REGULARITY OF THE COMPOSITION OPERATOR IN SPACES OF HÖLDER FUNCTIONS.

R. DE LA LLAVE [1] AND R. OBAYA [2]

ABSTRACT. We study the regularity of the composition operator $((f, g) \mapsto g \circ f)$ in spaces of Hölder differentiable functions. Depending on the smooth norms used to topologize f, g and their composition, the operator has different differentiability properties. We give complete and sharp results for the classical Hölder spaces of functions defined on geometrically well behaved open sets in Banach spaces. We also provide examples that show that the regularity conclusions are sharp and also that if the geometric conditions fail, even in finite dimensions, many elements of the theory of functions (smoothing, interpolation, extensions) can have somewhat unexpected properties.

1. INTRODUCTION.

The goal of this paper is to establish differentiability properties of composition in the classical spaces of Hölder functions.

The operator that to a pair of functions (f, g) associates $h = g \circ f$, (where $\circ$ denotes composition) can be considered as defined in function spaces equipped with different norms. Depending on the Banach spaces where we think f, g and h to be ranging, the composition operator may have different degrees of differentiability.

In this paper we consider the classical spaces of Hölder functions defined on open sets in a Banach space satisfying some geometric conditions and we give a complete characterization of the regularity of composition in all the spaces of this scale; that is, we provide counterexamples for all the regularities that we do not establish. We also give examples that show that if the geometric properties in our assumptions do not hold, there are different pathologies that may arise and the classical Hölder spaces may fail to have smoothing operators – that approximate functions in a controlled way by smoother ones –, satisfy interpolation inequalities or have consistent ways of extending them to a larger domain.

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 satisfying some geometric properties. The methods and the results we present also carry out to manifolds and it is easy to lift them to this case. In finite dimension, it is sometimes possible to obtain similar results using characterizations of Hölder spaces by approximation properties by analytic functions [36], [16], [31] but this does not seem to generalize easily to Banach spaces.

[1]Department of Mathematics, University of Texas at Austin, Austin TX 78712, USA

[2]Departamento de Matemática Aplicada a la Ingeniería, ETSII, Universidad de Valladolid, Valladolid, Spain.

As intermediate steps we will also prove elementary versions of interpolation inequalities that work in infinite dimensional domains with good geometric properties.

There are several motivations for the study we undertake in this paper. It seemed to us that often the papers in the literature developed the results they needed from scratch, and hence there was great deal of repetition while, at the same time, there was not a systematic survey; indeed, the precise boundaries were not well known and several important phenomena were not even appreciated at all.

Let us mention just a few applications, which we have been drawn to at some time or another. The list does not intend to be exhaustive or to reflect any other criterion than the history of the authors of this paper and we would welcome any information about new uses.

A first motivation is the study of diffeomorphism groups when the group operation is composition. These groups, besides being the natural symmetry groups of differentiable topology, have been found to have an increasing number of applications. A classical one is a formulation of hydrodynamics and elasticity as geometry in the group of diffeomorphisms (see [7], [3] and references there to previous work). Other applications to mathematics include the study of classifying spaces of foliations and, more modernly, to string theory. An application that has been important for us is the theorem that the commutator subgroups of many groups of diffeomorphisms are simple [10], and that many diffeomorphism groups are their own commutator [27]. One of our original motivations was the study of factorizations and decomposition theorems in spheres [25], which in turn was motivated by inverse problems in optics and in geodesic fields [21].

Unfortunately, as it will become apparent in this paper, the composition operation is not very regular, so the standard methods of Lie theory, even though they are a useful strategic guide, often have to be supplemented by hard analysis methods (notably implicit functions theorems of Nash-Moser type) that involve scales of spaces; hence, one needs to use the properties of composition in scales of spaces. With that in mind, we have also included some material on the existence of smoothing operators that are an indispensable tool. As it turns out, those are closely related to the interpolation inequalities that we need to use often.

Another motivation is dynamical systems. In this field, we often have to consider repeated compositions. Objects that remain invariant under the dynamics – and hence provide useful landmarks – are often found solving functional equations that involve composition. Having available sharp regularity results for the composition allows us to obtain in a clean way smoothness results on invariant manifolds. Indeed, it seems that this was one of the original motivations of [15]. Similarly, the proof of [29] and the appendix by J. Mather to [35] show that problems such as structural stability can be formulated as implicit function theorems whose unknowns are homeomorphisms. Some recent results such as the differentiability of the homeomorphisms appearing in structural stability with respect to the maps involved [24] or the differentiability of entropy [18] can be rewritten as functional equations involving the composition operator. Once one obtains differentiability of the conjugacy, it is possible to study the differentiability with respect to parameters of SRB measures, a problem considered recently by D. Ruelle [34]. With the results developed here, this seems to lead to some improvements in the regularities stated in those papers.

We also mention the recent work of [32] on triviality cocycles over an Anosov system taking values in groups of diffeomorphism of the n-torus $\mathbb{T}^n$. As it was discovered in the mentioned paper, one can use their results to prove rigidity of some non-hyperbolic actions. By using systematically the point of view of diffeomorphisms groups, it is possible to improve the regularity and generalize the results to other manifolds [23].

Let us also point out that there are other natural spaces that, presumably, deserve study similar to that presented here. For example, spaces of diffeomorphisms with a Sobolev topology. Some beginnings of this are done in the literature. For some applications of groups of diffeomorphisms with Sobolev topologies and some of the regularity properties of composition, we refer to [7]. Similar remarks apply to the Λ_r spaces (which agree with the C^r spaces considered here for $r \notin \mathbb{N}$). At the end, we also compare with other theories of functions, such as Whitney theory, which allows us to define functions on closed sets. Of course, for functions defined on closed sets with possibly empty interior, composition is delicate to define. We call attention to the paper [38], which considers two-dimensional hydrodynamics with initial data and borderline smoothness. The main technical tool there are very accurate estimates for composition in Besov spaces.

Let us finish this introduction by paying homage to previous surveys in the field that, even if restricted to integer regularities, have been influential and useful [15], [8] and also to the program sketched in [33]. A recent survey, including other useful spaces in partial differential equations is [2].

2. Basic definitions and notation.

Let E, F be Banach spaces. We will denote the Banach space of bounded linear functions from E to F by $L(E, F)$ and, in general, $L_n(E, F) = L(E, L_{n-1}(E, F))$. For each $x \in E$ and $r > 0$ we represent by $B_r^E(x) = \{y \in E | \, \|y - x\| < r\}$ the ball of center x and radius r, and briefly $B_r^E = B_r^E(0)$. Besides, we omit the vector space in the notation when it does not lead to confusion.

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

$$\lim_{\varepsilon \to 0} \frac{\|f(x + \varepsilon h) - f(x) - \varepsilon df(x)h\|}{\varepsilon} = 0 \tag{2.1}$$

for every $h \in E$ and uniformly with respect to $h \in B_1(0)$. We say that f is *differentiable in* U if f is differentiable at every point $x \in U$. We say that f is *of class* C^1 if it is differentiable and the mapping $df : U \to L(E, F)$, $x \to df(x)$ is continuous for the topologies of the norm.

Proceeding by induction, we define $d^n f$ to be the differential of $d^{n-1} f$, ($d^0 f = f$, $d^1 f = df$) and we say that a function f is *n times continuously differentiable* if $df : U \to L(E, F)$ is $(n - 1)$ times continuously differentiable.

We will denote by $L_n(E_1, \ldots, E_n, F)$ the space of bounded n-multilinear functions having arguments in $E_1, \ldots, E_n$ and values in F. It is standard to identify $L_n(E, F)$ with the space of bounded multilinear functions with n arguments in E taking values in F. We will think of derivatives as multilinear functions.

Definition 2.1. We say that $f \in C^n(U, F)$ if it is n times continuously differentiable and $\|d^n f(x)\|$ is uniformly bounded for all $x \in U$. In that case, we set

$$\|f\|_{C^n} = \sup_{0 \le k \le n} \sup_x \left\| d^k f(x) \right\| . \tag{2.2}$$

We say that $f \in C^n_u(U,F)$ if $f \in C^n(U,F)$ and the multilinear maps $d^j f : U \to L_j(E,F)$, $j = 1, \dots, n$ are uniformly continuous.

We recall that $C^n(U,F)$ and $C^n_u(U,F)$ endowed with the above norm are Banach spaces. We omit the elementary and well known proof. Notice that, according to our definition, there could be functions which are n times continuously differentiable but not in $C^n(U,F)$. We say that $f \in C^n(\overline{U},F)$ if $f \in C^n(U,F)$ and there exist functions $\overline{f}_j \in C(\overline{U}, L_j(E,F))$ for $j = 0, 1, \dots, n$ such that $\overline{f}_j = d^j f$ on U. It is easy to verify $C^n_u(U,F) \subset C^n(\overline{U},F)$.

In Banach spaces, the given notion of Fréchet differentiability coincides with the original one presented in [9]. Other weaker notions of differentiability have been also proposed and frequently used. Similarly, one could have considered other topologies for the set of the successive derivatives but they will not be taken into account here. Fréchet differentiability in locally convex spaces requires the existence of the directional derivatives (2.1) uniformly on the bounded sets. In this case, the corresponding space $L(E,F)$ is endowed with the topology of the uniform convergence on bounded sets. A review of this question can be found in [4].

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 $g \circ f \in C^n(U,G)$. Similarly, if $f \in C^n_u(U,F)$ and $g \in C^n_u(V,G)$ then also $g \circ f \in C^n_u(U,G)$. By a repeated application of the chain rule, we can find constants $c_{k,j_1,\dots,j_k}$ such that

$$d^n(g \circ f)(x) = \sum_{k=1}^{n} \sum_{j_1+\cdots+j_k=n} c_{k,j_1,\dots,j_k} d^k g(f(x))\{d^{j_1} f(x), \dots, d^{j_k} f(x)\}\,. \tag{2.3}$$

The combinatorial coefficients $c_{k,j_1,\dots,j_k}$ can be computed explicitly. Then, the formula (2.3) is called Faa-di-Bruno formula (see e.g. [1] p. 3 1.4).

The question of whether non-trivial differentiable functions exist and whether functions of a certain class can be approximated by smother ones is somewhat subtle in infinite dimensional spaces. The two questions are related since a natural method for approximating continuous functions by smoother ones is the use of smooth cut-off functions. (These are functions that take the value one in one ball and vanish outside of another larger ball).

The existence of such functions is non-trivial in infinite dimensional Banach spaces and, hence, some of the constructions that involve using them (linearizations, invariant manifolds, etc.) need the assumption that they exist in the space we are considering.

For example, it was shown in [17] that the space of continuous functions in an interval does not admit a C^1 cut-off function. More generally, in [22] one can find the result that a separable Banach space admits a cut-off function if and only if its dual is separable.

It was shown by Moulis in [30] that if $U \subset E$ is an open set of a separable Hilbert space and F is a Banach space, then every $2(k-1)$ times differentiable function can be approximated by infinitely differentiable functions in the C^k-fine topology on U. Besides, every k times differentiable function can be approximated by infinitely differentiable functions on a dense subset of U (see [13]). Many other results on approximation of functions on Banach spaces can be found in [26], [14].

Given Banach spaces E and F, it is said that E is *finitely represented* in F ($E \ll F$ for short), if for each finite dimensional subspace E' of E and positive number ϵ, there is a subspace F' of F and an isomorphism l from E' to F' such that $\|l\| \|l^{-1}\| \leq 1+\epsilon$. A Banach space E is said to be *superreflexive* if $F \ll E$ implies that F is reflexive. It can be found in [37] that E is superreflexive if and only if for every open subset $U \subset F$ there exists a uniformly continuously differentiable function $\varphi : F \to \mathbb{R}$ with $U = \{x \in E |\ \varphi(x) \neq 0\}$. This assures the existence on E of sufficiently many non-trivial differentiable functions. This concept of superreflexive will play a role later.

In finite dimensions, in compact manifolds there are many other tools that can be used in approximating functions by smoother ones (Fourier series, convolution with smoothing kernels etc.) and it is possible to obtain rather detailed quantitative results. Since these results are needed in many applications – the analysis involving composition operators often requieres Nash Moser Techniques – we have summarized some of them in paragraph 3.3.

Remark 2.2. Many references use induction in regularity by means of the so-called tangent functor trick. Given $f : U \subset X \to V \subset Y$ we define $Tf : U \times X \to V \times Y$ by $Tf(x, v) = (f(x), df(x)v)$. It is easy to see that Tf is n times differentiable if and only if f is $n+1$ times differentiable. Moreover, the chain rule reads $Tf \circ Tg = T(f \circ g)$. Therefore, statements that are true for maps in sets of the form $U \times V$ can use induction in the regularity. Unfortunately for our purposes, the tangent trick gives conclusions that can be improved. Notice that the first component of Tf is more differentiable than the second. This information is lost if we consider only the regularity class of both components jointly. Notice also that in the chain rule for $d(f \circ g)$, the factor $(dg) \circ f$ involves composition of the least differentiable component of Tg with the most differentiable component of Tf. We do not get any factor in which the least differentiable component of Tg is composed with the least differentiable component of Tf. Hence, the estimates based only on the regularity of Tf and Tg are not optimal.

Now we turn to the task of defining Hölder regularities. The appropriate setting of Hölder regularities is that of metric spaces.

Unfortunately, given a connected open set of a Banach space, there are two natural distances. One is the distance inherited from the metric in the Banach space and the other is the distance given by the minimum length of arcs joining the two points and contained in the set. Both have geometrical significance but can be very different even in simple examples.

Example 2.3. Consider the set

$$U = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 4\} - \{(x, 0) |\ 0 \leq x \leq 1\}.$$

Then, defining $p_n = (1/2, 1/n)$, $q_n = (1/2, -1/n)$ we have that $d_{\mathbb{R}^2}(p_n, q_n)$ tends to 0. Nevertheless, the minimum length of an arc joining these two points – which we will henceforth denote by $d_U(p_n, q_n)$ – converges to 1.

Example 2.4. Consider the set

$$U = \{(x, y) \in \mathbb{R}^2 \mid |y - x \sin(x)| \leq 1/100\},$$

and the points $p_n = (2\pi n, 0) \in U$. It is obvious that $d_{\mathbb{R}^2}(p_n, p_{n+1}) = 2\pi$ but $d_U(p_n, p_{n+1}) \geq 2n$ for every n.

Note that these examples also show that the distance d_U between two points depends in a very strong way on the sets we consider them to belong. Moreover, if we restrict the set, the distance may grow and become singular (it is true that if $U \subset U', d_U(p,q) \geq d_{U'}(p,q))$. As we will see later, this complicates the function theory based on d_U. On the other hand, if we are going to apply the mean value theorem to obtain estimates on the derivatives, the only reasonable option is to use the distance based on the arc.

The compromise that we will adopt in this paper is to use the induced distance in the definition of Hölder regularity, but we restrict consideration to domains in which the two distances are equivalent. This choice allows us enough flexibility for the applications we have in mind. We will later make some remarks throughout the exposition about how to extend the results here to more general domains.

We will adopt the standard convention that 1^- is a number strictly smaller than 1 but strictly bigger than any number smaller than 1 and that it works in powers as 1. We will also adhere to the convention of using Greek letters to denote symbols in the range $(0, 1^-]$, n, m, k for integers, and letters such r, s, t for "sums" of symbols as above and integers. The main motivation for such an additive notation comes from the interpolation inequalities that we will discuss below.

The Hölder spaces we will be concerned with in this paper are:

Definition 2.5. If $r = n + \alpha$ with $n \in \mathbb{N}$ and $0 < \alpha \leq 1^-$, we define $C^r(U, F)$ as the space of $f \in C^n(U, F)$ satisfying

$$H_\alpha(d^n f) \equiv \sup_{x \neq y} \frac{\|d^n f(x) - d^n f(y)\|}{\|x - y\|^\alpha} < \infty \tag{2.4}$$

endowed with the norm

$$\|f\|_{C^r} = \sup(\|f\|_{C^n}, H_\alpha(d^n f)) \,. \tag{2.5}$$

We recall that $C^r(U, F)$ are Banach spaces. We leave the well known proof to the reader.

From the definitions (2.2) and (2.5), we obtain

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

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

If $n \in \mathbb{N}$ and $0 < \alpha \leq 1^-$, it is obvious that $C^{n+\alpha}(U, F) \subset C^n_u(U, F) \subset C^n(U, F)$. Every limit in the C^n-topology of functions taken from $C^{n+\alpha}(U, F)$ belongs to $C^n_u(U, F)$. In order to complete the notation for non integer exponents, we set $C^{n+\alpha}_u(U, F) = C^{n+\alpha}(U, F) \cap C^n_u(U, F)$.

Observe that Hölder regularity requires global estimates on the domain and hence admits a weaker local version. We say that $f \in C^r_{\mathrm{loc}}(U, F)$ if for every $x \in U$ there exits an open neighborhood $U(x)$ such that $f \in C^r(U(x), F)$.

Take $f \in C^{n+\alpha}_{\mathrm{loc}}(U, F)$ and $x \in U$. The function f admits a local Taylor expansion

$$f(x + h) = f(x) + \sum_{j=1}^{n} d^j f(x)\{h, \overset{(j)}{\ldots}, h\} + R_n(x, h)$$

with remainder term $\|R_n(x, h)\| \leq (1/n!) \|f\|_{C^{n+\alpha}} \|h\|^{n+\alpha}$.

We also recall that, when U is finite dimensional, it is natural to consider the so-called Λ_r spaces for many questions on approximation theory. The space $\Lambda_r(U)$ agrees with $C^r(U, \mathbb{R})$ defined above when $r \notin \mathbb{N}$. We say that $f \in \Lambda_k(U)$ if $f \in$

$C^{k-1}(U,\mathbb{R})$ and

$$\|f\|_{\Lambda_k} = \|f\|_{C^{k-1}} + \sup_{x-h,x,x+h\in U} \frac{\left\|d^k f(x+h) + d^k f(x-h) - 2d^k f(x)\right\|}{\|h\|} < \infty\,.$$

The Λ_r-spaces endowed with the above Zygmund norms are Banach spaces. We will not discuss them in these notes and we refer to [16], [36] and [39] for a detailed description of their properties. If the domain U has sufficiently regular boundary then $\|f\|_{\Lambda_r(U)}$ is equivalent with the norm

$$\inf\{\ \left\|\overline{f}\right\|_{\Lambda_r(\mathbb{R}^m)}\ \mid \overline{f}:\mathbb{R}^m\to\mathbb{R}\ \text{ and }\ \overline{f}|_U = f\ \}.$$

In that case only the study of $\Lambda_r(\mathbb{R}^m)$ is required, their elements can be obtained as convenient limits of infinitely differentiable functions. For a U with an irregular boundary, these assertions could fail.

3. Geometry of domains, elementary interpolation properties and preliminary bounds.

In this section we develop some elementary tools that allow us to control the norm of a function in the scale of Hölder spaces. They are associated with the celebrated interpolation inequalities. The simplest ones are valid in general domains and relate regularity classes with the same integer part. General interpolation inequalities require geometric conditions in the domain of definition of the functions and we construct examples of domains that violate them and for which these inequalities are false. These examples show that the existence of smoothing and extension operators on spaces of differentiable functions is affected by the geometric properties of the domain. The material reviewed here will result basic for future discussions of the composition operator.

3.1. **Elementary interpolation inequalities.** The following elementary statement, valid for general domains, provides the relation between Hölder spaces for exponents with the same integer part.

Proposition 3.1. *Let E, F be Banach spaces and $U\subset E$ an open subset. Consider $r=n+\alpha$, $s=n+\beta$, $t=n+\gamma$ with $n\in\mathbb{N}$, $0\le\alpha<\beta<\gamma\le 1^-$, and set $\mu=(t-s)/(t-r)$. Then $C^t(U,F)\subset C^s(U,F)$ and for every $f\in C^t(U,F)$, one has*

$$\|f\|_{C^s} \le M_\alpha\,\|f\|_{C^r}^{\mu}\,\|f\|_{C^t}^{1-\mu} \tag{3.1}$$

where the constant $M_\alpha=2$ if $\alpha=0$ and $M_\alpha=1$ if $\alpha>0$.

Proof. First, we will study both cases for $n=0$ and then we will establish the claim using induction on n. If $r=\alpha=0$ then $\beta=\gamma(1-\mu)$ and we have

$$\begin{aligned} H_\beta(f) &= \sup\nolimits_{x\neq y} \|f(x)-f(y)\|^{\mu}\,\frac{\|f(x)-f(y)\|^{1-\mu}}{\|x-y\|^{\gamma(1-\mu)}} \\ &\le 2^{\mu}\,\|f\|_{C^0}^{\mu}\,H_\gamma(f)^{1-\mu}. \end{aligned}$$

Hence $C^t(U,F)\subset C^s(U,F)$ and $\|f\|_{C^s}\le 2\,\|f\|_{C^r}^{\mu}\,\|f\|_{C^t}^{1-\mu}$ for every $f\in C^t(U,F)$. Besides, if $r=\alpha>0$ then

$$\begin{aligned} H_\beta(f) &= \sup\nolimits_{x\neq y} \frac{\|f(x)-f(y)\|^{\mu}}{\|x-y\|^{\alpha\mu}}\,\frac{\|f(x)-f(y)\|^{1-\mu}}{\|x-y\|^{\gamma(1-\mu)}} \\ &\le H_\alpha(f)^{\mu} H_\gamma(f)^{1-\mu}, \end{aligned}$$

which provides

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

for $n = 0$.

Using the inductive hypothesis we find, for $n \ge 1$,

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

and, therefore, the desired conclusion. □

In contrast with Proposition 3.1, interpolation inequalities for different n are more complicated and only hold under additional hypotheses on the domain.

Example 3.2. Let $U = \{(x,y) \in \mathbb{R}^2 \mid x > 0, y > 1, xy < 1\}$. We take φ as a C^∞ cut-off function with $\|\varphi'\|_{C^0} \le 2$ which takes the value 1 on $[-1/3, 1/3]$ and zero outside of $[-1,1]$. We define on U the functions

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

It is immediate to check that

$$\frac{\partial^j f_n}{\partial y^j}(x,y) = \log(1+x)\,\varphi^{j)}(y-n) \quad \text{for every } j \ge 0,$$

$$\frac{\partial^{i+j} f_n}{\partial x^i \partial y^j}(x,y) = \frac{(-1)^{i+1}(i-1)!}{(1+x)^i}\,\varphi^{j)}(y-n) \quad \text{for every } i \ge 1, j \ge 0.$$

Observe that the support of $\varphi(y-n)$ is centered around $y = n$, which means that $x \approx 1/n$. In consequence $\|f_n\|_{C^2}$ remains bounded and $\|f_n\|_{C^0}$ converges to zero. Since $\partial f_n/\partial x$ evaluated at $(x,y) = (1/2n, n)$ does not converge to zero, we conclude that there cannot be an inequality of the form

$$\|f_n\|_{C^1} \le M\,\|f_n\|_{C^0}^{1/2}\,\|f_n\|_{C^2}^{1/2}\,.$$

More generally, notice that there is not interpolation inequalities for the range of values $0 \le r < 1^- \le s \le t$.

Proposition 3.3. *Let $n \in \mathbb{N}$ and $0 \le \alpha < \beta \le 1^-$. If E, F are finite dimensional Banach spaces and $U \subset E$ is an open bounded set, then the embedding $i : C^{n+\beta}(U,F) \to C^{n+\alpha}(U,F)$ is compact.*

Proof. In the finite dimensional case, here considered, the successive derivatives of each function of $C^{n+\beta}(U,F)$ admit a continuous extension to $\overline{U}$. Thus, the bounded sets of $C^{n+\beta}(U,F)$ are compact in $C^{n+\alpha}(U,F)$. In effect, for $\alpha = 0$ this is a consequence of Ascoli-Arzelà theorem. For $0 < \alpha < \beta \le 1^-$ we use the previous result for $\alpha = 0$ and the known interpolation inequality (3.1). □

We also quote Proposition A3 from [20]. It generalizes the elementary fact that the pointwise limit of uniformly Hölder functions is also Hölder.

Proposition 3.4. *Let $n \in \mathbb{N}$ and $0 < \alpha \le 1^-$. Take a sequence $(f_k)_{k\in\mathbb{N}}$ of functions of $C^{n+\alpha}(U,F)$ with $\|f_k\|_{C^{n+\alpha}} \le M$ and assume that for each $x \in U$, $f_k(x)$ converges weakly to $f(x)$. Then,*

i) *f lies in $C^{n+\alpha}(U,F)$ and $\|f\|_{C^{n+\alpha}} \le M$,*
ii) *$d^j f_k(x)$ converges weakly to $d^j f(x)$ for every $x \in U$ and $1 \le j \le n$.*

We refer to the paper [19] for the proof. The theorem there is stated only for the exponent $\alpha = 1^-$, but the proof goes through without any modification for the statement here.

We call attention to the fact that, in contrast with the usual Ascoli-Arzelà theorem, the space is not necessarily separable but, on the contrary, the weak convergence of $f_k(x)$ is assumed for every $x \in U$. We also note that Proposition 3.4 can be rephrased as stating that the unit ball in $C^{n+\alpha}$ is sequentially closed when we give the functions the topology of the pointwise weak limit.

3.2. **Compensated domains.** Notice that in Proposition 3.1 we have not claimed that if $r = n + \alpha < s = m + \beta$ then $C^s(U, F) \subset C^r(U, F)$. The fact that this is not true in general can be seen in the following example.

Example 3.5. *Let U be the domain in Example 2.3 and d_U as there denote the minimum length of arcs in U joining two points. Consider the function*

$$f(x, y) = \varphi(d_U(\,(x, y),\ (1/2, 1/2))),$$

where φ is a C^∞ function of the real line that equals the identity in $[10^{-2}, \sqrt{2}/2 + 10^{-3}]$, is constant in $[\sqrt{2}/2 + 10^{-2}, \infty)$ and on $(-\infty, 10^{-3}]$ and, moreover, satisfies $|\varphi'(x)| \leq 1$. 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 domain U, $C^1(U) \not\subseteq C^\alpha(U)$ for any $0 < \alpha \leq 1^-$.

Proof. Let us describe somewhat informally why the function f is differentiable. The gradient of $d_U((x, y), (1/2, 1/2))$ (the differential is just the inner product with the gradient) is, roughly, the unit vector in the shortest line that joins the two points going through the point. More precisely, the expression for $d_U((x, y), (1/2, 1/2))$ is a piecewise analytic function that can be expressed as either a) the Euclidean distance of (x, y) to $(1/2, 1/2)$, b) the distance of (x, y) to $(0, 0)$ plus $\sqrt{2}/2$ (which is the distance of $(0, 0)$ to $(1/2, 1/2)$), or c) the distance of (x, y) to $(1, 1)$ plus $\sqrt{2}/2$ (the distance of $(1, 0)$ to $(1/2, 1/2)$). Note that even if the straight lines that go through $(1, 0)$ or $(0, 0)$ are not contained in the domain, they can be approximated by paths with arbitrarily closed length in the domain. The gradients will be, the unit vectors pointed respectively to $(1/2, 1/2)$ in the case of a), to $(0, 0)$ in the case of b), and to $(1, 0)$ in the case of c).

The function d_U will be continuously differentiable in the points where these unit vectors depend continuously on the point. This includes all the points in U except $(1/2, 1/2)$ and the boundary between the regions a), b), c). As it turns out, the boundaries between a) and b) and a) and c) are not a problem: the boundaries are precisely those points in the segment joining $(1/2, 1/2)$ to $(0, 0)$ and to $(1, 0)$. Nevertheless, the gradient of $d_U((x, y), (1/2, 1/2))$ is indeed discontinuous at $(1/2, 1/2)$ and at the boundary between b) and c), which is the segment $\{(1/2, \text{y}) \mid y \leq 0\}$.

The chain rule establishes that, except for these points, the function f is continuously differentiable. On the other hand, our choice of the function φ forces f to be constant on a neighborhood of these points so that the derivative is 0. ∎

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

$$f(x) - f(y) = \int_0^1 df(\gamma(s))\dot{\gamma}(s)ds \tag{3.2}$$

where γ is a differentiable path joining x and y. This equality gives $\|f(x) - f(y)\| \le \|f\|_{C^1} d_U(x,y)$. In the domain U of the example this could be much larger than $\|x - y\|$. Clearly, if we want to use arguments based on the mean value formula (3.2), the alternatives are either to consider always the path distance or to restrict ourselves to spaces in which the distance among points is commensurate to the path distance. In this paper we have chosen the second alternative. Hence, we isolate the property:

Definition 3.6. We say that an open set U is *compensated* if there exists a constant κ_U such that $d_U(x,y) \le \kappa_U \|x - y\|$ for every $x, y \in U$.

This definition only makes sense for connected set, however this is not a serious shortcoming. If $U = U_1 \cup \cdots U_n$ and $U_i \cap U_j = \emptyset$ for $i \ne j$, then $C^r(U) = C^r(U_1) \oplus \cdots \oplus C^r(U_n)$. Similarly for infinite unions of sets that can be decomposed in infinite disjoint unions but the pairs are at a distance bounded from below. This is general enough for the applications we have in mind.

If the domain U is compensated it follows from (3.2) that $C^{n+1^-}(U,F)$ is a closed subset of $C^{n+1}(U,F)$ and $\|f\|_{C^{n+1}} \le \kappa_U \|f\|_{C^{n+1^-}}$ for every $f \in C^{n+1}(U,F)$. More precisely, we can state

Proposition 3.7. *Let U be a compensated open set and take numbers $0 \le r \le s$. Then $C^s(U,F) \subset C^r(U,F)$ and, for every $f \in C^s(U,F)$, one has*

$$\|f\|_{C^r} \le 2\kappa_U \|f\|_{C^s} .$$

Proof. Write $r = n + \alpha$, $s = m + \beta$ where $n, m \in \mathbb{N}$ and $0 \le \alpha, \beta \le 1^-$. Let us fix a function $f \in C^s(U,F)$. If $n = m$ and $\alpha < \beta$ we have verified in Proposition 3.1 that $C^s(U,F) \subset C^r(U,F)$ and

$$\|f\|_{C^r} \le 2 \|f\|_{C^n}^{\alpha/\beta} \|f\|_{C^s}^{1-\alpha/\beta} \le 2 \|f\|_{C^s} .$$

If $n < m$ we obtain $C^s(U,F) \subset C^m(U,F) \subset C^{n+1^-}(U,F) \subset C^r(U,F)$ and

$$\|f\|_{C^r} \le 2 \|f\|_{C^{n+1^-}} \le 2\kappa_U \|f\|_{C^m} \le 2\kappa_U \|f\|_{C^s} ,$$

which is the desired result. □

Notice that, if we consider bounded domains in the finite dimensional case, the embedding $i : C^s(U,F) \to C^r(U,F)$ turns out to be compact for all the exponents $r < s$ with $(r,s) \ne (n+1^-, n+1)$ for any $n \in \mathbb{N}$.

When we study problems involving composition, we have to ensure that the intermediate ranges also have this property. However, C^∞ functions could transform compensated domains into uncompensated ones.

Example 3.8. Let $D = \{z \in \mathbb{C} \mid 0 < |z| < 1, 0 < \operatorname{Arg} z < \pi\}$, the map $z \to z^2$ transforms the compensated open D into the uncompensated domain $U = \{z \in \mathbb{C} \mid 0 < |z| < 1, 0 < \operatorname{Arg} z < 2\pi\}$, 3.3. **Smoothing operators, extensions and interpolation.** The failure of interpolation for certain domains has some consequences for the classical theory of smoothing and extension operators. The results summarized in this section apply to finite dimensional vector spaces and will not be used in the rest of the paper.

Take $E_0 = C(\mathbb{T}^m, \mathbb{R})$, where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ represents the one dimensional torus. For each $\tau > 0$ we consider the complex strip

$$U_\tau = \{x \in \mathbb{C}^m \mid |\Im x_i| \leq \tau,\ 1 \leq i \leq m\}$$

and the Banach space E_τ defined as the set of real holomorphic functions f on U_τ with period 1 in each variable endowed with the norm $\|f\|_\tau = \sup_{x \in U_\tau} |f(x)|$.

We introduce the concept of smoothing operators which is an abstract version of the usual controlled approximation of Hölder functions by smoother ones. The results stated below can be found in [28] and [39].

Theorem 3.9. *There exists an analytic smoothing in the family $(E_\tau)_{\tau \geq 0}$ with respect to $\{\Lambda_r(\mathbb{T}^m)\}_{r \geq 0}$ (and $\{C^r(\mathbb{T}^m, \mathbb{R})\}_{r \geq 0}$), i.e. a collection of linear continuous operators $S_\tau \in L(C(\mathbb{T}^m, \mathbb{R}), E_1)$ together with constants k_r, $0 < r < \infty$ such that for all $r > 0$ and $\tau > 0$ one has*

i) $\lim_{\tau \to \infty} \|(S_\tau - 1)f\|_{C^0} = 0$ *if* $f \in C(\mathbb{T}^m, \mathbb{R})$,
ii) $\|S_\tau f\|_{\tau^{-1}} \leq k_r \|f\|_{\Lambda_r}$ *if* $f \in \Lambda_r(\mathbb{T}^m)$,
iii) $\|(S_\sigma - S_\tau)f\|_{\sigma^{-1}} \leq \tau^{-r} k_r \|f\|_{\Lambda_r}$ *if* $f \in \Lambda_r(\mathbb{T}^m)$, $\sigma \geq \tau > 0$.

In addition, it is interesting to remark that the above quantitative estimates in the approximation by analytic functions characterize the Λ_r-spaces. Observe that relations (ii) and (iii) can be stated for functions $f \in C^r(\mathbb{T}^m, \mathbb{R})$ with their usual norms. The argument of the proof of Theorem 3.9, based on realizing $S\tau$ as a convolution with a smooth kernel that approximates the identity, also works on $\mathbb{R}^m$.

Theorem 3.10. *There exists a C^∞-smoothing in the family $\{C^r(\mathbb{T}^m, \mathbb{R})\}_{r \geq 0}$ i.e. a collection $(S_\tau)_{\tau > 0}$ of linear mappings $S_\tau : C(\mathbb{T}^m, \mathbb{R}) \to C^\infty(\mathbb{T}^m, \mathbb{R})$, together with constants $k_{r,s}$, $0 \leq r, s < \infty$ satisfying*

i) $\lim_{\tau \to \infty} \|(S_\tau - 1)f\|_{C^0} = 0$ *if* $f \in C(\mathbb{T}^m, \mathbb{R})$,
ii) $\|S_\tau f\|_{C^s} \leq \tau^{(s-r)} k_{r,s} \|f\|_{C^r}$ *if* $f \in C^r(\mathbb{T}^m, \mathbb{R})$, $0 \leq r \leq s$,
iii) $\|(S_\tau - 1)f\|_{C^r} \leq \tau^{-(s-r)} k_{r,s} \|f\|_{C^s}$ *if* $f \in C^s(\mathbb{T}^m, \mathbb{R})$, $0 \leq r \leq s$.

More generally, one can find in [39] that smoothing operators exist for the scale of spaces $\{C^{n+r}(K, \mathbb{R})\}_{r \geq 0}$, where $n \in \mathbb{N}$ and K is a compact manifold. The compactness of K becomes essential for the proof, e.g. $C^r(\mathbb{R}, \mathbb{R})$ is not dense in $C(\mathbb{R}, \mathbb{R})$ for any $r > 0$. The connection with interpolation comes from the following statement taken from [39].

Proposition 3.11. *Let us assume that the scale of Banach spaces $\{E_r\}_{r \geq 0}$ admits a family S_τ of smoothing operators satisfying (i)-(iii) of Theorem 3.10. If $0 \leq r \leq t$, $\mu \in (0,1)$ and $x \in E_t$ then, defining $s = \mu r + (1-\mu)t$, we have*

$$\|x\|_s \leq M_{r,s,t} \|x\|_r^\mu \cdot \|x\|_t^{(1-\mu)}\ .$$

Proof. For all $\tau > 0$, we have

$$\|x\|_s \leq \|S_\tau x\|_s + \|S_\tau x - x\|_s \leq \tau^{(s-r)} k_{r,s} \|x\|_r + \tau^{-(t-s)} k_{s,t} \|x\|_s\ .$$

Evaluating at $\tau = (\|x\|_t / \|x\|_r)^{1/(t-r)}$ we obtain the desired conclusion. □

In consequence the existence of C^∞-smoothing operators on spaces of differentiable functions is affected by the geometry of the compact set of definition.

We continue by recalling the definition of Hölder spaces on closed sets that is used in [36] formalization of Whitney's extension theorem in terms of function spaces.

Definition 3.12. Let $X \subset \mathbb{R}^m$ be a closed set and take $n \in \mathbb{N}$, $0 < \alpha \le 1^-$. We say that $f \in C^{n+\alpha}(X, \mathbb{R})$ if there exist functions $f^{(j)}$ for $j \in \mathbb{Z}^m$, $0 \le |j| \le n$, defined on X such that, if

$$f^{(j)}(x) = \sum_{|j+l| \le k} \frac{f^{(j+l)}(y)}{l!}(x-y)^l + R_j(x,y),$$

then

$$|f^{(j)}(x)| \le N \quad \text{and} \quad |R_j(x,y)| \le N|x-y|^{n-|j|+\alpha} \tag{3.3}$$

for every $0 \le |j| \le n$ and $x, y \in X$. Again the norm $\|f\|_{C^{n+\alpha}}$ is taken to be the smallest N satisfying the inequalities (3.3).

If the domain U is badly behaved, it is not true in general that the extension $\overline{f}$ on $\overline{U}$ of a function $f \in C^r(U, \mathbb{R})$ lies in the space $C^r(\overline{U}, \mathbb{R})$ above defined.

An extension operator $E_{X,Y,r}$ for closed sets $X \subset Y$ is a bounded linear operator from $C^r(X, \mathbb{R})$ to $C^r(Y, \mathbb{R})$ such that $E_{X,Y,r}x$ and x agree on X. We say that an operator $E_{X,Y,r}$ is is *consistent* in a set of indices P if, for $s, r \in P$ with $s > r$, $E_{X,Y,r}$ maps $C^s(X, \mathbb{R})$ into $C^s(Y, \mathbb{R})$ and is bounded.

We note that if there is a coherent extension, X inherits the interpolation properties of Y in every subset of P. In effect, with $r < t$, $\mu \in (0,1)$, $s = \mu r + (1-\mu)t$, denoting the extension operator just by E, we have

$$\begin{aligned} \|f\|_{C^s(X,\mathbb{R})} &\le M_{r,s,t}\, \|Ef\|^{\mu}_{C^r(Y,\mathbb{R})}\, \|Ef\|^{(1-\mu)}_{C^t(Y,\mathbb{R})} \\ &\le \|E\|^{\mu}_{C^r(X,\mathbb{R}),C^r(Y,\mathbb{R})}\, \|E\|^{(1-\mu)}_{C^t(X,\mathbb{R}),C^t(Y,\mathbb{R})}\, M_{r,s,t}\, \|f\|^{\mu}_{C^r(X,\mathbb{R})}\, \|f\|^{(1-\mu)}_{C^t(X,\mathbb{R})}\,. \end{aligned}$$

Consider the domain U defined in Example 3.2, it is well-known that the celebrated Whitney extension theorem (see e.g. [36] VI, 2) provides extension operators from $C^r(\overline{U}, \mathbb{R})$ to $C^r(\mathbb{R}^2, \mathbb{R})$ for every $r \in \mathbb{R} - \mathbb{Z}$. We call attention to the fact that, according to the above remarks, it produces consistent extensions only for the sets of regularities with the same integer part.

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 [36] VI 3).

4. The range of compositions.

In this section, we will study the regularity of the composition of two functions expressed in the scale of spaces that we have introduced before.

The following results, which estimate composition with linear and bilinear functions, will be used in the study of the successive derivatives of the composition.

Proposition 4.1. *Let $l \in L(F, G)$ and $t \ge 0$. Then $l_* : C^t(U, F) \to C^t(U, G)$ defined by $l_*(f) = l \circ f$ is a linear continuous map. Moreover, $\|l_*\| = \|l\|$.*

Proof. Take $t = k + \gamma$ with $k \in \mathbb{N}$ and $0 \le \gamma \le 1^-$. When $k = 0$ it is easy to verify that $l \circ f \in C^t(U, G)$ and

$$\|l \circ f\|_{C^t} \le \|l\| \, \|f\|_{C^t} \ .$$

We introduce $l_1 : L(E, F) \to L(E, G), \ h \to l \circ h$. For $k \ge 1$ we find that $d(l \circ f) = (l_1)_*(df)$. Thus the proof can be carried on immediately by induction. $\square$

Proposition 4.2. *Let $b \in L_2(F, G; H)$ and $t \ge 0$. Then $b_* : C^t(U, F) \times C^t(U, G) \to C^t(U, H)$ defined by $b_*(f, g)(x) = b(f(x), g(x))$ is a bilinear continuous map. Moreover, $\|b_*\| \le 2 \|b\|$.*

Proof. Take $t = k + \gamma$ with $k \in \mathbb{N}$ and $0 \le \gamma \le 1^-$. When $k = 0$ one has

$$\|b_*(f, g)\|_{C^t} \le 2 \|b\| \, \|f\|_{C^t} \|g\|_{C^t} \ .$$

We introduce $b_1 : L(E, F) \times G \to L(E, H), \ (h_1, z) \to [x \to b(h_1(x), z)]$ and $b_2 : F \times L(E, G) \to L(E, H), \ (y, h_2) \to [x \to b(y, h_2(x))]$. For $k \ge 1$ we obtain $d(b_*(f, g)) = (b_1)_*(df, g) + (b_2)_*(f, dg)$, which allows us to complete the proof. $\square$

We will use the above propositions in the special case where $b : L(F, G) \times L(G, H) \to L(F, H), \ (h_1, h_2) \to h_2 \circ h_1$ is defined by composition of linear maps, which provides

$$\begin{array}{rcl} b_* : C^t(U, L(F, G)) \times C^t(U, L(G, H)) & \longrightarrow & C^t(U, L(F, H)) \\ (f, g) & \longrightarrow & [x \to g(x) \circ f(x)] \,. \end{array}$$

If $F = K$ then $L(F, G) \equiv G$, $L(F, H) \equiv H$ and b becomes the evaluation map. In consequence,

$$\begin{array}{rcl} b_* : C^t(U, G) \times C^t(U, L(G, H)) & \longrightarrow & C^t(U, H) \\ (f, g) & \longrightarrow & [x \to g(x)(f(x))] \,. \end{array}$$

To simplify the notation in the above cases, we will denote the map $b_*(f, g)$ by $g \cdot f$.

The following result that characterizes the regularity of compositions will be the main result of this section.

Theorem 4.3. *Let E, F, G be Banach spaces and $U \subset E$, $V \subset F$ be compensated subsets. Let $r = n + \alpha$, $s = m + \beta$ with $m, n \in \mathbb{N}$ and $0 \le \alpha, \beta \le 1^-$. If $f \in C^r(U, F)$, $f(U) \subset V$ and $g \in C^s(V, G)$, then $g \circ f$ belongs to $C^t(U, G)$ where we take t as follows:*

i) *If $n = m = 0$, $t = rs$. In that case, we have*

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

ii) *If $\max(r, s) \ge 1$, $t = \min(r, s)$. In that case, we have*

ii.1) *if $r \ge 1$, $0 < s \le 1^-$,*

$$\|g \circ f\|_{C^s} \le \|g\|_{C^s} \|f\|_{C^1}^s \kappa_U^s + \|g\|_{C^0} \ ,$$

ii.2) *if $0 < r \le 1^-$, $s \ge 1$,*

$$\|g \circ f\|_{C^r} \le \kappa_V \|g\|_{C^1} \|f\|_{C^r} + \|g\|_{C^0} \ ,$$

ii.3) *if $r \ge 1$, $s \ge 1$, $\exists \, M_t \ge 1$ such that*

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

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

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

This establishes i) and i.1). In this case no geometrical conditions on the domain are required. The proof of 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 $k \in \mathbb{N}$, $k \ge 1$ and $0 \le \gamma \le 1^-$. We proceed by induction on k. For $k = 1$, $d(g \circ f) = dg \circ f \cdot df$. Applying Proposition 4.2 and the above assertion ii.1) we obtain

$$\begin{aligned}\|d(g \circ f)\|_{C^\gamma} &\le 2(\|dg\|_{C^\gamma} \|f\|_{C^1}^\gamma \kappa_U^\gamma + \|dg\|_{C^0}) \|df\|_{C^\gamma} \\ &\le 2 \|g\|_{C^{1+\gamma}} (1 + \kappa_U^\gamma \|f\|_{C^{1+\gamma}}^\gamma) \|f\|_{C^{1+\gamma}} .\end{aligned}$$

Since there is a constant $M \ge 1$ such that $M(1 + |x|^{1+\gamma}) \ge 2(1 + \kappa_U^\gamma |x|^\gamma)|x|$ for every $x \in \mathbb{R}$, we obtain

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

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

$$\begin{aligned}\|d(g \circ f)\|_{C^{t-1}} &\le 2 \|dg \circ f\|_{C^{t-1}} \|df\|_{C^{t-1}} \\ &\le 2M_{t-1} \|dg\|_{C^{t-1}} (1 + \|f\|_{C^{t-1}}^{t-1}) \|df\|_{C^{t-1}} \\ &\le 2M_{t-1} \|dg\|_{C^{t-1}} (1 + (2\kappa_U \|f\|_{C^t})^{t-1}) \|df\|_{C^{t-1}} ,\end{aligned}$$

according to Proposition 3.7. The same inequality used for the case $k = 1$ provides

$$\begin{aligned}\|g \circ f\|_{C^t} &\le \sup(\|g \circ f\|_{C^0} , \|d(g \circ f)\|_{C^{t-1}}) \\ &\le M_t \|g\|_{C^t} (1 + \|f\|_{C^t}^t) ,\end{aligned}$$

which completes the proof of Theorem 4.3. Actually assertion (ii.3) only needs that the domain U is compensated. ☐

The following examples show that the hypotheses in Theorem 4.3 are sharp.

Example 4.4. Let $\alpha > 0$ and $f_\alpha : [-1, 1] \to \mathbb{R}$, $x \to |x|^\alpha$. Then

i) if $0 < \alpha, \beta \le 1^-$, then $f_\alpha \circ f_\beta \in C^\gamma([-1, 1], \mathbb{R})$ if and only if $\gamma \le \alpha\beta$,

ii) if $\alpha/2 \notin \mathbb{N}$, then $f_\alpha \circ Id \in C^\gamma([-1, 1], \mathbb{R})$ if and only if $\gamma \le (\alpha - 1) + 1^-$.

5. Interpolation inequalities for different numbers of derivatives.

In this section, we develop interpolation inequalities which work for indices that straddle different integer parts. As we have shown before, these cannot hold for general domains. The proof of interpolation inequalities for spaces involving different numbers of derivatives was started in [11] and [19] for functions in the line. The method of proof is rather elementary and hence it is the basis for the extensions to general Banach spaces.

The idea of the proof is that if we restricted the function to a curve, we could use the classical argument for interpolation of the restriction. To conclude interpolations for the whole set, it suffices to find through every point in the set a rather well behaved family of curves whose tangent vectors at said point allow us to generate the complete vector space. This is the main idea for the definition of regular subset. The Example 3.2 shows that indeed such a hypothesis on the domain is needed.

Let E be a Banach space. If $D \subset E$, the absolutely convex cover of D is defined by $\Gamma(D) = \{\sum_{i=1}^n \rho_i x_i | \sum_{i=1}^n |\rho_i| \leq 1, x_i \in D\}$. The closure of $\Gamma(D)$, that we represent by $\overline{\Gamma}(D)$, is the closed absolutely convex cover of D. Note that this set can be also calculated in the weak topology.

Definition 5.1. We say that an open subset $U \subset E$ is *regular* if there exist constants $\kappa_U^* > 0$ and $0 < \delta < 1$ for which the following conditions hold:

i) for each $x \in U$ there exists a bounded set of directions $D_x \subset B_{\kappa_U^*}$ with $B_1 \subset \overline{\Gamma(D_x)}$,
ii) for each $x \in U$ and $h \in D_x$ there is a map $\psi \in C^{1+\delta}([0,1], U)$ such that $\psi(0) = x$, $\psi'(0) = h$ and $\|\psi\|_{C^{1+\delta}} < \kappa_U^*$.

We next show the abundance of this kind of sets. Let us consider $0 < \delta < 1$ and $f \in C_{\mathrm{loc}}^{1+\delta}(\mathbb{R}^m, \mathbb{R})$. Assume that $K = \{x \in \mathbb{R}^m \mid f(x) = 0\}$ is compact and $\mathrm{grad}(f(x)) \neq 0$ for every $x \in K$. Then the open set $U = \{x \in \mathbb{R}^m \mid f(x) < 0\}$ is regular. The following result also holds

Proposition 5.2. *Every convex and bounded open $U \subset E$ is compensated and regular.*

Proof. Since every convex set is compensated we next refer to the regularity of U. Let us fix a constant $\kappa_0 > 0$ with $U \subset B_{\kappa_0}$. We take $x_0 \in U$ and $0 < \rho < 1$ satisfying that $B_\rho(x_0) \subset U$. We set $D_x = \{(2/\rho)(y - x) | y \in B_\rho(x_0)\}$ and for each $y \in B_\rho(x_0)$, we define

$$\begin{array}{rccc} \psi_{x,y}: & [0,1] & \to & U \\ & t & \to & x + (\frac{2}{\rho}t - (\frac{2}{\rho} - 1)t^2)(y - x). \end{array}$$

Then $\|\psi_{x,y}\|_{C^{1+\delta}} \leq 4\kappa_0/\rho = \kappa_U^*$ for every $0 < \delta < 1$. Besides

$$B_1 \subset \{\frac{1}{\rho}(y - x) - \frac{1}{\rho}(x_0 - x) \big| \, y \in B_\rho(x_0)\} \subset \Gamma(D_x)$$

which shows that the domain U is regular. □

Example 5.3. We consider the Banach space of summable sequences

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

with the usual norm $\|\mathbf{x}\|_1 = \sum_{n=1}^\infty |x_n|$. Then

$$U = \{(x_n)_{n=1}^\infty \in l^1 | \sum_{n=1}^\infty \frac{|x_n|}{n!} < 1\}$$

defines a convex open subset of l^1 which is not regular. In consequence, the boundedness of the domain is an essential hypothesis in Proposition 5.2.

We establish the existence of interpolation inequalities in compensated regular domains.

Theorem 5.4. *Assume that the domain U is compensated and 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*

$$\|f\|_{C^s} \leq M_{r,t} \|f\|_{C^r}^\mu \|f\|_{C^t}^{1-\mu} . \tag{5.1}$$

Proof. Let $0 \le r < s < t \le n - 1 + 1^-$. We will proceed by induction on n. The interpolation inequality for $n = 1$ was established in Proposition 3.1. In the proof of the assertion for the integer $n + 1$ we distinguish the following cases:

i) For $0 \le r < s < t < n$, the result follows from the inductive argument. Proposition 3.7 also provides inequality (5.1) for the exponent $t = n$. If $n \le r < s < t < n + 1$, we apply again Proposition 3.1.

ii) Let $r = n-1+\alpha, s = n, t = n+\gamma$ with $0 < \gamma \le 1^-$. Consider $z \in C^{1+\gamma}(\mathbb{R}_+, F)$ and assume that there is $t_0 \in \mathbb{R}_+$ such that $\|z'(t_0)\| = \|z'\|_{C^0}$. Substituting z by $az(\rho t)$, if necessary, we can assume that $\|z'\|_{C^0} = 1$ and that $H_\gamma(z') = 1/2$.

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

$$\|R(t)\| \le \left\| \int_{t_0}^{t} (z'(s) - z'(t_0))ds \right\| \le |t - t_0|^{1+\gamma} H_\gamma(z') \,.$$

Now we have

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

This shows that $H_\alpha(z) \ge 1/2$ and therefore,

$$\|z'\|_{C^0} \le 2H_\alpha(z)^\mu H_\gamma(z')^{1-\mu}. \tag{5.2}$$

Notice that this inequality holds for z and $az(\rho t)$ simultaneously.

We assume that the domain U is regular for a constant $\kappa_U^* > 0$ and fix $f \in C^t(U; F)$. Under our conditions, the map $d^{n-1}f : U \to L_{n-1}(E, F)$ is differentiable and given $\epsilon > 0$ we can find $x \in U$ and $h \in D_x$ such that $\|d^n f\|_{C^0} - \epsilon \le \|d^n f(x)h\|$. In addition, there exists $0 < \delta \le \gamma$ and $\psi \in C^{1+\delta}([0, 1], U)$ with $\psi(0) = x$, $\psi'(0) = h$ and $\|\psi\|_{C^{1+\delta}} \le \kappa_U^*$.

We define $z : \mathbb{R}_+ \to L_{n-1}(E, F)$, $t \to d^{n-1}f(\psi(1 - e^{-t}))$ and take the real number $\nu = \delta/(1 + \delta - \alpha)$. It follows from Theorem 4.3 and relation 5.2 that $z \in C^{1+\delta}([0, 1], F)$ and

$$\|d^n f\|_{C^0} - \epsilon \le \|z'\|_{C^0} \le 2H_\alpha(z)^\nu H_\delta(z')^{1-\nu} \le M' \, \|f\|_{C^r}^\nu \, \|f\|_{C^{n+\delta}}^{1-\nu} \,.$$

This provides

$$\begin{aligned} \|f\|_{C^n} &\le M' \, \|f\|_{C^{n-1+\alpha}}^\nu \, \|f\|_{C^{n+\delta}}^{1-\nu} \\ &\le M \, \|f\|_{C^{n-1+\alpha}}^\nu \, \|f\|_{C^n}^{(1-\nu)(1-\delta/\gamma)} \, \|f\|_{C^{n+\gamma}}^{(1-\nu)\delta/\gamma} \,, \end{aligned}$$

according to the assertion (i). The interpolation inequality (5.1) is obtained by eliminating $\|f\|_{C^n}$ in the above expressions.

iii) The case $0 \le r < s = n < t = n + \gamma$ with $0 \le \gamma \le 1^-$ is similar to the last step of ii).

iv) Let $0 \le r < s < n < t = n + \gamma$ with $0 < \gamma \le 1^-$. Take $\nu' = (n - s)/(n - r)$ and $\nu = (t - n)/(t - r)$. Then i) implies

$$\|f\|_{C^s} \le M' \, \|f\|_{C^r}^{\nu'} \, \|f\|_{C^n}^{1-\nu'} \;; \tag{5.3}$$

besides, ii) yields

$$\|f\|_{C^n} \le M \, \|f\|_{C^r}^{\nu} \, \|f\|_{C^t}^{1-\nu} \,,$$

whose insertion in (5.3) provides the desired result.

v) The case $0 \le r < n < s < t = n + \gamma$ with $0 < \gamma \le 1^-$ is similar to iv).

This covers the study of all the possible choices of r, s, t with with $0 \le r < s < t \le n + 1^-$, and completes the proof of the inductive argument. ∎

In particular, Theorem 5.4 assures that the C^t-norm of the composition increases linearly in the C^t-norms of the factors for large t. This is the alternative to statement ii.2) of Proposition 4.3, which applies in compensated domains.

Proposition 5.5. *Assume that $U \subset E$, $V \subset F$ are compensated and regular subsets. Given $t \geq 1$, there is a constant $M_t > 0$ such that for every $f \in C^t(U,F)$ and $g \in C^t(V,G)$ with $f(U) \subset V$ one has*

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

Proof. Write $t = n + \alpha$, where $n \geq 1$ and $0 \leq \alpha \leq 1^-$. We will make use of the two following well-known results:

i) given $0 \leq \gamma \leq 1$ there is $M > 0$ satisfying $x^\gamma y^{1-\gamma} \leq M(x+y)$ for every $x, y \geq 0$.
ii) given $\kappa, C \geq 0$ there is $M > 0$ satisfying $C(1 + \kappa^\alpha x^\alpha)x^{n-1} \leq M(1 + x^{n+\alpha-1})$ for every $x \geq 0$.

By a repeated application of the chain rule we can find constants $c_{k,j_1,\dots,j_k}$ such that

$$d^m(g \circ f)(x) = \sum_{k=1}^{m} \sum_{j_1+\cdots j_k=m} c_{k,j_1,\cdots,j_k} d^k g(f(x))\{d^{j_1}f(x), \dots, d^{j_k}f(x)\}\,, \tag{5.4}$$

with $j_i \geq 1$ and $1 \leq m \leq n$. Using the interpolation inequalities for both f and g with $1 < k$ and $j_i < m$ we obtain

$$\|g\|_{C^k} \leq M'\,\|g\|_{C^1}^{(m-k)/(m-1)}\,\|g\|_{C^m}^{(k-1)/(m-1)}\,,$$
$$\|f\|_{C^j} \leq M''\,\|f\|_{C^1}^{(m-j)/(m-1)}\,\|f\|_{C^m}^{(j-1)/(m-1)}\,.$$

Now we deduce from the above relation i) that

$$\|d^m(g \circ f)\|_{C^0} \leq M\left(\sum_{k=1}^{n} \|f\|_{C^1}^{k-1}\right)(\|g\|_{C^1}\,\|f\|_{C^m} + \|g\|_{C^m}\,\|f\|_{C^1})\,.$$

Take $x \neq y$; for all the terms of the sum (5.4) one has

$$\begin{aligned}
&\left\|d^k g(f(x))\{d^{j_1}f(x), \dots, d^{j_k}f(x)\} - d^k g(f(y))\{d^{j_1}f(y), \dots, d^{j_k}f(y)\}\right\| \\
&\quad\leq \left\|(d^k g(f(x)) - d^k g(f(y)))\{d^{j_1}f(x), \cdots d^{j_k}f(x)\}\right\| \\
&\quad\quad+ \left\|\sum_{i=1}^{k} d^k g(f(y))\{d^{j_1}f(y), \dots, d^{j_i}f(y) - d^{j_i}f(x), \dots, d^{j_k}f(x)\}\right\| \\
&\quad\leq \|g\|_{C^{k+\alpha}}\,\|f\|_{C^1}^{\alpha}\,\kappa_U^\alpha\,\|x-y\|^\alpha\,\|f\|_{C^{j_1}} \cdots \|f\|_{C^{j_k}} \\
&\quad\quad+ \sum_{i=1}^{k} \|g\|_{C^k}\,\|f\|_{C^{j_1}} \cdots \|f\|_{C^{j_i+\alpha}}\,\|x-y\|^\alpha \cdots \|f\|_{C^{j_k}}\ .
\end{aligned}$$

Applying the interpolation inequalities (5.4) to each factor with $r = 1, t = n + \alpha$ and relation i), we finally get the bound

$$H_\alpha(d^n(g \circ f)) \leq M(1 + \kappa_U^\alpha\,\|f\|_{C^1}^\alpha)\left(\sum_{k=1}^{n} \|f\|_{C^1}^{k-1}\right)(\|g\|_{C^1}\,\|f\|_{C^t} + \|g\|_{C^t}\,\|f\|_{C^1})\,,$$

from which we deduce, thinking of ii) for the case $\|f\|_{C^1} \geq 1$, the existence of a constant $M_t > 0$ satisfying the claim. $\square$

6. Differentiability properties of composition.

Throughout this section, we will systematically study differentiability properties of the map

$$\begin{array}{rcc} \mathrm{Comp} : [\mathcal{U} \subset C^r(U,F)] \times C^s(V,G) & \longrightarrow & C^t(U,G) \\ (f,g) & \longrightarrow & \mathrm{Comp}(f,g) = g \circ f \end{array}$$

under the following hypotheses:

h.1) the exponents $r, s \geq 0$ and

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

h.2) $\mathcal{U} \subset C^r(U,F)$ is open and $\operatorname{dist}(f(U), V^c) > 0$ for every $f \in \mathcal{U}$,
h.3) $U \subset E$ and $V \subset F$ are open compensated sets.
h.4) for every $1 > \varepsilon > 0$, there is $\delta(\varepsilon) > 0$ and an open compensated $V_\varepsilon \subset V$ such that $\{y \in V \mid \operatorname{dist}(y, V^c) > \varepsilon\} \subset V_\varepsilon \subset \{y \in V \mid \operatorname{dist}(y, V^c) > \delta(\varepsilon)\}$.

Hypotheses h.1) and h.2) are needed to obtain a correct definition of composition on a open neighborhood in the f-component. Finally, h.3) and h.4) require geometrical conditions on the domains which simplify the behavior of Comp; in particular, they assure that small translations of differentiable paths joining points of each domain V_ε remain in V. Observe that all these conditions are satisfied in most of the applications. We will study regularity with respect to each of the factors, as a previous step to deduce the conditions for the joint continuity and differentiability of the map.

The following result is a direct consequence of Theorem 4.3.

Proposition 6.1. *Let $r, s \geq 0$ and $f \in C^r(U,F)$. Then the map*

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

is linear and continuous.

Now we consider the variation of the first factor.

Proposition 6.2. *Let $r, s, t \geq 0$ and $g \in C^s(V,G)$. The map*

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

is continuous in the following cases:

i) *for $t = 0$, when $r \geq 0,\ s > 0$,*
ii) *for $0 < t < 1^-$, when $s > t,\ r \geq t$ and $rs > t$,*
iii) *for $t = k + \gamma$ with $k \geq 1$ and $0 \leq \gamma < 1^-$, when $s > t$ and $r \geq t$,*
iv) *for $t = k + 1^-$ with $k \geq 0$, when $s > k+1$ and $r \geq t$.*

Besides, for every $f_1 \in \mathcal{U}$ there exist positive real numbers δ, ρ and M, irrespective of g, such that if $f_2 \in C^r(U,F)$ with $\|f_1 - f_2\|_{C^r} < \delta$ one has that $f_2 \in \mathcal{U}$ and

$$\|g \circ f_1 - g \circ f_2\|_{C^t} \leq M\, \|g\|_{C^s}\, \|f_1 - f_2\|^{\rho}_{C^r}\,. \tag{6.1}$$

In particular, ρ does not depend on f, i.e. $\rho = \rho(r,s,t)$, and $g_* \in C^{\rho}_{\mathrm{loc}}(\mathcal{U}, C^t(U,G))$.

Proof. We make use of different properties of Comp obtained in Sections 3 and 4. Since we expound in detail borderline cases, the exponents $0 < \rho \leq 1$ obtained in the proof are not always optimum; sometimes they can be improved using interpolation inequalities. We fix $f_1 \in \mathcal{U}$ for all what follows.

i) It is similar to i) of Theorem 4.3. Thus, if $0 < s \leq 1^-$ then

$$\|g_*(f_1) - g_*(f_2)\|_{C^0} \leq \|g\|_{C^s} \|f_1 - f_2\|_{C^0}^s . \tag{6.2}$$

ii.1) First we consider $0 < s \leq 1^-$. We can also assume that $0 < t < r \leq 1^-$.

From Lemma 3.1, inequality (6.2) and i.1) of Theorem 4.3 we deduce

$$\begin{aligned} \|g_*(f_1) - g_*(f_2)\|_{C^t} &= \|g \circ f_1 - g \circ f_2\|_{C^t} \\ &\leq 2 \|g \circ f_1 - g \circ f_2\|_{C^0}^{(rs-t)/(rs)} \|g \circ f_1 - g \circ f_2\|_{C^{rs}}^{t/(rs)} \\ &\leq 2 \|g\|_{C^s} [2 + \|f_1\|_{C^r}^s + \|f_2\|_{C^r}^s]^{t/(rs)} \|f_1 - f_2\|_{C^0}^{s-t/r} , \end{aligned} \tag{6.3}$$

which is the inequality (6.1) with $\rho(r,s,t) = s - t/r$. This same argument provides the proof for the exponent $s = 1$.

ii.2) We now consider the case $s = 1 + \beta$ with $0 < \beta \leq 1^-$ and $0 < t \leq r \leq 1^-$. Since $\operatorname{dist}(f_1(U), V^c) = \lambda > 0$, there exist, according to hypothesis h.2), a real positive number $\delta(\lambda)$ and an open compensated set V_λ such that $f_1(U) \subset V_\lambda \subset \{y \in V \mid \operatorname{dist}(y, V^c) > \delta\}$. We can also assume that $\operatorname{dist}(f_1(U), V_\lambda^c) > 0$. Take $x, y \in U$ and represent $z_1 = f_1(x), w_1 = f_1(y)$. For each $\varepsilon > 0$ there is a path γ_1 joining z_1 and w_1 in V_λ with

$$\int_0^1 \|\dot\gamma_1(t)\| \, dt \leq (\kappa_\lambda + \varepsilon) \|z_1 - w_1\| \quad (\text{where } \kappa_\lambda = \kappa_{V_\lambda}) .$$

Let $f_2 \in \mathcal{U}$ with $\|f_2 - f_1\|_r < \delta$. First of all we have

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

If we denote $z_2 = f_2(x)$ and $w_2 = f_2(y)$ then

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

is a path joining z_2 and w_2 in V. We next introduce

$$\begin{aligned} I_1 &= \int_0^1 \|[dg(\gamma_1(t)) - dg(\gamma_2(t))]\dot\gamma_1(t)\| \, dt \\ &\leq \|g\|_{C^{1+\beta}} (\|z_1 - z_2\| + \|w_1 - w_2\|)^\beta (\kappa_\lambda + \varepsilon) \|z_1 - w_1\| , \end{aligned}$$

$$I_2 = \int_0^1 \|dg(\gamma_2(t))(\dot\gamma_1(t) - \dot\gamma_2(t))\| \, dt \leq \|g\|_1 \|z_2 - z_1 + w_1 - w_2\| .$$

We obtain

$$\begin{aligned} &\|g(f_1(x)) - g(f_2(x)) - g(f_1(y)) + g(f_2(y))\| \\ &\qquad \leq \int_0^1 \|(dg(\gamma_1(t))\dot\gamma_1(t) - dg(\gamma_2(t))\dot\gamma_2(t)\| \, dt \leq I_1 + I_2 \\ &\qquad \leq 2 \|g\|_{C^{1+\beta}} \|f_1 - f_2\|_{C^0}^\beta \kappa_\lambda \|f_1\|_{C^r} \|x - y\|^r \\ &\qquad\quad + \|g\|_{C^1} \|f_1 - f_2\|_{C^r} \|x - y\|^r . \end{aligned}$$

This shows that

$$\|g \circ f_1 - g \circ f_2\|_{C^r} \le \|g\|_{C^{1+\beta}} \left(2\kappa_\lambda \|f_1\|_{C^r} \|f_1 - f_2\|_{C^0}^{\beta} + \kappa_V \|f_1 - f_2\|_{C^r}\right).$$

Hence there is a constant M, irrespective of g, such that

$$\|g \circ f_1 - g \circ f_2\|_{C^r} \le M \|g\|_{C^{1+\beta}} \|f_1 - f_2\|_{C^r}^{\beta}. \tag{6.4}$$

Since the same above constants can be taken in a complete neighborhood of f_1, we can deduce that $g_* \in C^{\beta}_{\mathrm{loc}}(\mathcal{U}, C^t(U,G))$.

iii) Assume that $t = r = k + \alpha$ and $s = k + \beta$ with $k \ge 1$ and $0 \le \alpha < \beta \le 1^-$. Take $\delta > 0$ such that if $f_2 \in C^r(U,F)$ and $\|f_1 - f_2\|_{C^r} < \delta$ then $f_2 \in \mathcal{U}$.

We proceed by induction in k. For the integer $k = 1$ we have

$$\|g_*(f_1) - g_*(f_2)\|_{C^0} \le \kappa_V \|g\|_{C^1} \|f_1 - f_2\|_{C^0}.$$

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

$$\begin{aligned}
&\|d(g \circ f_1) - d(g \circ f_2)\|_{C^\alpha} \\
&\quad = \|dg \circ f_1 \cdot d(f_1 - f_2) - (dg \circ f_2 - dg \circ f_1) \cdot df_2\|_{C^\alpha} \\
&\quad \le 2 \|dg \circ f_1\|_{C^\alpha} \|d(f_1 - f_2)\|_{C^\alpha} + 2 \|dg \circ f_2 - dg \circ f_1\|_{C^\alpha} \|df_2\|_{C^\alpha}.
\end{aligned} \tag{6.5}$$

On the other hand,

$$\begin{aligned}
&\|dg \circ f_2 - dg \circ f_1\|_{C^\alpha} \\
&\qquad \le 2 \|dg \circ f_2 - dg \circ f_1\|_{C^0}^{(\beta-\alpha)/\beta} \left(\|dg \circ f_2\|_{C^\beta} + \|dg \circ f_1\|_{C^\beta}\right)^{\alpha/\beta} \\
&\qquad \le 2 \|g\|_{C^{1+\beta}} \left(\|f_1\|_{C^1}^{\beta} \kappa_U^{\beta} + \|f_2\|_{C^1}^{\beta} \kappa_U^{\beta} + 2\right)^{\alpha/\beta} \|f_1 - f_2\|_{C^0}^{\beta-\alpha}.
\end{aligned}$$

By substitution in (6.5) we find a constant $M_{1+\alpha}$, irrespective of g, such that if $\|f_1 - f_2\|_{C^{1+\alpha}} < \delta$ then

$$\|g \circ f_1 - g \circ f_2\|_{C^{1+\alpha}} \le M_{1+\alpha} \|g\|_{C^{1+\beta}} \|f_1 - f_2\|_{C^{1+\alpha}}^{\beta-\alpha}.$$

Let us assume that an inequality of the type (6.1) with $\rho(r,s,t) = \beta - \alpha$ holds for every integer lower or equal to $k-1$, and analyze the composition for the integer k. We obtain

$$\|dg \circ f_2 - dg \circ f_1\|_{C^{t-1}} \le M_{t-1} \|g\|_{C^s} \|f_1 - f_2\|_{C^{t-1}}^{\beta-\alpha}.$$

From relation (6.5) and ii.3) of Theorem 4.3 we deduce that

$$\begin{aligned}
\|d(g \circ f_2) - d(g \circ f_1)\|_{C^{t-1}} \le 2M_t' \|g\|_{C^t} \left(1 + \|f\|_{C^{t-1}}^{t-1}\right) \|f_1 - f_2\|_{C^t} \\
+ M_{t-1} \|g\|_{C^s} \|f_1 - f_2\|_{C^{t-1}}^{\beta-\alpha}
\end{aligned}$$

which yields

$$\|g \circ f_2 - g \circ f_1\|_{C^t} \le M_t \|g\|_{C^s} \|f_1 - f_2\|_{C^t}^{\beta-\alpha}$$

for an adequate constant $M_t > 0$. This completes the proof of the assertion for the range of exponents considered above.

iv) Assume that $t = r = k + 1^-$ and $s = k + 1 + \beta$ with $k \ge 0$ and $0 < \beta \le 1^-$. The continuity of g_* for $k = 0$ is a consequence of the arguments used in ii). There exist constants δ and M such that if $\|f_1 - f_2\|_{C^r} < \delta$ then $f_2 \in \mathcal{U}$ and

$$\|g_*(f_1) - g_*(f_2)\|_{C^r} \le M \|g\|_{C^s} \|f_1 - f_2\|_{C^r}^{\beta}. \tag{6.6}$$

The proof of the inequality (6.6) for $k \geq 1$ can be derived by an inductive argument similar to the one detailed in iii). In this case, one has $\rho(r,s,t) = \beta$.

Notice that the hypothesis h.4) is an important ingredient in the proof of the assertions ii.2) and iv). On the contrary, no geometrical conditions on the domains are required in the simplest situations. $\square$

Definition 6.3. Let $\Omega \subset \mathbb{R}^3_+$ be the set of exponents (r,s,t) satisfying some of the conditions stated in i)-iv) of Proposition 6.2. We say that Ω is the *set of the exponents of continuity.*

Example 6.4. *Let $(r,s,t) \in \mathbb{R}^3_+ - \Omega$. There are open balls $U \subset E$, $V \subset F$ of Banach spaces and $g \in C^s(V,\mathbb{R})$ such that $g_* : \mathcal{U} \subset C^r(U,F) \longrightarrow C^t(U,\mathbb{R})$ is discontinuous.*

Proof. We first introduce several examples concerning non integer exponents. 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 $[-1/2, 1/2]$ and the value zero outside of $[-1,1]$. We set

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

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

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

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

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

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

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

$$\left| (g \circ f_n - g \circ Id)(\frac{-1}{n}) - (g \circ f_n - g \circ Id)(0) \right| = 2\frac{1}{n^\gamma},$$

hence $\|g \circ f_n - g \circ Id\|_{C^\gamma} \geq 2$. Besides, for $k \geq 1$

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

and

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

so that $\|g \circ f_n - g \circ Id\|_{C^t} > 2\,k!$. In consequence, there is not convergence in the C^t-topology.

We detail the following examples in infinite-dimensional Banach spaces. For $p \geq 1$ we introduce the usual space of summable sequences,

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

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 $0 < r \leq 1^-$ 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, $g_{k,n}(x) = \int_0^x g_{k-1,n}(s)\,ds$. For $k \in \mathbb{N},\ k \geq 0$ and $\sigma > 0$ big enough we define

$$\begin{array}{rccc} \mathbf{g}_k: & B_\sigma \subset l^p & \longrightarrow & \mathbb{R} \\ & (x_j)_{j\in\mathbb{N}} & \longrightarrow & \sum_{j=1}^{\infty} g_{k,j}(x_j) \end{array}$$

Given $\mathbf{x} = (x_j)_{j\in\mathbb{N}} \in B_\sigma$, there exist indices $j_1, \ldots, j_{m_p}$ with $m_p \leq (2\sigma)^p$ such that if $\mathbf{y} \in B_{1/2}(\mathbf{x})$ then $|y_j| < 1$ for $j \in \mathbb{N} - \{j_1, \ldots, j_{m_p}\}$. If we denote

$$\begin{array}{rccc} \pi_x: & l^p & \longrightarrow & \mathbb{R}^{m_p} \\ & (y_j)_{j\in\mathbb{N}} & \longrightarrow & (y_{j_1}, \ldots, y_{j_{m_p}}) \end{array}$$

and

$$\begin{array}{rccc} \mathbf{g}_{k,x}: & \mathbb{R}^{m_p} & \longrightarrow & \mathbb{R} \\ & (y_1, \ldots, y_{m_p}) & \longrightarrow & \sum_{i=1}^{m_p} g_{k,j_i}(y_i)\,, \end{array}$$

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

Let h be a C^∞ cut-off function which takes the value 1 on $[-1/3, 1/3]$ but zero outside of $(-2/3, 2/3)$, and set $f(x) = (1 + |x-1|^r)h(x-1)$. We define

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

and

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

The same previous argument shows that $\mathbf{h} \in C^\infty(B_{3/2} \subset l^1, l^p)$ (this map will be used in the study of the differentiability of $\mathbf{g}_k$). On the other hand,

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

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, if $x \in [1, 1 + (1/2n)^p)$ for n big enough,

$$\begin{aligned} \mathbf{g}_1 \circ \mathbf{f}_n(xe_n) - \mathbf{g}_1 \circ \mathbf{f}(xe_n) &= g_{1,n}\left(f(x) + \frac{1}{2n} \right) - g_{1,n}(f(x)) \\ &= \frac{|x-1|^r}{2} + \frac{1}{8n}\,, \end{aligned}$$

which provides $\|\mathbf{g}_1 \circ \mathbf{f}_n - \mathbf{g}_1 \circ \mathbf{f}\|_{C^r} \geq 1/2$. In this example there is not continuity for $s = 1$ and $t = r$. This completes ii).

Finally, if we consider $p = 1$ and $f(x) = xh(x-1)$ in the above relations, then $\mathbf{f} \in C^\infty(B_{3/2} \subset l^1, l^p))$ and $\|\mathbf{f}_n - \mathbf{f}\|_{C^r} \to 0$ for every $r > 0$. Take $x \in [1, 1 + 1/2n)$; therefore

$$\mathbf{g}_0 \circ \mathbf{f}_n(xe_n) - \mathbf{g}_0 \circ \mathbf{f}(xe_n) = \frac{1}{2}\,,$$

which shows that the continuity in i) requires the uniform continuity of the second factor g. Besides, in the same conditions, one has

$$\begin{aligned} d^k(\mathbf{g}_{k+1} \circ \mathbf{f}_n - \mathbf{g}_{k+1} \circ \mathbf{f})(xe_n)\{e_n, \overset{(k)}{\ldots}, e_n\} &= g_{1,n}(f(x) + \frac{1}{2n}) - g_{1,n}(f(x)) \\ &= \frac{x-1}{2} + \frac{1}{8n}\,, \end{aligned}$$

so that $\|\mathbf{g}_{k+1} \circ \mathbf{f}_n - \mathbf{g}_{k+1} \circ \mathbf{f}\|_{C^{k+1^-}} \geq 1/2$. In this example, there is not continuity for $s = k+1,\ t = k+1^-$ and neither for $s = k+1,\ t = k+1$. □

Remark 6.5. If F is a finite dimensional space, $\overline{f(U)}$ is a compact subset of V. Every continuous function on V is uniformly continuous on each compact neighborhood of $f(U)$. Notice that under the additional hypothesis that $g \in C^s_u(V, F)$ or F is finite dimensional, we can supplement the cases of continuity in Proposition (6.2) with the following ones:

v) $0 < t \leq 1^-$, $s = 1$ and $r \geq t$,
vi) $t = k$, $s = k$ and $r \geq k$,
vii) $t = k + 1^-$, $s = k+1$ and $r \geq k + 1^-$.

Definition 6.6. Let $\Omega_u \subset \mathbb{R}^3_+$ be the set of exponents (r, s, t) satisfying some of the conditions stated in i)-vii). We say that Ω_u is the *extended set of the exponents of continuity.*

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

Proposition 6.7. *Let $r, s, t \geq 0$ and $m \in \mathbb{N}$, $m \geq 1$ such that $(r, s-m, t) \in \Omega$. If $g \in C^s(V, G)$ then g_* is m times continuously differentiable. Besides, $g_* \in C^{m+\rho}_{\text{loc}}(\mathcal{U}, C^t(U, G))$ where $\rho = \rho(r, s-m, t)$.*

Proof. It is known that the maps $(d^j g)_* : \mathcal{U} \to C^t(U, L_j(F, G))$, $j = 1, \ldots, m$ are continuous. Moreover, given $f_1 \in \mathcal{U}$ there are positive constants M, λ, ρ such that $\mathcal{B}_\delta(f_1) = \{f_2 \in C^r(U, F) \mid \|f_1 - f_2\|_{C^r} < \delta\} \subset \mathcal{U}$ and, if $1 \leq j \leq m$, then $(d^j g)_* \in C^\rho(\mathcal{B}_\delta(f_1), C^t(U, L_j(F, G)))$ with

$$\left\|(d^j g)_*(f_1) - (d^j g)_*(f_2)\right\|_{C^t} \leq M \|g\|_{C^s} \|f_1 - f_2\|^\rho_{C^r}\,.$$

We will verify the existence and continuity of the map

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

defined by $d(g_*)(f)(h) = dg \circ f \cdot h$. For $\|h\|_{C^r} \le 1$ and ε small enough we have

$$\begin{aligned}
&\frac{1}{\varepsilon} \|g_*(f_1 + \varepsilon h) - g_*(f_1) - \varepsilon dg \circ f_1 \cdot h\|_{C^t} \\
&\qquad = \frac{1}{\varepsilon} \|g \circ (f_1 + \varepsilon h) - g \circ f_1 - \varepsilon dg \circ f_1 \cdot h\|_{C^t} \\
&\qquad = \left\| \int_0^1 [dg \circ (f_1 + \varepsilon s h) - dg \circ f_1] \cdot h \, ds \right\|_{C^t} \\
&\qquad \le \int_0^1 2 \|(dg)_*(f_1 + \varepsilon s h) - (dg)_*(f_1)\|_{C^t} \|h\|_{C^t} \, ds \\
&\qquad \le 4 M \kappa_U \|g\|_{C^s} \varepsilon^\rho .
\end{aligned}$$

In addition, from Proposition (4.2) we deduce that

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

is a continuous linear map and $d(g_*) = l \circ (dg)_*$. Besides Proposition (4.1) asserts that $d(g_*) \in C^\rho(\mathcal{B}_\delta(f_1), L(C^r(U,F), C^t(U,G)))$. Now a simple inductive argument allows us to complete the proof of the statement.

It is important to remark that the continuity of $(d^j g)_*$ for $j = 1, \dots, m$ is the only essential requirement to derive that g_* is m times continuously differentiable. □

Again Proposition 6.7 contains the best conditions of differentiability.

Example 6.8. *Let $(r, s-1, t) \in \mathbb{R}^3_+ - \Omega$. There are open balls $U \subset E$ and $V \subset F$ of Banach spaces and $g \in C^s(V, \mathbb{R})$ such that $g_* : \mathcal{U} \subset C^r(U,F) \longrightarrow C^t(U, \mathbb{R})$ is not differentiable.*

Proof. Set $g = x^{k+1}|x|^\gamma$ with $k \in \mathbb{N}$ and $0 < \gamma \le 1^-$. Thus $g \in C^{k+1+\gamma}((-2,2), \mathbb{R})$.

If $k = 0$ and $0 < \alpha \le 1^-$ then the map $g_* : C^\alpha((-1,1), \mathbb{R}) \to C^{\alpha\gamma}((-1,1), \mathbb{R})$ is not differentiable. In effect, choosing $f = |x|^\alpha$ and $h = 1$ we find for every $x \in (-1,1)$

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} (g_*(f + \varepsilon h)(x) - g_*(f)(x)) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} (g(|x|^\alpha + \varepsilon) - g(|x|^\alpha)) = g'(|x|^\alpha) .$$

However,

$$-\frac{1}{\varepsilon} (g(|x|^\alpha - \varepsilon) - g(|x|^\alpha) + \varepsilon g'(|x|^\alpha))\Big|_0^{x^\alpha = \varepsilon/2} = -\left(1 + \frac{\gamma}{2^\gamma}\right) \varepsilon^\gamma ,$$

which shows that the convergence fails in the C^t-topology.

If $k \ge 0$ and $r \ge \max(1, k+\gamma)$, the map $g_* : C^r((-1,1), \mathbb{R}) \to C^{k+\gamma}((-1,1), \mathbb{R})$ is not differentiable. Taking $f = Id$ and $h = 1$, we have for every $x \in (-1,1)$

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} (g_*(f + \varepsilon h)(x) - g_*(f)(x)) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} (g(x + \varepsilon) - g(x)) = g'(x) .$$

Nevertheless,

$$\begin{aligned}
&\frac{1}{\varepsilon} \left(g^{(k)}(x + \varepsilon) - g^{(k)}(x) - \varepsilon g^{(k+1)}(x) \right) \Big|_{x = -\varepsilon/2}^0 \\
&\qquad\qquad = (k+1+\gamma)(k+\gamma) \cdots (2+\gamma) \left(1 + \frac{\gamma}{2^\gamma}\right) \varepsilon^\gamma ;
\end{aligned}$$

thus there is not convergence in the $C^{k+\gamma}$-topology.

We now recover, from Example 6.4, the functions $(\mathbf{g}_k)_{k\geq 0}$, $\mathbf{f}$ and $\mathbf{h}$ defined on the l^p-spaces. In all the cases, if $x \in [1, 1 + (1/2n)^p)$ one has

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \left(\mathbf{g}_{k+1} \circ (\mathbf{f} + \varepsilon\mathbf{h})(xe_n) - \mathbf{g}_{k+1} \circ \mathbf{f}(xe_n)\right)$$
$$= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \left(g_{k+1,n}(f(x) + \varepsilon) - g_{k+1,n}(f(x)\right) = g_{k,n}(f(x) = \mathbf{g}_k \circ \mathbf{f}(xe_n)\,.$$

Given $0 < \varepsilon < 1/4$, we can find $n \in \mathbb{N}$ with $1/4 \leq n\varepsilon < 1/2$. In the following examples, once ε has been fixed, we refer to the values of n and x mentioned above.

For $0 < r \leq 1^-$ and $p = 1/r$, the map $(\mathbf{g}_2)_* : C^r(B_{3/2} \subset l^1, l^p) \to C^r(B_{3/2}, \mathbb{R})$ is not differentiable. Notice that

$$\frac{1}{\varepsilon} \left(\mathbf{g}_2 \circ (\mathbf{f} + \varepsilon\mathbf{h})(xe_n) - \mathbf{g}_2 \circ \mathbf{f}(xe_n) - \varepsilon\mathbf{g_1} \circ \mathbf{f}(xe_n)\right) = \frac{n}{6}\varepsilon(\varepsilon + 3|x-1|^r)\,,$$

and thus

$$\frac{1}{\varepsilon}\,\|\mathbf{g_2} \circ (\mathbf{f} + \varepsilon\mathbf{h}) - \mathbf{g_2} \circ \mathbf{f} - \varepsilon\mathbf{g_1} \circ \mathbf{f}\|_{C^r} \geq \frac{1}{8}\,.$$

Similarly, when $p = 1$ the map $(\mathbf{g}_1)_* : C^r(B_{3/2} \subset l^1, l^1) \to C^0(B_{3/2}, \mathbb{R})$, $r \geq 0$ is not differentiable. In fact,

$$\frac{1}{\varepsilon}\,\|\mathbf{g_1} \circ (\mathbf{f} + \varepsilon\mathbf{h}) - \mathbf{g_1} \circ \mathbf{f} - \varepsilon\mathbf{g_0} \circ \mathbf{f}\|_{C^0} \geq \frac{1}{8}\,.$$

Finally, the maps $(\mathbf{g}_{k+2})_* : C^r(B_{3/2} \subset l^1, l^1) \to C^{k+1}(B_{3/2}, \mathbb{R})$ for $r \geq k+1$ and $(\mathbf{g}_{k+2})_* : C^r(B_{3/2} \subset l^1, l^1) \to C^{k+1^-}(B_{3/2}, \mathbb{R})$ for $r \geq k+1^-$ are not differentiable. A simple evaluation provides

$$\frac{1}{\varepsilon} d^k \left(\mathbf{g}_{k+2} \circ (\mathbf{f} + \varepsilon\mathbf{h}) - \mathbf{g}_{k+2} \circ \mathbf{f} - \varepsilon\mathbf{g}_{k+1} \circ \mathbf{f}\right)(xe_n)\{e_n, \overset{(k)}{\ldots}, e_n\}$$
$$= \frac{1}{\varepsilon} \left(g_{2,n}(f(x) + \varepsilon) - g_{2,n}(f(x) - \varepsilon g_{1,n}(f(x))\right) = \frac{n}{6}\varepsilon(\varepsilon + 3(x-1))\,;$$

hence

$$\frac{1}{\varepsilon}\,\|\mathbf{g}_{k+2} \circ (\mathbf{f} + \varepsilon\mathbf{h}) - \mathbf{g}_{k+2} \circ \mathbf{f} - \varepsilon\mathbf{g}_{k+1} \circ \mathbf{f}\|_{C^{k+1^-}} \geq \frac{1}{8}\,.$$

This completes the analysis of all the cases stated in Proposition 6.2. ☐

Proposition 6.9. i) *Let $r, s, t \geq 0$ and $m \in \mathbb{N}$ such that $(r, s-m, t) \in \Omega$. Then the map*

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

is m times continuously differentiable.

ii) *Let $r, s, t \geq 0$ and $m \in \mathbb{N}$, $m \geq 1$ such that $(r, s-m, t) \in \Omega_u$. Then the map*

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

is $m-1$ times continuously differentiable.

Proof. Let us take $(r, s-m, t) \in \Omega$. We first assume $m = 0$ and verify that R is continuous. Let $f_1, f_2 \in \mathcal{U}$ with $\|f_1 - f_2\|_{C^r} \leq \delta$ small enough in order that inequality (6.1) applies; then

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

When $m \geq 1$ 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

$$\frac{1}{\varepsilon} \|R(f_1 + \varepsilon h) - R(f_1) - \varepsilon dR(f_1)(h)\|_{C^{s,t}}$$
$$= \sup_{\|g\|_{C^s} \leq 1} \frac{1}{\varepsilon} \|g \circ (f_1 + \varepsilon h) - g \circ f_1 - \varepsilon dg \circ f_1 \cdot h\|_{C^t} \leq 4M\kappa_U \varepsilon^\rho .$$

We consider

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

and

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

We know that R_1 is $m-1$ times differentiable, l is a continuous linear map and $dR = l \circ R_1$. As a consequence of Proposition 4.1 we obtain that R is m times continuously differentiable.

Finally, if $m \geq 1$ and $(r, s-m, t) \in \Omega_u$ then $(r, s-(m-1), t) \in \Omega$ and hence R is $m-1$ times continuously differentiable. □

Assuming that the hypotheses h.1)-h.4) hold, we summarize our conclusions:

Theorem 6.10. *Let r, s, t be positive numbers and $m \in \mathbb{N}$ such that $(r, s-m, t) \in \Omega$. The map*

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

is continuous and m times continuously differentiable. Besides, $\mathrm{Comp} \in C^{m+\rho}_{\mathrm{loc}}(\mathcal{U} \times C^s(V,G))$ *where $\rho = \rho(r, s-m, t)$.*

Proof. The continuity and locally Hölder behaviour is a consequence of Proposition 6.1 and Proposition 6.2. Let $m \geq 1$ and consider

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

Note that $d\,\mathrm{Comp}$ is the sum of Q and the map R defined above. Moreover, Q 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 and bounded. As a consequence of the induction process we deduce that the intermediate map is $C^{m-1+\rho}_{\mathrm{loc}}(\mathcal{U} \times C^{s-1}(V, L(F,G)))$. Thus, we obtain that the composition operator is m times continuously differentiable and belongs to $C^{m+\rho}_{\mathrm{loc}}(\mathcal{U} \times C^s(V,G))$. □

Besides, if $(f,g) \in \mathcal{U} \times \mathcal{C}^f(\mathcal{V}, \mathcal{G})$, then for every $j = 1, \ldots, m$ we obtain

$$d^j\,\mathrm{Comp}(f,g)\{h, k, \overset{(j)}{\ldots}, h, k\} = j\, d^{j-1} k \circ f\{h, \overset{(j-1)}{\ldots}, h\} + d^j g \circ f\{h, \overset{(j)}{\ldots}, h\}.$$

This provides, in a neighborhood of (f,g), the local Taylor expansion

$$\begin{aligned}(g+k)\circ(f+h) = g\circ f + \sum_{j=1}^{m} \frac{1}{(j-1)!} d^{j-1}k\circ f\cdot\{h,\overset{(j-1)}{\ldots},h\} \\ + \sum_{j=1}^{m}\frac{1}{j!} d^j g\circ f\cdot\{h,\overset{(j)}{\ldots},h\} + R_m(h,k)\,,\end{aligned} \tag{6.7}$$

whose remainder term satisfies $\|R_m(h,k)\|_{C^t} \le M(\|h\|_{C^r}^{m+\rho} + \|h\|_{C^r}^m \|k\|_{C^s})$.

Theorem 6.11. *Let r,s,t be positive numbers and $m\in\mathbb{N}$ with $(r,s-m,t)\in\Omega_u$.*

i) *The map*

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

is continuous and m times continuously differentiable.

ii) *If F is finite dimensional, the map*

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

is continuous and m times continuously differentiable.

The proof is similar to the previous statement and we omit the details. The Taylor expansion (6.7) is again valid but, under these conditions, the remainder term satisfies $\|R_m(h,k)\|_{C^t}\le M\,\|h\|_{C^r}^m\,(\tilde{k}(h)+\|k\|_{C^s})$ where the function $\tilde{k}=o(1)$.

The existence of sufficiently regular differentiable functions allows us to maintain the conditions of differentiability with simpler geometry of the domains. Assuming that the hypotheses h.1)-h.3) hold we can state

Proposition 6.12. *Let us assume that F is a superreflexive Banach space.*

i) *If $(r,s-m,t)\in\Omega_u$, the map* Comp : $[\mathcal{U}\subset C^r(U,F)]\times C_u^s(V,G)\to C^t(U,G)$ *is continuous and m times continuously differentiable.*

ii) *If F is finite dimensional and $(r,s-m,t)\in\Omega$, the map* Comp : $[\mathcal{U}\subset C^r(U,F)]\times C^s(V,G)\to C^t(U,G)$ *is continuous and m times continuously differentiable.*

Proof. We only give a sketch of the proof of statement i). It basically agrees with the corresponding to i) of Theorem 6.11, which is based in the previous study of the regularity with respect to the factors.

We first discuss the case $m=0$. Let $0<t\le 1^-$, $s=1$, $r\ge t$ and take $f_1\in C^r(U,F)$. It is known that $f_1(U)$ is a bounded subset of F and $\operatorname{dist}(f_1(U),V^c)=\lambda>0$. We consider the open sets $V_1=\cup_{x\in U}B^F_{\lambda/3}(f_1(x))$, $V_2=\cup_{x\in U}B^F_{2\lambda/3}(f_1(x))$, then $\overline{V}_1\subset V_2\subset V$. There exists a constant $\delta>0$ such that if $f\in C^r(U,F)$ and $\|f-f_1\|_{C^r}\le\delta$, then $f\in\mathcal{U}$ and $f(U)\subset V_1$. The Banach space F is superreflexive, this allows us to take a uniformly continuously differentiable function $\varphi:F\to\mathbb{R}$ such that $0\le\varphi\le1$, $\varphi|_{V_1}=1$ and $V_2=\{y\in F\,|\,\varphi(y)\ne0\}$ (see [37]). Since V_2 is bounded, it is easy to check that $\varphi\in C_u^1(F,G)$. Besides, for every $g\in C_u^1(V,G)$ one has that $\varphi\cdot g\in C_u^1(F,G)$ and $C_u^1(V,G)\to C_u^1(F,G)$, $g\to\varphi\cdot g$ is a bounded linear map. Finally, $(\varphi\cdot g)\circ f|_U=g\circ f|_U$ for every $f\in\mathcal{B}_\delta(f_1)$, which proves that Comp is continuous. Observe that for $0<t\le1^-$, $s>1$, $t\ge1$, i.e. $(r,s,t)\in\Omega$, all the maps g_* admit a common modulus of continuity when g varies in a bounded subset of $C^s(V,G)$. This result will be used in the study of the differentiability of Comp.

For the case $t = k+1^-$, $s = k+1$ and $r \geq k+1^-$ with $k \in \mathbb{N}$, we follow an inductive process, similar to the one detailed in iv) of Proposition 6.2, to deduce the continuity of Comp. In each one of steps we make use of the above arguments, based in a convenient extension to F of continuously differentiable functions on V. In addition, for $t = k+1^-$, $s > k+1$, $r \geq k+1$, i.e. $(r,s,t) \in \Omega$, we remark again that all the maps g_* admit a common modulus of continuity when g varies in a bounded subset of $C^s(V,G)$.

We now consider $m \geq 1$. The corresponding versions, under the present conditions, of Proposition 6.7 and Proposition 6.9 allow us to conclude that Comp is m times continuously differentiable.

On the other hand, hypothesis h.4) was not required to check the differentiability of Comp for the rest of values (r,s,t) with $(r, s-m, t) \in \Omega_u$; thus the previous proof is also valid for this case. $\square$

We finally remark that, since the considered sets of definition are open subsets of Banach spaces, there may be functions which are continuously differentiable but unbounded or with unbounded derivatives. In these cases, besides the topology induced by the supremum norm, there are other natural topologies, e.g., the Withney topology associated with the uniform convergence on closed bounded sets. This topology is a projective limit of the considered in this paper, and the results on continuity can be immediately lifted. In addition, Theorems 6.10 and 6.11 yield the starting point to complete the study of the differentiability of the composition in the mentioned topology.

7. 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 León.

References

[1] R. Abraham, J.Robbin, *Transversal Mappings and flows,* Benjamin, New York, 1967.
[2] J. Appel, P. Zabrejko, *Nonlinear superposition operator,* Cambridge University Press, 1990.
[3] V. I. Arnol'd, Sur la geometrie differentielle des groupes de Lie de dimension infinie et ses applications a l'hydrodynamique des fluides parfaits, *Ann. Inst. Fourier* **16** (1996), 319–361.
[4] V. I. Averbukh, O. G. Smolyanov, The various definitions of the derivative in linear topological spaces, *Russian Math. Surveys* **23**, 4 (1968), 67-113.
[5] A. Banyaga, Sur la structure du groupe des diffeomorphismes qui preservent une forme symplectique, *Comm. Math. Helv.* **53** (1978), 174–277.
[6] A.M. Bloch, H. Flaschka, T.S. Ratiu, The Toda PDE and the geometry of the diffeomorphism group of the annulus, *Fields Inst. Comm.* **7** (1992), 57–92.
[7] D.G. Ebin, J. Marsden, Groups of diffeomorphisms and the motion of an incompressible fluid, *Ann. of Math.* **92** (1970), 102–163.
[8] J. Eels, A setting for global analysis, *Bull. A.M.S.* **72** (1966), 751–807.
[9] M. Frechet, La notion de diffΘrentielle dans l'analyse gΘnΘral, *C. R. Acad. Sci. Paris* **180** (1925), 806-809.
[10] D.B.A. Epstein, The simplicity of certain groups of homeomorphisms, *Compos. Math.* **22** (1972), 165–173.
[11] J. Hadamard, Sur le module maximum d'une fonction et de ses derives, *Bull. Soc. Math. France* **42** (1898), 68–72.

[12] R. HAMILTON, The inverse function theorem of Nash and Moser, *Bull. A.M.S.* **7** (1982), 65–222.

[13] M. P. HEBLE, Aproximation of differentiable functions on a Hilbert space, *Contemporary Math.* **54** (1986), 17-33.

[14] M. P. HEBLE, *Approximation problems in analysis and probability* North Holland, 1989

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

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

[17] N. KURZWEIL On approximation in real Banach spaces, *Studia Math.* **14** (1954), 213–231

[18] A. KATOK, G. KNIEPER, M. POLLICOTT, H. WEISS, Differentiability and analyticity of topological entropy for Anosov and geodesic flows, *Inv. Math.* **92** (1989), 581–597.

[19] E. LANDAU, Einige ungleichungen für zweimal differentiarbare funktionen, *Proc. London Math. Soc.* **13** (1896), 43–58.

[20] O. E. LANFORD III, Bifurcation of periodic solutions into invariant tori: the work of Ruelle and Takens in *Lect. Notes in Math.* **322,** Springer-Verlag, 1973.

[21] J. LANGER, D.A. SINGER, Diffeomorphisms of the circle and geodesic fields in Riemmann surfaces of genus one, *Inv. Mat.* **69** (1982), 229–242.

[22] E. B. LEACH, J.H.M. WHITFIELD Differentiable functions and rough norms on Banach spaces, *Proc. AMS*, **33** (1972), 120–126

[23] R. DE LA LLAVE, Cohomology equations over Anosov systems in non conmmutative groups, including groups of diffeomorphisms, *Preprint.*

[24] R. DE LA LLAVE, J.M. MARCO, R. MORIYÓN, Canonical perturbation theories of Anosov diffeomorphisms and regularity results for the Livsic cohomology equation, *Ann. of Math.* **123** (1986), 537–611.

[25] R. DE LA LLAVE, R. OBAYA, Decomposition theorems for some groups of diffeomorphisms, *Preprint.*

[26] J. G. LLAVONA *Approximation of continuously differentiable functions.* North Holland, 1986

[27] J.N. MATHER, Commutator of diffeomorphisms, *Comment. Mat. Helv.* **49** (1974), 512–528.

[28] J. MOSER, A rapidly convergent iteration method and non-linear differential equations. I and II, *Ann. Sc. Norm. Pisa* **20** (1966), 226-315 and 499–535.

[29] J. MOSER, On a theorem of Anosov, *Jour. Diff. Eq.* **5** (1969), 411–440.

[30] N. MOULIS, Approximation de fonctions differentiables sur certains espaces de Banach, *Ann. Inst. Fourier* **21**, 4 (1971), 293-345.

[31] S. M. NIKOLSKII, *Approximation of functions of several variables and imbedding theorems* Springer Verlag, New York, 1971.

[32] V. NITICĂ, A. TÖRÖK, Cohomology of dynamical systems and rigidity of partially hyperbolic actions of higher-rank lattices, *Duke Math. Jour.* **79** (1995), 751–810.

[33] R. S. PALAIS, *Foundations of global non-linear analysis,* Benjamin, New York, 1968.

[34] D. RUELLE, Differentiation of SRB states, *Deposited in mp_arc* **# 96-499**.

[35] S. SMALE, Differentiable dynamical systems, *Bull. A.M.S.* **73** (1967), 747–817.

[36] E. STEIN, *Singular integrals and differentiability properties of functions*, Princeton University Press, 1970.

[37] K. SUNDARESAN, Geometry and nonlinear analysis in Banach spaces, *Pacific Journal of Math.* **102**, 2 (1982), 487-498.

[38] M. VISHIK, Hydrodynamics in two dimensions, Preprint

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