

## DECOMPOSITION THEOREMS FOR GROUPS OF DIFFEOMORPHISMS IN THE SPHERE

R. DE LA LLAVE AND R. OBAYA

**ABSTRACT.** We study the algebraic structure of several groups of differentiable diffeomorphisms in  $\mathbf{S}^n$ . We show that any given sufficiently smooth diffeomorphism can be written as the composition of a finite number of diffeomorphisms which are symmetric under reflection, essentially one-dimensional and about as differentiable as the given one.

### 1. INTRODUCTION

The goal of this paper is to prove several decomposition theorems for diffeomorphism groups.

These theorems, roughly, state that any element of the group can be written as a finite product of elements lying in a smaller subgroup endowed with special properties, such as symmetry.

In our case, the groups considered will be groups of differentiable diffeomorphisms of the sphere, and the subgroups into which we factor them will be groups of diffeomorphisms that commute with reflections across a plane and which are essentially one dimensional. The factors into which we can decompose a given map will be slightly less differentiable than the original one. When we consider groups of  $C^\infty$  diffeomorphisms, the factors are also  $C^\infty$ .

There are several motivations for the study of theorems of this type. For example, in [5] it is shown that a theorem of this type for the circle can be used to solve the inverse problem for scattering of geodesic fields in surfaces of genus one. This problem admits the following physical interpretation: The given diffeomorphism of a circle can be thought of as the transformation to be effected by a lens to be constructed; then, find the distribution of the refractive index that produces the desired effect. The problem is such that if it is solved for two diffeomorphisms, it is solved for their composition. Moreover, for symmetric mappings a simple construction works. Hence, the decomposition theorem shows that it can be solved for all mappings. For this application the loss of differentiability incurred in the factorization does not essentially affect the conclusions.

We also note that these theorems are analogues of the usual factorization theorems in Lie algebras, and they could be useful in the problem of computing representations of diffeomorphism groups [1].

If we consider these theorems as infinite dimensional versions of factorization theorems for Lie groups, one first difficulty is that for diffeomorphism groups, the

---

Received by the editors October 24, 1997.

1991 *Mathematics Subject Classification.* Primary 58D05, 57S25, 57S05.

*Key words and phrases.* Decomposition theorems, diffeomorphism groups.

corresponding Lie algebra – the Lie algebra of vector fields – is considerably worse behaved than those of finite dimensional Lie groups. For example, the exponential is not surjective in any neighborhood of the identity. Hence, it is not possible to apply the usual implicit function theorem in Banach spaces to obtain results in a small enough neighborhood of the origin using the “infinitesimal results” obtained from the Lie algebra. Nevertheless, we will show that it is possible to obtain results similar to those in the finite dimensional case by using appropriate “hard” implicit function theorems of the Nash–Moser type. However, one needs to reformulate the problem so that it has some group structure (in the sense of Zehnder [10]). This requires some assumptions on the structure of the manifold. (Of course, the existence of the reflections is included in the assumptions.) Manifolds that satisfy this assumptions are tori [6], and in this paper we will show that it also works for spheres. Presumably other manifolds could be included (certainly products of spheres and tori), but we have not pursued this matter.

In our proof we prove the theorem by induction on the dimension. The construction is somewhat delicate, since we want to obtain a loss of differentiability that is independent of the dimension, in spite of the fact that the number of inductive steps has to grow with the dimension.

## 2. ACKNOWLEDGMENTS

The work of R.L. has been supported by NSF grants. Also R.L. wants to acknowledge the hospitality enjoyed at Valladolid during several visits. The work of R.O. has been partially supported by Junta de Castilla y Leon. Letters from J. Langer and F. Biehn discussing the implications of these theorems were very useful to us, and also provided encouragement.

## 3. DECOMPOSITION THEOREM FOR SPHERES.

**3.1. Notation and statement of results.** We start this section by recalling some known results which show how to endow certain function spaces on compact sets with a differentiable manifold structure, necessarily of infinite dimension.

If  $l$  is an integer and  $\Omega$  is a compact set in  $\mathbb{R}^d$ , we will say that  $f : \Omega \rightarrow \mathbb{R}$  is  $C^l$  if it has continuous derivatives of order  $l$ . We will denote by  $C^l(\Omega)$  the space of  $C^l$  functions in  $\Omega$  endowed with the norm  $\| \cdot \|_{C^l(\Omega)}$  defined as the supremum of the derivatives of order up to  $l$ . It is well known that  $C^l(\Omega)$  is a Banach space.

If  $l$  is not an integer, we define, as usual,

$$(1) \quad \|f\|_{C^l(\Omega)} = \max(\sup_{x \neq y} \|D^l f(x) - D^l f(y)\| / |x - y|^{l-[l]}, \|f\|_{C^{[l]}(\Omega)}).$$

To cover Lipschitz conditions as well, we introduce the notation  $l^- = (l - 1) + 1^-$ . In this notation,  $1^-$  will be bigger than any number in  $[0, 1)$  but smaller than 1.

These notions of regularity can be lifted to geometric objects that can be expressed in coordinates by saying that an object is  $C^l$  if we can find coordinate patches that cover the manifold on which the expression is  $C^l$ .

Furthermore, if we fix a set of coordinate patches that cover the manifold, we can talk about the  $C^l$  distance between two objects by declaring it to be the maximum of the  $C^l$  distance between the coordinate expressions. A moment's reflection will show that the statement that a geometric object is  $C^l$  is independent of the coordinate patches chosen (that is, if the coordinate expressions in one patch are  $C^l$ , then they are  $C^l$  in all others).

Unfortunately, the  $C^l$  distances do depend on the patch. Even if it is easy to show that the distances obtained using two patches are equivalent, they do differ, and some statements (such as that certain operators acting on  $C^l$  functions are contractions) require specific choices of coordinate patches. When  $l$  is an integer, there is a geometrically natural notion of  $C^l$  distance in Riemannian geometry based on the notion of jets. (See e.g. [3], Ch. 2.) For non-integer  $l$ , since the definition of Hölder distance (1) requires comparing the values at different points, there are no geometrically natural notions of  $C^l$  distance. For our purposes, the dependence on a coordinate patch is not a shortcoming, since many analytical operations that we will need to perform require taking coordinates anyway, and we will always take a fixed coordinate system.

We will denote by  $\text{Diff}^l(M, N)$ , the set of diffeomorphisms of  $M$  into  $N$ , and when  $M = N$  we will simply write  $\text{Diff}^l(N)$ . When  $l \geq 1$  is finite, these sets can be given the structure of Banach manifolds as follows. Using charts, we can define a norm on the space of  $C^l$  vector fields on  $N$  that makes it a Banach space. Given a diffeomorphism  $f \in \text{Diff}^l(M, N)$  and a  $C^l$  vector field  $v$ , the map  $\text{Exp}(f, v) : M \rightarrow N$  defined by  $\text{Exp}(f, v)(x) = \exp_{f(x)} v(f(x))$  – where  $\exp$  denotes the Riemannian geometry exponential associated to a  $C^\infty$  metric – is a  $C^l$  mapping. The implicit function theorem shows that for sufficiently small  $v$ , this mapping is also a diffeomorphism and, moreover, any diffeomorphism in a  $C^l$  neighborhood of  $f$  can be written in this way. If  $l < 1$ , we cannot apply the implicit function theorem to conclude that  $\text{Exp}(f, v)$  is a diffeomorphism; nevertheless, it is possible to show that  $\text{Exp}$  covers a neighborhood of  $f$  in the space of  $C^l$  mappings. (A similar argument shows that  $\text{Diff}^\infty(M)$  is a Fréchet manifold.)

In  $\text{Diff}^l(M)$  it is natural to define the group operation of composition. This set is a topological group for integer exponents; this is the reason we formulate the statements in this paper only for this case. Unfortunately, when  $l$  is not an integer, this operation will be discontinuous. Moreover, when  $l$  is an integer, even if composition is continuous, it will not be differentiable. For properties of the composition operator on Hölder spaces we refer to [7], where there is a very systematic treatment of these properties.

We will also consider families of diffeomorphisms depending of a parameter. We will denote the parameters by subindices. We will say that  $f_\lambda$  is a  $C^l$  family of diffeomorphisms when for every fixed  $\lambda \in \Lambda$ , where  $\Lambda \subset \mathbb{R}^m$  is an open set or a compact manifold, the map  $x \rightarrow f_\lambda(x)$  is a  $C^l$  diffeomorphism and, moreover,  $(\lambda, x) \rightarrow f_\lambda(x)$  is also  $C^l$ .

The composition of diffeomorphisms can be extended in a natural way to families. We denote by  $g_\lambda \circ f_\lambda$  the family  $(\lambda, x) \rightarrow g(\lambda, f_\lambda(x))$ . The notation  $f_\lambda$  is somewhat ambiguous, since it denotes at the same time the family and the mapping for a concrete value of  $\lambda$ . The notation is, however, standard, and does not lead to confusion.

*Remark.* It is also possible to think of families as mappings from  $\Lambda \times M$  to  $\Lambda \times N$  such that the first variable remains unaltered. This point of view makes it clear that many results about composition proved for  $C^l$  mappings are true for  $C^l$  families.

Notice that the map  $\Lambda \rightarrow C^l(M, N)$  defined by  $\lambda \rightarrow f_\lambda$  is only  $C^0$ . There are many  $C^0$  mappings  $\Lambda \rightarrow C^l(M, N)$  which are not  $C^l$  families.

From now on, the manifold under consideration will be the  $n$ -dimensional sphere  $\mathbf{S}^n$ . We take its standard embedding in Euclidean space:

$$\mathbf{S}^n = \left\{ x = (x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \mid \sum_{i=1}^{n+1} x_i^2 = 1 \right\}.$$

If  $x$  is a point in  $\mathbf{S}^n$ , we will denote by  $\Pi_x$  the hyperplane passing through the origin and consisting of vectors orthogonal to  $x$ , and by  $\mathcal{M}_{x,y}$  the one-dimensional maximum circle going from  $x$  to  $-x$  and passing through the point  $y \in \mathbf{S}^n - \{x, -x\}$ . We will refer to those sets as *meridians*. We will also represent by  $\mathcal{P}_{x,\gamma}$ ,  $-1 \leq \gamma \leq 1$ , the subsets of  $\mathbf{S}^n$  whose points have a constant projection  $\gamma$  along the line joining the origin and the point  $x$ . The sets of these form will be referred to as *parallels*. Finally, we denote by  $\delta_x$  the reflection across the plane  $\Pi_x$ .

We call attention to the fact that, with our notation, meridians are always one-dimensional, but parallels in  $\mathbf{S}^n$  are copies of  $\mathbf{S}^{n-1}$ . Outside a neighborhood of the poles  $\{x, -x\}$ , meridians  $\mathcal{M}_{x,y}$  and parallels  $\mathcal{P}_{x,\gamma}$  constitute two transversal foliations.

We will denote by  $\text{Diff}_0^l(\mathbf{S}^n)$  the connected component of the identity of  $\text{Diff}^l(\mathbf{S}^n)$  and by  $\text{Sym}_x^l$  the subset of  $\text{Diff}^l(\mathbf{S}^n)$  commuting with  $\delta_x$ . The symbol  $\text{Sym}_{x,0}^l$  will stand for the connected component of  $\text{Sym}_x^l$  containing the identity.

Next we study the algebraic structure of the group of diffeomorphisms in  $\mathbf{S}^n$ . The main theorem in this paper is

**Theorem 3.1.** *Let  $n \in \mathbb{N}$ ,  $n \geq 2$ . For any  $l \in \mathbb{N} \cup \{\infty\}$ ,  $l \geq 7$ , we can find points  $y_1, \dots, y_{(n+1)^2}$  such that any  $f \in \text{Diff}_0^l(\mathbf{S}^n)$  can be written as  $f = f_1 \circ \dots \circ f_N$ , where each  $f_i \in \text{Sym}_{y_j,0}^{l-2}$  for some  $j = 1, \dots, (n+1)^2$ . (The number  $N$  may depend on  $f$ .)*

For  $n = 1$ , there exist points  $y_1, y_2$  such that any  $f \in \text{Diff}_0^l(\mathbf{S}^1)$  can be written as  $f = f_1 \circ \dots \circ f_N$  with each  $f_i \in \text{Sym}_{y_j,0}^{l-2}$  for some  $j = 1, 2$ . This theorem was proved in [5] for  $l = \infty$ . A different proof, valid in classes of finitely differentiable functions, was given in [6], where an analogue result for  $\mathbf{T}^n$  was also proved.

We also remark that, as a corollary of the proof of Theorem 3.1, we will obtain that the diffeomorphisms appearing as factors can be identified with families of diffeomorphisms of the circle.

**3.2. Proof of Theorem 3.1.** The idea of the proof can be understood visually in the case of  $\mathbf{S}^2$  imbedded in  $\mathbb{R}^3$ , and we will describe it informally in the next paragraphs. A more formal description in any number of dimensions will follow.

Since the two dimensional sphere is a familiar object, we will use geographical terms to describe it.

First, we note that, since all the diffeomorphisms in the connected component of the identity can be expressed as a finite product of diffeomorphisms in a neighborhood, it suffices to prove the result in any neighborhood of the identity. We will assume that it is sufficiently small that we can apply the implicit function theorem at certain places of the proof that we will detail below.

Second, we observe that the diffeomorphisms in such neighborhood can be expressed as the composition of a diffeomorphism that agrees with the identity in a neighborhood of the north-south poles and another one that agrees with the identity in neighborhoods around antipodal points of the equator (this is called a fragmentation lemma). Since these two situations are equivalent under a rotation, it suffices to consider the situation where a neighborhood of the poles is kept fixed.

Third, we note that, by the implicit function theorem, a diffeomorphism like the above can be decomposed into a diffeomorphism that preserves the meridians and another one that preserves the parallels.

We will show that, indeed, the diffeomorphism that moves along the meridians can be factored in the desired way. (This will require the hard implicit function theorem.)

For the diffeomorphism that preserves the parallels and agrees with the identity in neighborhoods of the poles, we note that we can think of it as a family of diffeomorphisms of the circle parametrized by the latitude, and that, provided we can prove a decomposition theorem for diffeomorphisms of the circle with fixed planes and with smooth dependence on parameters, the factors can again be thought of as diffeomorphisms of the sphere. This result of factorization with dependence on parameters of families of maps of the circle will follow from our results, but was also stated and proved in [6].

For higher dimensions the argument is very similar. The fragmentation will require more pairs of antipodal points, and the step described above reduces the problem to a problem of families of diffeomorphisms on the spheres of one dimension less. (The parallels are spheres of one dimension less than the original sphere.) The case of one dimension is already done. Lemmas 3.2, 3.3 and 3.4 will implement the fragmentation, decomposition of diffeomorphisms into diffeomorphisms preserving parallels and meridians and the decomposition of those preserving meridians. The latter one, Lemma 3.4, will require the use of hard implicit function theorems.

Now, we turn to a precise implementation of the above sketch.

As noted, it suffices to prove the result in an arbitrary small neighborhood of the identity in the  $C^l$ -topology. In effect, given any open neighborhood  $\mathcal{U}$  of the identity, any diffeomorphism in the connected component of the identity can be written as the composition of a finite number of diffeomorphisms in  $\mathcal{U}$ .

We will establish the local theorem for families of diffeomorphisms by induction on the dimension. We introduce in the inductive argument that the map  $\mathcal{F}$ , constructed in the proof, which assigns to each family of diffeomorphisms  $f_\lambda \in \mathcal{U}$  the factors  $f_{\lambda,i}$ ,  $i = 1, \dots, N$ , in a neighborhood of the identity in the  $C^{l-2}$  topology, is continuous, and that for each value of the parameter  $\lambda \in \Lambda$  with  $f_\lambda = \text{Id}$ , one obtains factors  $f_{\lambda,i} = \text{Id}$  for every  $i = 1, \dots, N$ . This result for the one dimensional case is stated explicitly in [6].

In the statement of the theorem the points  $y_1, \dots, y_{(n+1)^2}$  are left to our choice. We fix a real number  $\rho$  and take those  $y = (y_k)$  of  $\mathbf{S}^n$  whose components are defined by one of the following conditions:

- there is  $1 \leq i \leq n+1$  such that  $y_i = 1$  and  $y_k = 0$  for  $k \neq i$ , or
- there are  $1 \leq i, j \leq n+1$  such that  $y_i = \cos \rho$ ,  $y_j = \sin \rho$  and  $y_k = 0$  for  $k \neq i, j$ .

(However, it will be clear that an more careful selection of these points would allow us to reduce their number. We have not bothered with this issue.)

The choices of the angle of separation  $\rho$  that are convenient for the method of proof are those satisfying the so-called *Diophantine* conditions,

$$(2) \quad \left| \rho - \frac{p}{q} \right|^{-1} \leq Cq^{2+\delta}$$

for all pairs of integers  $p, q$  with  $q > 0$ .

It is well known that for  $\delta > 0$ , the numbers satisfying such inequalities have full Lebesgue measure. For  $\delta = 0$ , the set of numbers satisfying these inequalities is called the set of constant type numbers. Moreover, relation (2) with  $\delta = 0$  is equivalent to  $\rho$  having a bounded continued fraction expansion. All irrational numbers that satisfy a quadratic equation with integer coefficients are of constant type; in particular, constant type numbers are dense.

Even if all Diophantine  $\rho$ 's would lead to a theorem of the type we want and, for  $\delta$  small, with the conditions of differentiability required in the statement, constant type numbers will lead to the sharpest differentiability conclusions. From now on we will assume we have chosen a constant type number. The actual choice of a constant type number will affect the sizes of the neighborhoods where the local theorem holds, but not the differentiability conclusions.

In order to apply the induction process, we will decompose diffeomorphisms of  $\mathbf{S}^n$  in a neighborhood of Id into factors with the same order of differentiability that preserve lower dimensional spheres. We will start by proving a parametrized version of the well known *fragmentation lemma*.

**Lemma 3.2.** *Let  $M$  be a compact manifold, and  $\{K_i\}_{i=1}^n$  be a collection of compact subsets of  $M$  such that*

$$\bigcap_{i=1}^n K_i = \emptyset.$$

*Then, for each  $l \in \mathbb{N} \cup \{\infty\}$  we can find a neighborhood  $\mathcal{U}$  of the identity in the space of  $C^l$  families of diffeomorphisms, and a map  $\mathcal{F}_0$  that, to each family  $f_\lambda$  in  $\mathcal{U}$ , associates  $C^l$  families  $f_\lambda^1, \dots, f_\lambda^n$ , satisfying*

- i)  $f_\lambda = f_\lambda^1 \circ \dots \circ f_\lambda^n$ ,
- ii)  $f_\lambda^i|_{K_i} = \text{Id}$ ,
- iii) *if for some  $x \in M$  and  $\lambda \in \Lambda$  one has  $f_\lambda(x) = x$ , then  $f_\lambda^i(x) = x$  for every  $i = 1, \dots, n$ .*

*The mapping  $\mathcal{F}_0$  is continuous when we give its domain and its range the topology of  $C^l$  families.*

*Proof.* We will assume that  $\mathcal{U}$  is small enough so that  $f_\lambda(x)$  is within the injectivity radius of the exponential. We will also need to know that some intermediate steps in our construction lie in this domain. Since these are a finite number of conditions, it will be easy to check that they are satisfied for a non-trivial neighborhood.

Suppose first  $n = 2$ ; then we can find a  $C^\infty$ -function  $\phi$  such that

$$\phi|_{K_1} \equiv 1, \quad \phi|_{K_2} \equiv 0.$$

If  $f_\lambda$  can be represented as the exponential of the vector field  $F_\lambda = \text{Exp}^{-1}(f_\lambda)$ , then  $g_\lambda = \text{Exp}(\phi F_\lambda)$  restricted to  $K_2$  is the identity. Moreover, since  $f_\lambda(x) = x \Leftrightarrow \text{Exp}^{-1}(f_\lambda)(x) = 0$ , the set of fixed points of  $g_\lambda$  is bigger than that of  $f_\lambda$ . Provided that  $f_\lambda$  is sufficiently close to the identity – which depends only on the  $K$ 's and  $\phi$  – we have  $g_\lambda|_{K_1} = f_\lambda$  and  $f_\lambda^2|_{K_2} = \text{Id}$ .

Hence, setting  $f_\lambda^1 = f_\lambda \circ (g_\lambda)^{-1}$ , the claims of the lemma are satisfied. Since the choice of  $\phi$  depends only on the compacts  $K_1$  and  $K_2$ , the continuity follows. Notice that the previous argument works even when one of the compacts is the empty set.

To prove the general case, we will proceed by induction in  $n$ , the number of compact sets, starting from the already verified case  $n = 2$ . We will assume that the conclusion is true for  $n \leq n_0$ ,  $n_0 \geq 2$ , and will check it for  $n = n_0 + 1$ .

If  $K_1, \dots, K_{n_0+1}$  are compact sets with empty intersection, we can find open sets  $V_1, \dots, V_{n_0+1}$  such that  $K_i \subset V_i$  and  $\bigcap_{i=1}^{n_0+1} \overline{V}_i = \emptyset$ .

Set  $L_0 = \bigcap_{i=1}^{n_0} \overline{V}_i$  and  $L_{n_0+1} = K_{n_0+1}$ . Since  $L_0 \cap L_{n_0+1} = \emptyset$ , we apply the previous result with  $n_0 = 2$  to obtain families  $f_\lambda^0, f_\lambda^{n_0+1}$  which are the identity on  $L_0, L_{n_0+1}$  respectively, and with not fewer fixed points than  $f_\lambda$ , such that  $f_\lambda = f_\lambda^0 \circ f_\lambda^{n_0+1}$ .

We can now consider the compact sets  $L_i = K_i - \bigcap_{j=1}^{n_0} V_j$ . They have empty intersection because  $\bigcap_{i=1}^{n_0} K_i \subset \bigcap_{i=1}^{n_0} V_i$ , so we can apply the induction hypothesis to  $f_\lambda^0$  to write  $f_\lambda^0 = f_\lambda^1 \circ \dots \circ f_\lambda^{n_0}$ . Each one of the  $f_\lambda^i$  will be the identity on a neighborhood of  $K_{n_0+1}$ . This gives us the representation  $f_\lambda = f_\lambda^1 \circ \dots \circ f_\lambda^{n_0+1}$  satisfying all the conditions of the statement.

The continuity of  $\mathcal{F}_0$  is also a consequence of the induction process.

We emphasize that, since, properties i)-iii) do not determine  $\mathcal{F}_0$  uniquely, the claim of continuity refers only to the one constructed above.  $\square$

In the next lemma, we show how diffeomorphisms can be factored into simpler maps that conserve the parallels and the meridians, which are in essence diffeomorphisms of  $\mathbf{S}^{n-1}$  and  $\mathbf{S}^1$  respectively. We will henceforth refer to diffeomorphisms and families that preserve the parallels as *horizontal* and those that preserve the meridians as *vertical*.

**Lemma 3.3.** *Let  $x_0$  be a point in  $\mathbf{S}^n$ ,  $-x_0$  the antipodal point, and  $U, U'$  two open neighborhoods around  $x_0, -x_0$  respectively. Then, for every  $l \in \mathbb{N} \cup \{\infty\}$ , we can find a neighborhood  $\mathcal{U}$  of the identity in the space of  $C^l$  families of diffeomorphisms in such a way that, for any family  $f_\lambda$  in  $\mathcal{U}$  with  $f_\lambda|_U = \text{Id}$ ,  $f_\lambda|_{U'} = \text{Id}$ , there exist  $C^l$  families  $f_\lambda^h, f_\lambda^v$  satisfying:*

- i)  $f_\lambda = f_\lambda^h \circ f_\lambda^v$ ,
- ii)  $f_\lambda^v \mathcal{M}_{x_0, x} = \mathcal{M}_{x_0, x}$  for all  $x \in \mathbf{S}^n - \{x_0, -x_0\}$ ,
- iii)  $f_\lambda^h \mathcal{P}_{x, \gamma} = \mathcal{P}_{x, \gamma}$  for all  $\gamma \in [-1, 1]$ ,
- iv) if for some  $x \in \mathbf{S}^n$  and  $\lambda \in \Lambda$  one has  $f_\lambda(x) = x$ , then  $f_\lambda^h(x) = f_\lambda^v(x) = x$ .

*In particular,  $f_\lambda^h|_{U \cup U'} = f_\lambda^v|_{U \cup U'} = \text{Id}$ .*

*Moreover, if we restrict  $f_\lambda^h, f_\lambda^v$  to be in a neighborhood of the identity, they are unique, and the mapping  $\mathcal{F}_1$  that assigns to each  $f_\lambda$  the pair  $f_\lambda^h, f_\lambda^v$  is continuous when we give its domain and its range the topology of  $C^l$  families.*

*Proof.* For any  $x \in \mathbf{S}^n - \{x_0, -x_0\}$ , we represent by  $\gamma(x)$  the projection of  $x$  along the line joining the origin and the point  $x_0$ . Obviously, we should set  $f_\lambda^v(x)$  to be the point in the meridian of  $x$  and on the parallel of  $f_\lambda(x)$ . Then,  $f_\lambda^h$  would move along the parallels so as to get to  $f_\lambda(x)$ .

More formally, using the notation for parallels and meridians introduced in the previous section, we have

$$\begin{aligned} f_\lambda^v(x) &= \mathcal{M}_{x_0, x} \cap \mathcal{P}_{x_0, \gamma(f_\lambda(x))}, \\ f_\lambda^h(x) &= \mathcal{M}_{x_0, f_\lambda((f_\lambda^v)^{-1}(x))} \cap \mathcal{P}_{x_0, \gamma(x)}. \end{aligned}$$

Thanks to the uniform transversality of the meridian and parallel foliations outside a neighborhood of the poles, the usual implicit function theorem establishes that

$f_\lambda^v(x)$  and  $f_\lambda^h(x)$  depend jointly  $C^l$  on  $x$  and on the parameter  $\lambda$ , and consequently, they are  $C^l$  families of diffeomorphisms.  $\square$

The induction on the dimension to prove Theorem 3.1 will proceed as follows. For  $0 < \varepsilon < 1$ , we will consider the sets  $K_i^\varepsilon = \{x \in \mathbf{S}^n \mid |x_i| \leq \varepsilon\}$ . It is obvious that  $\bigcap_{i=1}^n K_i^\varepsilon = \emptyset$  when  $\varepsilon$  is sufficiently small.

Lemma 3.2 will allow us to write any diffeomorphism  $f$  as a composition of diffeomorphisms  $f_i$ ,  $1 \leq i \leq n+1$ , each of which is the identity restricted to  $K_i^\varepsilon$ . From now on we consider the case  $i=1$ ; it is clear that the same arguments could be applied to the rest of diffeomorphisms  $f_i$ ,  $i=2, \dots, n+1$ .

Furthermore, for  $f_1$  we will pick a polar axis going through the points  $x_\pm = (0, \dots, \pm 1)$ , and using Lemma 3.3 we will write  $f_1 = f^h \circ f^v$ , as a composition of a horizontal and a vertical diffeomorphism. If  $f_1$  is sufficiently close to the identity then  $f^h, f^v$  would be the identity in  $K_1^\varepsilon$ . We will apply the induction hypothesis to  $f^h$  and will show how to factor  $f^v$ .

To apply the induction hypothesis, we will think of  $f = f^h$  as a family of maps sending  $\mathbf{S}^{n-1}$  into  $\mathbf{S}^{n-1}$ , the parameter being the latitude of the parallel. That is, the map will be  $f^h(p, \gamma) = (\tilde{f}_\gamma(p), \gamma)$ , where  $\tilde{f}_\gamma$  is a diffeomorphism of  $\mathbf{S}^{n-1}$  for every  $\gamma \in [-1, 1]$ . Moreover, we have  $\tilde{f}_\gamma = \text{Id}$  on  $\mathbf{S}^{n-1}$  when  $\gamma$  is close enough to  $-1$  or  $1$ . Notice how we are led to consider families of diffeomorphisms even if we start with a single one. With hindsight, this is the reason why we have formulated the inductive hypothesis and the statements of the paper in terms of families.

We obtain  $\tilde{f}_\gamma = \tilde{f}_{\gamma,1} \circ \dots \circ \tilde{f}_{\gamma,N'}$  as a composition of  $C^{l-2}$  families of diffeomorphisms of  $\mathbf{S}^{n-1}$ . For each  $i = 1, \dots, N'$  there exists  $j \in \{1, \dots, n^2\}$  (or  $j \in \{1, 2\}$  when  $n=2$ ) such that  $\tilde{f}_{\gamma,i} \circ \delta_{y_j} = \delta_{y_j} \circ \tilde{f}_{\gamma,i}$  on  $\mathbf{S}^{n-1}$  for all the values of  $\gamma$ . Since the induction includes in its conclusions that the factors  $\tilde{f}_{\gamma,i}$  depend smoothly on the latitude and preserve the identity in a neighborhood of the poles, we conclude that these factors are really diffeomorphisms in  $\mathbf{S}^n$  commuting with reflections. Clearly, the above arguments apply without any change to families of horizontal diffeomorphisms.

We now proceed to establish the factorization of  $f^v$  into symmetric diffeomorphisms. In keeping with the above remarks, we also can take the latitude as a periodic variable and consider the parameters in a set  $\Lambda$ . Since, as we have argued before, on the boundaries of the set the diffeomorphisms become the identity, we can easily extend and consider  $\Lambda = \mathbf{T}^{m-2}$ . Note that, as we need to apply the inductive procedure more times, the number of parameters will keep changing. (The reason we denote our parameter space by  $\mathbf{T}^{m-2}$  is that, later, we will combine it with variables in  $\mathbf{T}^2$ ; hence, the combined space will be  $\mathbf{T}^m$ ).

**Lemma 3.4.** *For an  $l \in \mathbb{N} \cup \{\infty\}$ ,  $l \geq 7$ , we can find two points  $y_1, y_2$  in the sphere and a neighborhood  $\mathcal{U}$  of the identity in the space of  $C^l$  families of diffeomorphisms such that, for every family  $f_\lambda^v \in \mathcal{U}$  with*

$$(h.1) \quad f_\lambda^v \mathcal{M}_{x_+,x} = \mathcal{M}_{x_+,x} \quad \forall x \in \mathbf{S}^n - \{x_\pm\},$$

$$(h.2) \quad f_\lambda^v|_{K_1^\varepsilon} = \text{Id},$$

*there exist two  $C^{l-2}$  families  $f_\lambda^{v,1}, f_\lambda^{v,2}$  in a neighborhood of the identity, satisfying*

- i)  $f_\lambda^{v,1} \mathcal{M}_{x_+,x} = f_\lambda^{v,2} \mathcal{M}_{x_+,x} = \mathcal{M}_{x_+,x} \quad \forall x \in \mathbf{S}^n - \{x_\pm\},$
- ii)  $f_\lambda^{v,1} \circ \delta_{y_1} = \delta_{y_1} \circ f_\lambda^{v,1} \text{ and } f_\lambda^{v,2} \circ \delta_{y_2} = \delta_{y_2} \circ f_\lambda^{v,2},$
- iii)  $f_\lambda^v = f_\lambda^{v,1} \circ f_\lambda^{v,2},$



iv) If  $f_\lambda^v = \text{Id}$  for some  $\lambda$ , then  $f_\lambda^{v,1} = f_\lambda^{v,2} = \text{Id}$ .

Besides, the map  $\mathcal{F}_2$  which assigns to  $f_\lambda^v$  the factors  $f_\lambda^{v,1}$  and  $f_\lambda^{v,2}$  is continuous for the  $C^l$ -topology in its domain and the  $C^{l-2}$ -topology in its range.

We remark again that, since the pair  $f_\lambda^{v,1}, f_\lambda^{v,2}$  is not uniquely determined (even when  $y_1, y_2$  are fixed), the conclusions of continuity in the statement are valid only for the map  $\mathcal{F}_2$  constructed below.

*Proof.* We will take polar coordinates in  $\mathbf{S}^n - K_1^\varepsilon$  defined by

$$\begin{aligned} x_{n+1} &= \cos \pi \varphi, \\ x_n &= \sin \pi \varphi \cos \pi \omega_1, \\ x_{n-1} &= \sin \pi \varphi \sin \pi \omega_1 \cos \pi \omega_2, \\ &\vdots \\ x_3 &= \sin \pi \varphi \sin \pi \omega_1 \cdots \cos \pi \omega_{n-2}, \\ x_2 &= \sin \pi \varphi \sin \pi \omega_1 \cdots \sin \pi \omega_{n-2} \cos 2\pi \theta, \\ x_1 &= \sin \pi \varphi \sin \pi \omega_1 \cdots \sin \pi \omega_{n-2} \sin 2\pi \theta, \end{aligned}$$

for  $\varphi, \omega_1, \dots, \omega_{n-2}, \theta \in (0, 1)$ .

Notice that if  $|x_1| \geq \varepsilon$ , then  $|x_{n+1}|^2 \leq 1 - \varepsilon^2$ , so that  $\varphi$  is defined univocally, and, since  $|x_n|^2 + |x_{n+1}|^2 \leq 1 - \varepsilon^2$ ,  $\omega_1$  is outside of the critical values of  $\sin$ ; this same argument is applicable to  $\omega_2 \dots \omega_{n-2}$ . In others words, this polar coordinate system defines a smooth transformation  $\Psi$  which identifies  $\mathbf{S}^n - K_1^\varepsilon$  with an open subset  $V$  of  $\mathbf{T}^n$ . Let  $\delta$  be a small positive real number and  $V_\delta = (\delta, 1 - \delta)^{n-1} \times [0, 1] \subset \mathbf{T}^n$ . Then  $U_\delta = \Psi^{-1}(V_\delta)$  is an open subset of  $\mathbf{S}^n$  which contains  $\mathbf{S}^n - K_\varepsilon$ , where the above coordinate system still works. In particular, the images by the map  $\Psi^{-1}$  of the points of the plane  $\theta = 0$  lie in  $K_1^\varepsilon$ . It is clear that every diffeomorphism  $f$  of  $\mathbf{S}^n$  with  $f|_{K_1^\varepsilon} = \text{Id}$  can be extended by the identity to a diffeomorphism of  $\mathbf{T}^n$ .

We observe that, when we express the family  $f_\lambda^v$  of vertical diffeomorphisms of  $\mathbf{S}^n$  in these coordinates, the value of  $\varphi$  changes but the others coordinates remain unaltered. Therefore, a reflection across the plane  $\Pi_{y_1}$  with  $y_1 = (1, \dots, 0)$  is expressed as a change in sign for the  $\theta$  coordinate while all the others remain fixed. More generally, a reflection across  $\Pi_{y_2}$  with  $y_2 = (\cos \rho, \sin \rho, \dots, 0)$  can be expressed in the coordinates (3.2) by sending  $\theta$  into  $2\rho - \theta$  and leaving all the other variables unchanged.

If we take coordinates  $(\varphi, \omega, \theta)$ , the family  $f_\lambda^v$  will be written as

$$(\varphi, \omega, \theta) \rightarrow (\varphi + v_\lambda(\varphi, \omega, \theta), \omega, \theta),$$

and we will look for  $f_\lambda^{v,1} f_\lambda^{v,2}$  that can be written as

$$(\varphi, \omega, \theta) \rightarrow (\varphi + v_\lambda^i(\varphi, \omega, \theta), \omega, \theta)$$

with

$$\begin{aligned} (3) \quad l v_\lambda^1(\varphi, \omega, \theta) &= v_\lambda^1(\varphi, \omega, -\theta), \\ v_\lambda^2(\varphi, \omega, \theta) &= v_\lambda^2(\varphi, \omega, -\theta + 2\rho). \end{aligned}$$

The equation that  $v_\lambda^1$  and  $v_\lambda^2$  have to solve is, as a straightforward calculation shows,

$$(4) \quad v_\lambda(\varphi, \omega, \theta) = v_\lambda^1(\varphi + v_\lambda^2(\varphi, \omega, \theta), \omega, \theta) + v_\lambda^2(\varphi, \omega, \theta).$$

Notice that there is an open set of values of  $\omega$  for which all the points corresponding to  $(\varphi, \omega, \theta)$  are in  $K_1^\varepsilon$ ; hence  $v_\lambda(\varphi, \omega, \theta) = 0$ . For these values of  $w$ , equation (4) admits the trivial solution  $v_\lambda^1(\varphi, \omega, \theta) = v_\lambda^2(\varphi, \omega, \theta) = 0$ . Once Lemma 3.4 is

established, its assertion iv) would imply that  $f_\lambda^{v,1} = \text{Id}$ ,  $f_\lambda^{v,2} = \text{Id}$  for the points in the sphere satisfying this condition. (We will show a similar behavior in the variable  $\varphi$ ).

Clearly, the solutions of (4) can be extended by the identity to all the points of the sphere in  $K_1^\varepsilon$ , even those for which our coordinate system does not work. Under conditions of proximity to the identity of  $f_\lambda^v$ ,  $f_\lambda^{v,1}$ ,  $f_\lambda^{v,2}$ , these will be solutions of our original problem.

Since, except for these considerations,  $\omega$  does not enter either in (4) or in (3), we can think of  $\omega$  as parameters, and we will henceforth use only the symbol  $\lambda$  to denote all the parameters, both the external ones and the ones that correspond to the latitude.

We will carry out the proof of the proposition considering the family  $(f_\lambda^v)$  as a vertical diffeomorphism in  $\mathbf{T}^m$ , where we will solve the functional equation (4). Finally, we will show that the factors obtained are in fact diffeomorphism of  $\mathbf{S}^n$ .

Proceeding heuristically for the moment, the importance of the Diophantine condition already imposed on  $\rho$  becomes apparent if we consider the *linearized equations* obtained by expanding (4) formally in the unknowns and keeping only the linear terms. We obtain

$$(5) \quad v_\lambda(\varphi, \theta) = v_\lambda^1(\varphi, \theta) + v_\lambda^2(\varphi, \theta).$$

By expanding  $v_\lambda$ ,  $v_\lambda^1$ , and  $v_\lambda^2$  in Fourier coefficients in the variable  $\theta$ ,  $v_\lambda(\varphi, \theta) = \sum_k \hat{v}_{\lambda,k}(\varphi) e^{2\pi i k \theta}$  and  $v_\lambda^j(\varphi, \theta) = \sum_k \hat{v}_{\lambda,k}^j(\varphi) e^{2\pi i k \theta}$  for  $j = 1, 2$ , we see that (5) is equivalent to

$$(6) \quad \hat{v}_{\lambda,k}(\varphi) = \hat{v}_{\lambda,k}^1(\varphi) + \hat{v}_{\lambda,k}^2(\varphi).$$

Similarly, the symmetry conditions for  $v_\lambda^1$ ,  $v_\lambda^2$  can be expressed as

$$(7) \quad \hat{v}_{\lambda,-k}^1(\varphi) = \hat{v}_{\lambda,k}^1(\varphi), \quad \hat{v}_{\lambda,-k}^2(\varphi) e^{4\pi i \rho} = \hat{v}_{\lambda,k}^2(\varphi).$$

When  $k \neq 0$ , we can group the equations for  $k$ ,  $-k$  in (6), and, using (7), we obtain

$$\hat{v}_{\lambda,k}(\varphi) = \hat{v}_{\lambda,k}^1(\varphi) + \hat{v}_{\lambda,k}^2(\varphi), \quad \hat{v}_{\lambda,-k}(\varphi) = \hat{v}_{\lambda,k}^1(\varphi) + \hat{v}_{\lambda,k}^2(\varphi) e^{-4\pi i \rho}.$$

These equations admit the solutions

$$(8) \quad \begin{aligned} \hat{v}_{\lambda,k}^1(\varphi) &= (\hat{v}_{\lambda,k}(\varphi) e^{-4\pi i \rho k} - \hat{v}_{\lambda,-k}(\varphi)) / (e^{-4\pi i \rho k} - 1), \\ \hat{v}_{\lambda,k}^2(\varphi) &= (\hat{v}_{\lambda,-k}(\varphi) - \hat{v}_{\lambda,k}(\varphi)) / (e^{-4\pi i \rho k} - 1). \end{aligned}$$

When  $k = 0$ , there are many solutions. Since  $v_\lambda(\varphi, 0) = 0$  for every  $\lambda, \varphi \in \mathbf{T}^{m-1}$ , we choose

$$(9) \quad \hat{v}_{\lambda,0}^1(\varphi) = - \sum_{k \neq 0} \hat{v}_{\lambda,k}^1(\varphi), \quad \hat{v}_{\lambda,0}^2(\varphi) = - \sum_{k \neq 0} \hat{v}_{\lambda,k}^2(\varphi),$$

which yields  $v^1(\varphi, 0) = v^2(\varphi, 0) = 0$  for every  $\varphi \in \mathbf{T}^m$ .

In addition, if  $\varphi \in \mathbf{T}$  with  $|\sin(\pi\varphi)| < \varepsilon$ , then  $v_\lambda(\varphi, \theta) = 0$  for every  $\theta \in \mathbf{T}$ ; hence their Fourier coefficients are  $\hat{v}_{\lambda,k}(\varphi) = 0$  for every  $k$ . From (8) and (9) we deduce that  $v_\lambda^1(\varphi, \theta) = v_\lambda^2(\varphi, \theta) = 0$  for all  $\theta$ .

Notice that, if  $\rho$  were rational, the denominators in (8) would become zero; that is, there would be no solution for the equation. Even if  $\rho$  is irrational, the denominators would become arbitrarily close to zero. Nevertheless, if  $\rho$  is a number of constant type, we have the bounds

$$(10) \quad |(e^{-2\pi i k \rho} - 1)^{-1}| \leq C' |k|$$

which allow us to control the new Fourier expansions.

If we think of (4) as an equation in spaces of functions, expression (5) would be the derivative – in some sense that will have to be specified – at zero. The calculation shows that the derivative, even if invertible, is not bounded; hence, the usual implicit function theorem in Banach spaces does not apply.

We also emphasize that the above calculation shows only the existence of the inverse of the derivative at zero. If we had started making perturbations around non-zero  $v_\lambda^1$  and  $v_\lambda^2$ , we would have been lead to an equation that cannot be readily analyzed in terms of Fourier components.

The above heuristic discussion suggests that, among the several *hard* implicit function theorems which can cope with unbounded inverses, the one that will be useful for this problem is that of Zehnder [10], which requires existence of an inverse only at one point, but which also requires a *group* structure for the equation. We continue the proof by recalling Zehnder's theorem, which we have taken from [10] in the version we are going to apply.

**Definition 3.5.** We say that  $(X_\alpha)_{\alpha \geq 0}$  is a *family of regular spaces* when  $X_\alpha$  are real Banach spaces with norm  $\|\cdot\|_\alpha$  and, whenever for  $\alpha' \leq \alpha$  we have

$$X_0 \supset X_{\alpha'} \supset X_\alpha \supseteq X_\infty = \bigcap_{\alpha \geq 0} X_\alpha,$$

then

$$\|x\|_{\alpha'} \leq \|x\|_\alpha \quad \forall x \in X_\alpha.$$

Furthermore, it is possible to define on them a  $C^\infty$ -regularization, that is, a family  $(S_t)_{t>0}$  of linear mappings  $S_t : X_0 \rightarrow X_\infty$  and constants  $C(\alpha, \alpha')$  satisfying:

- i)  $\lim_{t \rightarrow \infty} \|S_t x - x\|_0 = 0 \quad \forall x \in X_0,$
- ii)  $\|S_t x\|_\alpha \leq t^{\alpha-\alpha'} C(\alpha, \alpha') \|x\|_{\alpha'} \quad \forall x \in X_{\alpha'}, \quad 0 \leq \alpha' \leq \alpha,$
- iii)  $\|S_t x - x\|_{\alpha'} \leq t^{\alpha'-\alpha} C(\alpha, \alpha') \|x\|_\alpha \quad \forall x \in X_\alpha, \quad 0 \leq \alpha' \leq \alpha.$

Notice that  $X_\infty$  is dense in  $X_0$  but not necessarily in  $X_\alpha$  for  $0 < \alpha$ ; besides, the closure of  $X_\infty$  in  $X_\alpha$  contains  $X_{\alpha'}$  for every  $\alpha < \alpha'$ . We now consider several families of regular spaces and denote by the same symbol their norms, which are probably different. However, this notation is standard and clear.

**Theorem 3.6.** Let  $(X_\alpha)_{\alpha \geq 0}$ ,  $(Y_\alpha)_{\alpha \geq 0}$ ,  $(Z_\alpha)_{\alpha \geq 0}$  be families of regular spaces, and let  $(x_0, y_0) \in X_\infty \times Y_\infty$ .

Take  $F : X_0 \times Y_0 \rightarrow Z_0$  with  $F(x_0, y_0) = 0$  and continuous restriction on the set  $B_0$ . (For any  $\alpha$ , we will write  $B_\alpha = \{(x, y) \in X_\alpha \times Y_\alpha \mid \|x - x_0\|_\alpha < 1/2, \|y - y_0\|_\alpha < 1/2\}$ .)

Assume that the following conditions are satisfied

(h.1)  $F(x, \cdot) : Y_0 \rightarrow Z_0$  is twice differentiable and

$$\|d_2 F(x, y)\| \leq M_0, \quad \|d_2^2 F(x, y)\| \leq M_0 \quad \forall x, y \in B_0.$$

(h.2)  $F$  is uniformly Lipschitz in  $X_0$ . That is, for all  $(x, y) \in B_0$ ,  $(x', y) \in B_0$ ,

$$\|F(x, y) - F(x', y)\|_0 \leq M_0 \|x - x'\|_0.$$

(h.3) The triple  $(F, x_0, y_0)$  is of order  $\infty$ . That is,

$$(h.3.1) \quad F(B_0 \cap X_\alpha \times Y_\alpha) \subset Z_\alpha, \quad 0 \leq \alpha,$$

- (h.3.2) *there exist constants  $M_\alpha$ ,  $1 \leq \alpha$ , such that if  $(x, y) \in X_\alpha \times Y_\alpha \cap B_1$  satisfies  $\|x - x_0\|_\alpha \leq K$ ,  $\|y - y_0\|_\alpha \leq K$ , then  $\|F(x, y)\|_\alpha \leq K M_\alpha$ .*
- (h.4) *There exists an approximate right-inverse of loss  $\sigma$ ,  $1 \leq \sigma$ . That is, for every  $\alpha \geq \sigma$  and  $(x, y) \in B_\sigma \cap (X_\alpha \times Y_\alpha)$  there exist a linear mapping  $\eta(x, y) \in \mathcal{L}(Z_\alpha, Y_{\alpha-\sigma})$  such that:*
- (h.4.1)  $\|\eta(x, y)z\|_0 \leq M_0\|z\|_\sigma$  for all  $z \in Z_\sigma$ ,
- (h.4.2)  $\|d_2F(x, y) \circ \eta(x, y) - 1\|z\|_0 \leq M_0\|F(x, y)\|_\sigma\|z\|_\sigma$  for all  $z \in Z_\sigma$ ,
- (h.4.3) *if  $\|x - x_0\|_\alpha \leq K$ ,  $\|y - y_0\|_\alpha \leq K$ , then  $\|\eta(x, y)(F(x, y))\|_{\alpha-\sigma} \leq K M_\alpha$ .*

Then, if  $\varepsilon$  is a small positive number and  $\lambda = 2\sigma + \varepsilon$ , we can find an open neighborhood  $D_\lambda = \{x \in X_\lambda \mid \|x - x_0\|_\lambda < C_\lambda\}$  and a mapping  $\psi : D_\lambda \rightarrow Y_\sigma$  such that

- i)  $F(x, \psi(x)) = 0$ ,
- ii)  $\|\psi(x) - y_0\|_\sigma \leq C_\lambda^{-1}\|x - x_0\|_\lambda$ ,
- iii) *there exist  $0 < \beta < 1$  and constants  $C_\tau$  for  $\lambda < \tau$ , such that for every  $x \in D_\lambda \cap X_\tau$  one has  $\psi(x) \in Y_{\beta(\tau-1)}$  and  $\|\psi(x) - y_0\|_{\beta(\tau-1)} \leq C_\tau\|x - x_0\|_\tau^\beta$ .*

In particular,  $\psi(D_\lambda \cap X_\infty) \subset Y_\infty$ . Moreover, if  $\eta(x, y)$  depends continuously on  $(x, y)$ , then the mapping  $\psi : D_\lambda \rightarrow Y_\sigma$  is continuous.

We next specify the spaces and functional to which the implicit function theorem will be applied to yield the main result. For each  $\mu \in \mathbf{T}$ , we denote by  $C_\mu^r(\mathbf{T}^m)$  the set of functions  $v$  of  $C^r(\mathbf{T}^m)$  satisfying

$$v_\lambda(\varphi, \theta) = v_\lambda(\varphi, -\theta + 2\mu), \quad \forall (\lambda, \varphi, \theta) \in \mathbf{T}^m.$$

Let us fix  $l \in \mathbb{N}$ ,  $l \geq 7$ , and take

$$X_\alpha = Z_\alpha = \{v \in C^{l-7+\alpha}(\mathbf{T}^m) \mid v_\lambda(\varphi, 0) = 0 \quad \forall (\lambda, \varphi) \in \mathbf{T}^{m-1}\}$$

and

$$Y_\alpha = \{(v^1, v^2) \in C_0^{l-5+\alpha}(\mathbf{T}^m) \times C_\rho^{l-5+\alpha}(\mathbf{T}^m) \mid v_\lambda^i(\varphi, 0) = 0, \forall (\lambda, \varphi) \in \mathbf{T}^{m-1}, i = 1, 2\}$$

for any  $\alpha \geq 0$ , with the corresponding Hölder norms, and define the operators

$$\begin{aligned} F : \quad X_\alpha \times Y_\alpha &\longrightarrow Z_\alpha, \\ (v, v^1, v^2) &\longrightarrow v_\lambda^1(\varphi + v_\lambda^2(\varphi, \theta), \theta) + v_\lambda^2(\varphi, \theta) - v_\lambda(\varphi, \theta). \end{aligned}$$

From Lemma 2.5 of [10] we can immediately show that  $(X_\alpha)_{\alpha \geq 0}$ ,  $(Y_\alpha)_{\alpha \geq 0}$  and  $(Z_\alpha)_{\alpha \geq 0}$  are families of regular spaces. Observe that we only use those diffeomorphisms in  $\mathbf{T}^m$  which are the identity on the plane  $\theta = 0$ , in order to guarantee the uniqueness of the decomposition into symmetric factors.

*Remark.* Consider a  $C^r$  vertical diffeomorphism of  $\mathbf{T}^m$  in a small neighborhood of the identity,  $f_\lambda(\varphi, \theta) = (\varphi + v_\lambda(\varphi, \theta), \theta)$  with  $r \geq 1$ . Then  $f_\lambda^{-1}$  is also a  $C^r$  vertical diffeomorphism  $f_\lambda^{-1}(\varphi, \theta) = (\varphi + u_\lambda(\varphi, \theta), \theta)$ . Besides, taking  $B_r = \{v \in C^r \mid \|v\|_{C^r} < 1/2\}$ , the map  $B_r \subset C^r(\mathbf{T}^m) \rightarrow C^{r'}(\mathbf{T}^m)$ ,  $v \rightarrow u$  is continuous for every  $r' < r$ . In the case that  $v \in C_\mu^r(\mathbf{T}^m)$  and  $v_\lambda(\varphi, 0) = 0$  for every  $(\lambda, \varphi) \in \mathbf{T}^{m-1}$ , it is obvious that  $u \in C_\mu^r(\mathbf{T}^m)$  and  $u_\lambda(\varphi, 0) = 0$  for every  $(\lambda, \varphi) \in \mathbf{T}^{m-1}$ .

We will show that the above  $F$  satisfy all the conditions of the previous statement around  $(0, 0, 0)$ . Hence, we will obtain the result by applying Zehnder's theorem.

Note that the application of Zehnder's theorem also provides the proof of Lemma 3.4 for  $C^\infty$  families of diffeomorphisms in the sphere.

The conditions (h.1) (h.2) (h.3) are differentiability properties of the composition. They follow from Theorems 4.3 and 6.11 in [7]. The most crucial hypothesis is that about the existence of an approximate right inverse of the linearized operator, which we now discuss.

We will start the construction of the exact inverse at  $(0, 0, 0)$ . To study its differentiability properties, we will use the following result of Rüssmann [9].

**Lemma 3.7.** *Let  $\rho$  be a constant type number and*

$$v(z) = \sum_{k \in \mathbb{Z}} \hat{v}_k e^{ikz}$$

*an analytic complex function in the strip  $U_\tau = \{z \in \mathbb{C} \mid |\operatorname{Im} z| < \tau\}$  with  $|v(z)| \leq M \forall z \in U_\tau$ . The function with mean value zero*

$$h(z) = \sum_{k \in \mathbb{Z} - \{0\}} \frac{\hat{v}_k}{e^{2\pi i k \rho} - 1} e^{ikz}$$

*is analytic in  $U_{\tau'}$  for every  $\tau' < \tau$ , and there is a constant  $\kappa > 0$ , depending only on  $\rho$ , such that*

$$(11) \quad |h(z)| \leq \frac{\kappa M}{\tau - \tau'} \quad \text{for } |\operatorname{Im} z| < \tau'.$$

We take a periodic real analytic function  $v_\lambda(\varphi, \theta)$  defined in the strip  $U_\tau^m$  with  $v_\lambda(\varphi, 0) = 0$  for every  $(\lambda, \varphi) \in \mathbf{T}^{m-1}$ , and consider the pair  $h_\lambda^1, h_\lambda^2$  whose Fourier coefficients in the variable  $\theta$ ,  $(\hat{h}_{\lambda, k}(\varphi))_{k \in \mathbb{Z}}$ , are given by the relations (8) when  $k \neq 0$  and by  $\hat{h}_{\lambda, 0}(\varphi) = 0$  for every  $\lambda, \varphi$ . Define

$$\begin{aligned} v_\lambda^1(\varphi, \theta) &= h_\lambda^1(\varphi, \theta) - h_\lambda^1(\varphi, 0), \\ v_\lambda^2(\varphi, \theta) &= h_\lambda^2(\varphi, \theta) - h_\lambda^2(\varphi, 0). \end{aligned}$$

Observe that the Fourier coefficients in the variable  $\theta$  of  $v_\lambda^1, v_\lambda^2$  are given by the equations (8) and (9). The functions  $h_\lambda^1, h_\lambda^2, v_\lambda^1, v_\lambda^2$  are analytic in the strip  $U_{\tau'}^m$  for every  $\tau' < \tau$ , and satisfy a inequality similar to (11) with maybe some other constants. Notice that if  $\rho$  is a constant type number, so is  $-2\rho$ , that

$$e^{-4\pi i \rho k} / (e^{-4\pi i \rho k} - 1) = 1 + 1 / (e^{-4\pi i \rho k} - 1),$$

and that

$$\sup_{\Im z \leq \tau} \left| \sum_k \hat{v}_k e^{2\pi i k z} \right| = \sup_{\Im z \leq \tau} \left| \sum_k \hat{v}_{-k} e^{2\pi i k z} \right|.$$

Also,  $v_\lambda(\varphi, \theta) = v_\lambda^1(\varphi, \theta) + v_\lambda^2(\varphi, \theta)$ .

From Lemma 3.7 it is also possible to deduce regularity properties of the above decomposition in the classes of finitely differentiable functions. The key is that the differentiability properties of functions can be read of from the size of analytic approximations on thin strips – see [8] –. The natural spaces on which to consider such characterizations are the  $\Lambda_r$  spaces endowed with the Zygmund norms. We just remark that  $\Lambda_r = C^r$  when  $r \notin \mathbb{N}$ , whereas for  $r \in \mathbb{N}$  we have  $C^r \subset \Lambda_r$ , and the inclusion is strict. Notice, however, that the definition of  $C^r$  functions makes sense on Banach spaces (for  $0 \leq r \leq 1^-$  it makes sense in metric spaces), whereas the characterization by analytic approximations is useful only for functions on  $\mathbb{R}^n$ .

A detailed argument along these lines can be found in [8] p. 528, [10], [4]. This allows us to state the following:

**Lemma 3.8.** *Let  $v \in C^r(\mathbf{T}^m)$ ,  $r \notin \mathbb{N}$ ,  $r > 1$ , with  $v_\lambda(\varphi, 0) = 0$  for every  $(\lambda, \varphi) \in \mathbf{T}^{m-1}$ . There are functions  $v^1 \in C_0^{r-1}(\mathbf{T}^m)$ ,  $v^2 \in C_\rho^{r-1}(\mathbf{T}^m)$  and a constant  $\kappa$  depending only on  $\rho$  and  $r$  such that*

- i)  $v = v^1 + v^2$ ,
- ii)  $v_\lambda^1(\varphi, 0) = v_\lambda^2(\varphi, 0) = 0$  for all  $(\lambda, \varphi) \in \mathbf{T}^{m-1}$ ,
- iii)  $\|v^1\|_{C^{r-1}}, \|v^2\|_{C^{r-1}} \leq \kappa \|v\|_{C^r}$ .

Notice that  $F : X_\alpha \times Y_\alpha \rightarrow Z_\alpha$  is differentiable with respect to  $Y$ 's components, and  $d_2F(0, 0, 0)(v^1, v^2) = v^1 + v^2$  in all the cases. Consequently, the conclusion of Lemma 3.8 provides a direct estimation of the exact inverse at  $(0, 0, 0)$  of  $d_2F$ . We obtain

**Lemma 3.9.** *Let  $\epsilon > 0$  be a small parameter and  $\sigma = 3 + \epsilon$ . For every  $\alpha \geq \sigma$  and  $v \in Z_\alpha$ , there is a unique element  $(v^1, v^2) \in Y_{\alpha-\sigma}$  such that  $v = v^1 + v^2$ . Moreover, the map*

$$\begin{array}{ccc} \eta : & Z_\alpha & \longrightarrow Y_{\alpha-\sigma}, \\ & v & \longrightarrow (v^1, v^2) \end{array}$$

*is continuous.*

(More precisely, the map  $\eta : Z_{\alpha-\epsilon/2} \rightarrow Y_{\alpha-\sigma}$  is also continuous). Now an approximate right inverse in a complete neighborhood of  $(0, 0, 0)$  is required. An important observation is that one exact right inverse can be obtained thanks to the group action. See Hamilton [2], pg. 198, and Zehnder [10], pg. 133.

Let us consider  $F_2 : Y_\alpha \rightarrow Z_\alpha$ ,  $(v^1, v^2) \rightarrow F(0, v^1, v^2)$ . It is obvious that  $d_2F(v', v^1, v^2) = dF_2(v^1, v^2)$ . We write  $B_\alpha^2 = \{(v^1, v^2) \in Y_\alpha \mid \|v^1\|_\alpha < 1/2, \|v^2\|_\alpha < 1/2\}$ .

**Lemma 3.10.** *For each  $\alpha \geq \sigma$ , there exists a continuous map  $\eta : B_\sigma^2 \cap Y_\alpha \rightarrow \mathcal{L}(Z_\alpha, Y_{\alpha-\sigma})$  such that*

$$dF_2(v^1, v^2) \circ \eta(v^1, v^2)(v) = v$$

*for every  $(v^1, v^2) \in B_\sigma^2 \cap Y_\alpha$  and  $v \in Z_\alpha$ .*

*Proof.* First we solve this question in the spaces of infinitely differentiable functions. The sets  $X_\infty, Y_\infty, Z_\infty$  can be understood as tangent spaces at the identity of groups of  $C^\infty$  diffeomorphisms in  $\mathbf{T}^m$ .

Let us fix  $(v^1, v^2) \in B_0^2 \cap Y_\infty$  (that is,  $\|v^i\|_{C^{l-5}} < 1/2, i = 1, 2$ ), with which we associate the following differentiable functions:

$$\begin{array}{ccc} \Phi_{(v^1, v^2)} : & Z_\infty & \longrightarrow Z_\infty, \\ & w & \longrightarrow v_\lambda^1(\varphi + w_\lambda(\varphi + v_\lambda^2(\varphi, \theta), \theta) + v_\lambda^2(\varphi, \theta), \theta) \\ & & \quad + w_\lambda(\varphi + v_\lambda^2(\varphi, \theta), \theta) + v_\lambda^2(\varphi, \theta), \end{array}$$

and

$$\begin{array}{ccc} \mathcal{C}_{(v^1, v^2)} : & Y_\infty & \longrightarrow Y_\infty, \\ & (\omega^1, \omega^2) & \longrightarrow (v_\lambda^1(\varphi + \omega_\lambda^1(\varphi, \theta), \theta) + \omega_\lambda^1(\varphi, \theta), \\ & & \quad \omega_\lambda^2(\varphi + v_\lambda^2(\varphi, \theta), \theta) + v_\lambda^2(\varphi, \theta)). \end{array}$$

Take the vertical diffeomorphisms  $f_\lambda^i(\varphi, \theta) = (\varphi + v_\lambda^i(\varphi, \theta), \theta)$  with inverses

$$(f_\lambda^i)^{-1}(\varphi, \theta) = (\varphi + u_\lambda^i(\varphi, \theta), \theta), \quad i = 1, 2.$$

Notice that  $F_2(v^1, v^2) = \Phi_{(v^1, v^2)}(0)$ ; also, we find the following relations :

$$(12) \quad \Phi_{(v^1, v^2)} \circ \Phi_{((u^1), (u^2))} = \text{Id},$$

$$(13) \quad \Phi_{(v^1, v^2)} \circ F_2 = F_2 \circ \mathcal{C}_{(v^1, v^2)}.$$

By differentiating (13) we obtain

$$d\Phi_{(v^1, v^2)}(0) \circ dF_2(0, 0) = dF_2(v^1, v^2) \circ d\mathcal{C}_{(v^1, v^2)}(0, 0);$$

therefore, from (12) and Lemma 3.9, we can deduce that

$$\eta(v^1, v^2) = d\mathcal{C}_{(v^1, v^2)}(0, 0) \circ \eta \circ d\Phi_{(u^1, u^2)}(F_2(v^1, v^2))$$

is a linear continuous map from  $Z_\infty$  to  $Y_\infty$  with

$$(14) \quad dF_2(v^1, v^2) \circ \eta(v^1, v^2)(v) = v$$

for every  $v \in Z_\infty$ .

We now focus our attention on the classes of finitely differentiable functions where we assume  $\Phi_{(v^1, v^2)}$  and  $\mathcal{C}_{(v^1, v^2)}$  are defined. It follows from Propositions 6.1, 6.2 and Theorem 6.10 in [7] that if  $(v^1, v^2) \in B_\sigma^2 \cap Y_\alpha$  then  $\eta(v^1, v^2) \in \mathcal{L}(Z_\alpha, Y_{\alpha-\sigma})$ , and, besides,  $\eta : B_\sigma^2 \cap Y_\alpha \rightarrow \mathcal{L}(Z_\alpha, Y_{\alpha-\sigma})$ ,  $(v^1, v^2) \mapsto \eta(v^1, v^2)$  is continuous.

Actually, a more precise analysis of the continuity of  $\eta$  allows us to state that, if  $(v^1, v^2) \in B_0^2 \cap Y_{\alpha-\epsilon/2}$ , then  $\eta(v^1, v^2) \in \mathcal{L}(Z_{\alpha-\epsilon/2}, Y_{\alpha-\sigma})$  and the map  $\eta : B_0^2 \cap Y_{\alpha-\epsilon/2} \rightarrow \mathcal{L}(Z_{\alpha-\epsilon/2}, Y_{\alpha-\sigma})$  still remains continuous.

Taking into account that the closures of  $B_0^2 \cap Y_\infty$  and  $Z_\infty$  in the natural topologies of  $Y_{\alpha-\epsilon/2}$  and  $Z_{\alpha-\epsilon/2}$  contain  $B_\sigma^2 \cap Y_\alpha$  and  $Z_\alpha$  respectively, we deduce from (14) that

$$dF_2(v^1, v^2) \circ \eta(v^1, v^2)(v) = v$$

for every  $(v^1, v^2) \in B_\sigma^2 \cap Y_\alpha$  and  $v \in Z_\alpha$ . This proves that  $\eta$  is an exact right-inverse of loss  $\sigma$  of  $d_2F$ .  $\square$

We have verified all the hypotheses of Theorem 3.6. Its conclusions provide the factorization of every  $C^l$  vertical diffeomorphism of  $\mathbf{T}^m$  in a neighborhood of  $\text{Id}$  into  $C^{l-2}$  factors commuting with reflections across the planes  $\theta = 0$  and  $\theta = \rho$  respectively. Under these conditions we can return to our original problem in the sphere.

Let us assume that one of the above diffeomorphisms on the torus comes from  $f_\lambda^v$ , a  $C^l$  family of vertical diffeomorphisms of  $\mathbf{S}^n$  with  $f_\lambda^v|_{K_\epsilon^1} = \text{Id}$ , and consider its decomposition  $f_\lambda^v = f_\lambda^{v,1} \circ f_\lambda^{v,2}$  according to the thesis of Theorem 3.6. We denote, as usual,  $f_\lambda^v(\varphi, \theta) = (\varphi + v_\lambda(\varphi, \theta), \theta)$ , the factors  $f_\lambda^{v,i}(\varphi, \theta) = (\varphi + v_\lambda^i(\varphi, \theta), \theta)$  and their inverses  $(f_\lambda^{v,i})^{-1}(\varphi, \theta) = (\varphi + u_\lambda^i(\varphi, \theta), \theta)$ ,  $i = 1, 2$ . Let us fix  $\lambda \in \mathbf{T}^{m-2}$  and take  $\varphi \in \mathbf{T}$  with  $|\sin(\pi\varphi)| \leq \epsilon$ . Then  $v_\lambda(\varphi, \theta) = 0$ , which implies that  $(f_\lambda^{v,1})^{-1}(\varphi, \theta) = (f_\lambda^{v,2})(\varphi, \theta)$  and hence  $u_\lambda^1(\varphi, \theta) = v_\lambda^2(\varphi, \theta)$  for every  $\theta \in \mathbf{T}$ . An analysis of this equality, based on the Fourier coefficients, given in (7) and (9), shows that  $u_\lambda^1(\varphi, \theta) = v_\lambda^2(\varphi, \theta) = 0$  and hence  $f_\lambda^{v,1}(\varphi, \theta) = f_\lambda^{v,2}(\varphi, \theta) = (\varphi, \theta)$  for every  $\theta \in \mathbf{T}$ . This same argument proves that if for some  $\lambda$  one has  $f_\lambda^v = \text{Id}$ , then also  $f_\lambda^{v,1} = f_\lambda^{v,2} = \text{Id}$ .

Thus, it is clear that  $f_\lambda^{v,1}, f_\lambda^{v,2}$  can be extended by the identity on the open subset of  $\mathbf{S}^n$  where our coordinate system does not work, and defines  $C^{l-2}$  families of vertical diffeomorphisms of  $\mathbf{S}^n$  satisfying all the conditions (i)–(iv) of Lemma 3.4. This finishes the induction process and completes the proof of Theorem 3.1.  $\square$

## REFERENCES

- [1] F. Bien, *Global representations of the diffeomorphism group of the circle*, Infinite dimensional Lie algebras and groups, Marseille 1988, World Scientific, 1989, pp. 89–107. MR **90j**:22020
- [2] R. Hamilton, *The inverse function theorem of Nash and Moser*, Bull. A.M.S., **7** (1982), 65–222. MR **83j**:58014
- [3] M. Hirsch, *Differential Topology*, Springer-Verlag, 1976. MR **56**:6669
- [4] S. Krantz, *Lipschitz spaces, smoothness of functions and approximation theory*, Exposition. Math. **3** (1983), 193–260. MR **86g**:41001
- [5] J. Langer, D.A. Singer, *Diffeomorphisms of the circle and geodesic fields on Riemann surfaces of genus one*, Invent. Mat. **69** (1982), 229–242. MR **84h**:58115
- [6] R. de la Llave, *Remarks on J. Langer and D.A. Singer decomposition theorem for diffeomorphisms of the circle*, Comm. Math. Phys., **104** (1986), 387–401. MR **87h**:58174
- [7] R. de la Llave, R. Obaya, *Regularity of the composition operator in spaces of Hölder functions*, Discrete Contin. Dynamical Systems, **5** (1999), 157–184. Preprint available from [http://www.ma.utexas.edu/mp\\_arc](http://www.ma.utexas.edu/mp_arc)
- [8] J. Moser, *A rapidly convergent iteration method and non-linear differential equations II*, Ann. Scuola Norm. Sup. Pisa, **20** (1966), 499–535. MR **34**:6280
- [9] H. Rüssmann, *On optimal estimates for the solutions of linear difference equations on the circle*, Celestial Mech., **14** (1976), 33–37. MR **56**:4306
- [10] E. Zehnder, *Generalized implicit function theorems with applications to some small divisor problems I*, Comm. Pure and Appl. Math., **28** (1975), 91–140. MR **52**:1764
- [11] E. Zehnder, *Generalized implicit function theorems with applications to some small divisor problems II*, Comm. Pure and Appl. Math., **29** (1976), 49–111. MR **54**:14001

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TEXAS AT AUSTIN, AUSTIN, TEXAS 78712  
*E-mail address*: `llave@math.utexas.edu`

DEPARTAMENTO MATEMÁTICA APLICADA A LA INGENIERÍA, ESCUELA SUPERIOR DE INGENIEROS INDUSTRIALES, UNIVERSIDAD DE VALLADOLID, 47011 VALLADOLID, SPAIN  
*E-mail address*: `rafoba@wmatem.eis.uva.es`