of its subgroup of diffeomorphisms that take each sphere $\{x\} \times S^2$ to a sphere $\{y\} \times S^2$ by an element of the isometry group $O(3)$ of S^2, where the function $x \mapsto y$ is an isometry of S^1, an element of $O(2)$. It is not hard to see that this subgroup is homeomorphic to the product $O(2) \times O(3) \times \Omega SO(3)$, this last factor being the space of smooth loops in $SO(3)$ based at the identity. This has the same homotopy type as the space of continuous loops. These loopspaces have $H_{2i}(\Omega SO(3); \mathbb{Z})$ nonzero for all i, so we conclude that $\mathrm{Diff}(S^1 \times S^2)$ is not homotopy equivalent to a Lie group. Diffeomorphism groups of surfaces and many irreducible 3-manifolds are known to be homotopy equivalent to Lie groups (often discrete groups in fact), and $S^1 \times S^2$ is the simplest manifold for which this is not true.

Via the Smale conjecture, proved in [H], the calculation of the homotopy type of $\mathrm{Diff}(S^1 \times S^2)$ reduces easily to a problem about making families of 2-spheres in $S^1 \times S^2$ disjoint. To state the problem in slightly more generality, let M^3 be a connected 3-manifold containing a sphere $S^2 \subset M^3$ that does not bound a ball in M^3, and let $S^2 \times [-1, 1] \subset M^3$ be a bicollar neighborhood of this S^2. Let $\mathcal{E}$ be the space of smooth embedding $f : S^2 \to M^3$ whose image does not bound a ball, and let $\mathcal{E}'$ be the subspace of embeddings f for which there exists $x \in [-1, 1]$ with $f(S^2)$ disjoint from $\{x\} \times S^2$. The result we need is:

Theorem. *If any two embedded* 2*-spheres in* M *that do not bound a ball are isotopic, then the inclusion* $\mathcal{E}' \hookrightarrow \mathcal{E}$ *is a homotopy equivalence.*

The hypothesis is satisfied for both S^2-bundles over S^1, and one can determine the homotopy type of $\mathrm{Diff}(M^3)$ for the nonorientable bundle just as for $S^1 \times S^2$. The only other 3-manifolds that satisfy the hypothesis of the theorem are connected sums of two irreducible 3-manifolds. For these the theorem implies that $\mathrm{Diff}(M^3)$ has the homotopy type of the subgroup leaving invariant a 2-sphere that splits M as a connected sum. This essentially reduces the calculation of the homotopy type of $\mathrm{Diff}(M^3)$ to the calculation for the two irreducible summands.

Proof: Since $\mathcal{E}$ and $\mathcal{E}'$ are homotopy equivalent to CW complexes, it suffices to show that the inclusion $\mathcal{E}' \hookrightarrow \mathcal{E}$ induces isomorphisms on all homotopy groups. Represent an element of $\pi_k(\mathcal{E},\mathcal{E}')$ by a smooth family of embeddings $f_t \in \mathcal{E}$ for $t \in D^k$, with $f_t \in \mathcal{E}'$ for $t \in \partial D^k$. Since $\mathcal{E}'$ is open in $\mathcal{E}$ there is a subdisk $D_0^k \subset \mathrm{int}\, D^k$ such that $f_t \in \mathcal{E}'$ for $t \in D^k - \mathrm{int}\, D_0^k$. Choose a basepoint $* \in S^2$ and let $p_t = f_t(*)$ and $M_t = f_t(S^2)$.

Our first task is to find a finite number of slices $N_i = \{x_i\}\times S^2 \subset [-1,1]\times S^2 \subset M^3$ together with closed k-balls $B_i \subset \mathrm{int}\, D^k$ such that:

(1) M_t is transverse to N_i for all $t \in B_i$.
(2) The interiors of the B_i's form an open cover of D_0^k.
(3) $N_i \neq N_j$ for $i \neq j$.
(4) $p_t \notin N_i$ for $t \in B_i$.

For fixed $t \in D_0^k$ the slices $\{x\}\times S^2$ that are transverse to M_t are dense in $[-1,1]\times S^2$ by Sard's theorem, so we may choose such a slice that is disjoint from p_t. This slice will remain transverse to M_t and disjoint from p_t for all nearby t as well, say for t in a ball B_t centered at t. By compactness the cover of D_0^k by the interiors of the balls B_t has a finite subcover, so we have a finite collection of balls B_i with corresponding slices N_i satisfying (1), (2), and (4). By a small perturbation of the slices N_i we can achieve (3) as well without affecting the other three conditions since if M_t is transverse to a slice then it is also transverse to all nearby slices.

Let C_t^i be the collection of circles of $M_t \cap N_i$ for $t \in B_i$. Thus C_t^i is a finite set. Let $C_t = \cup_i C_t^i$, the union over all i such that $t \in B_i$. Each circle $c_t \in C_t$ bounds a unique disk $D_M(c_t) \subset M_t - \{p_t\}$. Choose functions $\varphi_t : C_t \to (0,1)$ such that

(5) $\varphi_t(c_t) < \varphi_t(c_t')$ whenever $D_M(c_t) \subset D_M(c_t')$

with $\varphi_y(c_t)$ varying smoothly with $t \in B_i$ for $c_t \in C_t^i$. For example we could let $\varphi_t(c_t)$ be the area of the disk $f_t^{-1}(D_M(c_t)) \subset S^2$, where the total area of S^2 is normalized to be 1.

It will be very convenient to have one further condition satisfied:

(6) φ_t is injective on C_t^i for each i with $t \in B_i$.

To achieve this, first replace each N_i by $k+1$ nearby slices N_{ij}. If these are sufficiently close to N_i the conditions (1), (3), and (4) will still hold. Let C_t^{ij} be the set of circles of $M_t \cap N_{ij}$ and let C_t be the union of the C_t^{ij}'s. Choose functions $\varphi_t : C_t \to (0,1)$ satisfying (5) as before. For each N_{ij} there is a subset K_{ij} of B_i where φ_t is not injective on C_t^{ij}. After a small perturbation of the functions φ_t to make the graphs of the functions $t \mapsto \varphi_t(c_t)$ have general position intersections in $D^k \times (0,1)$ we may assume that each K_{ij} is a finite union of codimension one submanifolds of B_i, the submanifolds where the values of φ_t on two circles in C_t^{ij} coincide, and we may assume that all these codimension one submanifolds have general position intersections. In particular this implies that $\cap_j K_{ij}$ is empty, being a union of general position intersections of $k+1$ codimension one submanifolds of D^k. Thus for small enough open neighborhoods V_{ij} of K_{ij} in B_i we have $D_0^k = \cup_{i,j}\operatorname{int}(B_i - V_{ij})$. By construction φ_t is injective on C_t^{ij} for $t \in B_i - K_{ij}$. Now choose finitely many small balls B_{ijl} covering $B_i - V_{ij}$ disjoint from K_{ij}, with corresponding slices N_{ikl} near N_{ij} such that (1)-(4) hold for these. Each circle c_t in $M_t \cap N_{ij}$ determines a nearby circle c_t^l in $M_t \cap N_{ijl}$, and we choose for $\varphi_t(c_t^l)$ a value near $\varphi_t(c_t)$ such that (5) holds for the circles $c_t^1, c_t^2, \cdots$. (For example, we could obtain $\varphi_t(c_t^l)$ from $\varphi_t(c_t)$ by adding or subtracting some small constant times the area of the annular region between $f_t^{-1}(c_t)$ and $f_t^{-1}(c_t^l)$ in S^2.) With $\{B_{ijl}\}$ for $\{B_i\}$ and $\{N_{ijl}\}$ for $\{N_i\}$ we still have (1)-(5), and (6) holds since any two circles c_t^l in $M_t \cap N_{ijl}$ are associated to different c_t's in $M_t \cap N_{ij}$, and the φ_t values of these c_t's are different since B_{ijl} is disjoint from K_{ij}.

By compactness of the balls B_i there is an $\varepsilon > 0$ such that the conditions (5) and (6) take the stronger forms

(5_ε) $\varphi_t(c_t) < \varphi_t(c_t') - \varepsilon$ whenever $D_M(c_t) \subset D_M(c_t')$.

(6_ε) $|\varphi_t(c_t) - \varphi_t(c_t')| > \varepsilon$ for all pairs $c_t \neq c_t'$ in C_t^i.

After these preliminaries we now begin the construction of an isotopy M_{tu} of $M_t = M_{t0}$ which eliminates all the circles of C_t. First we describe the construction of M_{tu} for a fixed value of t, and then after this is done we will describe the small modifications needed to make M_{tu} depend continuously on t.

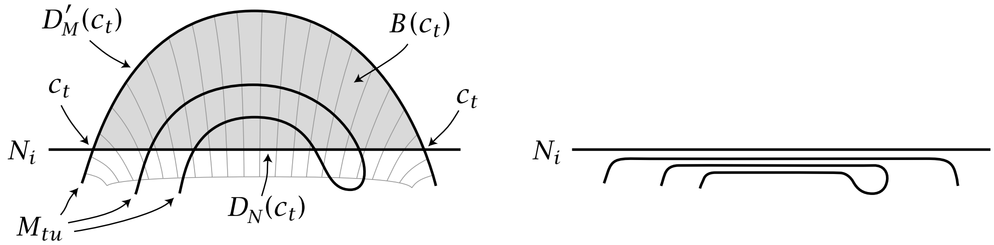
For fixed t, suppose inductively that for some $c_t \in C_t$ we have constructed M_{tu} for $u \leq \varphi_t(c_t)$ and the following conditions are satisfied:

(a) The isotopy M_{tu}, $u \leq \varphi_t(c_t)$, is stationary in a neighborhood of c_t.

(b) The isotopy M_{tu} for $u \leq \varphi_t(c_t)$ moves $D_M(c_t)$ to a disk $D_M'(c_t)$ with the property that $\operatorname{int}(D_M'(c_t)) \cap N_j = \varnothing$ for each j such that $t \in B_j$.

We call such a c_t a *primary* circle of C_t. Let N_i be the slice containing c_t. Since $D_M'(c_t) \cap N_i = c_t$, then since the manifold M^3 has exactly two isotopy classes of embedded spheres, exactly one of the two disks into which N_i is cut by c_t, say $D_N(c_t)$, is such that the 2-sphere $D_M'(c_t) \cup D_N(c_t)$ bounds a 3-ball in M^3. Denote this 3-ball by $B(c_t)$. Its boundary has a corner along c_t, with interior angle less than π rather than greater than π, otherwise N_i would be contained in the ball $B(c_t)$. Note that $B(c_t) \cap N_j = \varnothing$ for each $j \neq i$ with $t \in B_j$, since $\partial B(c_t) \cap N_j = \varnothing$ by (b).

The isotopy M_{tu} for $\varphi_t(c_t) \leq u \leq \varphi_t(c_t) + \varepsilon$ is now constructed to eliminate c_t by isotoping $D_M'(c_t)$ across $B(c_t)$ to $D_N(c_t)$ and a little beyond. If there are any other circles of C_t^i in $\operatorname{int}(D_N(c_t))$ remaining at time $u = \varphi_t(c_t)$, this isotopy also eliminates them, as indicated in the figure below which shows the analogous situation one dimension lower. We call such circles *secondary* circles.



$u = \varphi_t(c_t)$ $\qquad$ $u = \varphi_t(c_t) + \varepsilon$

Note that this isotopy eliminating the primary circle $c_t \in C_t^i$ during the u-interval $[\varphi_t(c_t), \varphi_t(c_t)+\varepsilon]$ does not change $M_{t,\varphi_t(c_t)} \cap N_j$ for any B_j containing t with $j \neq i$ since $B(c_t) \cap N_j = \varnothing$ for these slices.

If the interval $[\varphi_t(c_t), \varphi_t(c_t)+\varepsilon]$ overlaps another interval $[\varphi_t(c_t'), \varphi_t(c_t')+\varepsilon]$ for a primary circle $c_t' \in C_t^j$, then by (5_ε) the disks $D'_M(c_t)$ and $D'_M(c_t')$ are disjoint, and by (6_ε) we have $i \neq j$ so the disks $D_N(c_t)$ and $D_N(c_t')$ are disjoint. It then follows from (b) that the boundary spheres of the balls $B(c_t)$ and $B(c_t')$ are disjoint, and in fact that the balls themselves are disjoint. The two isotopies eliminating c_t and c_t' thus have disjoint supports and can be performed independently.

The process of eliminating circles of C_t can now be repeated inductively, to produce an isotopy M_{tu} for $0 \le u \le 1$ with the final M_{t1} disjoint from all N_i with $t \in B_i$.

It remains to make the isotopies M_{tu} depend continuously on t. The most obvious obstacle to continuity is the fact that as t leaves a ball B_i the circles of C_t^i are deleted from C_t and hence an isotopy eliminating a primary circle $c_t \in C_t^i$ during the u-interval $[\varphi_t(c_t), \varphi_t(c_t)+\varepsilon]$ is suddenly not performed. To fix this problem we introduce a tapering process. For each i let $B_i' \subset \operatorname{int} B_i$ be a concentric ball such that the interiors of the B_i''s still cover D_0^k. Let $\psi_i : B_i \to [0,1]$ be such that $\psi_i(\partial B_i) = 0$ and $\psi_i(B_i') = 1$. Then we refine the prescription for M_{tu} by specifying that for an isotopy eliminating a primary circle $c_t \in C_t^i$ during the u-interval $[\varphi_t(c_t), \varphi_t(c_t)+\varepsilon]$, only the portion of this isotopy with $u \le \psi_i(t)$ is to be used. Thus for $u > \psi_i(t)$ we simply forget about the slice N_i and the way M_{tu} intersects N_i. This creates no problems since isotopies eliminating primary circles of C_t^i have no effect on circles of C_t^j for $j \neq i$.

The other thing we need to do to make M_{tu} depend continuously on t is to arrange that the isotopies eliminating the primary circles $c_t \in C_t$ vary continuously with t. Let us first specify more precisely how these isotopies are to be constructed. For a neighborhood $N(D'_M(c_t))$ of $D'_M(c_t)$ in M_{tu} choose a collar $N(D'_M(c_t)) \times [0,1) \hookrightarrow M^3$

containing $B(c_t)$ and disjoint from the N_j's not containing c_t, with $N(D'_M(c_t))\times\{0\} = N(D'_M(c_t))$. Then the isotopy of $D'_M(c_t)$ just slides each point x along the arc $\{x\}\times[0,1)$. The family of isotopies M_{tu} will then be continuous if we can choose the collars to vary continuously with t.

After a small perturbation we may assume the graphs of all the functions φ_t and ψ_i have general position intersections. The projections of these intersections into D^k then give a stratification of D^k, whose strata are open manifolds of various dimensions. Consider the problem of constructing M_{tu} over a stratum, assuming inductively that the construction has already been made over strata of lower dimension, and in particular over the boundary of the stratum. The ordering of the circles of C_t by the functions φ_t is the same throughout the stratum. By induction we may assume M_{tu} has already been constructed for $u \le \varphi_t(c_t)$ for some primary circle c_t. As part of this construction we have already chosen collars on $N(D'_M(c_t))$ over the boundary of the stratum, and we wish to extend these collars over the stratum itself. The stratum can be obtained from its boundary by attaching a sequence of handles, so it suffices to construct collars over a handle $D^n\times D^{k-n}$ agreeing with given collars over $\partial D^n\times D^{k-n}$. First extend over a neighborhood of $\partial D^n\times D^{k-n}$ using isotopy extension. Call these collars *old collars*. Since a handle is a k-dimensional disk, collars over the handle itself exist by isotopy extension. Call these *new collars*. To make the old and new collars agree near $\partial D^n\times D^{k-n}$ we first push the old collars away from $N(D'_M(c_t))$ by sliding them along the $[0,1)$ factors of the new collars, compressing the interval $[0,1)$ into the subinterval $[\delta(t),1)$, where $\delta(t)$ goes from 0 on $\partial D^n\times D^{k-n}$ to a value near 1 as we move away from $\partial D^n\times D^{k-n}$, a value close enough to 1 so that the subcollars $N(D'_M(c_t))\times[0,1-\delta(t))$ contain $B(c_t)$. Then we can trim away the undesired parts of the old and new collars to create a continuously varying family of hybrid collars. (Details are left to the reader.)

Having the family of isotopies M_{tu} of the submanifolds M_t we can apply isotopy extension to get a family of isotopies f_{tu} of the embeddings f_t with $f_{tu}(S^2) = M_{tu}$. The balls B_i were chosen to be disjoint from ∂D^k so M_{tu} is independent of u for $t\in\partial D^k$, and we may assume the same is true of f_{tu}. Thus f_{tu} provides a homotopy of the given map $(D^k,\partial D^k)\to(\mathcal{E},\mathcal{E}')$ to a map with image in $\mathcal{E}'$, finishing the proof that $\pi_k(\mathcal{E},\mathcal{E}') = 0$ □

Corollary. *The map $O(2)\times O(3)\times\Omega SO(3)\to \mathrm{Diff}(S^1\times S^2)$ sending (α,β,γ) to the diffeomorphism $(x,y)\mapsto(\alpha(x),\beta(y)\gamma_x(y))$, is a homotopy equivalence.*

Proof: Let $G\subset \mathrm{Diff}(S^1\times S^2)$ be the subgroup described at the beginning of the paper, consisting of diffeomorphisms of the form $(x,y)\mapsto(\alpha(x),\beta_x(y))$ for $\alpha\in O(2)$ and $\beta_x\in O(3)$. To show that $\pi_k(\mathrm{Diff}(S^1\times S^2),G) = 0$ for all k, start with a family of diffeomorphisms $g_t: S^1\times S^2\to S^1\times S^2$ representing an element of this relative homotopy group. By the theorem we may assume the embeddings $f_t = g_t\,|\,\{x_0\}\times S^2$ are in $\mathcal{E}'$ for all t. The projection of $f_t(S^2)$ onto S^1 is then an arc varying continuously with t, so we may choose a point $x_t \in S^1$ outside this arc, also varying continuously with t. Thus we may view f_t as a family of embeddings of S^2 in the complement of $\{x_t\} \times S^2$, which we can identify with $S^2 \times (0,1)$. By the Smale conjecture the space of 2-spheres in $S^2 \times (0,1)$ that do not bound a ball is contractible, so we can deform the family g_t, staying fixed for $t \in \partial D^k$, so that $g_t(\{x_0\} \times S^2)$ is a sphere $\{y_t\} \times S^2$ for all t. Again by the Smale conjecture $\mathrm{Diff}(S^2 \times I)$ has the homotopy type of the subgroup of diffeomorphisms taking slices $S^2 \times \{x\}$ to slices $S^2 \times \{y\}$, so after a further deformation of g_t we may assume it has this property as well. Since the inclusion $O(3) \hookrightarrow \mathrm{Diff}(S^2)$ is a homotopy equivalence by [S] we can assume further that each restriction $g_t \mid \{x\} \times S^2$ lies in $O(3)$. Finally, by lifting a deformation retraction of $\mathrm{Diff}(S^1)$ into $O(2)$ we can deform g_t into G. All these deformations can be assumed to be fixed for $t \in \partial D^k$. Thus we have $\pi_k(\mathrm{Diff}(S^1 \times S^2), G) = 0$. (Note that this argument works also for the nonorientable S^2 bundle over S^1, with the appropriately modified definition of G.)

Projecting $S^1 \times S^2$ onto S^1 gives a homomorphism $G \to O(2)$ whose kernel K can be identified with the group of smooth maps $S^1 \to O(3)$. This homomorphism is a principal bundle, and it has a cross section, so it is a product bundle and G is homeomorphic to $O(2) \times K$. (Algebraically, G is only a semidirect product, not a product.) Since $O(3)$ is a group, the space K of smooth maps $S^1 \to O(3)$ is homeomorphic to the product of $O(3)$ with the space $\Omega SO(3)$ of smooth loops in $SO(3)$ based at the identity. □

The loopspace $\Omega SO(3)$ has two path components, and they are homotopy equivalent as is the case for all loopspaces. The path component consisting of homotopically trivial loops can be identified with ΩS^3 since S^3 is the universal cover of $SO(3)$. It is one of the standard applications of the Serre spectral sequence to compute that $H_i(\Omega S^3; \mathbb{Z})$ is $\mathbb{Z}$ for i even and 0 for i odd.

References

[H] A. Hatcher, A proof of the Smale conjecture, Ann. of Math. 117 (1983), 553-607.

[S] S. Smale, Diffeomorphisms of the 2-sphere, Proc. A.M.S. 10 (1959), 621-626.