Decomposition theorems for some groups of diffeomorphisms and regularity of the composition operator. *

R. de la Llave †‡
Dept. Math.
U. Texas at Austin
Austin TX 78712

R. Obaya
Dept. Matemática aplicada a la Ingeniería
Escuela Superior de Ingenieros Industriales
Univ. de Valladolid
47011 Valladolid
SPAIN

Abstract

We study the algebraic structure of several diffeomorphism groups in $\mathbf{S}^n$. We show that all diffeomorphisms can be written as a finite product of diffeomorphisms which have certain symmetries. Since the natural group operation among diffeomorphisms is the composition and Hölder regularities appear naturally in intermediate steps, we present an appendix in which optimal results about regularity of the composition operator are obtained when the spaces of diffeomorphisms are given the topology of the classical Hölder norms.

---

*This preprint is avaliable from the Math Physics preprints archive. Send e-mail to `mp_arcmath.utexas.edu` for details.

†Supported by N.S.F. grants

‡e-mail address: `llavemath.utexas.edu`

1 Introduction

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

These theorems, roughly, state that any diffeomorphism in the group can be written as a finite product of diffeomorphisms lying in a certain subgroup enjoying certain properties such as symmetry.

There are several motivations for the study of this theorems. For example, in [6] it is shown that one theorem of this type can be used to solve the inverse problem for scattering of geodesic fields in surfaces of genus one. It can also be argued that these theorems are analogues of the usual factorization theorems in Lie algebras and that stronger versions could be useful in the problem of computing representations of diffeomorphism groups [Bi].

If we consider these theorems as infinite dimensional versions of factorization theorems for Lie groups, one first difficulty is that for diffeomorphism groups, the 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"* using the Lie algebra. Nevertheles, we will show that it is possible to obtain similar results to those in the finite dimensional case by using appropriate *"hard"* implicit function theorems of the Nash–Moser type. This seems to be a natural generalization of the systematic methods of finite dimensional Lie group theory to groups of diffeomorphisms and we hope that they have wider applicabilty. e.g. when the groups of diffeomorphisms are required to preserve a certain geometric structure.

In a first chapter, we will consider decomposition theorems of the Langer – Singer type for diffeomorphisms on $\mathbf{S}^n$. Roughly, we show that 2^{n+1} subgroups of diffeomorphisms consisting of diffeomorphisms which commute with reflection generate the whole group of diffeomorphisms. That is, any diffeomorphism can be written as the composition of a finite number of factors, each of which commutes with reflactions across planes of codimension one. Moreover, the factors are essentially one dimensional. If we assume that the factor is differentiable sufficiently often, we can conclude that the factors are also severl times differentiable. the factors are

The method we use can obtain the previously known theorems for diffeomorphisms in $\mathbf{T}^n$ and obtain results for diffeomorphisms of $\mathbf{S}^n$. From the point of view of the theory of representations, it would be useful to have global analogues of the local decomposition theorems [Bi]. We produce some expamples that show that such global results would be difficult to obtain.

2 Acknowledgements

The work of R.L. has been supported by NSF. grants. Also R.L wants to ackwoledge the hospitality enjoyed at Valladolid during several visits. Letters from J. Langer and F. Bien discussing the implications of these theorems were very useful for us.

3 Decomposition theorem for spheres.

3.1 Notation and statement of results

We will consider the standard embedding of the sphere in Euclidean space $\mathbf{S}^{n-1} = \left\{\mathbf{x} \equiv (x_1, \cdots, x_n) \in \mathbb{R}^{n-1} | \sum_{i=1}^{n-1} x_i^2 = 1\right\}$

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

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

If k is not an integer, we define as usual

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

To cover Lipschitz conditions as well, we introduce the notation $k^- = (k-1) + 1^-$.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^k if we can find coordinate patches that cover the manifold on which the expression is C^k.

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

Unfortunately, the C^k 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^k functions are contractions require specific choices of coordinate patches.

For our purposes, this is not a shortcoming since many analytical operations that we will need to perform require taking coordinates anyway. We should however point out that, when k is an integer, there is a geometrically natural notion of C^k distance in Riemannian geometry based on the notion of jets. (See e.g. [4] Ch. 2)

For non-integer k, since the definition of Hölder distance 1 requires comparing the values at different points, there are no geometrically natural notions of C^k distance. Nevertheless, if we select coordinate charts covering the manifold, we can define a distance. In a compact manifold, even if the choice of different charts leads to different metrics, the resulting metrics are equivalent.

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

In $\mathrm{Diff}^k(M)$ it is natural to define the group operation of composition. Unfortunately, when k is not an integer, this operation will be discontinuous and when k is an integer, even when composition is continuous, it will not be differentiable. We refer to appendix A for a sharp results about the differentiability of composition in Hölder spaces.

We will denote by $\mathrm{Diff}^k_0(\mathbf{S}^{n-1})$ the connected component of the identity of $\mathrm{Diff}^k(\mathbf{S}^{n-1})$ and the subset of $\mathrm{Diff}^k(\mathbf{S}^{n-1})$ commuting with δ_x. $\mathrm{Sym}^k_{x,0}$ will be the connected component containing the identity.

We also have to consider families of diffeomorphisms depending of a parameter. We will denote the parameters by subindices. We will say that f_λ is a C^r 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 $x \to f_\lambda(x)$ is a C^r diffeomorphism and, moreover $(\lambda, x) \to f_\lambda(x)$ is also C^r.

The composition of diffeomorphisms can be extended in a natural way to families. We denote by $g_\lambda \circ f_\lambda$ the family $(\lambda, x) \to g(\lambda, f_\lambda(x))$.

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^r mappings are true for C^r families.

Notice also that the notation f_λ is somewhat ambiguous since it denotes at the same time the family and the mapping for a concrete value of λ. The notation is, however, standard use and it does not lead to confusion.

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

The main theorem in the paper will be

Theorem 3.1 *For any $n \in \mathbb{N}$ and any $k \geq 7$, $k \in \mathbb{N}$, we can find points $x_1, \cdots, x_{2^n}$ in such a way that any $f \in Diff^k_0(\mathbf{S}^{n-1})$ can be written $f = f_1 \circ f_2 \circ \cdots \circ f_N$ with each $f_i \in Sym^{k-2}_{x_j}$ for some j.*

For the cases of $n = 1$, $k = \infty$ this theorem was proved in [6]. Also for $n = 1$ and finite differentiability a different proof was given in [7]. The latter proof also provided with an analogue for $\mathbf{T}^n$.

We remark that as a corollary of the proof we will obtain that the diffeomorphisms appearing as factors can be identified with families of diffeomorphisms of the interval.

3.2 Proof of the theorem

We start by observing that it suffices to prove the result in an arbitrary small neighborhood of the identity. In effect, given any open neighborhood U of the identity, it is possible to write any diffeomorphism in the connected component of the identity as the composition of a finite number

of diffeomorphisms in U.

We will establish the local theorem by induction on the dimension. The result for $n = 1$ is stated explicitly in [7]

Our first goal will be to decompose diffeomorphisms of $\mathbf{S}^n$ into simpler diffeomorphisms that preserve lower dimensional spheres, so that we can apply induction.

It will turn out that the induction will proceed in such a way that we prove the theorem for each parallel and we consider the results as a diffeomorphism parametrized by the height.

We will start by proving a parametrized version of the well known *fragmentation lemma.*

Lemma 3.1 *Let M be a compact manifold, $n \in \mathbb{N}$ and $\{K_i\}_{i=1}^{n}$ be a collection of compact subsets of M such that*

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

Then, we can find a neighborhood $\mathcal{U}$ of the identity in the space of C^l and a map $\mathcal{F}_0$ that, to each family f_λ in $\mathcal{U}$, associates C^l families $f_\lambda^1, \cdots, f_\lambda^n$, such that

i) $f_\lambda = f_\lambda^1 \circ \cdots \circ f_\lambda^n \, .$

ii) $f_\lambda^i|_{K_i} = Id \, .$

iii) If for some x, $f_\lambda(x) = x$ for all λ, then $f_\lambda^i(x) = x$ for $i = 1, \ldots, n$ and all values of λ.

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

Remark. Properties i)-iii) do not determine $\mathcal{F}_0$ uniquely. So that the claim of continuity refers only to the one we construct.

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

We will discuss first the case $n = 2$. We can find a C^∞-function φ such that

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

If f_λ can be represented as the exponential of the vector field $F_\lambda = E^{-1}(f_\lambda)$, we see that $f_\lambda^2 = E(F_\lambda \varphi)$ will be the identity restricted to K_2. Moreover, since $f_\lambda(x) = x \iff E^{-1}(f_\lambda)(x) = 0$, the set of fixed points of f_λ^2 is bigger than those of f_λ. Provided that f_λ is sufficiently close to the identity – depending only on the K's and φ – then $f_\lambda^2|_{K_1} = f_\lambda$ and $f_\lambda^2|_{K_2} = Id$.

Hence, setting $f_\lambda^1 = f_\lambda \circ (f_\lambda^2)^{-1}$, the claims of the lemma are satisfied. Notice that since the choice of φ depends only on the compacts K_1 and K_2 the continuity follows.

Remark. Notice that the previous argument works even in the case that one of the compacts is the empty set!

To prove the general case, we will proceed by induction in the number of compacts set n. We will assume that it is true when $n \leq n_0$, $n_0 \geq 2$, and will prove it for $n = n_0 + 1$.

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

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

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

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

Lemma 3.2 *Let x be a point in $\mathbf{S}^n$, $-x$ the antipodal point and U, U' two open neighborhoods around x, $-x$ respectively. Then, for every $r \in \mathbb{N} \cup \{\infty\}$, we can find a neighborhood $\mathcal{U}$ of the identity in the space of C^r families of diffeomorphisms in such a way that for any family f_λ in $\mathcal{U}$, such that $f_\lambda|_U = Id$, $f_\lambda|_{U'} = Id$ we can find C^r families f^h_λ, f^v_λ verifying:*

i) $f_\lambda = f^h_\lambda \circ f^v_\lambda$.

ii) $f^v_\lambda \mathcal{M}_{x,\theta} = \mathcal{M}_{x,\theta}$.

iii) $f^h_\lambda \mathcal{P}_{x,\gamma} = \mathcal{P}_{x,\gamma}$.

iv) If $x \in \mathbf{S}^n$ and $f_\lambda(x) = x$ then $f^h_\lambda(x) = f^v_\lambda(x) = x$. In particular $f^h_\lambda|_{U\cup U'} = f^v_\lambda|_{U\cup U'} = Id$.

Moreover, if we restrict f^h_λ, f^v_λ to be in a neighborhood of the identity, they are unique, and the mapping $\mathcal{F}_1$ that assigns to each f_λ the f^h_λ, f^v_λ is continuous.

We will henceforth refer to diffeomorphisms and families that preserve the parallels as *horizontal* and those that preserve the meridians as *vertical*.

Proof. This lemma says that any family can be factored into a family that preserves the parallels and another one that preserves the meridians. Obviously, we should set $f^v_\lambda(p)$ to be the point in the meridian of p and on the parallel of $f_\lambda(p)$. Then, f^h_λ would move along the parallels so as to get to $f_\lambda(p)$.

More formally, using the notation for parallels and meridians introduced in 1, we have:

$$f^v_\lambda(p) = \mathcal{M}_{x,p} \cap \mathcal{P}_{x,f_\lambda(p)}$$
$$f^h_\lambda(p) = \mathcal{M}_{x,f_\lambda((f^v_\lambda)^{-1}(p))} \cap \mathcal{P}_{x,p}$$

Using the uniform transversality of the meridian and parallel foliations outside a neighborhood of the poles, the usual implicit function theorem, establishes that $f^v_\lambda(p)$ $f^h_\lambda(p)$ depend jointly C^r on p and on the parameters λ.

Remark. If $f_\lambda(\varphi,\theta) = (\varphi + v_\lambda(\varphi,\theta),\theta)$ denotes a vertical diffeomorphism in a small neighborhood of the identity with $v_\lambda \in C^r(\mathbf{T}^m)$, $r \geq 1$, then f_λ^{-1} is also a vertical diffeomorphism $f_\lambda^{-1}(\varphi,\theta) = (\varphi + \mathbf{v}_\lambda(\varphi,\theta),\theta)$ and $\mathbf{v}_\lambda \in C^r(\mathbf{T}^m)$

The induction in the dimension to prove the main theorem 3.1 will proceed as follows. For ε sufficiently small, we will consider the sets $K_i^\varepsilon = \{x \in \mathbf{S}^n \,/\, |x_i| \leq \varepsilon\}$.

Using Lemma 3.1 we will be able to write any diffeomorphism f as composition of diffeomorphisms f_i, each of which is the identity restricted to K_i^ε.

Furthermore, for f_i we will pick a polar axis going through the points $x_i = \pm 1$, $x_j = 0$, $j \neq i$ and using Lemma 3.2 will write it $f_i = f_i^h \circ f_i^v$.

If f_i was sufficiently close to the identity then f_i^h, f_i^v would be the identity in K_i^ε.

We will show how to factor f_i^v and will apply the induction hypothesis to f_i^h.

To apply the induction hypothesis, we will think of f_i^h as a family of maps sending $\mathbf{S}^{n-1}$ into $\mathbf{S}^{n-1}$, the parameter being the height of the parallel.

Notice how we are lead to consider families of diffeomorphisms even if we start with a single one. With hindsight, this is the reason why we formulated Lemma 3.1 and Lemma 3.2 in terms of families.

Clearly, if the induction result obtains in its conclusions that the factors are a smooth family, this will imply that the factors are really diffeomorphisms in $\mathbf{S}^n$.

We now proceed to establish the factorization of f_i^v into symmetric diffeomorphisms.

To simplify the notation and without any loss of generality, we will assume that $i = 1$ and that the axis is going through the points $x_\pm = (0,\ldots,\pm 1)$. We also can think about the height as a periodic variable and consider the parameters in a set $\Lambda = \mathbf{T}^{m-2}$.

Lemma 3.3 *We can find two points y_1, y_2 in the sphere such that there is a neighborhood $\mathcal{U}$ of the identity in the space of C^r $r \geq 7$ families of diffeomorphisms such that for every family $f_\lambda^v \in \mathcal{U}$ and satisfying* *i)* $f_\lambda^v \mathcal{M}_{x_+,\theta} = \mathcal{M}_{x_+,\theta} \qquad \forall \theta.$

ii) $f_\lambda^v|_{K_1^\varepsilon} = Id,$

then, we can find two C^{r-2} *families* $f_\lambda^{v,1}$, $f_\lambda^{v,2}$ *satisfying*

a) $f_\lambda^{v,1} \mathcal{M}_{x_+,\theta} = f_\lambda^{v,2} \mathcal{M}_{x_+,\theta} = \mathcal{M}_{x_+,\theta} \qquad \forall \theta.$

b) $f_\lambda^{v,1} \circ \Pi_{y_1} = \Pi_{y_1} \circ f_\lambda^{v,1}.$

$f_\lambda^{v,2} \circ \Pi_{y_2} = \Pi_{y_2} \circ f_\lambda^{v,2}.$

c) $f_\lambda^v = f_\lambda^{v,1} \circ f_\lambda^{v,2}.$

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

$$\begin{array}{rcl}
x_{n+1} & = & \cos 2\pi\varphi \\
x_n & = & \sin 2\pi\varphi \ \cos 2\pi w_1 \\
x_{n-1} & = & \sin 2\pi\varphi \ \sin 2\pi w_1 \ \cos 2\pi w_2 \\
\vdots & & \quad \vdots \\
x_3 & = & \sin 2\pi\varphi \ \sin 2\pi w_1 \cdots \sin 2\pi w_{n-2} \\
x_2 & = & \sin 2\pi\varphi \ \sin 2\pi w_1 \cdots \sin 2\pi w_{n-2} \ \cos 2\pi\theta \\
x_1 & = & \sin 2\pi\varphi \ \sin 2\pi w_1 \cdots \sin 2\pi w_{n-2} \ \sin 2\pi\theta
\end{array}$$

for φ, $w_1 , \dots , w_{n-2} \in (0, 1/2)$, $\theta \in (0, 1)$.

Notice that if $|x_1| \geq \varepsilon$, then $|x_{n+1}| \leq 1-\varepsilon'$ so that φ is defined univocally, and, since $|x_n|^2 + |x_{n+1}|^2 \leq 1 - \varepsilon^2$, w_1 is likewise outside of the critical values of sin and the same argument is applicable to $w_2 , \dots , w_{n-2}$. In others words, this polar coordinate system is a good coordinate system admitting a differentiable inverse defined in a open set $V \subset \mathbf{T}^{n+1}$.

We observe that when we express a vertical diffeomorphism in these coordinates, the value of φ changes but the others coordinates remain unaltered. Similarly, a reflection across the plane Π_{y_1} with $y_1 = (1, 0, \dots , 0)$ is expressed as a change in sign for the θ coordinate while all the others remain fixed. More generally, a reflection across Π_{y_2} with $y_2 = (\cos\rho, \sin\rho, 0, \dots , 0)$ can be expressed by sending θ into $2\rho - \theta$ and leaving all the other variables unchanged.

If we take coordinates (φ, θ, w) the family f^v_λ will be written as

$$(\varphi, \theta, w) \to (\varphi + v_\lambda(\varphi, \theta, w), \theta, w)$$

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

$$(\varphi, \theta, w) \to (\varphi + v^i_\lambda(\varphi, \theta, w), \theta, w)$$

with

$$\begin{aligned} v^1_\lambda(\varphi, \theta, w) &= v^1_\lambda(\varphi, -\theta, w) \\ v^2_\lambda(\varphi, \theta, w) &= v^2_\lambda(\varphi, -\theta + 2\rho, w) \end{aligned} \tag{2}$$

The equation that v^1_λ v^2_λ have to solve is, as an straightforward calculation shows:

$$v_\lambda(\varphi, \theta, w) = v^1_\lambda(\varphi + v^2_\lambda(\varphi, \theta, w), \theta, w) + v^2_\lambda(\varphi, \theta, w) \tag{3}$$

Notice that there is an open set of values of w for which the point correspond (φ, θ, w) is in K^ε_1 hence $v_\lambda(\varphi, \theta, w) = 0$. For these values of w, (3) admits the trivial solution $v^1_\lambda(\varphi, \theta, w) = v^2_\lambda(\varphi, \theta, w) = 0$. This implies that $f^{v,1}_\lambda = Id$, $f^{v,2}_\lambda = Id$ for the points in the sphere satisfying this condition.

Clearly, its solutions can be extended by the identity to all points of the sphere in K^ε_1 even for those for which our coordinate system does not work. Under conditions of proximity to the identity of f^v_λ, $f^{v,1}_\lambda$, $f^{v,2}_\lambda$ these will be solutions of our original problem.

Since, except for these considerations w does not enter either (3) or in (2) we can consider these variables as parameters in the problem in the same way as λ. We will, hence assume that λ refers both to the parameters in the family and to w and omit w from the notation.

In the enunciate of the theorem y_1, y_2 were left at our choice. The choices of ρ – the angle between the planes passing through the origin and y_1, y_2 respectively– that are acceptable for the theorem are those satisfying the so-called *Diophantine* conditions.

Those are conditions of the form

$$\left| \rho - \frac{p}{q} \right|^{-1} \leq C q^{2+\delta} \tag{4}$$

It is well known that for $\delta > 0$, the numbers satisfying inequalities of this form have full measure. For $\delta = 0$, the set of numbers satisfying these inequalities is called the of constant type numbers. (4) is equivalent to ρ having a bounded continued fraction expansion. All irrational numbers that satisfy a quadratic equation with integer coefficients are constant type. In particular constant type numbers are dense.

Even if all Diophantine ρ's would lead to a theorem of the type we want and for δ small with the conditions of differentiability required in the statement, constant type numbers will lead to the sharpest differentiability conclusions and we will, from now on assume we have chosen a constant type number. The actual choice will affect the sizes of the neighborhoods but not the differentiability properties.

Proceeding heuristically for the moment, the importance of the Diophantine properties becomes apparent if we consider the *linearized equations* obtained by expanding formally (3) in the unknowns and keeping only the linear terms. We obtain:

$$v_\lambda(\varphi, \theta) = v^1_\lambda(\varphi, \theta) + v^2_\lambda(\varphi, \theta) \tag{5}$$

We expand in Fourier coefficients in θ $v_\lambda(\varphi, \theta) = \sum_k \hat{v}_{\lambda,k}(\varphi) e^{2\pi i k\theta}$ and analogously for v^1_λ, v^2_λ and obtain that (5) is equivalent to:

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

Similarly, the symmetry conditions for v^1_λ, v^2_λ can be expressed as:

$$\begin{aligned} \hat{v}^1_{\lambda,-k}(\varphi) &= \hat{v}^1_{\lambda,k}(\varphi) \\ \hat{v}^2_{\lambda,-k}(\varphi) e^{4\pi i \rho} &= \hat{v}^2_{\lambda,-k}(\varphi) \end{aligned} \tag{7}$$

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

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

When $k \neq 0$, these equations admit the solutions:

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

When $k = 0$, there are many solutions. We just choose:

$$\hat{v}^1_{\lambda,0}(\varphi) = \hat{v}^2_{\lambda,0}(\varphi) = \frac{1}{2}\hat{v}_{\lambda,0}(\varphi) \tag{9}$$

Notice that, if ρ were rational, the denominators would become zero, that is, there would be no solution for the equation. Even if ρ is irrational the denominators would become arbitrarily close to zero. Nevertheless, if ρ is a number of constant type, we have the bounds

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

that allow us to control the new Fourier expansions.

If we think of (3) as an equation in spaces of functions, (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 calculation above shows only existence of the inverse of the derivative zero. If we had started making perturbations around a non-zero, v^1_λ, v^2_λ 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 can cope with unbounded inverses, the one that will be useful for this problem is that of Zehnder [13] which requires existence of 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 [13] in the version we are going to apply.

Definition 3.1 *We say $(X_\alpha)_{\alpha\geq 0}$ is a family of regular spaces when X_α are real Banach spaces with norm $\|\ \ \|_\alpha$ and, whenever for $\alpha' \geq \alpha$ we have:*

$$X_0 \supset X_\alpha \supset X_{\alpha'} \supseteq X_\infty = \bigcap_{\alpha\geq 0} X_\alpha$$
$$\|u\|_\alpha \leq \|u\|_{\alpha'} \quad \forall u \in X_{\alpha'}$$

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

satisfying:

$$i) \quad \lim_{t\to\infty} \|S_t u - u\|_0 = 0 \quad \forall u \in X_0$$

$$ii) \quad \|S_t u\|_\alpha \le t^{\alpha-\alpha'} C(\alpha,\alpha') \|u\|_{\alpha'} \quad \forall u \in X_{\alpha'} \quad 0 \le \alpha' \le \alpha$$

$$iii) \quad \|S_t u - u\|_{\alpha'} \le t^{\alpha'-\alpha} C(\alpha,\alpha') \|u\|_\alpha \quad \forall u \in X_\alpha \quad 0 \le \alpha \le \alpha'$$

We now consider several families of regular spaces and denote by the same symbol their norms which are probably different. However the confusion is not possible because the kind of vector that we are evaluating indicate the norm we are using.

Theorem 3.2 *Let X_α, Y_α, Z_α be families of regular spaces.*

Let $F : X_0 \times Y_0 \to Z_0$ be such that $F(x_0, y_0) = 0$ and continuous on a set B_0. (For any α, we will write $B_\alpha = \{x, y \in X_0 \times Y_0 \ / \ \|x - x_0\|_\alpha \le 1\,,\ \|y - y_0\|_\alpha \le 1\}$.)

Assume

(H.1.) *$F(x, \cdot) : Y_0 \to Z_0$ is twice differentiable and*

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

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

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

(H.3.) *The triple (F, x_0, y_0) is of order ∞. That is*

(H.3.1.) $(x_0, y_0) \in X_\infty \times Y_\infty$.

(H.3.2.) $F(B_0 \cap X_\alpha \times Y_\alpha) \subset Z_\alpha \quad 0 \le \alpha$.

(H.3.3.) *There exist constants M_α, $1 \le \alpha$ such that if $(x, y) \in X_\alpha \times Y_\alpha \cap B_1$ satisfies $\|x - x_0\|_\alpha \le K$, $\|y - y_0\| \le K$, then $\|F(x, y)\|_\alpha \le K\, M_\alpha$.*

(H.4.) *There exist an approximate right-inverse of loss γ, $1 \le \gamma$. That is, for every $(x, y) \in B_\alpha$ there exist a linear mapping $\eta(x, y) \in \mathcal{L}(Z_\alpha, Y_{\alpha-\gamma})$ such that:*

(H.4.1.) $\|\eta(x, y) z\|_0 \le M_0 \|z\|_\gamma$.

(H.4.2.) $\|d_2F(x,y)\,\eta(x,y)-1)z\|_0 \le M_0\|F(x,y)\|_\gamma\,\|z\|_\gamma$.

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

a) $F(x,\psi(x)) = 0$.

b) $\|\psi(x)-y_0\|_\gamma \le C_\lambda^{-1}\|x-x_0\|_\lambda$.

In particular $\psi(D_\lambda \cap X_\infty) \subset Y_\infty$.

Moreover, if $\eta(x,y)$ depends continuously on (x,y), then the mapping $\psi : D_\lambda \to Y_\gamma$ is continuous.

To apply the theorem to solve (3) we have to choose an appropriate functional and appropriate spaces of families functions.

We will consider Banach spaces of C^r-families of mappings. Let $\delta \in \mathbb{R}$, we stand by $C^r_\delta(\mathbf{T}^m)$ the space of functions $v \in C^r(\mathbf{T}^m)$ which verify:

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

Let us fix $s \ge 0$ and a small parameter $\varepsilon > 0$. We take $X_\alpha = Z_\alpha = C^{s+\alpha}(\mathbf{T}^m)$, $Y_\alpha = C_0^{s+\alpha+2+\varepsilon}(\mathbf{T}^m) \times C_\rho^{s+\alpha+2+\varepsilon}(\mathbf{T}^m)$ with the corresponding Hölder norms and define the operators

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

We will show that F verify all the conditions of the previous statement around $(0,0,0)$. In our case Zehnder's theorem is not anymore than a hidden inverse function theorem.

Most of its conditions are deduced in the Appendix A. The hypothesis (H.1.) (H.2.) (H.3.) are immediate consequences of XXXXXX .

The most crucial hypothesis is that about existence of an approximate right-inverse of the linearized operator, which we now discuss.

We will start the construction of such approximate inverse $(0,0,0)$. To study its differentiability properties, we will use the following result proved in Rüssmann [10].

Lemma 3.4 *Let ρ be a constant type number and*

$$h(z) = \sum_{k=-\infty}^{\infty} \hat{h}_k e^{ikz}$$

an analytic complex function in $B_r = \{z \in \mathbf{C} \,/\, |\mathrm{Im} z| < r\}$ with mean value zero and $|h(z)| \leq M \ \forall z \in B_r$. If we take

$$u(z) = \sum_{k=-\infty}^{\infty} \frac{\hat{h}_k}{e^{2\pi i k\rho - 1}} e^{ikz}$$

there is a constant K such that if $r' < r$ then

$$|u(z)| \leq \frac{K\, M}{r - r'} \qquad \text{for} \quad |\mathrm{Im} z| < r'. \tag{11}$$

Notice that $F : X_\alpha \times Y_\alpha \to Z_\alpha$ is differentiable with respect the $Y's$ components and $d_2F(0,0,0)(v_\lambda^1, v_\lambda^2) = v_\lambda^1 + v_\lambda^2$ in all the cases.

If we take a periodic analytic function $v_\lambda(\varphi, \theta)$ defined in the strip U_r^m and consider the pair v_λ^1, v_λ^2 defined by the equations (8), (9) that verify $v_\lambda = v_\lambda^1 + v_\lambda^2$ then both functions are analytic in $U_{r'}^m$ for $r' < r$ and satisfies a inequality similar to (11) with may be some other constants. (Notice that if ρ is a constant type number, so is -2ρ, $e^{-4\pi i\rho k}/(e^{-4\pi i\rho k} - 1) = 1 + 1/(e^{-4\pi i\rho k} - 1)$ and that $\sup_{\Im x \leq r} |\sum_k \hat{h}_k e^{2\pi i k x}| = \sup_{\Im x \leq r} |\sum_k \hat{h}_{-k} e^{2\pi i k x}|$) From lemma 3.4, it is possible to deduce regularity properties of the inverse difference operator 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 Λ_r spaces (a modern survey of the properties of these spaces, including the characterizations by approximation properties is [12] We just remark that, when $r \notin \mathbb{N}$, $\Lambda_r = C^r$. When $r \in \mathbb{N}$, $C^r \subset \Lambda_r$ 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, [13] We have:

Lemma 3.5 *Let ρ be a constant type number and*

$$h(z) = \sum_{k=-\infty}^{\infty} \hat{h}_k e^{ikz}$$

a C^r function, $r \notin \mathbb{N}$ $r > 1$, with mean value zero. The function:

$$u(z) = \sum_{k=-\infty}^{\infty} \frac{\hat{h}_k}{e^{2\pi i k\rho - 1}} e^{ikz}$$

satisfies:

$$||u||_{C^{r-1}} \leq K||h||_{C^r}$$

where K is a constant depending only on ρ.

Using 3.5 to estimate the solution of the approximate inverse that we are considering, we obtain:

Lemma 3.6 *Let $\gamma > 3$ and $\alpha \geq \gamma$. For every $v \in Z_\alpha$ there exist functions $(v^1, v^2) \in Y_{\alpha-\gamma}$ such that $v = v^1 + v^2$. Moreover the map*

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

is continuous.

Now a 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 [3], pg 198], Zehnder [13], pg. 133].

We proceed heuristically in the following. Consider for instance a smooth manifold $\mathcal{M}$, which locally can be identified with a Banach space, and a differentiable action $\Phi : \mathcal{G} \times \mathcal{M} \to \mathcal{M}$ where $\mathcal{G}$ is an infinite dimensional group. Assume that

i) $F : \mathcal{G} \to \mathcal{M}$ is defined as $F(g) = \Phi(g, m_0)$ for $m_0 \in \mathcal{M}$.

ii) $dF(id) : G = T_{id}\mathcal{G} \to T_{m_0}\mathcal{M}$ is surjective with right inverse η.

If $L_g : \mathcal{G} \to \mathcal{G}$ denotes the left multiplication by g on $\mathcal{G}$ and $\Phi_g : \mathcal{M} \to \mathcal{M}$ denotes the action of g on $\mathcal{M}$ then

$$\Phi_g \circ F(h) = F \circ L_g(h) \tag{12}$$

for every $g, h \in \mathcal{G}$, which yields

$$d\Phi_g(m_0) \circ dF(id) = dF(g) \circ dL_g(id) \tag{13}$$

Hence the map $\eta(g) = dL_g(id) \circ \eta \circ d\Phi_{g^{-1}}(F(g))$ defines a right-inverse of $dF(g)$.

We now make precise the spaces and the domains in which the above calculations are valid.

Lemma 3.7 *Let $\alpha \geq \gamma$ and take $(v'_\lambda, v^1_\lambda, v^2_\lambda) \in B_\alpha$. There exists $\eta(v^1_\lambda, v^2_\lambda) \in \mathcal{L}(Z_\alpha, Y_{\alpha-\gamma})$ such that*

$$d_2F(v'_\lambda, v^1_\lambda, v^2_\lambda) \circ \eta(v^1_\lambda, v^2_\lambda)(v_\lambda) = v_\lambda$$

for every $v_\lambda \in Z_\alpha$.

Proof. Let us consider $F_2 : Y_\alpha \to Z_\alpha$, $(v^1_\lambda, v^2_\lambda) \to F(0, v^1_\lambda, v^2_\lambda)$. It is obvious that $d_2F(v'_\lambda, v^1_\lambda, v^2_\lambda) = dF_2(v^1_\lambda, v^2_\lambda)$.

We can realize the sets $X_\alpha, Y_\alpha, Z_\alpha$ as the corresponding tangent spaces at the identity of groups of diffeomorphisms in $\mathbf{T}^m$, where the functional equation defined in lemma (3.3,c) is the composition.

Let $(v^1_\lambda, v^2_\lambda) \in Y_\alpha$ and $v_\lambda = F_2(v^1_\lambda, v^2_\lambda)$. In our problem the group operation and differentiable action which appear in (13) are defined by the relations

$$\begin{array}{rccl} \Phi : & Z_\alpha & \longrightarrow & Z_\alpha \\ & v_\lambda & \longrightarrow & \begin{array}{l} \mathbf{v}^1_\lambda(\varphi + v_\lambda(\varphi + \mathbf{v}^2_\lambda(\varphi,\theta),\theta) + \mathbf{v}^2_\lambda(\varphi,\theta),\theta) + \\ v_\lambda(\varphi + \mathbf{v}^2_\lambda(\varphi,\theta),\theta) + \mathbf{v}^2_\lambda(\varphi,\theta). \end{array} \end{array}$$

and

$$\begin{array}{rccl} L : & Y_{\alpha-\gamma} & \longrightarrow & Y_{\alpha-\gamma} \\ & (w^1_\lambda, w^2_\lambda) & \longrightarrow & \begin{array}{l} (v^1_\lambda(\varphi + w^1_\lambda(\varphi,\theta),\theta) + w^1_\lambda(\varphi,\theta), \\ w^2_\lambda(\varphi + v^2_\lambda(\varphi,\theta),\theta) + v^2_\lambda(\varphi,\theta)). \end{array} \end{array}$$

It follows from proposition (A.8) that Φ is continuously differentiable and $d\Phi(v_\lambda) \in \mathcal{L}(Z_\alpha, Z_\alpha)$ and

$$d\Phi(v_\lambda)(w_\lambda) = w_\lambda(\varphi + \mathbf{v}^2_\lambda(\varphi,\theta),\theta) +$$

$$\frac{\partial \mathbf{v}^1_\lambda}{\partial \varphi}(\varphi + v_\lambda(\varphi + \mathbf{v}^2_\lambda(\varphi,\theta),\theta) w_\lambda(\varphi + \mathbf{v}^2_\lambda(\varphi,\theta),\theta).$$

Similarly, L is continuously differentiable, $dL(0) \in \mathcal{L}(Y_{\alpha-\gamma}, Y_{\alpha-\gamma})$ and

$$dL(0)(w_\lambda^1, w_\lambda^2) = (\frac{\partial v_\lambda^1}{\partial \varphi}(\varphi, \theta) w_\lambda^1(\varphi, \theta) + w_\lambda^1(\varphi, \theta), w_\lambda^2(\varphi + v_\lambda^2(\varphi, \theta), \theta))$$

The map $\eta(v_\lambda^1, v_\lambda^2) = dL(0) \circ \eta \circ d\Phi(v_\lambda) \in \mathcal{L}(Z_\alpha, Y_{\alpha-\gamma})$ and defines a right-inverse of dF_2.

We have verified all the hypotheses of theorem (3.2). Therefore, given $v_\lambda \in C^s(\mathbf{T}^m)$,$s > 6$ close to 0 there are $(v_\lambda^1, v_\lambda^2) \in C_0^{s-1-\varepsilon}(\mathbf{T}^m) \times C_\rho^{s-1-\varepsilon}(\mathbf{T}^m)$ such that $F(v_\lambda, v_\lambda^1, v_\lambda^2) = 0$, and hence the conclusions of the lemma (3.3) hold.

Therefore, we have shown that in a C^r neighborhood of the identity $r > 6$ all diffeomorphisms can be factored. Unfortunately, the set of C^r diffeomorphisms is not a group unless $r \in \mathbb{N}$. This is the reason why the statement of 3.1 is made using only C^7 regularity.

A Differentiability of composition in Hölder spaces

A.1 Introduction

In this appendix we collect some theorems and counterexamples that clarify the differentiability properties of composition in the classical spaces of Hölder functions.

We have emphasized the methods based on elementary estimates, triangle inequality , mean value theorem and the like. Therefore, the results presented here will be valid in a great generality - e.g., open sets of Banach spaces. In finite dimensions, it is sometimes possible to obtain similar results using characterizations of Hölder spaces by approximation properties by analytic functions [11], [12], [9]. As intermediate steps we will also prove elementary versions of interpolation inequalities that work in infinite dimensional domains with good geometric properties.

Since the composition operator is the basic operator of dynamical systems, it is clear that the study of this operator acting on different functional spaces should be very useful. The systematic study of differentiability properties of the composition operator in the spaces of continuously differentiable

functions can be found in [5]. In order to deduce good properties from this operator uniform continuity of the factors is required. Irwin formulates his version on compact sets where continuity automatically implies uniform continuity.
In this appendix, we have studied only classical Hölder spaces. Since Hölder character assures uniform continuity we consider maps defined on open subsets. Notice that some of the results for these spaces, even on compact sets, are different from the straightforward extensions of those for continuously differentiable functions. Unfortunately, the literature abounds with confusions caused by this fact. Notice also that, as stated in 1, many of the results discussed here carry over to manifolds.

A.2 Basic definitions and notation

Let E, F be Banach spaces. We will denote the space of bounded linear functions from E to F by $\mathcal{L}(E,F)$ and in general $\mathcal{L}_n(E,F) = \mathcal{L}(E, \mathcal{L}_{n-1}(E,F))$.

Let $U \subset E$ be an open set. A function $f : U \to F$ is Frechet differentiable at x_0 if we can find a bounded linear function $df(x_0)$ such that

$$\lim_{h\to 0} \frac{\| f(x_0+h) - f(x_0) - df(x_0)h \|}{\| h \|} = 0$$

We say that f is differentiable in U if f is differentiable at every point $x_0 \in U$. We say that f is of class C^1 if it is differentiable and the mapping $df : U \to \mathcal{L}(E,F)$ that to each $x_0 \in U$ associates $df(x_0)$ is continuous.

Proceeding by induction, we define $d^n f$ to be the differential of $d^{n-1} f$ ($d^1 f = df$) and we say that a function f is n times continuously differentiable, $df : U \to \mathcal{L}(E,F)$ is $(n-1)$ times continuously differentiable.

We will denote by $\mathcal{L}_n(E_1, \ldots, E_n, F)$ the space of n-multilinear functions having arguments in $E_1, \ldots, E_n$ and values in F. It is standard that we can identify $\mathcal{L}_n(E,F)$ with the space of bounded multilinear functions with n arguments in E taking values in F. We will think of derivatives as multilinear functions.

We will say that $f \in C^n(U,F)$ if f is n times continuously differentiable and $\| d^n f(x) \|$ is uniformly bounded by $x \in U$. In that case, we denote

$$\| f \|_{C^n} = \sup_{0\le k\le n} \quad \sup_x \| d^k f(x) \| \ .$$

Notice that, according to our definition there could be functions which are n times continuously differentiable but not in $C^n(U,F)$.

If $f \in C^n(U,F)$, V is an open subset of F with $f(U) \subset V$ and $g \in C^n(V,G)$ then $g \circ f \in C^n(U,G)$. Moreover by a repeated application of the chain rule we can find constants $c_{n,j_1,\ldots,j_n}$ such that

$$d^n(g \circ f)(x) = \sum_{k=1}^{n} c_{k,j_1,\ldots,j_n} d^k g(f(x))\{d^{j_1} f(x), \ldots, d^{j_k} f(x)\}$$

Remark. Weaker notions of differentiability spaces have been proposed and the one we consider here is sometimes called "strongly differentiable functions." Moreover, if the manifold under consideration is not compact, - e.g., $\mathbb{R}^d$ or a bounded subset of an infinite dimensional Banach space - there may be functions which are continuously differentiable but whose derivative is unbounded. In those cases, besides the topology induced by the supremum norm we consider here, there are other natural topologies, e.g., the Whitney topology associated to uniform convergence on compact sets. This topology is an inductive limit of the ones we consider here and, hence, the results of continuity can be immediately lifted. Since this topology does not come from a norm only weak concepts of differentiability can be introduced for composition and similar operators, when we consider such topologies.

Definition A.1 *If $r = n + \alpha$ with $n \in \mathbb{N}$ and $0 < \alpha \leq 1^-$ we say that $f \in C^r(U,F)$ if $f \in C^n(U,F)$ and moreover we have*

$$H_\alpha(d^n f) = \sup_{x \neq y} \frac{\| d^n f(x) - d^n f(y) \|}{\| x - y \|^\alpha} < \infty$$

With the introduction of

$$\| f \|_{C^r} = \sup(\| f \|_{C^n}, H_\alpha(d^n f))$$

$C^r(U,F)$ is a Banach space.

Notice that with this definition, we have

$$\| f \|_{C^r} \leq \sup(\| f \|_{C^0}, \| df \|_{C^{r-1}})$$

this makes it easy to carry out several inductions and hence, this definition is preferable to other equivalent norms.

Remark. If $U = U_1 \cup U_2$, $U_1 \cap U_2 \neq \emptyset$, we can write $C^r(U) = C^r(U_1) \oplus C^r(U_2)$ and it is possible to study $C^r(U_i)$ separately. Similarly, the study of composition operators can be broken up into the study of the components. So, we will henceforth assume that the domains in which we study the functions are connected by arcs.

A.3 Geometry of domains and Interpolation properties

The following are immediate.

Proposition A.1 *Let $0 \leq \alpha < \beta \leq 1^-$, and $n \in \mathbb{N}$. Then*

i) $C^{n+\beta}(U) \subset C^{n+\alpha}(U) \subset C^n(U)$.

ii) $C^{n+\beta}(U)$ is a closed subset of $C^{n+\alpha}(U)$.

Remark. For most spaces useful in analysis the inclusion in Proposition A.1 is strict. It suffices to consider $|x|^{\alpha+n}$ on [0,1] which is in $C^{n+\alpha}$ but not on $C^{n+\beta}$.

In the following lemmas we develop some interpolation properties of the norms that we will use later in the discussions of the composition operator.

Lemma A.1 *Let $r = n + \alpha$, $s = n + \beta$, $t = n + \gamma$ with $0 \leq \alpha < \beta < \gamma \leq 1$ and $\mu = (t - s)/(t - r)$. There is a constant $M_{r,t}$ such that if $f \in C^t(U, F)$, then*

$$\| f \|_{C^s} \leq M_{r,t} \| f \|_{C^r}^{\mu} \| f \|_{C^t}^{1-\mu} \tag{14}$$

Proof: Notice that $1 - \mu = (s - r)/(t - r)$ and $s = \mu r + (1 - \mu)t$. We will establish the claim using induction on n. We will first describe the case $n = 0$

If we define for $\beta \in (0, 1^-]$:

$$H_\beta(f) = \sup \frac{\| f(x) - f(y) \|}{\| x - y \|^\beta}$$

we have

$$H_\beta = \sup \frac{\| f(x) - f(y) \|^\mu}{\| x - y \|^{\alpha\mu}} \sup \frac{\| f(x) - f(y) \|^{1-\mu}}{\| x - y \|^{\gamma(1-\mu)}} = H_r(f)^\mu H_t(f)^{1-\mu}.$$

From here it is obvious that

$$\| f \|_{C^s} \leq M_{r,t} \| f \|_{C^r}^{\mu} \| f \|_{C^t}^{1-\mu}$$

where

$$M_{r,t} = \begin{cases} 1 & \text{if} \quad 0 < \alpha < \gamma \leq 1^- \\ 2 & \text{if} \quad 0 = \alpha < \gamma \leq 1^- \end{cases}$$

For $n \geq 1$ using the inductive hypothesis we find

$$\| df \|_{C^{s-1}} \leq M_{r,t} \| df \|_{C^{r-1}}^{\mu} \| df \|_{C^{t-1}}^{1-\mu} \leq M_{r,t} \| f \|_{C^r}^{\mu} \| f \|_{C^t}^{1-\mu}$$

and thus

$$\| f \|_{C^s} \leq M_{r,t} \| f \|_{C^r}^{\mu} \| f \|_{C^t}^{1-\mu}$$

Notice that $\ln \| f \|_{C^r}$ is a convex , therefore continuous, function in each open interval $(n, n+1)$ with $n \in \mathbb{N}$.

Proposition A.2 *If $0 \leq \alpha < \beta \leq 1^-$ and U is separable then bounded sets of $C^{n+\beta}(U)$ are compact in $C^{n+\alpha}(U)$.*

Proof. For $\alpha = 0$ this is a consequence of Ascoli-Arzela theorem.

For $\alpha > 0$ we use the result for $\alpha = 0$ and the interpolation formula (14).

Notice that in Proposition (A.1) we have not claimed that if $r = n + \alpha < s = m + \beta$ then $C^s \subset C^r$.

The fact that this is not unfortunately true in general can be seen in the following.

Example A.1 Let $D = \{z \in C / |z| < 1\}$, $S = \{z \in C / \Im z = 0, \ 0 \leq \Re z \leq 1\}$ and $U = D - S$.

Denote by $dist_U(z, z')$ the infimum of the length of smooth paths contained in U joining z to z' and define

$$f(x, y) = dist_U(x + iy, \frac{1}{2} + \frac{i}{2})$$

Then, f is continuously differentiable in U and $\| df \|_{C^0} \leq 1$. Nevertheless, there are no constants $k > 0, \alpha > 0$ such that

$$|f(x, y) - f(x', y')| \leq k |(x, y) - (x', y')|^{\alpha}$$

Hence, in the space U $C^1(U) \not\subseteq C^{\alpha}(U)$ for any $0 < \alpha \leq 1^-$.

In many respects this example is a "pathology." Its origins are clear if one tries to use the mean value theorem.

$$f(\mathbf{x}) - f(\mathbf{y}) = \int_0^1 df(\gamma(s))\dot{\gamma}(s)ds \tag{15}$$

where γ is a path joining $\mathbf{x}, \mathbf{y}$.

This immediately gives $|f(\mathbf{x}) - f(\mathbf{y})| \leq \| f \|_{C^1} \ dist_U(\mathbf{x}, \mathbf{y})$ where the symbol $dist_U(\mathbf{x}, \mathbf{y})$ denotes the infimum of the lengths of paths joining $\mathbf{x}$ to $\mathbf{y}$. In the space U of the example this could be much larger than $|\mathbf{x} - \mathbf{y}|$. Clearly, if we want to use arguments based on the mean value formula (15) we have to restrict ourselves to spaces in which $dist_U(x, y)$ is commensurate to $|(x, y)|$.

When we consider problems involving composition, we have to ensure that the intermediate ranges also have this property.

Definition A.2 *balanced We say that an open set U is balanced if there exists a constant k_U such that $d_U(x, y) \leq k_U \| x - y \| \ \forall x, y \in U$.*

This property depends on the shape of U. For example, a convex set of a Banach space is balanced or more generally, a connected set which is a finite union of convex sets, is balanced. Notice that, as stated, the definition only makes sense for connected sets. This is not a shortcoming in view of the fact that , as remarked in the introduction, it suffices to consider spaces of function on connected sets. Notice also that a subset of a balanced subset needs not be balanced.

Unfortunately, C^∞ functions could transform balanced domains into unbalanced ones.

Example A.2 Let $D_+ = \{z \in C/|z| < 1, \Im z > 0\}$ the map $z \to z^2$ transforms the balanced open D_+ on the unbalanced U defined in the example (A.1)

The following results are useful in the representation of the sucesive derivatives of the composition. We omit their proofs which are routine

Proposition A.3 *Let $l \in L(F, G)$. Then $l_\star : C^r(U, F) \to C^r(U, G)$ defined by $l_\star f = l \circ f$ is linear and bounded. Moreover, $\| l_\star \| = \| l \|$.*

Proposition A.4 *Let $r, s, t \in \mathbb{R}$, $r \geq t, s \geq t$. Let $b \in \mathcal{L}_2(F, G; H)$. Then the map $b_\star : C^r(U, F) \times C^s(U, G) \to C^t(U, G)$ defined by $b_\star(f, g)(x) = b(f(x), g(x))$ is continuous and bilinear. Moreover $\| b_\star \| \leq 2 \| b \|$.*

We specially employ these relations when

$$\begin{array}{rcc} b : \mathcal{L}(F, G) \times \mathcal{L}(G, H) & \longrightarrow & \mathcal{L}(F, H) \\ (l, m) & \longrightarrow & l \circ m \end{array}$$

represents the composition of linear operators. If $F = K$ then $\mathcal{L}(F, G) = G$, $\mathcal{L}(F, H) = H$. We will stand by $g \cdot f$ the map $b_\star(f, g)$ defined as

$$\begin{array}{rcc} b : C^r(U, G) \times C^s(U, \mathcal{L}(G, H)) & \longrightarrow & C^t(U, H) \\ (f, g) & \longrightarrow & [x \to g(x)(f(x))] \end{array}$$

Lemma A.2 *Let F, G be Banach spaces $U \subset F, V \subset G$ be balanced subsets. Let $r = n + \alpha, s = m + \beta$ with $m, n \in \mathbb{N}, \alpha, \beta \in [0, 1^-]$. If $f \in C^r(U, F)$, $f(U) \subset V, g \in C^s(V, G)$ then $g \circ f$ belongs to $C^t(U, G)$ where we can take t as follows:*

i) If $n = m = 0, t = rs$

Moreover

i.1) $\| g \circ f \|_{C^{rs}} \leq \| g \|_{C^s} \| f \|_{C^r}^s + \| g \|_{C^0}$

ii) If $\max(r, s) \geq 1, t = \min(r, s)$

Moreover

ii.1) if $r \geq 1, 0 < s \leq 1^-$

$$\| g \circ f \|_{C^s} \leq \| g \|_{C^s} \| f \|_{C^1}^s \; k_U^s + \| g \|_{C^0}$$

ii.2) If $0 < r \leq 1^-, s \geq 1$

$$\| g \circ f \|_{C^r} \leq \| g \|_{C^1} \| f \|_{\dot{C}^r} \; k_V + \| g \|_{C^0}$$

ii.3) If $r \geq 1, s \geq 1, \exists M_t$ *such that*

$$\| g \circ f \|_{C^t} \leq M_t \| g \|_t \, (1 + \| f \|_t^t)$$

Proof. If $n = m = 0$, we have

$$\| g \circ f(x) - g \circ f(y) \| \leq \| g \|_{C^\alpha} \| f(x) - f(y) \|^\alpha \leq \| g \|_{C^\alpha} \| f \|_{C^\beta} \| x - y \|^{\alpha\beta}$$

This establishes i), i.1). The cases ii.1) and ii.2) can be obtained in a similar way.

To verify the last case, we assume that $s = r = t = k + \gamma$ with $k, \in \mathbb{N}$ $k \geq 1$ and $\gamma \in [0, 1^-]$. We proceed by induction in k.

For $k = 1$, $d(g \circ f) = dg \circ f \cdot df$. Applying Proposition [A.4] and the i) of Proposition [A.2] we obtain

$$\| dg \circ f \|_{C^\gamma} \leq k_U^\gamma \| dg \|_{C^\gamma} \| f \|_{C^1}^\gamma + \| dg \|_{C^0} \leq$$

$$M \| g \|_{C^{1+\gamma}} (1 + \| f \|_{C^{1+\gamma}}^{1+\gamma})$$

Let us assume that the above inequality holds until the integer $k - 1$ and consider $t = k + \gamma$. Then by induction we have that $d(g \circ f) = dg \circ f \cdot df \in C^{t-1}(U)$ and

$$\| d(g \circ f) \|_{C^{t-1}} \leq 2 \| dg \circ f \|_{C^{t-1}} \| df \|_{C^{t-1}} \leq$$

$$2M \| dg \|_{C^{t-1}} (1 + \| df \|_{C^{t-1}}^{t-1}) \| df \|_{C^{t-1}}$$

since $M_t(1 + |x|^t) \geq M(1 + |x|^{t-1})|x|$ for some $M_t > 1$, we obtain

$$\| g \circ f \|_{C^t} \leq \sup(\| g \circ f \|_{C^0}, \| d(g \circ f) \|_{C^{t-1}}) \leq M_t \| g \|_{C^t} (1 + \| f \|_{C^t}^t)$$

This completes the proof of Lemma A.2

The following examples show that the hypotheses in proposition (A.2) are sharp.

Example A.3 Let $\alpha > 0$ and $f_\alpha : [-1, 1] \to \mathbb{R},\ x \to |x|^\alpha$
i) If $0 < \alpha, \beta \leq 1^-$, then $f_\alpha \circ f_\beta \in C^\gamma([-1, 1], \mathbb{R})$ if and only if $\gamma \leq \alpha \cdot \beta$.
ii) If $\alpha/2 \notin \mathbb{N}$ then $f_\alpha \circ Id \in C^\gamma([-1, 1], \mathbb{R})$ if and only if $\gamma \leq (\alpha - 1) + 1^-$.

As a corollary of proposition (A.2) we obtain that if U is a balanced open set then the formula (14) also holds for $t = n+1$ taking

$$M_{r,t} = \begin{cases} k_U & \text{if } 0 < \alpha < \gamma = 1 \\ \max(2, k_U) & \text{if } 0 = \alpha < \gamma = 1 \end{cases}$$

However, unless we impose some condition on the domain, these inequalities are not possible for exponents with different number of derivatives as shown in the following example.

Example A.4 Let $U = \{(x,y) \in \mathbb{R}^2 / x > 0, y > 1, xy < 1\}$. We take φ as a C^∞ cut-off function with $\| \varphi' \|_{C^0} \leq 4$ which takes the value 1 on $[-\frac{1}{3}, \frac{1}{3}]$ and zero outside of $[-1, 1]$. We define on U the following functions

$$f_n(x,y) = \log(1+x)\varphi(y-n)$$

For $n \in \mathbb{N}$ big enough we find

$$\begin{aligned} \tfrac{\partial F_n}{\partial x}(x,y) &= \tfrac{1}{1+x}\varphi(y-n) \\ \tfrac{\partial F_n}{\partial y}(x,y) &= \log(1+x)\varphi'(y-n) \\ \tfrac{\partial^2 F_n}{\partial x^2} &= -\tfrac{1}{(1+x)^2}\varphi(y-n) \\ \tfrac{\partial^2 F_n}{\partial y^2} &= \log(1+x)\varphi''(y-n) \\ \tfrac{\partial^2 F_n}{\partial x \partial y} &= \tfrac{1}{1+x}\varphi'(y-n) \end{aligned}$$

Observe that since the support of $\varphi(y-n)$ is centereed around $y = n$, which in the domain we are considering implies $x \approx 1/n$ we have:

$$\begin{aligned} &|\tfrac{\partial F_n}{\partial x}| = \mathcal{O}(1); |\tfrac{\partial F_n}{\partial y}| = \mathcal{O}(1/n); \\ &|\tfrac{\partial^2 F_n}{\partial x^2}| = \mathcal{O}(1); |\tfrac{\partial^2 F}{\partial y^2}| = \mathcal{O}(1/n); \quad |\tfrac{\partial^2 F}{\partial x \partial y}| = \mathcal{O}(1) \\ &||F||_{C^0} = \mathcal{O}(1/n) \end{aligned}$$

This implies that $\| F \|_{C^2}$ remains bounded and also $||F||_{C^0}$ converges to zero.

Since $\frac{\partial F_n}{\partial x}$ evaluated at $x = \frac{1}{zn}, y = n$ does not converge to zero, we conclude that there cannot be an inequality of the form

$$||F_n||_{C^1} \leq M ||F_n||_{C^0}^{1/2} ||F_n||_{C^2}^{1/2}$$

More generally, we observe that for any $s \in [1^-,\ 1+1^-)$ $||F_n||_{C^s}$ does not tend to zero. For any $r \in [0, 1^-)$ $||F_n||_{C^r}$ tends to zero and for any $t \leq 2$, $||F_n||_{C^t}$ remains bounded.

Nevertheless, as it is well known, interpolation inequalities for spaces of functions defined on $\mathbb{R}^n$ can be proved those ranges of regularities (see e.g. [11]).

This shows that it should be neccessary to include some conditions in the domain U if interpolation inequalities are to hold.

Remark We observe that there is an abstract argument (see [13]) that shows that if there are smoothing operators, then interpolation inequalities hold. Hence, we conclude that it is impossible to define smooting operators for Hölder spaces of functions in U.

Notice that U in Example A.4 is a subset of $\mathbb{R}^2$ and that there are interpolation formulas for $C^\alpha(\mathbb{R}^2)$ when α is not an integer (see e.g. $[K_r]$). Hence, the fact that there are interpolation formulas is for C^α spaces is not inherited by subsets.

This is somewhat surprising since a naive argument would suggest that we can apply a Whitney extension theorem to extend our function F in U to a function in $\mathbb{R}^2$, apply interpolation to the extended functions and then restrict. The reason why this naive argument is wrong is that the extension operators used in Whitney theorem depend on the regularity considered. In [11] VI, 2 one can find a construction of extension operators ξ_k which extend C^α for $k\alpha < k+1$. Example A.4 shows that it is impossible to find a unique extension operator working simultaneously in C^{α_1} C^{α_2} when the integer parts of α_1 and α_2 are not the same.

This situation is in contrast with regularities measured in other scales of spaces - e.g. Sobolev spaces L^p_k where it is possible to find extension operators that work for all the scale of spaces simultaneously (see. [11] VI 3).

In the following, we will develop son conditions that guarantee that one can obtain interpolation inequalities. We do not know whether condition A.3 on U is neccessary for the validity of interpolation inequalities in U . We do not know conditions which ensure that there is a unique extension operator for all C^α regularities.

Definition A.3 *Let U a balanced open set and $0<\beta\leq 1^-$. We say that U is β -regular if there is a constant M_0 such that for every $x_0\in U$ and $h\in E$ with $\| h \|=1$ we can find $\gamma\in C^{1+\beta}([0,1],U)$ with $\gamma(0)=x_0$, $\gamma'(0)=\tau h$ with $\tau=1$ or -1 and $\| \gamma \|_{C^{1+\beta}}\leq M_0$.*

We say that U is regular if it is β-regular for some β.

Notice that the union of β-regular subsets is also β-regular.

The relevance of regularity lies in the following lemma.

Lemma A.3 *Assume that U is regular. Let r,s,t be positive numbers with $0\leq r<s<t$ and $\mu=(t-s)/(t-r)$. There is a constant $M_{r,t}$ such that if $f\in C^t(U,F)$, then*

$$\| f \|_{C^s}\leq M_{r,t} \| f \|_{C^r}^{\mu}\| f \|_{C^t}^{1-\mu} \tag{16}$$

Proof: We suppose that U is a β-regular open. First we consider $r=n-1+\alpha$, $s=n$, $t=n+\gamma$ with $\gamma\leq\beta$.

Let $n=1$. Let $z\in C^t(\mathbb{R}_+,F)$ and assume that there is $t_0\in\mathbb{R}_+$ with $\| z' \|_{C^0}=\| z'(t_0) \|$.

Substituting z by $az(\delta t)$ if neccessary we can assume that $\| z'(t_0) \|=\| z' \|_{C^0}=1$, and $H_\beta(z')=\frac{1}{2}$.

Taking $R(t)=z(t)-z(t_0)-z'(t_0)(t-t_0)$ then,

$$\| R(t) \|\leq\| \int_{t_0}^{t}(z'(s)-z'(0))ds \|\leq |t-t_0|^{1+\beta}H_\gamma(z').$$

Now we have

$$\| z(t_0+1)-z(t_0) \|=\| z'(t_0)+R(t_0+1) \|\geq \frac{1}{2}.$$

This shows that $H_\alpha(z)\geq\frac{1}{2}$ and therefore,

$$\| z' \|_{C^0}\leq 2H_\alpha(z)^\mu H_\gamma(z')^{1-\mu}.$$

Let us suppose that $\| df \|_{C^0}=1$. For every small ϵ there exists $x_0\in U$, $h\in E$ with $\| h \|=1$ such that $1-\epsilon\leq\| df(x_0)h \|\leq 1$.

We define $z(t)=f(\gamma_{x_0,h}(1-e^{-t}))$.Then

$$\| df \|_{C^0}-\epsilon\leq\| z' \|_{C^0}\leq 2H_\alpha(z)^\mu H_\gamma(z')^{1-\mu}\leq M \| f \|_{C^r}^{\mu}\| f \|_{C^t}^{1-\mu}$$

for every ϵ. Thus we obtain

$$\| f \|_{C^s} \leq M \| f \|_{C^r}^{\mu} \| f \|_{C^t}^{1-\mu} \tag{17}$$

When $n > 1$ we can deduce the inequality (17) by induction.

To obtain the complete result we only have to apply the previous lemma.

Using the inequality (16) and arguing by induction as in the finite dimensional case, considered in [3], we obtain:

Lemma A.4 *Assume that $U \subset F$, $V \subset G$ are regular subsets. Let $t \geq 1$, $f \in C^t(U, F)$, $g \in C^t(V, G)$ with $f(U) \subset V$. If $\| f \|_{C^1} \leq K$ there exist a constant M_t, depending on K such that*

$$\| g \circ f \|_{C^t} \leq M_t(\| g \|_{C^t} + \| g \|_{C^1} \| f \|_{C^t})$$

A.4 Differentiability properties of composition

We will systematically study differentiability properties of

$$\mathrm{comp}: \; C^r(U, F) \times C^s(V, G) \to C^t(U, G)$$

First we will study continuity and differentiability with respect to each of the components and then, join continuity and differentiability.

In order to obtain $g \circ f$ we need $f(U) \subset V$. We assume this condition in proposition (A.5). To consider this composition as a function of the first factor it is necessary that $d(f(U), V^c) \geq \lambda > 0$ to make it possible in a complete neighborhood $\mathcal{U}$ of f. We assume this condition from lemma A.6 on.

The following is a direct consequence of lemma (A.2)

Proposition A.5 *Let $r, s \geq 0$ and*

$$t \leq \begin{cases} rs & \text{if } \; 0 \leq s < 1, 0 \leq r < 1 \\ \min(r, s) & \text{if } \; 1 \leq \max(r, s) \end{cases}$$

Let $f \in C^r(U, F)$. Then the map

$$\begin{array}{rcl} f^* : C^s(V, G) & \longrightarrow & C^t(U, G) \\ g & \longrightarrow & g \circ f \end{array}$$

is linear and continuous .

Now we consider the variation of the first factor.

Proposition A.6 *Let $r, s, t \geq 0$, U, V balanced subsets. and $g \in C^s(V, G)$ Let $\mathcal{U} \subset C^r(U, F)$ be such that for every $f \in \mathcal{U}$ we have $d(f(U, v^c) > \epsilon > 0$.*

The map

$$\begin{array}{rcl} g_* : \mathcal{U} \subset C^r(U, F) & \longrightarrow & C^t(U, G) \\ f & \longrightarrow & g \circ f \end{array}$$

is continuous in the following cases

i) for $t = 0$ when $r \geq 0$, $s > 0$.

ii) for $0 < t < 1^-$ when $s > t$, $r \geq t$, $rs > t$.

iii) for $t = k + \gamma$ with $k \geq 1$, $\gamma \in [0, 1)$ when $s > t$, $r \geq t$.

iv) for $t = k + 1^-$ with $k \geq 0$ when $s > k + 1$, $r \geq t$

Proof: i) It is similar to (i) of proposition (A.2). Thus if $0 \leq s \leq 1^-$ then

$$\| g_*(f_1) - g_*(f_2) \|_{C^0} \leq \| g \|_{C^s} \| f_1 - f_2 \|_{C^0}^s$$

ii.1) First we consider that $0 < s \leq 1^-$. We can also assume $0 < t < r < 1$. Using lemma (A.1) and (i1) of proposition (A.2) we have

$$\begin{gathered} \| g_*(f_1) - g_*(f_2) \|_{C^t} = \| g \circ f_1 - g \circ f_2 \|_{C^t} \leq \\ 2 \| g \circ f_1 - g \circ f_2 \|_{C^0}^{\frac{rs-t}{rs}} \| g \circ f_1 - g \circ f_2 \|_{C^{rs}}^{\frac{t}{rs}} \leq \\ 2 \| g \|_{C^s} [2 + \| f_1 \|_{C^r}^s + \| f_2 \|_{C^r}^s]^{\frac{t}{rs}} \| f_1 - f_2 \|_{C^0}^{s - \frac{t}{r}} \end{gathered} \tag{18}$$

ii.2) We assume that $s = 1 + \beta$ with $0 < \beta < 1$ and $0 \leq t \leq r < 1$.

Let $f_1 \in \mathcal{U}$ and $\lambda = d(f_1(U), V^c) > 0$. We take $x, y \in U$ and make $z_1 = f_1(x), w_1 = f_1(y)$. For $\epsilon > 0$ fixed there is a path γ_1 joining z_1 to w_1 such that

$$g(z_1) - g(w_1) = \int_0^1 dg(\gamma_1(t)) \dot{\gamma}_1(t) dt$$

and

$$\int_0^1 \| \dot{\gamma}_1(t) \| dt \leq (k_V + \epsilon) \| z_1 - w_1 \|$$

Let $f_2 \in \mathcal{U}$ with $\| f_2 - f_1 \|_r < \lambda$. If $z_2 = f_2(x), w_2 = f_2(y)$ then

$$\gamma_2(t) = \gamma_1(t) + (1-t)(z_2 - z_1) + t(w_2 - w_1)$$

is a path contained in U joining z_2 to w_2 and

$$g(z_2) - g(w_2) = \int_0^1 dg(\gamma_2(t))\dot{\gamma}_2(t)dt$$

We take

$$S_1 = \| \int_0^1 [dg(\gamma_1(t)) - dg(\gamma_2(t))]\dot{\gamma}_1(t)dt \| \leq$$

$$\| g \|_{C^{1+\beta}} (\| z_1 - z_2 \| + \| w_1 - w_2 \|)^\beta (k_V + \epsilon) \| z_1 - w_1 \|$$

$$S_2 = \| \int_0^1 dg(\gamma_2(t))(\dot{\gamma}_1(t) - \dot{\gamma}_2(t))dt \| \leq \| g \|_1 \| z_2 - z_1 + w_1 - w_2 \|$$

thus

$$\| g(f_1(x)) - g(f_2(x)) - g(f_1(y)) + g(f_2(y)) \| \leq S_1 + S_2 \leq$$

$$2 \| g \|_{C^{1+\beta}} \| f_1 - f_2 \|_{C^0}^\beta k_V \| f_1 \|_{C^r} \| x - y \|^r + \| g \|_{C^1} \| f_1 - f_2 \|_{C^r} \| x - y \|^r$$

This shows that

$$\| g \circ f_1 - g \circ f_2 \|_{C^r} \leq 2k_V \| g \|_{C^{1+\beta}} (\| f_1 \|_{C^r} \| f_1 - f_2 \|_{C^0}^\beta + \| f_1 - f_2 \|_{C^r})$$

hence there is a constant M such that

$$\| g \circ f_1 - g \circ f_2 \|_{C^r} \leq M \| g \|_{C^{1+\beta}} \| f_1 - f_2 \|_{C^r}^\beta \tag{19}$$

iii) Let us assume that $t = r = k + \alpha, s = k + \beta$ with $0 \leq \alpha < \beta < 1$. The case $\alpha = 0$ is included in [5] , so we consider $\alpha > 0$.

Let $\lambda > 0$ and $f_1, f_2 \in \mathcal{U}$ with $\| f_i \|_r < \lambda \quad i = 1, 2$. For $k = 1$ we have

$$\| g_*(f_1) - g_*(f_2) \|_{C^0} \leq \| g \|_{C^1} k_V \| f_1 - f_2 \|_{C^0}$$

Moreover $d(g \circ f) \in C^{r-1}(U, \mathcal{L}(E, G))$ and

$$\| d(g \circ f_1) - d(g \circ f_2) \|_{C^\alpha} = \| dg \circ f_1 \cdot d(f_1 - f_2) - (dg \circ f_2 - dg \circ f_1) \cdot df_2 \|_{C^\alpha} \leq$$

$$2 \| dg \circ f_1 \|_{C^\alpha} \| d(f_1 - f_2) \|_{C^\alpha} + 2 \| dg \circ f_2 - dg \circ f_1 \|_{C^\alpha} \| df_2 \|_{C^\alpha} \tag{20}$$

On the other hand

$$\| dg\circ f_2 - dg\circ f_1 \|_{C^\alpha} \leq 2 \| dg\circ f_2 - dg\circ f_1 \|_{C^0}^{\frac{s-r}{\beta}} (\| dg\circ f_2 \|_{C^\beta} + \| dg\circ f_1 \|_{C^\beta})^{\frac{\alpha}{\beta}} \leq$$

$$2 \| g \|_{C^{1+\beta}} (\| f_1 \|_{C^1}^{\beta} k_U^{\beta} + \| f_2 \|_{C^1}^{\beta} k_U^{\beta} + 2)^{\frac{\alpha}{\beta}} \| f_1 - f_2 \|_{C^0}^{s-r} \qquad (21)$$

Substituting (21) in (20) we find a constant M such that

$$\| g\circ f_1 - g\circ f_2 \|_{C^t} \leq M \| g \|_{C^{1+\beta}} \| f_1 - f_2 \|_{C^r}^{s-r} \qquad (22)$$

The verification of the inequality (22) for every n is a routine induction argument.

iv) Is a consequence of the arguments used in (ii) and (iii).

Definition A.4 *Let $\Omega \subset \mathbb{R}_+^3$ be the set of indices (r,s,t) which satisfy some of the conditions stated in proposition A.6. We say that Ω is the set of the exponents of continuity*

Proposition A.7 *Let $(r,s,t) \in \mathbb{R}^3 - \Omega$ and $\rho > 0$. There are Banach spaces E, F and $g \in C^s(B_{2\rho} \subset F, \mathbb{R})$ such that*

$$g_* : \mathcal{U} \subset C^r(B_\rho \subset E, F) \longrightarrow C^t(B_\rho, \mathbb{R})$$

is discontinuous.

The continuity in (i) requires the uniform continuity of g. In the remaining cases if $(r,s,t) \notin \Omega$ it satisfies some of the conditions considered below.

If $0 < s \leq 1^-, 0 < r \leq 1^-$ we take φ to be a C^∞ cut-off function which takes the value 1 on $[-\frac{1}{2}, \frac{1}{2}]$ and the value zero outside of $[-1,1]$. We set

$$g(x) = \sum_{i=6}^{\infty} |x - \frac{1}{i}|^s \varphi(2^i(x - \frac{1}{i}))(-1)^i.$$

Since each of the terms in the sum has disjoint support, the sum converges. It is clear that it defines a C^s function out of zero. At zero $g(0) = 0$ and since

$$2^{-s(i-1)} \leq \sup_x |x - \frac{1}{i}|^s \varphi(2^i(x - \frac{1}{i})) \leq 2^{si},$$

$$2^{-s(i-1)} \le \sup_{x\in[\frac{1}{i}-2^{-i},\frac{1}{i}+2^{-i}]} |g(x)| \le 2^{si}$$

It follows that $|g(x)-g(0)|/|x|^s$ is bounded and hence g is C^s:

We take $f_n(x) = |x|^r + \frac{1}{n}$ and $f_0(x) = |x|^r$. Clearly, $\| f_n - f_0 \|_{C^r} \to 0$. Let $n > m$ be integers, m, n of different parity. Then $|g\circ f_n - g\circ f_m|(x) = 2|x|^{rs}$ when $x \in [-2^{\frac{-n-1}{r}}, 2^{\frac{-n-1}{r}}]$ so that $\| g\circ f_n - g\circ f_m \|_{C^{rs}} \ge 2$. Hence, it is impossible that $g\circ f_n$ converges in C^{rs}.

If we consider in (ii) or (iii) the option $s = t$, $r \ge 1$ the continuity also fails. This fact is well known for integer exponents and we will sharpen it below. If $s = t = k+\gamma$ with $k \in \mathbb{N}, 0 < \gamma \le 1^-$, it suffices to take $g = x^k|x|^\gamma, f_n = x + \frac{1}{n}$. Clearly $\| f_n - Id \|_{C^r} \to 0$ for any r.

$$(g\circ f_n)^{(k)}(x) = |x+\frac{1}{n}|^\gamma (k+\gamma)\cdots(\gamma+1)$$

Hence,

$$|(g\circ f_n - g\circ Id)^{(k)}(\frac{-1}{n}) - (g\circ f_n - g\circ Id)^{(k)}(0)| = 2\frac{1}{n^\gamma}(k+\gamma)\cdots(\gamma+1)$$

so that $\| g\circ f_n - g\circ Id \|_{C^t} \ge K > 0$.

We deal the other cases in infinite-dimensional Banach spaces. For $p \ge 1$ we introduce the usual space of summable sequences

$$l^p = \{\mathbf{x} = (x_n)_{n=1}^\infty \in \mathbb{R}^{\mathbb{N}} / \sum_{n=1}^\infty |x_n|^p < \infty\}$$

with the norm

$$\| \mathbf{x} \| = (\sum_{n=1}^\infty |x_n|^p)^{1/p}$$

which converts it into a Banach space. We denote by e_n the sequence with all the components 0 except the n-th one which takes the value 1.

First, we take $0 < r < 1$ and $p = 1/r$. Let us consider the real functions

$$g_{0,n}(x) = n(x-1)\chi_{[1,1+1/n]} + \chi_{[1+1/n,\infty]}$$

and in general $g_{k,n}(x) = \int_0^x g_{k-1,n}(s)ds$. We define

$$\begin{array}{rcc} \mathbf{g_k} : \ B_3 \subset l^p & \longrightarrow & \mathbb{R} \\ (x_j)_{j\in\mathbb{N}} & \longrightarrow & \sum_{j=1}^\infty g_{k,j}(x_j) \end{array}$$

Given $\mathbf{x} = (x_j)_{j\in\mathbb{N}} \in B_3(\mathbf{x})$ there exist indices $j_1, \dots, j_6$ such that if $\mathbf{y} \in B_{1/2}(\mathbf{x})$ then $|y_j| < 1$ for $j \in \mathbb{N} - \{j_1, \dots, j_6\}$. If we denote

$$\pi_x : \ l^p \to \mathbb{R}^6, \ \ \mathbf{x} \to (x_{j_1}, \dots, x_{j_6})$$

and

$$\mathbf{g}_{k,x} : \ \mathbb{R}^6 \to \mathbb{R}, \ \ (x_1, \dots, x_6) \to \sum_{i=1}^{6} g_{k,j_i}(x_i)$$

then $\mathbf{g}_k(\mathbf{y}) = \mathbf{g}_{k,x} \circ \pi_x(\mathbf{y})$, for every $\mathbf{y} \in B_{1/2}(\mathbf{x})$. Hence $\mathbf{g}_k \in C^k(B_3, \mathbb{R})$.

Let φ be a cut-off function which takes the value 1 on $[\frac{-1}{3}, \frac{1}{3}]$ but zero outside of $[\frac{-2}{3}, \frac{2}{3}]$ and set $h(x) = (1 + |x-1|^r)\varphi(x-1)$. We define

$$\begin{array}{rccc} \mathbf{f} : & B_{3/2} \subset l^1 & \longrightarrow & l^p \\ & (x_j)_{j\in\mathbb{N}} & \longrightarrow & (x_j)_{j\in\mathbb{N}} \end{array}$$

Then

$$\| \mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y}) \| = y(\sum_{j=1}^{\infty} |h(x_j) - h(y_j)|^p)^{1/p} \geq \| h \|_{C^r} \| \mathbf{x} - \mathbf{y} \|^r$$

and $\mathbf{f}_n = \mathbf{f} + e_n/2n$ satisfies this same inequality for every $n \in \mathbb{N}$. It is obvious that $\mathbf{f}, \mathbf{f}_n \in C^r(B_{3/2} \subset l^1, l^p)$ and $\| \mathbf{f}_n - \mathbf{f} \|_{C^r} \to 0$.

However for n big enough if $x \in [1, 1 + (1/2n)^{-r}]$ then

$$|\mathbf{g}_1 \circ \mathbf{f}_n(xe_n) - \mathbf{g}_1 \circ \mathbf{f}(xe_n)| = |x-1|^r/2 + 1/8n$$

so that $\| \mathbf{g}_1 \circ \mathbf{f}_n - \mathbf{g}_1 \circ \mathbf{f} \|_{C^r} \geq K > 0$. In this example there is not continuity for $s = 1$ and $t = r$. This completes (ii).

Finally if we consider $r = 1$ and $h(x) = x\varphi(x-1)$ in the above relations then $\| \mathbf{g}_{k+1} \circ \mathbf{f}_n - \mathbf{g}_{k+1} \circ \mathbf{f} \|_{C^{k+1^-}} \geq K > 0$. In this example there is not continuity for $s = k+1$, $t = k + 1^-$.

Remark If F is a finite dimensional space then $\overline{f(U)}$ is a compact subset of V. Every continuous function on V is uniformly continuous on $f(U)$. For this reason under the additional hypothesis that F is finite dimensional, we can supplement the cases in proposition (A.6) with

v) $0 < t = r \leq 1^-$, $s = 1$.

vi) $t = s = k$, $r \geq k$.

vii) $s=k+1$, $t=k+1^-$, $r\geq k+1^-$.

We will denote the corresponding set of exponents of continuity by Ω'.

Proposition A.8 *Let $r,s,t\geq 0$ and $n\in\mathbb{N}$ such that $(r,s-n,t)\in\Omega$. If $g\in C^s(V,G)$ then g_* is n times differentiable.*

Proof: In these conditions

$$(d^n g)_*:\mathcal{U}\longrightarrow C^t(U,\mathcal{L}_k(F,G))$$

is continuous. Moreover if $f_0\in\mathcal{U}$ there are positive constants M,λ,ρ such that

$$\| (d^n g)_*(f_1)-(d^n g)_*(f_2)\|_{C^t}\leq M \ \| g\|_{C^s}\| f_1-f_2\|^{\rho}_{C^r} \tag{23}$$

when $f_1,f_2\in\mathcal{U}_\lambda=\{f\in\mathcal{U}/\ \| f_2-f_1\|_{C^r}<\lambda$. In consequence $d^n(g_*)\in C^\rho(\mathcal{U}_\lambda,C^t(U,\mathcal{L}_k(F,G))$.

We will show that g_* is n times continuously differentiable and

$$\| d^k(g_*)(f_1)-d^k(g_*)(f_2)\|_{C^{r,t}}\leq M' \ \| g\|_{C^s}\| f_1-f_2\|^{\rho}_{C^r} \tag{24}$$

for an adequate constant M'.

Let us consider $n=1$. We will verify that

$$d(g_*):\mathcal{U}\longrightarrow \mathcal{L}(C^r(U,F),C^t(U,G))$$

exists is defined by $d(g_*)(f)[h]=dg\circ f\cdot h$ and is continuous.

For $\| h\|_{C^r}\leq 1$ and ϵ small enough we have

$$\| \frac{g_*(f+\epsilon h)-g_*(f)-\epsilon dg\circ f\cdot h}{\epsilon}\|_{C^t}=$$

$$\| \frac{g\circ(f+\epsilon h)-g\circ f-\epsilon dg\circ f\cdot h}{\epsilon}\|_{C^t}=\| \int_0^1[dg\circ(f+\epsilon sh)-dg\circ f]\cdot hds\ \|_{C^t}\leq$$

$$\int_0^1 2\ \| (dg)_*(f+\epsilon sh)-(dg)_*(f)\|_{C^t}\| h\|_{C^t}\ ds\leq 2M\ \| g\|_{C^s}\ \epsilon^\rho$$

Moreover

$$\begin{array}{rlcl} l: & C^t(U,\mathcal{L}(F,G)) & \longrightarrow & \mathcal{L}(C^r(U,G),C^t(U,G)) \\ & u & \longrightarrow & l(u)[\eta]=[x\to u(x)\cdot\eta(x)] \end{array}$$

is a continuous linear map and $d(g_*)=l\circ(dg)_*$. From proposition (A.3) and proposition(A.4) we deduce that $d(g_*)\in C^\rho(\mathcal{U}_\lambda,\mathcal{L}(C^r(U,G),C^t(U,G))$ and obtain the inequality (24). Now the induction carries on immediately. Again proposition (A.8) contains the best conditions on differentiability.

Example A.5 *If* $(r, s-1, t) \notin \Omega$ *then there is* $g \in C^s((-2,2), \mathbb{R})$ *such that* $g_* : C^r((-1,1), \mathbb{R}) \to C^t((-1,1), \mathbb{R})$ *is not differentiable.*

We only deal with the cases (ii) and (iii) of lemma A.6.

Let us consider $g = x^{n+1}|x|^\alpha$ with $n \in \mathbb{N}$, $0 < \alpha < 1$. If $n \geq 1$, the map $g_* : C^{n+1}((-1,1), \mathbb{R}) \to C^{n+\alpha}((-1,1), \mathbb{R})$ is continuous but not differentiable. In effect taking $f = Id$ and $h = 1$, then for every $x \in (-1,1)$

$$\lim_{\epsilon \to 0} \frac{g_*(f + \epsilon h)(x) - g_*(f)(x)}{\epsilon} = \lim_{\epsilon \to 0} \frac{g(x+\epsilon) - g(x)}{\epsilon} = g'(x) \qquad (25)$$

However

$$\frac{1}{\epsilon}[g^{(n)}(x+\epsilon) - g^{(n)}(x) - \epsilon g^{(n+1)}(x)] \,|_{x=\frac{-\epsilon}{2}}^{0} =$$
$$(n+1+\alpha)(n+\alpha)\cdots(2+\alpha)[1 + \frac{\alpha}{2^\alpha}]\epsilon^\alpha$$

This shows that the convergence fails in $C^{n+\alpha}$ - norm.

On the other hand if $n = 0$, $0 < \beta < 1$ and we take $r = \beta$, $t = \alpha\beta$ then $g_* : C^r((-1,1), \mathbb{R}) \to C^t((-1,1), \mathbb{R})$ is not differentiable . Choosing $f = |x|^\beta, h = 1$ we find for every $x \in (-1,1)$

$$\lim_{\epsilon \to 0} \frac{g_*(f + \epsilon h)(x) - g_*(f)(x)}{\epsilon} = \lim_{\epsilon \to 0} \frac{g(|x|^r + \epsilon) - g(|x|^r)}{\epsilon} = g'(|x|^r) \qquad (26)$$

However

$$\frac{1}{-\epsilon}[g(|x|^r - \epsilon) - g(|x|^r) + \epsilon g'(|x|^r)]|_0^{x^r = \epsilon/2} = -[1 + \frac{\alpha}{2^\alpha}]\epsilon^\alpha$$

this shows that the convergence fails in C^t - norm.

Now we treat the differentiability of the composition operator in the space product. We transfer its analysis to the partial derivatives

Lemma A.5 *Let* $r, s, t \geq 0$ *and* $k \in \mathbb{N}$ *such that* $(r, s-k, t) \in \Omega$. *Then*

i) The map

$$\begin{array}{rcl} R : \mathcal{U} \subset C^r(U, F) & \longrightarrow & L(C^s(V, G), C^t(U, G)) \\ f & \longrightarrow & f^* \end{array}$$

is k times continuously differentiable.

Proof: i) We assume $k=0$. Let $f_1, f_2 \in \mathcal{U}$ with $\| f_1 - f_2 \|_r \leq \lambda$ small enough in order that inequality (23) applies, then

$$\| R(f_1) - R(f_2) \|_{C^{s,r}} = \sup_{\|g\|_{C^s} \leq 1} \| g \circ f_1 - g \circ f_2 \|_{C^t} \leq M \| f_1 - f_2 \|_{C^r}^{\rho}$$

When $k \geq 1$ we complete the proof by induction. We will show that R is continuously differentiable and

$$\begin{array}{rcl} dR : \mathcal{U} \subset C^r(U,F) & \longrightarrow & L(C^r(U,F), L(C^s(V,G), C^t(U,G))) \\ f & \longrightarrow & dR(f)(h) = [g \to dg \circ f \cdot h] \end{array}$$

Indeed if $\| h \|_{C^r} \leq 1$ and ϵ is a small parameter then

$$\| \frac{R(f+\epsilon h) - R(f) - \epsilon dR(f)(h)}{\epsilon} \|_{C^{s,t}} =$$

$$\sup_{\|g\|_{C^s} \leq 1} \frac{1}{\epsilon} \| g \circ (f + \epsilon h) - g \circ f - \epsilon dg \circ f \cdot h \|_{C^t} \leq M \epsilon^{\rho}$$

We consider

$$\begin{array}{rcl} R_1 : \mathcal{U} \subset C^r(U,F) & \longrightarrow & L(C^{s-1}(V, L(F,G)), C^t(U, L(F,G)) \\ f & \longrightarrow & f^* \end{array}$$

$$\begin{array}{rcl} l : L(C^{s-1}(V, L(F,G)), C^t(U, L(F,G)) & \longrightarrow & L(C^r(U,F), L(C^s(V,G), C^t(U,G))) \\ u & \longrightarrow & l(u)(h) = [g \to u(dg) \cdot h] \end{array}$$

We know that R_1 is $k-1$ times differentiable, l is a continuous linear map and $dR = l \circ R_1$. As a consequence of lemma [A.3] we otain that R is k times continuously differentiable.
We summarize our conclusions

Theorem A.1 *Let r, s, t be positive numbers and $K \in \mathbb{N}$ such that $(r, s-k, t) \in \Omega$. The map*

$$\begin{array}{rcl} comp: \ \mathcal{U} \subset C^r(U,F)] \times C^s(V;G) & \longrightarrow & C^t(U,G) \\ (f,g) & \longrightarrow & g \circ f \end{array}$$

is continuous and k times continuously differentiable.

Proof: The continuity is a consequence of lemma A.5 and lemma A.6. Let $k \geq 1$ and consider

$$\begin{array}{rcl} L : [\mathcal{U} \subset C^r(U,F)] \times C^s(V,G) & \longrightarrow & L(C^r(U,F), C^t(U,G)) \\ (f,g) & \longrightarrow & L(f,g) = [h \to dg \circ f \cdot h] \end{array}$$

Note that $d$comp is the sum of L and the map R defined above. Moreover L is the compositions of the maps

$$(f,g) \longrightarrow (f,dg) \longrightarrow dg \circ f \longrightarrow [h \to dg \circ f \cdot h]$$

The first and the third maps are linear. As a consequence of the induction processes we deduce that the intermediate map is $k-1$ times continuously differentiable. Thus we obtain that the composition operator is k times continuously differentiable.

References

[1] J. Eels *A setting for global analysis* *Bull. A.M.S.* **72** 751–807 (1966)

[2] F. Bien *Representations of* $\mathrm{Diff}(S^1)$ *Preprint*

[3] R. Hamilton *The inverse function theorem of Nash and Moser* *Bull. A.M.S.* **7** 65-222 (1982)

[4] M. Hirsch *Differential Topology* Springer Verlag (1976)

[5] M.C. Irwin *On the smoothness of the composition map* *Quart. Jour. Math.* **23** 113–133 (1972)

[6] J. Langer, D.A. Singer *Diffeomorphisms of the circle and geodesic fields on Riemann surfaces of genus one* Inv. Mat. **69** 229-242 (1982)

[7] R. de la Llave *Remarks on J. Langer and D.A. Singer decomposition theorem for diffeomorphisms of the circle* Comm. Math. *Phys.* **104** 387-401 (1986)

[8] J. Moser *A rapidley convergent iteration method and non-linear differential equations II* *Ann. Sc. Norm. Pisa* **20** 499-535 (1966)

[9] S. M. Nikolskii *Approximation of functions of several variables and imbedding theorems* Springer Verlag, New York (1971)

[10] H. Rüssmann *On optimal estimates for the solutions of linear difference equatiopns on the circle* *Cel. Mech.* **14** 33–37 (1976)

[11] E. Stein *Singular integrals and differentiability properties of functions* Princeton University Press (1970)

[12] S. Krantz *Lipschitz spaces, smoothness of functions and approximation theory* *Expo. Mat.* **3** 193–260 (1983)

[13] E. Zehnder *Generalized implicit function theorems with applications to some small divisor problems I* Comm. Pure and Appl. *Math.* **28** 91–140 (1975)

[14] E. Zehnder *Generalized implicit function theorems with applications to some small divisor problems II* Comm. Pure and *Appl. Math.* **29** 49–111 (1976)