THE PARAMETERIZATION METHOD FOR INVARIANT MANIFOLDS I: MANIFOLDS ASSOCIATED TO NON-RESONANT SUBSPACES

XAVIER CABRÉ, ERNEST FONTICH, AND RAFAEL DE LA LLAVE

Abstract. We introduce a method to prove existence of invariant manifolds and, at the same time study their dynamics. As a first application, we consider the dynamical system given by a C^r map F in a Banach space X close to a fixed point: $F(x) = Ax + N(x)$, A linear, $N(0) = 0$, $DN(0) = 0$. We show that if X_1 is an invariant subspace of A and A satisfies certain spectral properties, then there exists a unique C^r manifold which is invariant under F and tangent to X_1. The method of proof also provides information about the dynamics on the manifold. In addition, we prove regularity results on dependence on parameters.

When X_1 corresponds to spectral subspaces associated to sets of the spectrum contained in disks around the origin or their complement, we recover the classical (strong) (un)stable manifold theorems. Our theorems, however, apply to other invariant spaces. Indeed, we do not require X_1 to be an spectral subspace or even to have a complement invariant under A.

1. Introduction and statements of results

The main goal of this paper is to develop a method to prove existence and regularity of invariant manifolds for dynamical systems. We call it the parameterization method, and we use it to prove a variety of invariant manifold results. We establish optimal regularity for the invariant objects, as well as regularity with respect to dependence on parameters. As a particular case, our results generalize the classical stable and unstable invariant manifold theorems in the neighborhood of a fixed point.

We consider a C^r map F on a Banach space X such that $F(0) = 0$. We want to study aspects of the dynamics in a neighborhood of the fixed point. The heuristics is that, in small neighborhoods, the map is very similar to its linear part. One can hope that subspaces X_1 invariant under the linearization have nonlinear counterparts: smooth manifolds tangent to X_1 at 0 which are invariant under the map F.

Roughly speaking, the method consists on trying to find at the same time a parameterization of the manifold and the dynamics restricted to it. The parameterization has to satisfy a functional equation that expresses that its range is

Date: Oct 8 2001.

invariant and that it semi-conjugates the dynamics to the dynamics of a simpler map (in some cases linear). More precisely, we look for maps $K : U_1 \subset X_1 \to X$ (the parameterization) and $R : X_1 \to X_1$ satisfying the functional equation

$$F \circ K = K \circ R \qquad \text{in } U_1 \ .$$

Recall that X_1 is a subspace of X invariant under $DF(0)$. Then, the previous equation guarantees that $K(U_1)$ is an invariant manifold for F.

We will show that, provided that the spectrum of $DF(0)$ satisfies certain non-resonance conditions, one can indeed solve the functional equation above and find these invariant manifolds. This will extend previous results from [dlL97]. We point out that [dlL97] contains examples which show that the non-resonance conditions are necessary for existence.

In [dlL97], the invariant spaces for $DF(0)$ for which the corresponding manifolds invariant under F were constructed need not be spectral subspaces. Nevertheless, it was required there that they had an invariant complement. In this paper we only require that there is a complement, but it need not be invariant under $DF(0)$. In this way, we can associate, for example, invariant manifolds to the spaces corresponding to eigendirections in a non-trivial Jordan block.

As a technical improvement over [dlL97], in the present paper we obtain invariant manifolds of class C^r whenever the map F is C^r –rather than just $C^{r-1+\mathrm{Lip}}$ manifolds as in [dlL97]. This improvement has also been obtained in [ElB98].

Moreover, we also study the regularity of the dependence with respect to parameters and obtain rather sharp results. Indeed, both technical improvements rely on the same idea of recursively obtaining candidates for the derivative and showing that they indeed are derivatives.

Even if our results are already novel for finite dimensions, we work in the generality of Banach spaces since the use of finite dimensions does not simplify much the proofs (except the ones on sharp differentiability results). Moreover, the use of infinite dimensional systems leads to interesting applications in dynamical systems. A construction of [HP70] shows that one can reduce the existence of finite dimensional invariant objects (such as foliations) to the existence of invariant manifolds for certain operators defined on infinite dimensional spaces of sections; see Section 2 for these questions.

The parameterization method that we present has several advantages (in contrast with the graph transform method, for instance):

(i) We are lead to consider functional equations given as fixed points of operators which are differentiable in spaces of smooth functions. This allows to obtain very simply smooth dependence on parameters by applying the standard implicit function theorem. The regularity obtained in this way is almost optimal.

This is in contrast with the operators whose fixed points give the invariant manifolds in the graph transform method. These graph transform operators are only a contraction (or differentiable) in C^r norms when considered in subsets of spaces of C^r functions consisting of C^{r+1} functions.

(ii) The method lends itself very well to numerical implementations.

The papers [FR81] and [FG92] use this method to deal with one dimensional manifolds. Numerical work for higher dimensional maps has been done in [BK98], that undertook the task of systematically computing Taylor expansions of invariant manifolds.

We note that the method of proof leads easily to *a posteriori* estimates (see Remark 13 in Section 3.4). That is, if we can find functions which solve our equation approximately, we can assure that there are true solutions whose distance form the computed solution is bounded by the residual of the solution. Preliminary implementations seem to show that the method is quite stable. Perhaps the numerical stability is related to good functional-analytic properties.

(iii) The method is geometrically natural.

That is, the equations we use can be expressed in a coordinate free way (they are just semi-conjugacy equations). Hence, they carry through geometric structures such as symplectic forms, volume forms, etc. We have not explored these possibilities here, but we hope to come to them in future work.

(iv) The method provides a global representation of the manifold.

The solutions of the equation that we study provide a representation of the invariant manifold which covers the whole manifold. When the map is entire (e.g. polynomial), the functions representing the manifold are entire.

This is in contrast with the graph transform methods, which only provide a representation of the part of the manifold which is a graph and, hence they fail when the manifold starts to fold. From the numerical point of view, this avoids a complicated step of *globalization*.

(v) The method can be used to study a variety of invariant objects, both in finite and in infinite dimensional Banach spaces (see Section 2).

One of the applications of the non-resonant manifolds that we produce is to make sense of the so called *slow manifolds* near a fixed point. That is, manifolds associated to the slowest eigenvalues. They are important in applications since one can argue –indeed we justify to a certain extent– that the eigenvalues closest to the unit circle are the ones which get suppressed the slowest and, hence, dominate the asymptotics of the convergence. This is particularly relevant in chemical kinetics since often one has large sets of equations and the eigenvalues of fixed points are widely different.

Indeed, we think that our results provide a rigorous foundation for some geometric methods to study long term asymptotics that have been used in the applied literature. We just quote some recent references in chemical kinetics: [Fra98], [Smo91], [PR93], [WMD96]. For the use of slow manifolds in chemical reactions in a more dynamical context, see [Bar88]. The name slow manifolds seems to have been used for related –but perhaps similar– objects in atmospheric science (see [Lor92] for a review).

It is interesting to note that there is another possible construction of slow manifolds, namely Irwin's pseudostable manifolds (see [dlLW95] for a modern treatment). As it is discussed in detail in [dlL97], it turns out that our smooth manifolds and the Irwin's one may not coincide even in systems in which both of them can be defined. Irwin's manifolds are unique under conditions of global behavior, while the ones considered in the present paper are unique under smoothness conditions. There are many –indeed they are C^1 generic– systems in which it is impossible to satisfy at the same time the global conditions and the smoothness conditions. Hence, the Irwin method and the one here may produce different invariant manifolds.

Roughly, the situation of slow manifolds can be summarized as saying that we can find a number $r_0 \leq r_1$ depending only on the spectrum of the linearization such that:

- There are infinitely many manifolds which are $C^{r_0-\delta}$.
- There is at most one $C^{r_1+\delta}$.
- In case that there is a $C^{r_1+\delta}$ manifold, it is as smooth as the map.
- There are non-resonance conditions that imply that there is a $C^{r_1+\delta}$ manifold.
- The smooth manifold –if it exits– depends smoothly on parameters.

Indeed, we think that this explains certain discrepancies in the numerical literature in chemical kinetics. We believe that the method of Fraser computes the smooth manifolds described here, while the method of Maas-Pope computes Irwin manifolds. See [Fra88], [MP92], [Fra98] and a more detailed discussion in Appendix B.

We warn the reader interested only in finite dimensional results and not interested in the optimal regularity of the invariant manifold that, in this case, the proof of our main result just requires reading Lemmas 3.1, 3.3 and 3.4.

The dependence on parameters is considered in detail in Section 5, where optimal regularity is obtained both for Theorem 1.1 and 1.2. See Theorems 5.2 and 5.7 respectively.

Several of the spectral results that we need are quite easy and well known when our Banach spaces are finite dimensional. In order to facilitate the exposition for those readers interested only in finite dimensional systems, we have relegated the proofs in general Banach spaces to Appendix A.

We finish the article with a historical section (Appendix B), where we describe some early works related to ours. We also comment on some literature which deals with computations of the invariant manifolds and applications to chemical kinetics and fluid dynamics.

1.1. **Some notation and conventions used.**

1.1.1. *Complexification.* Since we use some spectral theory, we will need that some operators and other objects are defined on a complex Banach space. If the system we are interested in is defined on a real Banach space, we use the well known device of complexification. That is, if X is a real space we can consider the space $\tilde{X} = X \oplus iX$ with the obvious multiplication by a complex number.

Given a multilinear (in particular, a linear) operator A in X, we can extend it to $\tilde{X}$ in a canonical way by defining $\tilde{A}(x_1 + iy_1, \cdots, x_n + iy_n)$ as the result of expanding all the sums and taking all the i's out (i.e., proceeding as if it was multilinear). It is routine and well known how to check that the resulting operator in $\tilde{X}$ is multilinear when we consider it as an operator in the complex space $\tilde{X}$.

Of course, a multilinear operator in $\tilde{X}$ which has range in the real subspace for given data in the real subspace can also be considered as an operator in X. It turns out that the extension and the restriction are inverse operations. We will make some remarks along the proofs about this very general and well known procedure.

1.1.2. *Spectrum.* The spectrum of a linear operator A in X will be denoted by $\mathrm{Spec}(A)$. We emphasize that even if X is a real Banach space, $\mathrm{Spec}(A)$ denotes indeed the spectrum of the complex extension $\tilde{A}$ of A, and hence $\mathrm{Spec}(A)$ is a compact subset of $\mathbb{C}$.

Given two sets Λ and Σ of complex numbers, we use the notation

$$\Lambda\Sigma = \{\lambda\sigma \mid \lambda \in \Lambda, \sigma \in \Sigma\} \subset \mathbb{C}$$

and

$$\Lambda^n = \{\lambda_1\lambda_2 \cdots \lambda_n \mid \lambda_1, \ldots, \lambda_n \in \Lambda\} \subset \mathbb{C}\ .$$

A polynomial P defined on X taking values in Y is a function from X to Y of the form $P = \sum_{i=1}^{n} P_i$, where P_i is the restriction to the diagonal of a multilinear map of degree i from $X \times \cdots \times X$ to Y. We will then say that P is a polynomial of degree not larger than n (or simply of degree n).

1.1.3. *Spaces of functions.* Given X, Y Banach spaces and $U \subset X$ an open set, $C^r(U, Y)$ is the set of functions $f : U \to Y$ which are r times continuously differentiable (in the strong sense) and which have all derivatives up to order r bounded on U. It is a Banach space with the norm

$$\|f\|_{C^r(U,Y)} = \max\left\{\|f\|_{C^0(U)}, \|Df\|_{C^0(U)}, \ldots, \|D^r f\|_{C^0(U)}\right\}\ , \tag{1}$$

where for functions g taking values in a Banach space, we write

$$\|g\|_{C^0(U)} = \sup_{x\in U} |g(x)|\ .$$

The space $C^\infty(U, Y)$ consists of those functions which are r times continuously differentiable for every $r \in \mathbb{N}$.

We also consider the space $C^\omega(U, Y)$ of bounded analytic functions defined in a complex neighborhood U, equipped with the supremum norm (see Section 3.3 for details). When U, Y are clear from the context, we will suppress them from the notation.

1.2. **Statement of results.** The first main result of the paper is the following:

Theorem 1.1. *Let X be a real or complex Banach space, U an open set of X, $0 \in U$, and let $F : U \to X$, $F(0) = 0$, be a C^r map, with $r \in \mathbb{N} \cup \{\infty, \omega\}$.*

Let $A = DF(0)$, $N(x) = F(x) - Ax$, and $X = X_1 \oplus X_2$ be a direct sum decomposition into closed subspaces. Denote by π_1, π_2 the corresponding projections.

Assume:

0) *F is a local diffeomorphism. In particular, A is invertible.*
1) *The space X_1 is invariant under A. That is*
$$AX_1 \subset X_1 \ .$$
Let $A_1 = \pi_1 A|_{X_1}$, $A_2 = \pi_2 A|_{X_2}$ and $B = \pi_1 A|_{X_2}$. Hence, we have $A = \begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix}$ with respect to the above decomposition. Assume:
2) $\operatorname{Spec}(A_1) \subset \{z \in \mathbb{C} \,|\, |z| < 1\}$.
3) $0 \notin \operatorname{Spec}(A_2)$.

Let $L \geq 1$ be an integer such that
$$\big(\operatorname{Spec}(A_1)\big)^{L+1} \operatorname{Spec}(A^{-1}) \subset \{z \in \mathbb{C} \,|\, |z| < 1\} \ , \tag{2}$$
and assume that:
4) $\big(\operatorname{Spec}(A_1)\big)^i \cap \operatorname{Spec}(A_2) = \emptyset$ *for every integer i with $2 \leq i \leq L$ (in case that $L \geq 2$).*
5) $L + 1 \leq r$.

Then,

a) *We can find a polynomial map $R : X_1 \to X_1$ with*
$$R(0) = 0 \ , \quad DR(0) = A_1 \ , \tag{3}$$
of degree not larger than L, and a C^r map $K : U_1 \subset X_1 \to X$, where U_1 is an open neighborhood of 0, such that
$$F \circ K = K \circ R \tag{4}$$
holds in U_1, and
$$K(0) = 0 \ , \tag{5}$$
$$\pi_1 DK(0) = \mathrm{Id} \ , \quad \pi_2 DK(0) = 0 \ . \tag{6}$$
In particular, $K(U_1)$ is a C^r manifold invariant under F and tangent to X_1 at 0.

b) *In case that we further assume*
$$\big(\operatorname{Spec}(A_1)\big)^i \cap \operatorname{Spec}(A_1) = \emptyset \qquad \textit{for every integer } i \textit{ with } 2 \leq i \leq L \ , \tag{7}$$
then we can choose R in a) above to be linear. More generally, if
$$\big(\operatorname{Spec}(A_1)\big)^i \cap \operatorname{Spec}(A_1) = \emptyset \qquad \textit{for every integer } i \textit{ with } M \leq i \leq L \ , \tag{8}$$

then we can choose R in a) above to be a polynomial of degree not larger $M - 1$.

c) *The C^r manifold produced in a) is unique among C^{L+1} locally invariant manifolds tangent to X_1 at 0. That is, every two C^{L+1} locally invariant manifolds will coincide in a neighborhood of 0 in X. Note that the parameterization K and the map R need not be unique; it is the manifold $K(U_1)$ which is unique.*

d) *The following parameter version of the result also holds. Let Λ be an open subset of a Banach space Λ_0 and let $F : \Lambda \times U \subset \Lambda_0 \times X \to X$, that we will write as $F_\lambda(x)$. Assume that $F \in C^r(\Lambda \times U, X)$ (i.e., F is jointly C^r in its two arguments), that the subspace X_1 (which is independent of λ) is invariant under $A_\lambda := D_x F_\lambda(0)$ for every λ, and that hypotheses 0)–5) above are satisfied uniformly in λ.*

Then, we can find K_λ and R_λ satisfying a) and b) for every λ, and such that K_λ is C^{r-L-1} and R_λ is C^{r-L} jointly in their two arguments.

As we will see in Remark 6, a result on invariant manifolds for flows can be deduced rigorously from Theorem 1.1 using time t maps and the uniqueness result.

Equation (4) ensures that the range of K is an invariant manifold under F. By (6), it is tangent to X_1 at the origin. We will see that the composition $K \circ R$ in equation (4) is well defined, since R will send certain neighborhoods of 0 in X_1 into themselves (indeed, balls centered at zero of sufficiently small radius for an appropriate norm in which A_1 is a contraction).

Note that, in both finite and infinite dimensional cases, assumption 2) on the spectrum of A_1 guarantees the existence of an integer $L \geq 1$ for which (2) holds. We also point out that, by hypothesis 5), we always have $r \geq 2$ (r is the order of differentiability of the map F).

It is appropriate to call condition 4) a non-resonance condition since, in the finite dimensional setting, it amounts to the fact that every product of at most L eigenvalues of A_1 is not an eigenvalue of A_2.

One situation in which condition 4) is satisfied automatically is when X_1 corresponds to the spectral subspace associated to a closed disk of radius $\rho < 1$. Then, if A_1 is invertible, our theorem produces the strong stable manifolds for invertible maps. The differentiability assumptions required by the theorem are, however, stronger than those in the classical proofs, due to hypotheses (2) and 5). This is to be expected since we also obtain information on the dynamics (weaker differentiability assumptions will be addressed in Theorem 1.2). Indeed, we obtain that the dynamics on the invariant manifold is semi-conjugated to the dynamics of R that is just a polynomial, which under hypothesis (7), becomes a linear map.

Since A is invertible, passing to the inverse, we can establish the existence of strong unstable manifolds. Theorem 1.1 applies, however, to other invariant subspaces. In this respect, note that the theorem does not require X_1 to be an spectral subspace or even to have a complement invariant under A (see the examples in Remark 3).

Remark 1. The existence of the map K in conclusion a) of Theorem 1.1 also holds if one assumes, instead of the non-resonances 4), the existence of polynomials R and $K^{\leq}$ of degree not larger than L satisfying (3), (5), (6) and $F \circ K^{\leq}(x) = K^{\leq} \circ R(x) + o(|x|^L)$. It may happen that there exist several R, $K^{\leq}$ satisfying the above mentioned conditions. Associated to each such solution $K^{\leq}$ there will be a unique C^{L+1} invariant manifold $K(U)$ with K such that its L-jet is $K^{\leq}$, i.e. $D^j K(0) = D^j K^{\leq}(0)$ for each $j \in \{0, \dots, L\}$.

As we will see in Section 3.1, the non-resonance conditions 4) ensure that certain linear equations can be solved (see Lemma 3.1). However, even if there exist resonances, it could happen that these equations can nevertheless be solved for the system at hand. Indeed, we could use some conditions weaker than 4), but at the price of imposing restrictions on the non-linear part. When X is finite-dimensional, these restrictions are submanifolds of finite codimension and, if we have enough parameters to select, we can get the invariant manifolds for some special values of the parameters. When the map preserves a geometric structure (e.g. volume, contact or symplectic) or a symmetry, resonances often exist, but at the same time, there are usually cancellations that allow the equations to be solved.

Remark 2. Recall that, since A is invertible, the equality

$$A^{-1} - \lambda \,\mathrm{Id} = -\lambda A^{-1}(A - \lambda^{-1}\,\mathrm{Id})$$

implies that the spectrum of A^{-1} is exactly $\{\lambda^{-1} \mid \lambda \in \mathrm{Spec}(A)\}$. Therefore, (2) and conditions 2) and 4) in Theorem 1.1 imply

$$\big(\mathrm{Spec}(A_1)\big)^i \cap \mathrm{Spec}(A_2) = \emptyset, \qquad \forall\, i \geq 2 \,.$$

Similarly, if in addition (7) holds then

$$\big(\mathrm{Spec}(A_1)\big)^i \cap \mathrm{Spec}(A_1) = \emptyset, \qquad \forall\, i \geq 2 \,.$$

Our formulation in the theorem makes clear that these are really a finite number of conditions since, for $i > L$, they are satisfied automatically.

Remark 3. We emphasize that the spectrum of A_1 and that of A_2 need not be disjoint, since condition 4) is only required for powers bigger or equal than 2. For example, the theorem applies to

$$A = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}, \quad X_1 = \{(x^1, 0)\}, \;\; X_2 = \{(0, x^2)\} \;,$$

and also to

$$A = \begin{pmatrix} 1/2 & 1 \\ 0 & 1/2 \end{pmatrix}, \quad X_1 = \{(x^1, 0)\}, \;\; X_2 = \{(0, x^2)\} \;.$$

A more sophisticated example where our results apply is

$$A = \begin{pmatrix} 1/2 & 1 & & & \\ 0 & 1/2 & & & \\ & & 1/3 & & \\ & & & 1/5 & 1 \\ & & & & 1/5 \end{pmatrix} .$$

Then, denoting by E_i the i^{th} coordinate axis, we could associate invariant manifolds to E_1, E_3, E_4, $E_1 \oplus E_2$, $E_4 \oplus E_5$, or to sums of these spaces, e.g., $E_1 \oplus E_4$, $E_1 \oplus E_2 \oplus E_4$, etc.

If X is finite dimensional and A is invertible then A_1 and A_2 are also invertible. Instead, when X is infinite dimensional, the fact that A is invertible does not ensure that A_1 and A_2 are also invertible (see Example A.2). Nevertheless, from the invertibility of A, which is equivalent to the unique solvability of the system

$$\begin{aligned} &A_1 x^1 + B x^2 = y^1 \\ &A_2 x^2 = y^2 \ , \end{aligned}$$

we deduce that A_1 is injective and A_2 is onto. Moreover we have that if two of the linear maps A_1, A_2, A are invertible then the third one is also invertible. In particular, under the hypotheses of Theorem 1.1 we have that A_1 is invertible, since A and A_2 are assumed to be invertible.

On the other hand, we have that

$$A^{-1} = \begin{pmatrix} A_1^{-1} & -A_1^{-1} B A_2^{-1} \\ 0 & A_2^{-1} \end{pmatrix} .$$

Using the formula

$$\rho(A) = \max\big(\rho(A_1), \rho(A_2)\big) \tag{9}$$

which is proved in Appendix A, applied to A^{-1} instead of A, we deduce

$$\rho(A^{-1}) = \max\big(\rho(A_1^{-1}), \rho(A_2^{-1})\big) \geq \rho(A_2^{-1}) \ . \tag{10}$$

The last inequality leads to the following fact:

$$\begin{aligned} &\big(\operatorname{Spec}(A_1)\big)^{L+1} \operatorname{Spec}(A^{-1}) \subset \{|z| < 1\} \\ &\qquad \Longrightarrow \big(\operatorname{Spec}(A_1)\big)^{L+1} \operatorname{Spec}(A_2^{-1}) \subset \{|z| < 1\} \ . \end{aligned} \tag{11}$$

Remark 4. Note that, in finite dimensions, (11) is obvious because then,

$$\operatorname{Spec}(A_2^{-1}) \subset \operatorname{Spec}(A^{-1}) \ . \tag{12}$$

In infinite dimensions, (12) could be false. See examples in Section A. The argument presented for (11) is true both for finite and infinite dimensions.

Theorem 1.1 does not cover exactly the main theorem in [dlL97] since the definition of L, which in Theorem 1.1 is given by (2), in [dlL97] is

$$\big(\operatorname{Spec}(A_1)\big)^{L+1}\operatorname{Spec}(A_2^{-1})\subset\{z\in\mathbb{C}\mid |z|<1\}\ . \tag{13}$$

Therefore, by (11), the exponent L of Theorem 1.1 is larger than that of [dlL97]. In particular, assumption 5) of the present paper requires more differentiability for the map F than the corresponding result in [dlL97]. This stronger assumption is reasonable since the conclusions of Theorem 1.1 are also stronger than those in [dlL97] –because they include information, through semi-conjugacy, about the dynamics on the invariant manifold (given by R).

The shortcoming mentioned above is remedied in the following result. Theorem 1.2 below is a strict generalization of the main result in [dlL97] since it makes exactly the same differentiability assumptions, obtains the same conclusions, and improves the differentiability result for K. Moreover, we obtain uniqueness conclusions. Another way that Theorem 1.2 improves on the main result of [dlL97] is that it does not require the decomposition $X = X_1 \oplus X_2$ to be invariant under A –a result that seems not to appear in any reference on invariant manifolds that we are aware of.

Theorem 1.2. *Assume hypotheses 0) - 5) of Theorem 1.1 except that* (2) *is replaced by*

$$\big(\operatorname{Spec}(A_1)\big)^{L+1}\operatorname{Spec}(A_2^{-1})\subset\{z\in\mathbb{C}\,|\,|z|<1\}\ ,\qquad L\geq 1\ . \tag{14}$$

Then,

a) *We can find a C^r map K and a C^r map R satisfying* (4)*. Moreover we can choose K in such a way that*

$$\pi_1 K = \mathrm{Id}\,. \tag{15}$$

b) **b1)** *Furthermore, the C^r manifold produced is the unique C^{L+1} locally invariant manifold tangent to X_1 at* 0*.*

In fact, the following stronger result holds.

b2) *K is the unique* (*locally around the origin*) *solution of* (4) *in the class of Lipschitz functions $K : U_1 \subset X_1 \to X$ of the form $K = (\mathrm{Id}, w_L + h)$, with w_L being a polynomial of degree L such that $w_L(0) = 0$ and $Dw_L(0) = 0$, and with $\sup_{x\in U_1}(|h(x)|/|x|^{L+1}) < \infty$.*

c) *Let Λ be an open subset of a Banach space Λ_0 and let $F : \Lambda\times U \subset \Lambda_0\times X \to X$, that we will write as $F_\lambda(x)$. Assume that $F \in C^r(\Lambda \times U, X)$* (*i.e., F is jointly C^r in its two arguments*)*, that the subspace X_1* (*which is independent of λ*) *is invariant under $A_\lambda := D_xF_\lambda(0)$ for every λ, and that hypotheses 0)–5) above are satisfied uniformly in λ.*

Then the unique solution K_λ is C^{r-L-1} jointly in their two arguments.

Remark 5. Following the proof of this theorem one sees that we only use that A is invertible (hypothesis 0) to get, from the spectral conditions on A, the existence of a norm in X, equivalent to the original one, such that the following condition is satisfied: $\|A_2^{-1}\| \, \|A_1\|^{L+1} < 1$. Actually, this norm condition is the technical condition we use through the rest of the proof. In finite dimensional spaces the spectral condition and the norm condition are equivalent, but this property does not seem to be true in general. We prefer to put the spectral condition on the statement because it is intrinsic. However the conclusions of Theorem 1.2 also hold if we do not assume that A is invertible and instead of (14) we assume that $\|A_2^{-1}\| \, \|A_1\|^{L+1} < 1$.

Remark 6. Note that one of the consequences of the uniqueness conclusions of Theorem 1.1 and Theorem 1.2 is that these theorems apply also to flows with a fixed point at zero. That is, we can associate a manifold which is invariant under all the elements of the flow to every subspace that is invariant under all the elements of the linearized flow and that satisfies the non-resonance conditions for one element of the flow.

For, if we denote by $\{S_t\}_{t\in\mathbb{R}}$ a flow of class C^r, and S_{t_0} satisfies either the conditions of Theorem 1.1 or the ones of Theorem 1.2 for some t_0, we know that there exists a manifold W tangent to the given subspace, say X_1, and such that $S_{t_0}(W) \subset W$.

We claim that, for any s, $S_s(W)$ is invariant by S_{t_0} and that $T_0 S_s(W) = X_1$. Indeed,

$$S_{t_0} \circ S_s(W) = S_s \circ S_{t_0}(W) \subset S_s(W) \ .$$

Moreover, since by assumption we know that $DS_s(0)X_1 = X_1$, then

$$T_0 S_s(W) = DS_s(0) T_0 W = X_1 \ .$$

Hence, we obtain that the manifold $S_s(W)$ satisfies the conclusions of Theorem 1.1 or 1.2 for the map $F = S_{t_0}$ and hence, by uniqueness, $S_s(W) = W$.

In the following section we indicate how to obtain other types of invariant objects using Theorems 1.1 and 1.2. For the proofs of these two theorems (that start in Section 3), one can skip Section 2, which is devoted to applications.

2. Invariant manifolds for normally hyperbolic manifolds and invariant foliations

The previous construction of invariant manifolds associated to non-resonant subspaces can be lifted to construct invariant manifolds near another invariant manifold that generalize the stable manifolds of normally hyperbolic invariant manifolds. It can also be used to construct invariant foliations or invariant prefoliations.

Even if the existence of these objects can be obtained by following the steps of the proof in the present paper, most of them can also be obtained immediately –through a device used in [HP70]– by applying Theorems 1.1 and 1.2 to an appropriate map between certain infinite dimensional Banach spaces. This was indeed one of the motivations to formulate our results in general Banach spaces.

The main idea is, given a diffeomorphism, to consider its action on spaces of continuous sections. In the following, we will detail the construction.

Let M be a C^∞ Riemannian manifold, $N \subset M$ a C^1 submanifold (an important case is $N = M$), and $f : M \longrightarrow M$ a C^r diffeomorphism, $r \geq 1$. Assume that $f(N) = N$. Following [HP70], we define an operator $\mathcal{L}_f$, acting on sufficiently small sections v of TM defined on N (i.e., $v \in C^0(N, TM)$), by

$$[\mathcal{L}_f v](x) = \exp_x^{-1} f(\exp_{f^{-1}(x)} v(f^{-1}(x))), \tag{16}$$

where exp is the exponential map of Riemannian geometry associated to a C^∞ metric. This operator is well defined since $f(N) = N$.

In a more suggestive way, we write formula (16) as

$$[\mathcal{L}_f v](f(y)) = f(y + v(y)) - f(y) \tag{17}$$

where, of course, by the sum of a point and a vector, we mean the exponential. In the particular case that M is a torus with the flat metric, (17) agrees with the usual sum of vectors in the torus.

We recall that

$$\mathcal{L}_f : U \subset C^0(N, TM) \longrightarrow C^0(N, TM)$$

is C^{r-1}, even when N is a Banach manifold (here, U is an open set formed by sufficiently small sections).

If N is a compact submanifold, then $\mathcal{L}_f$ is C^r (see [dlLO99] for these questions about regularity of composition operators). Throughout the rest of this section, we will assume that M and, therefore, N are finite dimensional manifolds.

Note that we also have

$$\begin{aligned} \mathcal{L}_f(0) &= 0 \\ D\mathcal{L}_f(0) &= f_* \end{aligned}$$

where $f_* \Delta\sigma = Tf \circ \Delta\sigma \circ f^{-1}$.

We also recall the following result from [Mat68] (see also [CL99]):

Theorem 2.1. *The map* $f_* : C^0(N, TM) \longrightarrow C^0(N, TM)$ *satisfies*

$$\operatorname{Spec}(f_*, C^0(N, TM)) \subset \bigcup_{i=1}^{n} \{z \in \mathbb{C} \mid \lambda_i^- \leq |z| \leq \lambda_i^+\} \tag{18}$$

for some integer n, *and some reals* $\lambda_i^- \leq \lambda_i^+ < \lambda_{i+1}^- \leq \lambda_{i+1}^+$, $i = 1, \ldots, n-1$.

Moreover, we can find a continuous decomposition

$$T_x M = \bigoplus_{i=1}^{n} E_x^i \tag{19}$$

such that

$$v \in E_x^i \quad \Leftrightarrow \quad \begin{cases} |Df^n(x)v| \leq c(\lambda_i^+ + \varepsilon)^n |v|, & n \geq 0, \\ |Df^n(x)v| \leq c(\lambda_i^- - \varepsilon)^n |v|, & n \leq 0, \end{cases}$$

for some constant c and some $\varepsilon > 0$ sufficiently small. Finally, if we denote by Π^i the spectral projections of f_* corresponding to $\{z \in \mathbb{C} \mid \lambda_i^- \le |z| \le \lambda_i^+\}$ and by P_x^i the projections corresponding to the bundles above, we then have

$$[\Pi^i v](x) = P_x^i v(x).$$

Remark 7. Even if $C^0(N, M)$ is a real Banach space, there is a canonical way to complexify it and to extend f_* to this complex space. Then, the spectrum of the operator refers to the complexified operator. Nevertheless, the spectral projections are real Banach spaces; see [Mat68] for more details.

Remark 8. Another theorem of [Mat68], which we will not use, in the formulation of our results, shows that if $f_{|N}$ is such that aperiodic orbits are dense, then $\mathrm{Spec}(f_*, C^0(N,TM))$ is indeed a union of annuli as in (18). Also, it is shown that, when $M = N$, f is Anosov if and only if $1 \notin \mathrm{Spec}(f_*, C^0(M,TM))$. Note that there cannot be more spectral gaps than the dimension of TM, since the existence of spectral gaps implies the existence of subbundles.

Remark 9. Even if the decomposition produced in Theorem 2.1 is only claimed there to be continuous, using the invariant section theorem [HP70] it is possible to show that it is actually more regular. The regularity depends on the numbers λ_i.

Similarly, the theory of normally hyperbolic manifolds shows that if $T_x N \subset \bigcup_{i=\alpha}^{\beta} E_x^i$, then the manifold N is actually more regular than C^1.

If we apply our Theorem 1.1 to $\mathcal{L}_f$, we obtain:

Theorem 2.2. *Let f and $\mathcal{L}_f$ be as described previously. Assume that $\mathcal{L}_f$ is C^r and that $\Sigma \subset \{1, \ldots, n\}$ is a subset such that*

1) $\lambda_i^+ < 1$ *for every* $i \in \Sigma$.
2) *There exists* $L \ge 1$ *such that* $(\lambda_\Sigma^+)^{L+1}(\lambda_1^-)^{-1} < 1$, *where* $\lambda_\Sigma^+ \equiv \max_{i\in\Sigma} \lambda_i^+$.
3) $$\Big(\bigcup_{i\in\Sigma}[\lambda_i^-, \lambda_i^+]\Big)^j \cap \Big(\bigcup_{i\notin\Sigma}[\lambda_i^-, \lambda_i^+]\Big) = \emptyset \quad \textit{for every } 2 \le j \le L\ .$$
4) $L + 1 \le r$.

Then, we can find a C^r manifold $W^\Sigma \subset C^0(N, TM)$ invariant under $\mathcal{L}_f$.

We now define

$$W_x^\Sigma = \{x + v(x) \mid v \in W^\Sigma\}.$$

The invariance of W^Σ under $\mathcal{L}_f$, and the fact that by definition of $\mathcal{L}_f$ we have

$$f(x + v(x)) = f(x) + \mathcal{L}_f v(f(x)), \tag{20}$$

imply that

$$f(W_x^\Sigma) \subset W_{f(x)}^\Sigma. \tag{21}$$

Moreover, since W_x^Σ is the image under the exponential map of W^Σ, then it is a C^r manifold. The dependence $x \mapsto W_x^\Sigma$ is C^0.

Therefore, we have produced a continuous family of C^r manifolds which is invariant under f in the sense of (21).

We also note that if $\lambda_\Sigma^+ \equiv \max_{i\in\Sigma} \lambda_i^+$, $\lambda_\Sigma^- \equiv \min_{i\in\Sigma} \lambda_i^-$ and we assume

$$c(\lambda_\Sigma^- - \varepsilon)^n |v| \le |\mathcal{L}_f^n v| \le c(\lambda_\Sigma^+ + \varepsilon)^n |v|, \qquad \text{for all } n \ge 0, v \in W^\Sigma \ ,$$

then

$$c(\lambda_\Sigma^- - \varepsilon)^n \le d(f^n(x), f^n(y)) \le c(\lambda_\Sigma^+ + \varepsilon)^n, \qquad \text{for all } n \ge 0, y \in W_x^\Sigma. \tag{22}$$

In case that $N \subset M$ and that Σ corresponds to the part of the spectrum closest to the origin, the previous construction reduces to the strong stable manifold for invariant submanifolds. In this case, (22) is not only consequence of $y \in W_x^\Sigma$, but also

$$d(f^n(x), f^n(y)) \le c(\lambda_\Sigma^+ + \varepsilon)^n, \qquad \text{for all } n \ge 0$$

is sufficient for $y \in W_x^\Sigma$.

Hence, when Σ is the part of the spectrum closest to the origin, the relation between x and y given by (22) is equivalent to $x \in W_y^\Sigma$, and also to $y \in W_x^\Sigma$. Since this is clearly an equivalence relation between x and y, we obtain that W_x^Σ constitute a foliation. This is the strong stable foliation.

When we take Σ to consist just of one band, a similar result proved by a different method can be found in [Pes73].

When Σ contains intermediate components, then (22) may not be equivalent to $x \in W_y^\Sigma$. Indeed, in [JPdlL95] it is shown that the W_x^Σ may fail to be a foliation.

In some cases ($M = \mathbb{T}^d$) it is possible to use other constructions [dlLW95] to associate invariant manifolds to invariant subspaces. The Irwin construction does indeed lead to a foliation. Nevertheless, the leaves are not very smooth.

The above construction of invariant foliations admits several extensions:

- As shown in [HP70], the map $\mathcal{L}_f$ can be defined even when N is a much more general set than a manifold. This allows to show that W_x^Σ is a manifold even when N is not a manifold.
- The map $x \mapsto W_x^\Sigma$ is more regular than just continuous. One can indeed show that it is Hölder continuous for some exponent.

In spite of the fact that the regularity consequences of the above method are not optimal, we hope that the painless way to construct these geometric objects out of our main theorem may serve as motivation for the study of these objects.

3. Proof of Theorem 1.1

The proof will proceed by showing first that, due to the non-resonance conditions, we can solve equation (4),

$$F \circ K = K \circ R \ ,$$

order by order in the sense of power series. We will then show that, once we have reached a high enough order (actually L), we can fix the polynomials $K^{\leq}$ and R so obtained and reduce the equation for the higher order terms to a fixed point problem for $K^{>}$ (where $K = K^{\leq} + K^{>}$), which can be solved by appealing to the standard contraction mapping theorem. This procedure will lead to the loss of one derivative in the conclusions, and we will need a separate argument to recover it.

In numerical applications, the computation to high enough order may provide sufficient precision and the iteration leading to a fixed point may be readily implementable. We emphasize the fact that one of the consequences of our analysis is that if we start the iteration leading to the fixed point with a polynomial approximation of an order which is not high enough, then the procedure may converge to a different, less differentiable solution.

It will also follow from the analysis that we present that if we compute functions which solve $F \circ K = K \circ R$ with good enough accuracy, then there is a true solution near by and the distance from the true solution to the computed one is bounded by the error incurred in solving the equation. Hence, the contraction mapping fixed point theorem leads to *a posteriori* error estimates for numerical methods (see Remark 13).

This section deals with the proof of the existence statements a) and b) of Theorem 1.1. The uniqueness statement c) of the theorem is a particular case of the uniqueness result of Theorem 1.2, which is proved later in Section 4.1 (recall that if all assumptions of Theorem 1.1 are satisfied then the ones of Theorem 1.2 are also satisfied for the same L and r). The dependence on parameters, statement d) in Theorem 1.1, will be postponed to Section 5, where we will prove a more general result.

Along the proofs we will need some results on spectral properties of operators in Banach spaces, that we present in Appendix A. In the finite dimensional case, these results are rather simple and well-known.

Sections 3.1 to 3.4 establish the existence of a solution $K \in C^{r-1}$ of (4). The proof that K is indeed C^r is given in Section 3.5 and consists in studying the equation satisfied by the first derivative or differential DK. Such equation has strong analogies with the one satisfied by K, (4). Knowing that K is already C^{r-1}, we will prove the existence of a solution $G \in C^{r-1}$ to the equation for DK. Then we will have that $DK = G$, by a uniqueness property. Hence, $DK = G \in C^{r-1}$ and we will conclude $K \in C^r$.

To simplify the proofs, we scale the maps involved in the equations. Following standard practice, given a real number $\delta > 0$ and a map H, we consider $H^\delta(x) = \frac{1}{\delta} H(\delta x)$. Note that (4) holds in the ball of radius δ if and only if

$$F^\delta \circ K^\delta = K^\delta \circ R^\delta$$

holds in the ball of radius 1. Moreover,

$$F^\delta = A + N^\delta \ ,$$

where N^δ satisfies $N^\delta(0) = 0$, $DN^\delta(0) = 0$ and that $\|N^\delta\|_{C^r}$ is arbitrarily small in the ball of radius 3 if we take δ sufficiently small. We also note that this change of scale does not affect conclusions (3), (5) and (6).

Therefore, rather than considering small balls, we will assume that $\|N\|_{C^r}$ is sufficiently small in the ball of radius 1. This has the advantage that it simplifies the spaces and also some arguments used in the proofs.

3.1. **Formal solution.** In this section we show that, under the non-resonance hypotheses, we can find polynomials $K^{\leq}$ and R satisfying

$$F \circ K^{\leq}(x) = K^{\leq} \circ R(x) + o(|x|^L) \ .$$

For this, we will use the device of complexification. The precise result is the following:

Lemma 3.1. *Assume that X is a real or complex Banach space, hypotheses 0), 1), 3) and 4) of Theorem 1.1, and $r \geq L$. Then,*

a) *We can find polynomials $K^{\leq} = \sum_{i=1}^{L} K_i$ and $R = \sum_{i=1}^{L} R_i$ of degree not larger than L, where K_i and R_i are multilinear maps of degree i, satisfying*

$$F \circ K^{\leq}(x) = K^{\leq} \circ R(x) + o(|x|^L) \ ,$$

and (3), (5), (6)*, i.e., $R(0) = 0$, $DR(0) = A_1$, $K^{\leq}(0) = 0$, $DK^{\leq}(0) =$* $(\mathrm{Id}, 0)$.

b) *If we further assume that*

$$\big(\mathrm{Spec}(A_1)\big)^j \cap \mathrm{Spec}(A_1) = \emptyset \qquad \textit{for some } j \textit{ with } 2 \leq j \leq L \ , \tag{23}$$

then we can choose $R_j = 0$.

c) *Let Λ be an open subset of a Banach space Λ_0 and let $F : \Lambda \times U \subset \Lambda_0 \times X \to X$, that we will write as $F_\lambda(x)$. Assume that the subspace X_1 (which is independent of λ) is invariant under $A_\lambda := D_x F_\lambda(0)$ for every λ, and that hypotheses 0), 1), 3), 4) above are satisfied uniformly in λ. Suppose that the map $\lambda \mapsto D_x^i F_\lambda(0)$ is C^l for every $i \leq L$. Then it is possible to find mappings $\lambda \mapsto K_i$ and $\lambda \mapsto R_i$ of class C^l. In particular, if $F_\lambda(x)$ is C^r in (λ, x) then it is possible to find K_i and R_i to be C^{r-i} in λ.*

Remark 10. In Remark 2 we have pointed out that, under all hypotheses of Theorem 1.1, the non-resonances assumptions 4) of Lemma 3.1 hold up to order r, and not only L. Hence, by Lemma 3.1 applied with L replaced by r, we can obtain polynomials K of degree not larger than r and R of degree not larger than L, such that $F \circ K(x) = K \circ R(x) + o(|x|^r)$.

Therefore, under the non-resonance conditions of Theorems 1.1 and 1.2 and provided that F is C^∞, the calculations can be continued up to any order. This is indeed a practical algorithm that has been implemented. Using Lemma 3.4 one can develop a-posteriori bounds for the computation (see Remark 13).

Remark 11. We will obtain K_i and R_i in a recursive way. Since the solution of the recursive equations will involve a right hand side that depends on N –the nonlinear part of $F-$ and we can assume that N is sufficiently C^L small (by the scaling procedure described above), we can ensure that all the coefficients of the polynomials $K^{\leq}$ and R of degree 2 or higher are arbitrarily close to zero. In particular the polynomial $K^{\leq}$, considered as a function, is arbitrarily close to the immersion of X_1 into X for any smooth norm on functions defined in the unit ball. Similarly, R is arbitrarily close to A_1. This remark plays an essential role in the calculations of the next sections and will be used throughout.

The main ingredient in the proof of Lemma 3.1 is the following:

Proposition 3.2. *Let X and Y be complex Banach spaces. Denote by $\mathbf{M}_n$ the space of n-multilinear maps on X taking values in Y, and by $\mathbf{S}_n$ the space of symmetric n-multilinear maps on X taking values in Y.*

Given bounded linear operators $A : X \to X$ and $B : Y \to Y$, consider the operators $\mathcal{L}_B$, $\mathcal{R}^k_A$ and $\mathcal{L}_{n,A,B}$ acting on $\mathbf{M}_n$ by

$$\begin{aligned}(\mathcal{L}_B\, M)(x_1,\dots,x_n) &= B\,M(x_1,\dots,x_n),\\ (\mathcal{R}^k_A\, M)(x_1,\dots,x_n) &= M(x_1,\dots x_{k-1},Ax_k,x_{k+1},\dots,x_n),\\ (\mathcal{L}_{n,A,B}M)(x_1,\dots,x_n) &= BM(Ax_1,\dots,Ax_n)\ .\end{aligned} \tag{24}$$

Note that $\mathcal{L}_{n,A,B}$ also acts on $\mathbf{S}_n$ by the same formula.

Then,

$$\begin{aligned}\operatorname{Spec}(\mathcal{L}_B,\mathbf{M}_n) &\subset \operatorname{Spec}(B,Y)\ ,\\ \operatorname{Spec}(\mathcal{R}^k_A,\mathbf{M}_n) &\subset \operatorname{Spec}(A,X)\ ,\\ \operatorname{Spec}(\mathcal{L}_{n,A,B},\mathbf{S}_n) \subset \operatorname{Spec}(\mathcal{L}_{n,A,B},\mathbf{M}_n) &\subset \operatorname{Spec}(B,Y)\big(\operatorname{Spec}(A,X)\big)^n\ .\end{aligned} \tag{25}$$

In case that X and Y are finite dimensional (or more generally, that the spectra of A and B are the closure of the set of their eigenvalues), all inclusions in (25) *are equalities.*

We give a proof of Proposition 3.2 in Appendix A, which essentially follows the one in [dlLW97].

The operators in (24) are called Sylvester operators in [BK98]. Their solution is crucial in the numerical study of stable invariant manifolds. We point out that there are examples arising naturally in dynamical systems in which inclusions (25) are strict (see [dlL98]).

Proof of Lemma 3.1. We first present the proof for complex Banach spaces, where we can use with ease spectral theory. At the end of the proof we will discuss the changes needed to deal with real Banach spaces.

Equating the derivatives of both sides of (4) evaluated at zero, we obtain that

$$AK_1 = K_1R_1\ .$$

We see that this equation is indeed satisfied if we choose K_1, R_1 as in Theorem 1.1, i.e., $K_1 = (\mathrm{Id}, 0)$ and $R_1 = A_1$. However, in general this is not the only possible choice!

Equating derivatives of order i at zero in (4) for $i > 1$, we obtain

$$AK_i + \Gamma_i = K_i A_1^{\otimes i} + K_1 R_i \ , \tag{26}$$

where Γ_i is a polynomial expression in K_j, R_j (with $j \leq i-1$) and in the derivatives of F at zero up to order i.

We study the system of equations (26) by induction on i, by considering (26) as an equation to be solved for K_i and R_i once $K_1, \dots, K_{i-1}$, $R_1, \dots, R_{i-1}$ –and therefore Γ_i– are known. So, we turn our efforts into studying the solvability of (26) considered as an equation for K_i and R_i when all the other terms are known.

Taking projections into X_1 and X_2, and using the notation $K_i^1 = \pi_1 K_i$, $K_i^2 = \pi_2 K_i$, etc., we see that (26) is equivalent to

$$A_1 K_i^1 - K_i^1 A_1^{\otimes i} - R_i = -\Gamma_i^1 - BK_i^2, \tag{27}$$

$$A_2 K_i^2 - K_i^2 A_1^{\otimes i} = -\Gamma_i^2 \ . \tag{28}$$

With the notation (24), equations (27) and (28) can be written as

$$\Big(\mathcal{L}_{A_1} - \mathcal{L}_{i,A_1,\mathrm{Id}}\Big) K_i^1 - R_i = -\Gamma_i^1 - BK_i^2 \ , \tag{29}$$

$$\Big(\mathcal{L}_{A_2} - \mathcal{L}_{i,A_1,\mathrm{Id}}\Big) K_i^2 = -\Gamma_i^2 \ . \tag{30}$$

The crux of the problem is the second equation (30). Once equation (30) is solved, we see that the first one can be solved, e.g., by taking $K_i^1 = 0$ and R_i in such a way that it matches all the rest. The fact that (28), that is (30), can be uniquely solved is a consequence of Proposition 3.2. Indeed, since by hypothesis 3) A_2 is invertible, we can write

$$\mathcal{L}_{A_2} - \mathcal{L}_{i,A_1,\mathrm{Id}} = \mathcal{L}_{A_2}\Big(\mathrm{Id} - \mathcal{L}_{i,A_1,A_2^{-1}}\Big)$$

Now, hypothesis 4) of Theorem 1.1 and Proposition 3.2 imply that both $\mathcal{L}_{A_2}$ and $\mathrm{Id} - \mathcal{L}_{i,A_1,A_2^{-1}}$ are invertible.

Now we turn to statement b) of the lemma. If we add the condition

$$\big(\operatorname{Spec}(A_1)\big)^j \cap \operatorname{Spec}(A_1) = \emptyset$$

at the level $i = j$, then we can choose $R_j = 0$ and solve uniquely K_j^1 in the first equation, (29), by writing

$$\mathcal{L}_{A_1} - \mathcal{L}_{j,A_1,\mathrm{Id}} = \mathcal{L}_{A_1}\Big(\mathrm{Id} - \mathcal{L}_{j,A_1,A_1^{-1}}\Big)$$

and using Proposition 3.2 (recall that A_1 is invertible since A and A_2 are assumed to be invertible).

To prove c), we note that the above procedures ensure the differentiability on parameters that we claimed for the coefficients of $K^{\leq}$ and R. Here we need to make the same choices (for all the values of λ) on which terms of equations (27) and (28) are eliminated (see Remark 12).

It is clear that the mappings that to A_1, A_2 associate $\mathcal{L}_{i,A_1,A_2}$ are analytic (in fact polynomial) so that if $A_{1,\lambda}$, $A_{2,\lambda}$ are C^l on λ then

$$\mathcal{L}_{A_{1,\lambda}},\ \mathcal{L}_{A_{2,\lambda}},\ \mathcal{L}_{i,A_{1,\lambda},A_{2,\lambda}^{-1}} \text{ and } \mathcal{L}_{i,A_{1,\lambda},A_{1,\lambda}^{-1}}$$

are also C^l on parameters as maps from Λ to $L(\mathbf{M}_i, \mathbf{M}_i)$. Therefore, if the right hand sides of (27) and (28) are C^l on parameters, then, K_i^1, K_i^2 and R_i are also C^l on parameters.

Finally, in case that our Banach space is real, we can reduce ourselves to the complex case by using the well known device of complexification. First note that the result will be proved if we use, in place of F, its r Taylor approximation

$$F^{[\leq r]}(x) = \sum_{j=0}^{r} \frac{1}{j!} D^j F(0) x^{\otimes j} \ .$$

Now, the space X can be complexified to $\tilde{X} = X + iX$ and the function $F^{[\leq r]}$, since it is a polynomial, can be complexified to a function in $\tilde{X}$. Note that the operators $\mathcal{L}_{i,A,B}$ behave well under complexification, that is, $\mathcal{L}_{i,\tilde{A},\tilde{B}} = \widetilde{\mathcal{L}_{i,A,B}}$. Then, it follows by induction that if all the $\tilde{R}_i$, $\tilde{K}_i$ preserve the real subspace X for $i \leq i_0$, then, the same is true for $\Gamma^1_{i_0}$, $\Gamma^2_{i_0}$ and therefore is also true for K_i and R_i chosen according to the prescriptions we have made explicit. ☐

As we have said in Remark 1, even in the case that the non-resonance assumptions were not satisfied, we could find solutions of the recursive equations (27) and (28) provided that we assumed conditions on their right hand sides. In the finite dimensional case, these conditions happen in a set of finite codimension.

Note however that, in case that the non-resonance conditions are not satisfied, the solution of K_i that we find will not be unique. In some cases, one is able to use this freedom to ensure that the equations of higher order will satisfy the solvability conditions. Hence, the codimension of the maps possessing invariant manifolds may be much smaller than what a naive count of parameters would give. We will not pursue this line of research here since it seems that it is best done for concrete examples.

Remark 12. Concerning the dependence on parameters, since we have many choices on which terms are eliminated with R and which with $K^{\leq}$, we could make different choices for different values of λ, and then obtain that the resulting $K^{\leq}$ and R are quite irregular with respect to λ. The regularity claimed in Lemma 3.1 is, of course, for the procedure that we described in which the same choices of which terms to eliminate are made for all the values of λ.

3.2. **Formulation of the problem as a fixed point problem.** From now on, $K^{\leq}$ and R are the polynomials of degree not larger than L obtained in Lemma 3.1, and we write

$$K = K^{\leq} + K^{>} \ .$$

We will show that it is possible to find $K^{>}$ such that $D^i K^{>}(0) = 0$ for $i \le L$, and such that $K = K^{\le} + K^{>}$ satisfies (4), which can be written as

$$AK^{\le} + AK^{>} + N \circ (K^{\le} + K^{>}) = K^{\le} \circ R + K^{>} \circ R$$

or in the equivalent form

$$AK^{>} - K^{>} \circ R = -N \circ (K^{\le} + K^{>}) - AK^{\le} + K^{\le} \circ R \ . \tag{31}$$

Note that the way that $K^{\le}$ and R are determined ensures that, since $K^{>}$ vanishes at 0 up to order L, all derivatives up to order L of the right hand side of (31) also vanish at the origin. Hence, we will consider (31) as a functional equation for C^r functions $K^{>}$ whose first L derivatives vanish at 0.

In Appendix A we prove that it is possible to substitute the norm in X by an equivalent one for which

$$\|A_1\| < 1 \ , \quad \|B\| \text{ is as small as necessary,}$$

and

$$\|A^{-1}\| \|A_1\|^{L+1} < 1 \quad \text{for Theorem 1.1}$$

or

$$\|A_2^{-1}\| \|A_1\|^{L+1} < 1 \quad \text{for Theorem 1.2.}$$

We use this norm throughout the rest of the paper.

By Remark 11 ue know that we can assume $K^{\le}$ to be arbitrarily close to the immersion of X_1 into X, and R to be arbitrarily close to A_1. Note also that, by taking the scaling parameter δ sufficiently small, we can also assume that $\|R^\delta - DR^\delta(0)\|_{C^r}$ is sufficiently small in the ball of radius 1. Hence, in order to solve equation (31), we may assume that R is approximately linear and, since $DR(0) = A_1$, R may be assumed to be a contraction which maps the unit ball into a ball of radius smaller than 1.

We also take the scaling parameter such that

$$\|DR\|_{C^0(B_1)} \le \|A_1\| + \varepsilon < 1 \ , \tag{32}$$

with ε small enough such that

$$\begin{aligned} \|DR\|^{L+1}_{C^0(B_1)} \|A^{-1}\| & \\ &\le (\|A_1\| + \varepsilon)^{L+1} \|A^{-1}\| \\ &< 1 \ . \end{aligned}$$

Finally, from the fact that R is a polynomial of degree L arbitrarily close to A_1, we easily deduce that

$$|D^k R^j(x)| \le C_k (\|A_1\| + \varepsilon)^j \tag{33}$$

with C_k independent on j (these bounds were already considered in page 574 of [dlLMM86] and in Lemma 5.4 of [BdlLW96]).

If we have a parameter family that satisfies the assumptions with uniform constants, it is possible to choose a scaling that reaches the above smallness conditions for all values of the parameters. In such a case, it is easy to see that if the parameter family is C^l when N is given the C^r topology, so is the scaled family.

3.3. **Study of the linearized problem.** Since N may be assumed to be small, to solve equation (31) we first study the linear operator

$$\mathcal{S}H = AH - H \circ R \tag{34}$$

acting on functions $H : B_1 \subset X_1 \longrightarrow X$, where B_1 is the unit ball and H belongs to the Banach space $\Gamma_{s,l}$ defined as follows. Given a Banach space Y (that in this subsection will coincide with X), given $s \in \mathbb{N} \cup \{\omega\}$ and $l \in \mathbb{N}$ with $s \geq l$, we consider

$$\begin{aligned}\Gamma_{s,l} \quad = \quad & \{H : B_1 \subset X_1 \longrightarrow Y \mid H \in C^s(B_1), D^k H(0) = 0 \text{ for } 0 \leq k \leq l,\\ & \sup_{x \in B_1} \big(|D^l H(x)|/|x|\big) < \infty\}\end{aligned}$$

equipped with the norm

$$\|H\|_{\Gamma_{s,l}} := \max\Big\{\|H\|_{C^0(B_1)}, \dots, \|D^s H\|_{C^0(B_1)}, \sup_{x\in B_1} \big(|D^l H(x)|/|x|\big)\Big\}$$

if $s \in \mathbb{N}$, and with the norm

$$\|H\|_{\Gamma_{\omega,l}} := \|D^{l+1}H\|_{C^0(B_1)}$$

if $s = \omega$. We emphasize that, in the case $s = \omega$, we take the supremum norm on B_1 where, if the space X is real then B_1 is the unit ball of $X_1 \oplus iX_1$ in the complexified space $X \oplus iX$. Note here that if X is real and a function is (real) analytic in a neighborhood of 0 in X, then it can be extended to be (complex) analytic in a ball around 0 in the complexified space $X \oplus iX$. Then, choosing the scaling parameter δ small enough, we may (and do) assume that such ball in $X \oplus iX$ is the unit ball. It is also clear that $\|D^{l+1}H\|_{C^0(B_1)}$ is a norm in $\Gamma_{\omega,l}$, since these functions have all derivatives at 0 up to order l equal to zero.

We have that $\Gamma_{s,l}$ is a Banach space. Note that the term

$$\sup_{x\in B_1} \big(|D^l H(x)|/|x|\big)$$

included in the definition of $\|\cdot\|_{\Gamma_{s,l}}$ is relevant only whenever $s = l$ and it could be omitted (since $D^l H(0) = 0$) when $s > l$.

The rest of this section is dedicated to proving the following result. Here $Y = X$ in the definition of $\Gamma_{s,L}$.

Lemma 3.3. *Under the assumptions of Theorem 1.1 and under the standing assumptions arranged by scaling at the beginning of Section 3, if $s \in \mathbb{N} \cup \{\omega\}$ then $\mathcal{S} : \Gamma_{r-1,L} \longrightarrow \Gamma_{r-1,L}$ is a bounded invertible operator. Moreover, $\|S^{-1}\|$ can be bounded by a constant independent of the scaling parameter. Obviously, if $r = \omega$ then $\Gamma_{r-1,L} = \Gamma_{\omega,L}$.*

This result is a simplified version of Theorem 5.1 in [BdlLW96].

Proof of Lemma 3.3. Note that $r-1 \geq L$. It is easy to verify $\mathcal{S}$ is a bounded operator from $\Gamma_{r-1,L}$ into itself.

Next, we need to show that given $\eta \in \Gamma_{r-1,L}$, we can find a unique $H \in \Gamma_{r-1,L}$ such that

$$\mathcal{S}H = \eta. \tag{35}$$

We will also see that $\|H\|_{\Gamma_{r-1,L}} \leq C\|\eta\|_{\Gamma_{r-1,L}}$.

The equation (35) for H is equivalent to

$$H = A^{-1}H \circ R + A^{-1}\eta \ . \tag{36}$$

We claim that the solution of (36) is given by

$$H = \sum_{j=0}^{\infty} A^{-(j+1)}\eta \circ R^j \ . \tag{37}$$

To establish the claim, we will show that the series in (37) converges absolutely in $\Gamma_{r-1,L}$ and that

$$\sum_{j=0}^{\infty} \|A^{-(j+1)}\eta \circ R^j\|_{\Gamma_{r-1,L}} \leq C\|\eta\|_{\Gamma_{r-1,L}}, \qquad \forall \eta \in \Gamma_{r-1,L} \ . \tag{38}$$

In particular, when substituted in (36), we can rearrange the terms and show that H is indeed a solution.

The previous fact also establishes the uniqueness of solution. Indeed, if $\eta = 0$, by (36) we have $H = A^{-1}H \circ R$, and hence $H = A^{-j}H \circ R^j$ for every $j \geq 1$. But $\|A^{-j}H \circ R^j\|_{\Gamma_{r-1,L}} \to 0$ as $j \to \infty$, by (38) applied with $\eta = H$. We conclude that $H = 0$.

We now establish (38) when $n \in \mathbb{N}$. Since $\eta \in \Gamma_{r-1,L}$, we have $|D^L\eta(\xi)| \leq \|\eta\|_{\Gamma_{r-1,L}}|\xi|$ for $\xi \in B_1$, and $\eta(0) = 0, \ldots, D^L\eta(0) = 0$. Hence, by Taylor's formula, we have that $|\eta(y)| \leq C\|\eta\|_{\Gamma_{r-1,L}}|y|^{L+1}$, for $y \in B_1$. Moreover, $\mathrm{Lip}(R) \leq \|A_1\| + \varepsilon$ where $\varepsilon > 0$ is small by the scaling argument. We conclude that

$$|\eta \circ R^j(x)| \leq C\|\eta\|_{\Gamma_{r-1,L}}(\|A_1\| + \varepsilon)^{(L+1)j}|x|^{L+1} \ .$$

Since $\|A^{-1}\|(\|A_1\| + \varepsilon)^{L+1} < 1$, we deduce that the (37) converges absolutely in the C^0 norm.

Now we turn to the estimates in $\Gamma_{r-1,L}$. By the Faa-di-Bruno formula, we have

$$D^k(\eta \circ R^j) = \sum_{i=0}^{k} \sum_{\substack{1 \leq k_1, \ldots, k_i \leq k \\ k_1 + \cdots + k_i = k}} \sigma^{i,k}_{k_1,\ldots,k_i}([D^i\eta] \circ R^j)\, D^{k_1}R^j \cdots D^{k_i}R^j, \tag{39}$$

where $\sigma^{i,k}_{k_1,\ldots,k_i}$ is an explicit combinatorial coefficient. We recall that, by (33),

$$|D^kR^j(x)| \leq C(\|A_1\| + \varepsilon)^j, \qquad x \in B_1 \ ,$$

where C is independent of j. Moreover, for $i \le r-1$, we have

$$\begin{aligned} |[D^i\eta] \circ R^j(x)| &\le C\|\eta\|_{\Gamma_{r-1,L}} |R^j(x)|^{(L+1-i)_+} \\ &\le C\|\eta\|_{\Gamma_{r-1,L}} (\|A_1\| + \varepsilon)^{j(L+1-i)_+} |x|^{(L+1-i)_+} \end{aligned} \tag{40}$$

where we have used the notation $t_+ = \max(t, 0)$ to treat simultaneously the cases $i \le L+1$ and $i > L+1$.

Using these bounds and (39), we deduce that

$$\begin{aligned} |D^k[A^{-(j+1)}\eta \circ R^j](x)| &\le C\|\eta\|_{\Gamma_{r-1,L}} \sum_{i=1}^{k} \|A^{-1}\|^j \\ &\quad \cdot (\|A_1\| + \varepsilon)^{j[(L+1-i)_+ + i]} |x|^{(L+1-i)_+} \\ &\le C\|\eta\|_{\Gamma_{r-1,L}} [\|A^{-1}\|(\|A_1\| + \varepsilon)^{L+1}]^j |x|^{(L+1-i)_+} \ . \end{aligned}$$

Taking the supremum in B_1 of this expression for $k \le r-1$, and also the supremum of the L-derivative divided by $|x|$, we conclude (38). This establishes the result when $r \in \mathbb{N}$.

In the analytic case $r = \omega$, we first note that, using $\eta \in \Gamma_{\omega,L}$ and Taylor's theorem, we have $\|D^i\eta\|_{C^0(B_1)} \le C\|D^{L+1}\eta\|_{C^0(B_1)} = C\|\eta\|_{\Gamma_{\omega,L}}$ for every $0 \le i \le L+1$. Hence, by (39) with $k = L+1$,

$$\|D^{L+1}[A^{-(j+1)}\eta \circ R^j\|_{C^0(B_1)} \le C\|\eta\|_{\Gamma_{\omega,L}} [\|A^{-1}\|(\|A_1\| + \varepsilon)^{(L+1)}]^j$$

with C independent of j. We therefore conclude (38). ☐

3.4. **Solution of the fixed point problem.** We want to solve equation (31), that can be rewritten using the operator $\mathcal{S}$ introduced in (34) as

$$\mathcal{S}K^> = -N \circ (K^\le + K^>) - AK^\le + K^\le \circ R \ .$$

We have shown in Lemma 3.3 that $\mathcal{S}$ is invertible in $\Gamma_{r-1,L}$. Note that the way we determined $K^\le$ and R in Lemma 3.1 ensures that $-N \circ (K^\le + K^>) - AK^\le + K^\le \circ R$ vanishes up to order L at the origin whenever $K^> \in \Gamma_{r-1,L}$. Hence, we can rewrite $F \circ K = K \circ R$ as

$$K^> = \mathcal{T}(K^>) \ ,$$

where $\mathcal{T}$ is defined by

$$\mathcal{T}(K^>) = \mathcal{S}^{-1}\big[-N \circ (K^\le + K^>) - AK^\le + K^\le \circ R \,\big] \ . \tag{41}$$

Since we are assuming that N is C^r small, it will be easy to show (this is the content of Lemma 3.4 below) that $\mathcal{T}$ is a contraction when $K^>$ is given the $\Gamma_{r-1,L}$ topology. This will establish Theorem 1.1 for $r \in \mathbb{N} \cup \{\omega\}$, but with one less derivative in the conclusions when $r \in \mathbb{N}$. The C^∞ case is treated at the end of this subsection. Finally, for $r \in \mathbb{N}$ a separate argument (given in Section 3.5) will allow us to recover the last derivative and finish the proof of parts a) and b) of Theorem 1.1 as stated.

Lemma 3.4. *Under the hypotheses of Theorem 1.1 and under the standing assumptions arranged by scaling at the beginning of Section 3, if* $s \in \mathbb{N} \cup \{\omega\}$ *then* $\mathcal{T}$ *sends the closed unit ball* $\bar{B}_1^{r-1}$ *of* $\Gamma_{r-1,L}$ *into itself, and it is a contraction in* $\bar{B}_1^{r-1}$ *with the* $\Gamma_{r-1,L}$ *norm. Therefore* $\mathcal{T}$ *has a fixed point* $K^>$ *in the closed unit ball of* $\Gamma_{r-1,L}$.

Equation $K^> = \mathcal{T}(K^>)$ can also be studied in a very concise way by appealing to the standard implicit function theorem in Banach spaces. Even if this leads to a shorter proof (see, e.g., the expository work [CFdlL00]) it gives less differentiability for the solution. Of course, the standard proof of the implicit function theorem reduces to a contraction mapping theorem.

Remark 13. Out of the proof Lemma 3.4, we will obtain rather explicit bounds for the contraction constant which are rather good, especially for analytic functions.

We can use these bounds to obtain *a posteriori* estimates for the validity of numerical calculations. Note that from

$$\mathrm{Lip}(\mathcal{T}) \le \kappa < 1 \quad \text{and} \quad \|\mathcal{T}(K_0^>) - K_0^>\| \le \delta$$

for some computed $K_0^>$, one can deduce that there exists a fixed point $K^>$ satisfying

$$\|K^> - K_0^>\| \le \delta/(1-\kappa) \ .$$

Proof of Lemma 3.4. By Remark 11 we can assume that $K^{\le}$ is arbitrarily close to the identity and R is arbitrarily close to A_1. Therefore, if $K^>$ is in the closed unit ball of $\Gamma_{r-1,L}$, the image of the unit ball B_1 in X_1 by $K^{\le} + K^>$ is contained in the ball of radius 3. Hence, the composition $N \circ (K^{\le} + K^>)$ is well defined and of class C^{r-1}.

We first show the contraction property for $\mathcal{T}$. For this, note that the derivative of $\mathcal{T}$ guessed by manipulating the system formally is

$$D\mathcal{T}(K^>)\Delta = -\mathcal{S}^{-1} DN \circ (K^{\le} + K^>)\Delta \ , \tag{42}$$

but the notation $D\mathcal{T}$ is only formal and should not be taken to imply that $D\mathcal{T}$ is a derivative in the Fréchet sense (see Remark 14 below). Nevertheless, for $K^>$ and $K^> + \Delta$ in the closed unit ball of $\Gamma_{r-1,L}$, we trivially have the finite increments formula

$$\begin{aligned} \mathcal{T}(K^> + \Delta) - \mathcal{T}(K^>) &= \int_0^1 \frac{d}{ds}\Big[\mathcal{T}(K^> + s\Delta)\Big]\, ds \\ &= -\int_0^1 \mathcal{S}^{-1} DN \circ (K^{\le} + K^> + s\Delta)\Delta \, ds \ . \end{aligned} \tag{43}$$

Taking derivatives of this expression up to order r and using that $\mathcal{S}^{-1}$ maps $\Gamma_{r-1,L}$ into itself, we deduce that $\mathcal{T}(K^> + \Delta) - \mathcal{T}(K^>) \in \Gamma_{r-1,L}$. Next, for $r \in \mathbb{N}$ we take derivatives of (43) up to order $r-1$ and their supremum in B_1, and also the supremum of the L-derivative at x divided by $|x|$. When $r = \omega$, we simply take the supremum of (43) on B_1. In both cases, we deduce

$$\|\mathcal{T}(K^> + \Delta) - \mathcal{T}(K^>)\|_{\Gamma_{r-1,L}} \le C\|N\|_{C^r}\|\Delta\|_{\Gamma_{r-1,L}} \ .$$

Using that $\|N\|_{C^r}$ is small, this proves the contraction property for $\mathcal{T}$ with the $\Gamma_{r-1,L}$ norm.

It remains to prove that $\mathcal{T}$ sends the closed unit ball of $\Gamma_{r-1,L}$ into itself. For this, note that if $\|K^>\|_{\Gamma_{r-1,L}} \le 1$ then

$$\begin{aligned}\mathcal{T}(K^>) &= \mathcal{T}(0) + \{\mathcal{T}(K^>) - \mathcal{T}(0)\} \\ &= \{\mathcal{S}^{-1}[K^\le \circ R - F \circ K^\le]\} + \{\mathcal{T}(K^>) - \mathcal{T}(0)\} \ .\end{aligned}$$

By Lemmas 3.1 and 3.3, we know that the first term in the last expression has derivatives at 0 up to order L equal to zero, and that it has small $\Gamma_{r-1,L}$ norm after scaling. On the other hand, in the previous proof of the contraction property we have seen that $\mathcal{T}(K^>) - \mathcal{T}(0)$ has derivatives at 0 up to order L equal to zero, and $\|\mathcal{T}(K^>) - \mathcal{T}(0)\|_{\Gamma_{r-1,L}} \le \nu\|K^>\|_{\Gamma_{r-1,L}} \le \nu < 1$. We conclude that $\mathcal{T}$ maps the closed unit ball of $\Gamma_{r-1,L}$ into itself. □

Remark 14. The reason why (43) does not imply that $\mathcal{T}$ is differentiable is because the map $K^> \mapsto DN \circ (K^\le + K^>)$ in (42) could fail to be continuous as a function of $K^>$ (some examples of this are constructed in [dlLO99]). In finite dimensional spaces, or more generally if D^rN is uniformly continuous (e.g., if it is $C^{r+\delta}$), it can be shown that $\mathcal{T}$ is C^1 on $\Gamma_{r-1,L}$.

We also emphasize that the operator $\mathcal{T}$ in (41) is better behaved than the operator appearing in the usual graph transform proofs (see Section 4 below and also the pedagogical expositions in [LI73] and [LI83]), which usually involves the composition of a function with an expression that depends on the function itself. The graph transform operator cannot be differentiable in any C^r space whenever $r \in \mathbb{N}$, the essential reason being that the model map $K \in C^r \mapsto K \circ K \in C^r$ is not differentiable when $r \in \mathbb{N}$. Several problems will be overcome in the proof of Theorem 1.2.

Finally, we deal with the case $F \in C^\infty$. According to Lemma 3.4, for any $r \in \mathbb{N}$ with $r \ge L+1$, if $\|N\|_{C^r}$ is sufficiently small there is a C^{r-1} invariant manifold parameterized by K. Note that given r and r', if the required smallness conditions are simultaneously verified then the parameterization K coincides for both values (since both parameterizations are a fixed point of the same contraction). Looking at F before the scaling is made, this amounts to saying that there exists ρ_r such that K is C^{r-1} in the ball of radius ρ_r of X_1.

In particular K is defined and C^1 in the ball of radius ρ_1. Now, the key point is that equation $F \circ K = K \circ R$ leads to

$$K = F^{-j} \circ K \circ R^j \qquad \text{for every } j \ge 1 \ . \tag{44}$$

Therefore, since R is a contraction, this equality (with j large) allows to recover K in the ball of radius ρ_1 from K restricted to the ball of radius ρ_r. Therefore K is C^{r-1} in a fixed ball, for every $r \in \mathbb{N}$. That is, K is in C^∞.

3.5. **Sharp regularity.** Here we improve the previous result to obtain the C^r –not just C^{r-1}– differentiability for $K^>$ claimed in Theorem 1.1. Therefore, in this section we always have $r \in \mathbb{N}$.

Proposition 3.5. *The function $K^>$ produced in Lemma 3.4 (under perhaps stronger smallness conditions on $\|N\|_{C^r}$) is C^r.*

The proof of this result is based on differentiating the equation for $K^> \in \Gamma_{r-1,L}$,

$$AK^> - K^> \circ R = -N \circ (K^\leq + K^>) - AK^\leq + K^\leq \circ R \ ,$$

to obtain

$$\begin{aligned} ADK^> - DK^> \circ R\ DR = -DN{\circ}(K^\leq + K^>)(DK^\leq + DK^>) \\ - ADK^\leq + DK^\leq \circ R\ DR \ , \end{aligned} \tag{45}$$

with $DK^> : B_1 \subset X_1 \longrightarrow \mathcal{L}(X_1, X)$. Taking $Y = \mathcal{L}(X_1, X)$ in the definition of the spaces $\Gamma_{s,l}$ in Section 3.3, we study equation (45) for $G := DK^> \in \Gamma_{r-2,L-1}$, that we can rewrite as

$$\tilde{\mathcal{S}} DK^> = \tilde{\mathcal{T}} DK^> + U \ , \tag{46}$$

where

$$\tilde{\mathcal{S}} G = AG - G \circ R\ DR \ ,$$

$$\tilde{\mathcal{T}} G = -DN \circ (K^\leq + K^>)\ G$$

and

$$U = -DN \circ (K^\leq + K^>) DK^\leq - ADK^\leq + DK^\leq \circ R\ DR \ . \tag{47}$$

Note that in the definitions of $\tilde{\mathcal{T}}$ and U we take $K^> \in \Gamma_{r-1,L}$ to be the solution found in the previous section and, in this way, we look at $\tilde{\mathcal{T}}$ as a linear operator acting on G.

We will show that equation $\tilde{\mathcal{S}} G_0 = \tilde{\mathcal{T}} G_0 + U$ admits a solution $G_0 \in C^{r-1}$. Then, by a uniqueness property and (46), we will deduce that $DK^> = G_0 \in C^{r-1}$ and hence $K^> \in C^r$. The argument will be based on the following:

Lemma 3.6. *Under the hypotheses of Theorem 1.1 and under the standing assumptions arranged by scaling at the beginning of Section 3, if $s \in \mathbb{N}$ and $L - 1 \leq s \leq r - 1$ then $\tilde{\mathcal{S}}$ and $\tilde{\mathcal{T}}$ are bounded linear operators from $\Gamma_{s,L-1}$ into itself. Moreover, taking $\|N\|_{C^r}$ sufficiently small, $\tilde{\mathcal{S}}$ is invertible and $\|\tilde{\mathcal{S}}^{-1}\| \|\tilde{\mathcal{T}}\| < 1$.*

Using this lemma with $s = r - 2 \geq L - 1$ and with $s = r - 1$, we can finish the

Proof of Proposition 3.5. Since $DK^> \in \Gamma_{r-2,L-1}$, Lemma 3.6 applied with $s = r - 2$ gives that $\tilde{\mathcal{S}} DK^>$ and $\tilde{\mathcal{T}} DK^>$ also belong to $\Gamma_{r-2,L-1}$. Hence, since $DK^>$ is a solution of (46), i.e. $\tilde{\mathcal{S}}DK^> - \tilde{\mathcal{T}}DK^> = U$, we deduce that $U \in \Gamma_{r-2,L-1}$ and

$$(\mathrm{Id} - \tilde{\mathcal{S}}^{-1}\tilde{\mathcal{T}})DK^> = \tilde{\mathcal{S}}^{-1}U \ .$$

Recall that $\|\tilde{\mathcal{S}}^{-1}\tilde{\mathcal{T}}\| \le \|\tilde{\mathcal{S}}^{-1}\|\|\tilde{\mathcal{T}}\| < 1$ by the lemma. Hence

$$DK^> = \sum_{j=0}^{\infty} (\tilde{\mathcal{S}}^{-1}\tilde{\mathcal{T}})^j \tilde{\mathcal{S}}^{-1}U \qquad \text{in } \Gamma_{r-2,L-1} \ . \tag{48}$$

Now, by the definition (47) of U we have that $U \in \Gamma_{r-1,L-1}$. Moreover, $\Gamma_{r-1,L-1} \subset \Gamma_{r-2,L-1}$ and, by Lemma 3.6 applied now with $s = r-1$, the operators $\tilde{\mathcal{S}}^{-1}$ and $\tilde{\mathcal{T}}$ send $\Gamma_{r-1,L-1}$ into itself. Since the series (48) is convergent in $\Gamma_{r-1,L-1}$, we conclude that $DK^> \in \Gamma_{r-1,L-1} \subset C^{r-1}$ and hence that $K^> \in C^r$. □

Finally we give the

Proof of Lemma 3.6. The statements about the operator $\tilde{\mathcal{T}}G = -DN \circ (K^{\le} + K^>)\ G$ are easily proved. Indeed, if $G \in \Gamma_{s,L-1}$ with $L-1 \le s \le r-1$, then $\tilde{\mathcal{T}}G \in C^s$ and its derivatives at the origin up to order $L-1$ vanish. In addition $|D^{L-1}(\tilde{\mathcal{T}}G)(x)|/|x|$ is bounded, and hence $\tilde{\mathcal{T}}G \in \Gamma_{s,L-1}$. Moreover, $\|\tilde{\mathcal{T}}\|$ is small if $\|N\|_{C^r}$ is sufficiently small (just note that, as pointed out in the previous section, $K^{\le} + K^>$ remains in the ball of radius 3 of $\Gamma_{r-1,L}$ independently of the scaling).

We now study the operator $\tilde{\mathcal{S}}G = AG - G \circ R\ DR$, which is clearly bounded from $\Gamma_{s,L-1}$ into itself. Given $\eta \in \Gamma_{s,L-1}$ we need to establish the existence and uniqueness of a solution in $\Gamma_{s,L-1}$ for the equation

$$AG - G \circ R\ DR = \eta \ . \tag{49}$$

For this, we proceed as in the proof of Lemma 3.3 for the operator $\mathcal{S}$. It is clear that the series

$$G = \sum_{j=0}^{\infty} A^{-(j+1)} \eta \circ R^j\ DR^j$$

is formally the solution of (49). To finish the proof we only need to show that

$$\sum_{j=0}^{\infty} \|A^{-(j+1)} \eta \circ R^j\ DR^j\|_{\Gamma_{s,L-1}} \le C\|\eta\|_{\Gamma_{s,L-1}} \tag{50}$$

for some constant C independent of the scaling.

For $0 \le l \le s$, we have

$$
\begin{aligned}
&|D^l(A^{-(j+1)}\eta \circ R^j \; DR^j)(x)| \\
&\qquad \le C\|A^{-1}\|^j \sum_{k=0}^{l} |D^k(\eta \circ R^j)(x)||D^{(l-k)} DR^j(x)| \\
&\qquad \le C\|A^{-1}\|^j(\|A_1\|+\varepsilon)^j \sum_{k=0}^{l} |D^k(\eta \circ R^j)(x)| \; , \qquad (51)
\end{aligned}
$$

by (33). Since $\eta \in \Gamma_{s,L-1}$, for $0 \le i \le s$ we have $|D^i\eta(y)| \le C\|\eta\|_{\Gamma_{s,L-1}}|y|^{(L-i)_+}$, and hence

$$
|(D^i\eta)\circ R^j(x)| \le C\|\eta\|_{\Gamma_{s,L-1}}(\|A_1\|+\varepsilon)^{j(L-i)_+}|x|^{(L-i)_+} \; .
$$

Combined with the Faa-di-Bruno formula (39) and the bound (33) , this leads to

$$
\begin{aligned}
|D^k(\eta \circ R^j)(x)| &\le C\|\eta\|_{\Gamma_{s,L-1}} \sum_{i=0}^{k} (\|A_1\|+\varepsilon)^{j[(L-i)_+ + i]}|x|^{(L-i)_+} \\
&\le C\|\eta\|_{\Gamma_{s,L-1}}(\|A_1\|+\varepsilon)^{jL}|x|^{(L-k)_+} \; .
\end{aligned}
$$

Using (51), we finally arrive at

$$
|D^l(A^{-(j+1)}\eta \circ R^j \; DR^j)(x)| \le C\|\eta\|_{\Gamma_{s,L-1}} \left[\|A^{-1}\|(\|A_1\|+\varepsilon)^{(L+1)}\right]^j |x|^{(L-l)_+} \; .
$$

Taking the supremum of this expression in B_1 for $0 \le l \le s$, and also the supremum of the $(L-1)$-derivative divided by $|x|$, we conclude (50). $\square$

4. Proof of Theorem 1.2

Equation $F \circ K = K \circ R$ can be written in components as

$$
\begin{aligned}
A_1K^1 + N_1 \circ K + BK^2 &= K^1 \circ R \; , \\
A_2K^2 + N_2 \circ K &= K^2 \circ R \; .
\end{aligned}
$$

If we now decide to solve this equation by setting $K^1 = \mathrm{Id}$ and then determine R from the first equation, we obtain that the second one becomes

$$
A_2K^2 + N_2 \circ (\mathrm{Id}, K^2) = K^2 \circ \Big(A_1 + N_1 \circ (\mathrm{Id}, K^2) + BK^2\Big)
$$

which, since A_2 is invertible, we can rewrite as a fixed point problem

$$
K^2(x) = A_2^{-1}\Big[K^2\Big(A_1x + N_1(x, K^2(x)) + BK^2(x)\Big) - N_2(x, K^2(x))\Big] \; . \qquad (52)
$$

The reader will recognize immediately that, when $B = 0$, this is the customary functional equation that appears in the graph transform methods (see e.g. [LI83]). In our case, besides including the term with $B \ne 0$, in addition we are not assuming that the decomposition $X = X_1 \oplus X_2$ corresponds to spectral projections in a disk and its complement.

Note that if there is any manifold tangent to X_1, by the implicit function theorem it can be written in a unique way as the graph of a function from X_1 to

X_2. Then, the usual manipulations in the graph transform method lead to the fact that this function must satisfy equation (52) (see the uniqueness argument at the end of Section 4.1 below).

The study of the equation under non-resonance conditions were undertaken for $B = 0$ in [dlL97] by reducing the non-linear terms to a particularly simple form. In this paper we will follow a different route which follows closely both numerical implementations and the proof of Theorem 1.1. We will use, therefore, some lemmas of the previous section.

In a first step, we use Lemma 3.1 to show that, under the non-resonance conditions included in the theorem and having prescribed $K^2(0) = 0$ and $DK^2(0) = 0$, it is possible to determine uniquely $D^iK^2(0)$ for $2 \le i \le L$ from the requirement that the first L derivatives at zero of (52) have to match.

The second step shows that the fixed point equation $K^2 = \mathcal{N}(K^2)$, with $\mathcal{N}$ defined by

$$[\mathcal{N}K^2](x) = A_2^{-1}\left[K^2\left(A_1x + N_1(x, K^2(x)) + BK^2(x)\right) - N_2(x, K^2(x))\right] , \tag{53}$$

has a unique solution in the class of Lipschitz functions claimed in Theorem 1.2.

In a third step, we will show that there is a fixed point of the operator $\mathcal{N}$ of class C^{r-1}. We point out that, unfortunately, the operator $\mathcal{N}$ defined in (53) is not differentiable acting on C^l spaces. Nevertheless, as in the methods based on graph transform, we will be able to show that the operator leaves invariant a set of functions with bounded derivatives of order up to r and that it is a contraction on the $\Gamma_{r-1,L}$ norm there –for this we will consider the associated operator acting on $K^{>,2} = K^2 - K^{\le,2}$, where $K^{\le,2}$ will be the polynomial of degree L given by Lemma 3.1.

A fourth step will improve the regularity of the function to conclude that it is C^r. A fifth step, that we postpone to Section 5, will establish differentiability with respect to parameters.

4.1. **Uniqueness of the solution of the fixed point equation.** To simplify notation, we write $K^2 = w : B_1 \subset X_1 \longrightarrow X_2$. We need to solve equation $w = \mathcal{N}(w)$, where $\mathcal{N}$ is given by (53) and that, with this notation, becomes

$$\mathcal{N}(w) = A_2^{-1}\left[w \circ \psi_w - N_2 \circ (\mathrm{Id}, w)\right] \tag{54}$$

with

$$\psi_w = A_1 + N_1 \circ (\mathrm{Id}, w) + Bw \ . \tag{55}$$

We start solving $w = \mathcal{N}(w)$ up to order L. In the proof of Lemma 3.1 we have seen that there exists a unique polynomial $K^{\le} = \sum_{i=1}^{L} K_i$ of degree L such that $K_1 = (\mathrm{Id}, 0)$, $\pi_1 K^{\le} = K^{\le,1} = \mathrm{Id}$ (that is, $K_i^1 = 0$ for $2 \le i \le L$) and such that $F \circ K^{\le}(x) = K^{\le} \circ R(x) + o(|x|^L)$ for some polynomial R of degree L with $R_1 = A_1$. We know that, having set $K^{\le,1} = \mathrm{Id}$, $F \circ K^{\le}(x) = K^{\le} \circ R(x) + o(|x|^L)$

is equivalent to $K^{\leq,2}(x) = \mathcal{N}(K^{\leq,2})(x) + o(|x|^L)$. Hence, setting

$$w^{\leq} = \pi_2 K^{\leq} = K^{\leq,2} \ ,$$

we have:

Lemma 4.1. *Assume that the non-resonance conditions of Theorem 1.2 are satisfied. Then, there exists a unique polynomial $w^{\leq}$ of degree L such that*

$$w^{\leq}(x) = \mathcal{N}(w^{\leq})(x) + o(|x|^L) \ , \tag{56}$$

$$w^{\leq}(0) = 0 \ , \qquad Dw^{\leq}(0) = 0 \ . \tag{57}$$

Moreover, with a suitable scaling of F we can get $w^{\leq}$ as small as we want. Finally, if F_λ is C^r in (λ, x), then $w^{\leq}_\lambda$ is C^{r-L}.

Note that, in particular, all the C^L functions w that satisfy

$$w(x) = \mathcal{N}(w)(x) + o(|x|^L)$$

as well as (57) must have their derivatives at zero up to order L equal to those of $w^{\leq}$. Hence, defining $w^{>}$ by

$$w = w^{\leq} + w^{>}$$

(where $w^{\leq}$ is the polynomial produced in Lemma 4.1 above), we have that $w^{>}$ and its derivatives up to order L vanish at 0, and the fixed point equation $w = \mathcal{N}(w)$ becomes

$$w^{>} = \mathcal{M}(w^{>}) \ , \tag{58}$$

where

$$\mathcal{M}(w^{>}) = -w^{\leq} + \mathcal{N}(w^{\leq} + w^{>}) \ . \tag{59}$$

Before dealing with the existence of invariant manifolds, we prove in this section the uniqueness statement of Theorem 1.2. As a consequence we will also obtain the uniqueness result of Theorem 1.1. The key ingredient is the following:

Lemma 4.2. *Under the hypotheses of Theorem 1.2 (and with $\|B\|$ and $\|N\|_{C^r}$ small enough after scaling), equation $w^{>} = \mathcal{M}(w^{>})$ has at most one (in a local sense around the origin) Lipschitz solution $w^{>} : B_1 \subset X_1 \to X_2$ such that $\sup_{x \in B_1}(|w^{>}(x)|/|x|^{L+1}) < \infty$.*

Proof. Consider two solutions $w_i = w^{\leq} + w_i^{>}$, $i = 1, 2$. Note that $w_2^{>} - w_1^{>} = w_2 - w_1$ and $\mathcal{M}(w_2^{>}) - \mathcal{M}(w_1^{>}) = \mathcal{N}(w_2) - \mathcal{N}(w_1)$. We introduce the quantity

$$[w^{>}]_{L+1} = \sup_{x \in B_1} \frac{|w^{>}(x)|}{|x|^{L+1}} \ .$$

Note that, if we take $\|B\|$ and $\|N\|_{C^1}$ small enough (depending here on $[w_2^{>}]_{L+1}$), we have that

$$|\psi_{w_2}(x)| \leq (\|A_1\| + \varepsilon)|x|, \tag{60}$$

with ε small such that $\|A_2^{-1}\|(\|A_1\|+\varepsilon)^{L+1} < 1$. Using the bound (60), and (54) (55), we have:

$$
\begin{aligned}
[\mathcal{M}(w_2^{>}) - \mathcal{M}(w_1^{>})]_{L+1} &= [\mathcal{N}(w_2) - \mathcal{N}(w_1)]_{L+1} \\
&\le \|A_2^{-1}\| \sup_{x\in B_1} |w^{\le}(\psi_{w_2}(x)) - w^{\le}(\psi_{w_1}(x))|/|x|^{L+1} \\
&\quad + \|A_2^{-1}\| \sup_{x\in B_1} \Big\{ |w_2^{>}(\psi_{w_2}(x)) - w_1^{>}(\psi_{w_2}(x))| \\
&\qquad + |w_1^{>}(\psi_{w_2}(x)) - w_1^{>}(\psi_{w_1}(x))| \Big\} / |x|^{L+1} \\
&\quad + \|A_2^{-1}\| \sup_{x\in B_1} \big| N_2(x, w^{\le}(x) + w_2^{>}(x)) \\
&\qquad - N_2(x, w^{\le}(x) + w_1^{>}(x)) \big| / |x|^{L+1} \\
&\le \|A_2^{-1}\| (\mathrm{Lip}\, w^{\le})(\mathrm{Lip}\, N_1 + \|B\|)[w_2^{>} - w_1^{>}]_{L+1} \\
&\quad + \|A_2^{-1}\| \Big([w_2^{>} - w_1^{>}]_{L+1} |\psi_{w_2}(x)|^{L+1}/|x|^{L+1} \\
&\qquad + (\mathrm{Lip}\, w_1^{>})(\mathrm{Lip}\, N_1 + \|B\|)[w_2^{>} - w_1^{>}]_{L+1} \Big) \\
&\quad + \|A_2^{-1}\| (\mathrm{Lip}\, N_2)[w_2^{>} - w_1^{>}]_{L+1} \\
&\le \Big(\varepsilon + \|A_2^{-1}\|(\|A_1\| + \varepsilon)^{L+1} \Big) [w_2^{>} - w_1^{>}]_{L+1} \\
&\le \nu [w_2^{>} - w_1^{>}]_{L+1}
\end{aligned}
$$

for some constant $\nu < 1$, under the smallness assumptions (that here may depend on $\mathrm{Lip}\, w_1^{>}$ and also, as already pointed out, on $[w_2^{>}]_{L+1}$). This proves that $[w_2^{>} - w_1^{>}]_{L+1} = 0$ and therefore that $w_1^{>} = w_2^{>}$. □

Proof of the uniqueness statements in Theorems 1.1 and 1.2. We start with the second part b2) of Theorem 1.2. After a scaling, we may assume that the solution $K = (\mathrm{Id}, w_L + h)$ of (4) in the statement is defined in the unit ball B_1 of X_1. Since $\sup_{x\in B_1}(|h(x)|/|x|^{L+1}) < \infty$ and w_L is a polynomial of degree L, Lemma 4.1 and the remarks preceding it imply that w_L must coincide with the polynomial $w^{\le}$ of Lemma 4.1. We deduce that h is a solution of $h = \mathcal{M}(h)$ –recall also that equation (4) is equivalent to the graph transform equation when $\pi_1 K = \mathrm{Id}$. Now the conclusion follows from Lemma 4.2.

We can now deduce easily that there is a unique (locally around 0) C^{L+1} invariant manifold tangent to X_1 at 0, as stated in Theorems 1.1 and 1.2. Indeed, let $G = (G^1, G^2)$ be a C^{L+1} parameterization of the manifold. By the tangency condition, we must have $G = (\mathrm{Id}, 0) + O(x^2)$. Therefore, we can represent uniquely (locally near zero) the invariant manifold as a graph of a C^{L+1} function $H : U_1 \subset X_1 \longrightarrow X_2$. For this, we simply take $H = G^2 \circ (G^1)^{-1}$. Now, the graph transform equation

$$F^2 \circ (\mathrm{Id}, H) = H \circ F^1 \circ (\mathrm{Id}, H)$$

is a consequence of the invariance assumption. In particular, (4) also holds for $K = (\mathrm{Id}, H)$ and $R = F^1 \circ K$. Hence, uniqueness of H, and therefore uniqueness of the manifold, follows from the proof of b2) in Theorem 1.2 given above. For this, we take w_L to be the Taylor expansion of degree L of H at 0, and $h = H - w_L$ –which clearly satisfies $\sup_{x\in B_1}(|h(x)|/|x|^{L+1}) < \infty$.

To end the proof of uniqueness, we note that from local uniqueness we get uniqueness, since by formula (44) we can recover the whole manifold. $\square$

4.2. **Existence of solution of the fixed point equation.** The proof will be similar to that in [dlL97], which in turn is somewhat related to the proofs based on graph transform methods. The main difference with the standard graph transform methods is that we consider our operator acting on Banach spaces of functions whose first derivatives are prescribed. Another difference with [dlL97] is that we obtain the optimal regularity by following a method very similar to the one that we have used in the proof of Proposition 3.5.

Recall that $L \geq 1$, $r \geq L+1$ and consider $w^> \in \Gamma_{r,L}$, the Banach space defined with $Y = X_2$ in Section 3.3. We then have:

Proposition 4.3. *Let $\mathcal{M}$ be defined as in* (59) *and $r \in \mathbb{N} \cup \{\omega\}$. Under appropriate smallness conditions on $\|B\|$ and $\|N\|_{C^r}$, the map $\mathcal{M}$ sends the closed unit ball $\bar{B}_1^r$ of $\Gamma_{r,L}$ into itself, and it is a contraction in $\bar{B}_1^r$ with the $\Gamma_{r-1,L}$ norm. That is*:

$$\|\mathcal{M}(w^>)\|_{\Gamma_{r,L}} \leq 1 \quad \textit{if } \|w^>\|_{\Gamma_{r,L}} \leq 1$$

and, for some constant $\nu < 1$,

$$\begin{aligned} \|\mathcal{M}(w_2^>) - &\mathcal{M}(w_1^>)\|_{\Gamma_{r-1,L}} \\ &\leq \nu \|w_2^> - w_1^>\|_{\Gamma_{r-1,L}} \quad \textit{if } \|w_2^>\|_{\Gamma_{r,L}} \leq 1 \textit{ and } \|w_1^>\|_{\Gamma_{r,L}} \leq 1 \ . \end{aligned} \tag{61}$$

In particular, equation $w^> = \mathcal{M}(w^>)$ admits a C^{r-1} solution which belongs to the closed unit ball of $\Gamma_{r-1,L}$.

From standard results (see [LI73], Lemma 2.5), the fixed point in $\Gamma_{r-1,L}$ of the previous proposition is, indeed, $C^{r-1+\mathrm{Lipschitz}}$ (due to the uniform $\Gamma_{r,L}$ bound on the sequence of iterates). In the next section, we will prove the C^r regularity result.

Proof of Proposition 4.3. The last statement of the proposition follows easily from the rest. Indeed, starting with $w^> = 0$, all the iterations $\mathcal{M}^k(0)$ remain in the closed unit ball of $\Gamma_{r,L}$ and hence, by (61), they converge in the $\Gamma_{r-1,L}$ norm to a solution (which belongs to the closed unit ball of $\Gamma_{r-1,L}$ since all iterations $\mathcal{M}^k(0)$ stay in such ball).

Next, note that

$$\begin{aligned} D[\mathcal{N}(w)] =& A_2^{-1}(Dw \circ \psi_w) D\psi_w - A_2^{-1} D_1 N_2 \circ (\mathrm{Id}, w) \\ & - A_2^{-1}[D_2 N_2 \circ (\mathrm{Id}, w)] Dw \end{aligned} \tag{62}$$

and

$$D\psi_w(x) = A_1 + D_1N_1(x, w(x)) + [D_2N_1(x, w(x)) + B]Dw(x) \ .$$

Proceeding by induction, it is not hard to show that, for $2 \le i \le r$, we have

$$\begin{aligned} D^i[\mathcal{N}(w)] \quad &= \quad A_2^{-1}(D^i w \circ \psi_w) D\psi_w^{\otimes i} \\ &\quad +A_2^{-1}(Dw \circ \psi_w)[D_2N_1 \circ (\mathrm{Id}, w) + B]D^i w \\ &\quad -A_2^{-1}[D_2N_2 \circ (\mathrm{Id}, w)]D^i w + V_i \end{aligned} \tag{63}$$

where V_i is a polynomial expression of $w, Dw, \ldots, D^{i-1}w$ involving as coefficients A, B and derivatives of N. All terms in V_i contain at least one derivative $D^j w$ and at least one factor $D^j N$, $1 \le j \le i$, or B (to see this, note that for $j \ge 2$, every term in the expression for $D^j\psi_w$ contains at least one of such factors).

We start showing the contraction property (61). For $w^> = w_1^>$ and $w^> + \Delta = w_2^>$ in the closed unit ball of $\Gamma_{r,L}$, we have

$$\mathcal{M}(w^> + \Delta) - \mathcal{M}(w^>) = \mathcal{N}(w + \Delta) - \mathcal{N}(w) = \int_0^1 \frac{d}{ds}[\mathcal{N}(w + s\Delta)] \ ds \ .$$

Since

$$\psi_{w+s\Delta}(x) = A_1 x + N_1(x, (w + s\Delta)(x)) + B[w + s\Delta](x)$$

and

$$\mathcal{N}(w + s\Delta)(x) = A_2^{-1}\Big\{(w + s\Delta)(\psi_{w+s\Delta}(x)) - N_2(x, (w + s\Delta)(x))\Big\} \ ,$$

we deduce that

$$\begin{aligned} &\mathcal{M}(w^> + \Delta) - \mathcal{M}(w^>) \\ &\qquad = \int_0^1 ds \ A_2^{-1}\Big\{\Delta \circ \psi_{w+s\Delta} \\ &\qquad\qquad + [D(w + s\Delta) \circ \psi_{w+s\Delta}][D_2N_1 \circ (\mathrm{Id}, w + s\Delta) + B]\Delta \\ &\qquad\qquad - D_2N_2 \circ (\mathrm{Id}, w + s\Delta)\Delta\Big\} \ . \end{aligned}$$

Therefore, for $0 \le i \le r - 1$, we have

$$D^i[\mathcal{M}(w^> + \Delta) - \mathcal{M}(w^>)] = \int_0^1 ds \left\{A_2^{-1}[(D^i\Delta) \circ \psi_{w+s\Delta}]D\psi_{w+s\Delta}^{\otimes i} + W_i\right\} \ , \tag{64}$$

where W_i is a polynomial expression in the derivatives of Δ up to order i. Every term in this polynomial contains at least one derivative of Δ up to order $i \le r-1$, and at least one factor which is B or a derivative of N up to order $i + 1 \le r$. The terms also include factors involving the derivatives of $w + s\Delta$ up to order $i+1 \le r$, which are all bounded by $\|w^{\le}\|_{C^r} + 1 \le \varepsilon + 1 \le 2$ since we are assuming $\|w^>\|_{\Gamma_{r,L}} \le 1$ and $\|w^> + \Delta\|_{\Gamma_{r,L}} \le 1$.

We use the notation $\mathcal{I} = \mathcal{M}(w^> + \Delta) - \mathcal{M}(w^>)$. For $r \in \mathbb{N}$, from (64) and $|D^L\Delta(y)| \le \|\Delta\|_{\Gamma_{r-1,L}}|y|$ if $y \in B_1$, we deduce for $x \in B_1$

$$|D^L\mathcal{I}(x)|/|x| \le [\|A_2^{-1}\|(\|A_1\| + \varepsilon)^{L+1} + \varepsilon]\|\Delta\|_{\Gamma_{r-1,L}} \le \nu\|\Delta\|_{\Gamma_{r-1,L}} \tag{65}$$

for some constant $\nu < 1$. For $L + 1 \le i \le r - 1$ (in case that $L + 1 < r$), we get for $x \in B_1$,

$$\begin{aligned}|D^i\mathcal{I}(x)| &\le [\|A_2^{-1}\|(\|A_1\| + \varepsilon)^{L+1} + \varepsilon]\|\Delta\|_{\Gamma_{r-1,L}}\\ &\le \nu\|\Delta\|_{\Gamma_{r-1,L}}.\end{aligned}$$

Note also that for $i \le L$, (64) gives that $D^i\mathcal{I}(0) = 0$. Hence, by Taylor's formula, $\|D^i\mathcal{I}\|_{C^0} \le \frac{1}{(L-i)!}\|D^L\mathcal{I}\|_{C^0} \le \nu\|\Delta\|_{\Gamma_{r-1,L}}$, where we have used (65) in the last inequality.

Taking the supremum of all these quantities, we conclude that $\|\mathcal{I}\|_{\Gamma_{r-1,L}} \le \nu\|\Delta\|_{\Gamma_{r-1,L}}$, and therefore (61).

In case $r = \omega$, we use (64) with $i = L + 1$. To bound all terms, note that $\|D^j\Delta\|_{C^0(B_1)} \le C\|D^{L+1}\Delta\|_{C^0(B_1)} = C\|\Delta\|_{\Gamma_{\omega,L}}$ for every $0 \le j \le L + 1$ (by Taylor's theorem), and also that

$$\begin{aligned}\|[D^{L+2}(w + s\Delta)]\circ\psi_{w+s\Delta}\|_{C^0(B_1)} &\le C\|D^{L+2}(w + s\Delta)\|_{C^0(B_{\|A_1\|+\varepsilon})}\\ &\le C\|D^{L+1}(w + s\Delta)\|_{C^0(B_1)} = C\|w + s\Delta\|_{\Gamma_{\omega,L}}\\ &\le C\end{aligned}$$

(by the Cauchy estimates).

We conclude that

$$\begin{aligned}\|\mathcal{I}\|_{\Gamma_{\omega,L}} &= \|D^{L+1}\mathcal{I}\|_{C^0(B_1)}\\ &\le [\|A_2^{-1}\|(\|A_1\| + \varepsilon)^{L+1} + \varepsilon]\|\Delta\|_{\Gamma_{\omega,L}} = \nu\|\Delta\|_{\Gamma_{\omega,L}}\end{aligned}$$

with $\nu < 1$. We therefore have the contraction property (61).

It only remains to prove that $\mathcal{M}$ maps the closed unit ball of $\Gamma_{r,L}$ into itself. For this, note that if $\|w^>\|_{\Gamma_{r,L}} \le 1$ then

$$\begin{aligned}\mathcal{M}(w^>) &= \mathcal{M}(0) + [\mathcal{M}(w^>) - \mathcal{M}(0)]\\ &= [\mathcal{N}(w^\le) - w^\le] + [\mathcal{M}(w^>) - \mathcal{M}(0)]\ .\end{aligned}$$

By Lemma 4.1, we know that $\mathcal{N}(w^\le) - w^\le$ has derivatives at 0 up to order L equal to zero, and that $\|\mathcal{N}(w^\le) - w^\le\|_{\Gamma_{r-1,L}}$ is small after scaling. On the other hand, in the previous proof of the contraction property we have seen that $\mathcal{M}(w^>) - \mathcal{M}(0)$ has derivatives at 0 up to order L equal to zero, and that $\|\mathcal{M}(w^>) - \mathcal{M}(0)\|_{\Gamma_{r-1,L}} \le \nu\|w^>\|_{\Gamma_{r-1,L}} \le \nu$.

Adding these two bounds, we have $\|\mathcal{M}(w^>)\|_{\Gamma_{r-1,L}} \leq \varepsilon + \nu$. Finally, when $r \in \mathbb{N}$, (63) leads to

$$\begin{aligned}\|D^r\mathcal{M}(w^>)\|_{C^0} &\leq \|D^r w^{\leq}\|_{C^0} + \|D^r\mathcal{N}(w)\|_{C^0}\\ &\leq \varepsilon + [\|A_2^{-1}\|(\|A_1\|+\varepsilon)^{L+1}+\varepsilon]\|D^r w\|_{C^0}\\ &\leq \varepsilon + \nu(\|D^r w^{\leq}\|_{C^0} + \|w^>\|_{\Gamma_{r,L}})\\ &\leq \varepsilon + \nu(\varepsilon+1)\ .\end{aligned}$$

Taking ε small enough, we conclude that $\mathcal{M}(w^>) \in \Gamma_{r,L}$ and $\|\mathcal{M}(w^>)\|_{\Gamma_{r,L}} \leq 1$. □

When $r = \omega$, the previous proposition establishes statement a) of Theorem 1.2. The proposition also leads to the result for $r = \infty$, using the argument presented for Theorem 1.1 at the end of Section 3.4. Finally, next section recovers the last derivative when $r \in \mathbb{N}$.

4.3. **Sharp regularity.** To establish the optimal C^r regularity we will use a method very similar to the one of Section 3.5 for Theorem 1.1. In this section we always have $r \in \mathbb{N}$.

Proposition 4.4. *The solution $w^>$ produced in Proposition 4.3 (under perhaps stronger smallness conditions on $\|B\|$ and $\|N\|_{C^r}$) is C^r.*

As in Section 3.5, to establish this result we differentiate the equation $w^> = \mathcal{M}(w^>)$ to find a fixed point equation for the unknown $H^> := Dw^> : B_1 \subset X_1 \longrightarrow \mathcal{L}(X_1, X_2)$, in which we consider $w^> \in C^{r-1}$ (the solution of Proposition 4.3) as a given function (and not as an unknown) whenever it appears in the equation for $H^>$ without being differentiated. This "freezing" technique will make the fixed point equation for $H^>$ simpler than the equation of the preceding section for $w^>$. Indeed, here we will obtain a bilinear equation for $H^>$.

Differentiating the equation $w = \mathcal{N}(w)$ satisfied by the solution $w = w^{\leq} + w^> \in C^{r-1}$ of Proposition 4.3 and using (62), we obtain

$$Dw = A_2^{-1}\Big\{(Dw \circ \psi_w)D\psi_w - D_1N_2 \circ (\mathrm{Id}, w) - [D_2N_2 \circ (\mathrm{Id}, w)]Dw\Big\}\ ,$$

where

$$\psi_w(x) = A_1x + N_1(x, w(x)) + Bw(x) \tag{66}$$

and

$$D\psi_w(x) = A_1 + D_1N_1(x, w(x)) + [D_2N_1(x, w(x)) + B]Dw(x)\ .$$

We obtain that $H_0 := Dw$ is a solution of

$$H = \tilde{\mathcal{N}}(H)\ ,$$

where we define, with ψ_w given by (66),

$$\begin{aligned}\tilde{\mathcal{N}}(H)(x) =& A_2^{-1}\Big\{H(\psi_w(x))\Big[A_1 + D_1N_1(x,w(x)) \\ &+ \Big(D_2N_1(x,w(x)) + B\Big)H(x)\Big] \\ &- D_1N_2(x,w(x)) - D_2N_2(x,w(x))H(x)\Big\} \ ,\end{aligned} \tag{67}$$

where we emphasize that here w, and hence ψ_w, are fixed (hence independent of H) and given by the solution $w = w^{\leq} + w^{>}$ of Proposition 4.3.

Let us consider $H^{\leq} = Dw^{\leq}$ fixed, and write $H = H^{\leq} + H^{>}$. We need to solve the fixed point equation $H^{>} = \tilde{\mathcal{M}}(H^{>})$, where

$$\tilde{\mathcal{M}}(H^{>}) = -H^{\leq} + \tilde{\mathcal{N}}(H^{\leq} + H^{>}) \ . \tag{68}$$

We take $H^{>} : B_1 \subset X_1 \longrightarrow \mathcal{L}(X_1, X_2) = Y$ in the space $\Gamma_{s,L-1}$ defined in Section 3.3, with either $s = r-2$ or $s = r-1$. We know that $H_0^{>} := Dw^{>} \in \Gamma_{r-2,L-1}$ is a solution of $H^{>} = \tilde{\mathcal{M}}(H^{>})$. It is easy to check that $\|H_0^{>}\|_{\Gamma_{r-2,L-1}} \leq 1$, since we know that $\|w^{>}\|_{\Gamma_{r-1,L}} \leq 1$. Proposition 4.4 will follow easily combining this fact and the following result.

Lemma 4.5. *Under appropriate smallness conditions on $\|B\|$ and $\|N\|_{C^r}$, if $r-2 \leq s \leq r-1$ then $\tilde{\mathcal{M}}$ maps the closed unit ball $\bar{B}_1^s$ of $\Gamma_{s,L-1}$ into itself, and it is a contraction in $\bar{B}_1^s$ with the $\Gamma_{s,L-1}$ norm.*

Applying this lemma with $s = r-1$ we obtain a solution $H_1^{>} \in \Gamma_{r-1,L-1}$ such that $\|H_1^{>}\|_{\Gamma_{r-1,L-1}} \leq 1$. In particular, $\|H_1^{>}\|_{\Gamma_{r-2,L-1}} \leq 1$ and therefore, by the contraction property of the lemma with $s = r-2$, both solutions $H_0^{>}$ and $H_1^{>}$ in the closed unit ball of $\Gamma_{r-2,L-1}$ must agree. Hence $Dw^{>} = H_0^{>} = H_1^{>} \in C^{r-1}$, and we conclude that $w^{>} \in C^r$.

Proof of Lemma 4.5. First, note that $\tilde{\mathcal{N}}$ and $\tilde{\mathcal{M}}$ map functions of class C^s into functions of class C^s whenever $s \leq r-1$.

We start showing the contraction property. For $H^{>}$ and $H^{>} + \Delta$ in the closed unit ball of $\Gamma_{s,L-1}$, we have

$$\begin{aligned}&[\tilde{\mathcal{M}}(H^{>} + \Delta) - \tilde{\mathcal{M}}(H^{>})](x) = [\tilde{\mathcal{N}}(H+\Delta) - \tilde{\mathcal{N}}(H)](x) \\ &\quad= \int_0^1 \frac{d}{ds}[\tilde{\mathcal{N}}(H + s\Delta)(x)]\, ds \\ &\quad= \int_0^1 ds\ A_2^{-1}\Big\{\Delta(\psi_w(x)) \\ &\qquad \cdot \Big[A_1 + D_1N_1(x,w(x)) + \Big(D_2N_1(x,w(x)) + B\Big)(H+s\Delta)(x)\Big] \\ &\qquad + (H+s\Delta)(\psi_w(x))\Big(D_2N_1(x,w(x)) + B\Big)\Delta(x) \\ &\qquad - D_2N_2(x,w(x))\Delta(x)\Big\} \ .\end{aligned} \tag{69}$$

Therefore, for $0 \le i \le s$, we have

$$D^i[\tilde{\mathcal{M}}(H^> + \Delta) - \tilde{\mathcal{M}}(H^>)] = \int_0^1 ds \left\{ A_2^{-1}[(D^i\Delta) \circ \psi_w] D\psi_w^{\otimes i} \otimes A_1 + U_i \right\} , \tag{70}$$

where U_i is a polynomial expression in the derivatives of Δ up to order i. Every term in U_i contain at least one derivative $D^j\Delta$ and at least one factor which is B or a derivative of N up to order $i+1 \le r$.

Using the notation $\mathcal{J} = \tilde{\mathcal{M}}(H^> + \Delta) - \tilde{\mathcal{M}}(H^>)$, (70) and $|D^{L-1}\Delta(y)| \le \|\Delta\|_{\Gamma_{s,L-1}}|y|$ for $y \in B_1$, we deduce, for some constant $\nu < 1$,

$$\begin{aligned} |D^{L-1}\mathcal{J}(x)|/|x| &\le \|A_2^{-1}\|(\|A_1\| + \varepsilon)^L |D^{L-1}\Delta(\psi_w(x))| + \varepsilon\|\Delta\|_{\Gamma_{s,L-1}} \\ &\le [\|A_2^{-1}\|(\|A_1\| + \varepsilon)^{L+1} + \varepsilon]\|\Delta\|_{\Gamma_{s,L-1}} \le \nu\|\Delta\|_{\Gamma_{s,L-1}} \quad \text{for } x \in B_1 \end{aligned} \tag{71}$$

and, for $L \le i \le s$ (in case $L-1 < s$),

$$|D^i\mathcal{J}(x)| \le [\|A_2^{-1}\|(\|A_1\| + \varepsilon)^{L+1} + \varepsilon]\|\Delta\|_{\Gamma_{s,L-1}} \le \nu\|\Delta\|_{\Gamma_{s,L-1}} \quad \text{for } x \in B_1 \ .$$

Note also that for $i \le L-1$, (70) gives that $D^i\mathcal{J}(0) = 0$. Hence, by Taylor's formula, $\|D^i\mathcal{J}\|_{C^0} \le \frac{1}{(L-1-i)!}\|D^{L-1}\mathcal{J}\|_{C^0} \le \nu\|\Delta\|_{\Gamma_{s,L-1}}$, where we have used (71) in the last inequality.

Taking the supremum of all these quantities, we conclude that $\|\mathcal{J}\|_{\Gamma_{s,L-1}} \le \nu\|\Delta\|_{\Gamma_{s,L-1}}$, and therefore the contraction property claimed in the lemma.

It only remains to prove that $\tilde{\mathcal{M}}$ maps the closed unit ball of $\Gamma_{s,L-1}$ into itself. For this, note that if $\|H^>\|_{\Gamma_{s,L-1}} \le 1$ then

$$\tilde{\mathcal{M}}(H^>) = Dw^> + [\tilde{\mathcal{M}}(0) - \tilde{\mathcal{M}}(Dw^>)] + [\tilde{\mathcal{M}}(H^>) - \tilde{\mathcal{M}}(0)] \ , \tag{72}$$

since, by definition (68), $Dw^> - \tilde{\mathcal{M}}(Dw^>) = Dw - \tilde{\mathcal{N}}(Dw) = 0$.

Expression (72) is useful in order to check that $\tilde{\mathcal{M}}(H^>)$ has all derivatives at 0 up to order $L-1$ equal to zero. Indeed, we know that the term $Dw^>$ in (72) belongs to $\Gamma_{r-2,L-1}$ and has small norm after scaling (and after freezing $K^>$). On the other hand, the previous proof of the contraction property gives that the two last terms of (72) also belong to $\Gamma_{r-2,L-1}$, and have $\Gamma_{r-2,L-1}$-norm bounded by $\nu\|Dw^>\|_{\Gamma_{r-2,L-1}} \le \varepsilon$ and $\nu\|H^>\|_{\Gamma_{r-2,L-1}} \le \nu$, respectively. Adding the three bounds, we conclude $\|\tilde{\mathcal{M}}(H^>)\|_{\Gamma_{r-2,L-1}} \le \varepsilon + \nu$. This concludes the proof for the case $s = r-2$.

In the case $s = r-1$, since $\|G\|_{\Gamma_{r-1,L-1}} = \max(\|G\|_{\Gamma_{r-2,L-1}}, \|D^{r-1}G\|_{C^0})$, it only remains to prove the bound

$$\|D^{r-1}[\tilde{\mathcal{M}}(H^>)]\|_{C^0} \le 1 \qquad \text{if } \|H^>\|_{\Gamma_{r-1,L-1}} \le 1 \ .$$

To establish it, we use definition (68) and that $H^\le = Dw^\le$ is a polynomial of degree $L-1 < r-1$ to deduce $D^{r-1}[\tilde{\mathcal{M}}(H^>)] = D^{r-1}[\tilde{\mathcal{N}}(H)]$, where $H = H^\le + H^>$. We also note that $\|H\|_{C^{r-1}} \le \varepsilon + \|H^>\|_{C^{r-1}} \le \varepsilon + 1$. Hence, differentiating (67) $r-1$ times and using the smallness of $\|B\|$ and $\|N\|_{C^r}$, we conclude

$$\begin{aligned}
\|D^{r-1}[\tilde{\mathcal{M}}(H^>)]\|_{C^0} &= \|D^{r-1}[\tilde{\mathcal{N}}(H)]\|_{C^0} \\
&\le \Big(\|A_2^{-1}\|(\|A_1\|+\varepsilon)^r+\varepsilon\Big)\|H\|_{C^{r-1}} \\
&\le \Big(\|A_2^{-1}\|(\|A_1\|+\varepsilon)^r+\varepsilon\Big)(\varepsilon+1) \\
&\le 1
\end{aligned} \tag{73}$$

if we take ε small enough. □

5. Dependence on parameters

The main results of this section are Theorem 5.2 and Theorem 5.7, which imply the conclusions about smooth dependence on parameters stated in the last parts of Theorem 1.1 and of Theorem 1.2, respectively. They provide optimal differentiability results in certain classes of regularities to be defined below and do not require any extra non-resonance assumptions.

The method of proof of Theorem 5.2 will be very similar to the one of the proof of optimal regularity in Theorems 1.1 and 1.2, that is, first we prove that the $K^>$ produced in Section 3 has almost optimal regularity with respect to the variables (λ, x), and then, studying the equation which the derivative $D_x K^>$ satisfies, we get the optimal regularity for $K^>$ stated in the Theorems.

This kind of proof does not work for Theorem 5.7 because the parameter version of the operators $\mathcal{N}$ and $\mathcal{M}$ used in the proof of Theorem 1.2 are not contractive in the usual norms of the spaces of differentiable functions. However, it is possible that a clever choice of a norm with different weights for the partial derivatives $D_\lambda^i D_x^j$ would make $\mathcal{M}$ a contraction, but we have not explored systematically this possibility.

The proof of Theorem 5.7 follows a different route than the one outlined above. We hope that having two methods of proof available may help to study other theorems.

In the proof of Theorem 5.7 that we present, we inductively prove that the derivatives in the parameter and in the variable exist and are continuous and bounded, provided that derivatives of *lower order* exist and are continuous and bounded.

The precise definition of derivatives of lower order is somewhat intricate (they are derivatives on two variables, so the ordering has to involve two indices) and will be motivated by the induction proof and the structure of the operator $\mathcal{M}$, which involves composition in the x variable. That is, we will define a suitable order on the sets of indices of derivatives for functions of two variables, in such a way that we can use the functional equation to express derivatives of a certain index in terms of derivatives with indices which are smaller in the indicated order.

First we will show that whenever N has enough partial derivatives, we can express $D^i_\lambda D^j_x \mathcal{M}(w^>)$ as

$$D^i_\lambda D^j_x \mathcal{M}(w^>) = \mathcal{A}(w^>) D^i_\lambda D^j_x w^> + \mathcal{B}(w^>) \ ,$$

where $\mathcal{A}(w^>)$ is a linear operator and $\mathcal{B}(w^>)$ is an expression that involves only derivatives of lower order (indeed, this calculation will motivate the definition of what lower order means). Starting with $w^>_0 = 0$, we consider the iteration $w^>_n = \mathcal{M}^n(0)$. We will show that $\mathcal{A}_n := \mathcal{A}(w^>_n)$ is a contractive operator. From this it will follow that, when $D^k_\lambda D^l_x \mathcal{M}^n(0)$ converge uniformly on compact sets as $n \to \infty$ for all (k,l) of lower order than (i,j), then $D^i_\lambda D^j_x \mathcal{M}^n(0)$ also converges uniformly on compact sets.

From the finite increment formula for the derivatives of the form $D^{i-1}_\lambda D^j_x w^>_n$ or $D^i_\lambda D^{j-1}_x w^>_n$, we will recover the finite increment formula for their limits as $n \to \infty$. This will give us that the limit of $D^i_\lambda D^j_x w^>_n$ as n tends to infinity is indeed a true derivative.

We will encounter the difficulty that in infinite dimensions, continuity on compact sets does not imply continuity and boundedness on a ball, and we will need to use an extra argument. We will show that the derivative which we have obtained as a limit satisfies a certain functional equation. By studying this equation, it will be possible to show that, when the derivatives of lower order are continuous and bounded, then the limit of $D^i_\lambda D^j_x \mathcal{M}^n(0)$ is also continuous.

The induction starts with the regularity results that we have already proved, and stops when the conditions requiring enough derivatives for N are violated. The initial step of the induction is the fact (already established) that $\mathcal{M}^n(0)$ and $D_x \mathcal{M}^n(0)$ converge to $w^>$ and $D_x w^>$, respectively.

To carry on the induction process, it is important to define spaces of regular functions which are adapted to the argument. These classes must require the existence of those derivatives needed to make the inductive argument work.

We now detail the classes which we have found useful.

We consider functions $f : \Lambda \times U \to Y$, where Λ and U are open sets in two Banach spaces and Y is another Banach space.

We say that f is $C^{l,m}$ when $D^i_\lambda D^j_x f(\lambda, x)$ exist, are continuous and bounded for $0 \le i \le l$, $0 \le j \le m$. We consider the space of such functions endowed with the topology given by the supremum of all the derivatives above.

We say that f is jointly C^r when it is a C^r mapping from $\Lambda \times U \to Y$ with bounded derivatives up to order r. This is equivalent to the existence, continuity and boundedness of $D^i_\lambda D^j_x f(\lambda, x)$ for $0 \le i + j \le r$.

More generally, if $\Sigma \subset (\mathbb{Z}^+)^2$ is such that $(i,j) \in \Sigma$ and $\tilde{\imath} \le i, \tilde{\jmath} \le j$ implies $(\tilde{\imath}, \tilde{\jmath}) \in \Sigma$, we denote C^Σ the set of functions f for which $D^i_\lambda D^j_x f$ exists, is continuous and bounded for every $(i,j) \in \Sigma$. We consider C^Σ endowed with the norm

$$||f||_{C^\Sigma(B)} = \sup_{(i,j) \in \Sigma, (\lambda,x) \in B} |D^i_\lambda D^j_x f(\lambda, x)| \ ,$$

which makes it a Banach space. When the set B is understood from the context, we will suppress it from the notation.

It is clear that if $\Sigma' \subset \Sigma$ then $C^{\Sigma} \subset C^{\Sigma'}$. In particular

$$\begin{aligned} C^{l+r} &\subset C^{l,r} \ , \\ C^{r,r} &\subset C^{r} \ , \\ C^{l+1,r} &\subset C^{l,r} \ , \\ C^{l,r+1} &\subset C^{l,r} \ , \end{aligned}$$

and, except in trivial cases such as zero dimensional spaces or $r = 0$ in the second formula, these inclusions are strict.

Now we turn to the definition of derivatives of *lower order*. The operators $\mathcal{T}$ and $\mathcal{M}$ involve compositions. If f_λ and g_λ are families of maps, when we take derivatives of the composition $h_\lambda = f_\lambda \circ g_\lambda$ with respect to parameters, we are forced to take derivatives of f_λ with respect to the variable (and not only with respect to the parameter). Hence, more derivatives with respect to the parameter require also more derivatives with respect to the space variables. Quantitative versions of this heuristic argument are made precise in Lemma 5.3 and Lemma 5.11. This motivates the following definition:

Definition 5.1. *Given $(i,j) \in (\mathbb{Z}^+)^2$, we define*

$$\Sigma_{i,j} = \{(a,b) \in (\mathbb{Z}^+)^2 \mid \ a+b \le i+j, \, a \le i\} \ .$$

We say that the derivative $D^k_\lambda D^l_x$ is of lower order than the derivative $D^i_\lambda D^j_x$ if $(k,l) \in \Sigma_{i,j}$.

We will see that the classes $C^{\Sigma_{i,j}}$ are well adapted to the study of regularity with respect to parameters, since one can use them to make induction arguments.

Remark 15. As we have seen in Section 3, in general R depends on λ. When we consider the equation

$$F_\lambda \circ (K^{\le}_\lambda + K^{>}) = (K^{\le}_\lambda + K^{>}) \circ R_\lambda$$

and we try to solve it applying the implicit function theorem, we are led to consider the operator $\mathcal{P} : \Lambda \times \Gamma_{r,L} \longrightarrow \Gamma_{r,L}$ defined by

$$\mathcal{P}(\lambda, K^{>}) = F_\lambda \circ (K^{\le}_\lambda + K^{>}) - (K^{\le}_\lambda + K^{>}) \circ R_\lambda \ .$$

We already know that for each λ_0 such that the hypotheses of Theorem 1.1 are satisfied, there exists $K^{>}_{\lambda_0}$ such that $\mathcal{P}(\lambda_0, K^{>}_{\lambda_0}) = 0$.

Here we are interested in the regularity of $K^{>}_\lambda$ with respect to (λ, x). The following technical difficulty appears. The map $K^{>} \mapsto K^{>} \circ R_\lambda$ can not be differentiable from C^r to C^r spaces for any $r < \omega$. Therefore it is not possible to apply the implicit function theorem in a straightforward way. The situation changes if it happens that R does not depend on λ. For instance this is the case when A_1 does not depend on λ and the non-resonance assumption

$$\big(\mathrm{Spec}(A_1)\big)^i \cap \mathrm{Spec}(A_1) = \emptyset \qquad \text{for every integer } i \text{ with } 2 \le i \le L \ ,$$

which appeared in condition b) of Theorem 1.1 (and, also in b) of Lemma 3.1), is satisfied. This situation occurs when we consider only the effect of nonlinear terms in the results of the problem. Of course, there could be other situations when the dependence of the coefficients of R in λ just happens to vanish. In this case $\mathcal{P}$ can be made differentiable, of course assuming that F_λ is C^s and $K^>$ is C^r with $s > r$. In this respect, we call attention that [dlLO99] contains a classification of the cases where the composition operator among Hölder functions (and differentiable functions) is differentiable. In these cases, one can obtain regularity results very quickly using the implicit function theorem, but these results are one derivative short of optimal.

In the following we use the convention to represent functions G defined in subsets of $\Lambda \times X_1$ or $\Lambda \times X$ by $G_\lambda(x) = G(\lambda, x)$. Therefore, we will write G, G_λ or $G_\lambda(x)$ according to the context. In particular, in formulas involving composition in the x variable, where λ acts as a parameter, we will write $G_\lambda \circ H_\lambda$ for $G(\lambda, H(\lambda, \cdot))$.

For typographical reasons, when a map already has a subindex we will not write explicitly the dependence on λ. That is, we will write A_1, A_2, N_1 and N_2 for $A_{1,\lambda}$, $A_{2,\lambda}$, $N_{1,\lambda}$ and $N_{2,\lambda}$ respectively.

5.1. **Dependence on parameters for Theorem 1.1.** The following is the main result of this section. We assume that F has derivatives $D_\lambda^a D_x^b F$ with (a, b) in a set $\Sigma_{i,j}$ and we get that the parameterization K given by Theorem 1.1 has derivatives $D_\lambda^a D_x^b K$ with (a, b) in the same set. Below we will show that this immediately implies the result on differentiability with respect to parameters claimed in Theorem 1.1.

Theorem 5.2. *Let Λ_0 and X be Banach spaces, Λ an open set of Λ_0, and U an open set of X, with $0 \in U$. Let $F : \Lambda \times U \longrightarrow X$ be a family of maps. Assume that:*

1) *For some $(i, j) \in (\mathbb{Z}^+)^2$,*
$$F \in C^{\Sigma_{i,j}}$$
where $\Sigma_{i,j}$ is the set introduced in Definition 5.1 and $C^{\Sigma_{i,j}}$ denotes the set of functions which have all the derivatives of orders in $\Sigma_{i,j}$ defined, continuous and bounded.
2) 2.1) *For all $\lambda \in \Lambda$, $D_x F_\lambda(0)$ preserves the space X_1, which is independent of λ.*
   2.2) *For all values $\lambda \in \Lambda$, conditions 0)–4) of Theorem 1.1 are satisfied with uniform constants.*
3) $j \geq L + 1$.

Then, the family of maps $K : \Lambda \times U_1 \longrightarrow X$ given by Theorem 1.1 belongs to $C^{\Sigma_{i,j}}$.

We now show that the regularity with respect to parameters claimed in d) of Theorem 1.1 follows from Theorem 5.2. Indeed, first observe that $C^r \subset$

$C^{\Sigma_{r-L-1,L+1}} \subset C^{r-L-1}$. Hence, the hypothesis of regularity on parameters in Theorem 1.1 imply that $F \in C^{\Sigma_{r-L-1,L+1}}$ and, applying Theorem 5.2, we obtain that $K \in C^{\Sigma_{r-L-1,L+1}} \subset C^{r-L-1}$.

Remark 16. Condition 2.1) is not a serious restriction. If $X_{1,\lambda}$ depends smoothly on λ, one can make a change of coordinates in such a way that $X_{1,\lambda}$ becomes independent of λ and then apply Theorem 5.2 to the transformed family of maps.

To prove Theorem 5.2 we need a series of lemmas. We begin by recalling that for every given value $\lambda \in \Lambda$, the parameterization $K^{>}_{\lambda}$ given by Theorem 1.1 is the fixed point of the operator $\mathcal{T}$ defined by

$$(\mathcal{T}K^{>})_\lambda = (\mathcal{S}^{-1}H)_\lambda = \sum_{k=0}^{\infty} A_\lambda^{-(k+1)} H_\lambda \circ R_\lambda^k \ ,$$

where

$$H_\lambda = -N_\lambda \circ (K^{\le}_\lambda + K^{>}_\lambda) - A_\lambda K^{\le}_\lambda + K^{\le}_\lambda \circ R_\lambda$$

and $K^{\le}_\lambda$ and R_λ are the polynomials of degree L given by Section 3.1. Under the assumptions of the theorem, these polynomials are C^i with respect to λ.

We assume that we have scaled the map in such a way that $\|N\|_{C^{\Sigma_{i,j}}}$, $\|K^{\le} - \mathrm{Id}\,\|_{C^{i,j}}$ and $\|R - A_1\|_{C^{i,j}}$ are as small as we need, uniformly in $\lambda \in \Lambda$.

The scheme of the proof of Theorem 5.2 will be the same as the one of the proofs of existence and sharp regularity of K in Section 3. The main difference is that here we will work in spaces of differentiable functions jointly with respect to the space variable and the parameter.

To describe in some detail the structure of the expression of the derivatives of the composition in terms of the derivatives of the functions, we introduce the following auxiliary sets of indices, closely related to $\Sigma_{i,j}$,

$$\begin{aligned} \Sigma^*_{i,0} &= \{(a,b) \in (\mathbb{Z}^+)^2 \mid \ a+b \le i,\, b \ge 1\} \cup \{(i,0)\} \ , \\ \Sigma^*_{i,j} &= \{(a,b) \in (\mathbb{Z}^+)^2 \mid \ a+b \le i+j,\, a \le i,\, b \ge 1\} \qquad \text{if } j \ge 1 \ , \end{aligned} \tag{74}$$

$$\begin{aligned} \tilde\Sigma_{i,0} &= \{(a,b) \in (\mathbb{Z}^+)^2 \mid \ a+b \le i\} \ , \\ \tilde\Sigma_{i,j} &= \Sigma^*_{i,j} \qquad\qquad \text{if } j \ge 1 \ , \end{aligned} \tag{75}$$

For $s \ge l$ let

$$\begin{aligned} \|H\|_{\Gamma_{i,s,l}} := \max\Big\{ &\max_{a,b \in \Sigma_{i,s}} \|D^a_\lambda D^b_x H\|_{C^0(\Lambda\times B_1)}, \\ &\max_{0 \le a \le i} \sup_{(\lambda,x)\in\Lambda\times B_1} \big(|D^a_\lambda D^l_x H_\lambda(x)|/|x|\big)\Big\} \end{aligned} \tag{76}$$

and

$$\begin{aligned} \Gamma_{i,s,l} = \{&H : \Lambda \times B_1 \subset \Lambda_0 \times X_1 \longrightarrow Y \mid H \in C^{\Sigma_{i,s}}(\Lambda \times B_1), \\ &D^a_\lambda D^b_x H_\lambda(0) = 0 \text{ for } 0 \le a \le i,\ 0 \le b \le l,\ \|H\|_{\Gamma_{i,s,l}} < \infty\} \ . \end{aligned}$$

The norm $\|\cdot\|_{\Gamma_{i,j,l}}$ makes $\Gamma_{i,j,l}$ a Banach space. As we remarked for the space $\Gamma_{i,l}$ in Section 3.3, here the terms $\sup_{(\lambda,x)\in\Lambda\times B_1}\left(|D^a_\lambda D^l_x H_\lambda(x)|/|x|\right)$ are only relevant whenever $s=l$ and they could be omitted (since $D^a_\lambda D^l_x H_\lambda(0)=0$) when $s>l$.

Note that if $H\in\Gamma_{i,s,l}$ and $(a,b)\in\Sigma_{i,s}$, by Taylor's theorem and the definition of $\|\cdot\|_{\Gamma_{i,j,l}}$ we have

$$|D^a_\lambda D^b_x H_\lambda(x)| \le \frac{1}{(l-b)_+!}\|H\|_{\Gamma_{i,s,l}}|x|^{(l-b+1)_+}, \qquad (\lambda,x)\in\Lambda\times B_1\ .$$

Lemma 5.3. *With the notation above, assuming that $H\in C^{\Sigma_{i,j}}$, $R\in C^{i,j}$ and $A_\lambda\in C^{i,\infty}$, we have that*

$$D^i_\lambda D^j_x\left[A_\lambda^{-(k+1)}H_\lambda\circ R^k_\lambda\right]$$

is a polynomial expression that involves only derivatives $D^m_\lambda A_\lambda^{-(k+1)}$ with $0\le m\le i$, derivatives $D^a_\lambda D^b_x H_\lambda$ evaluated at R^k_λ with $(a,b)\in\tilde\Sigma_{i,j}$, and derivatives $D^\alpha_\lambda D^\beta_x R^k_\lambda$ with $0\le\alpha\le i$, $0\le\beta\le j$, except for $(\alpha,\beta)=(0,0)$.

Moreover, each term in the polynomial expression contains:

- *A factor $D^m_\lambda A_\lambda^{-(k+1)}$*
- *A factor $D^a_\lambda D^b_x H_\lambda\circ R^k_\lambda$*
- *A factor*
$$D^{i_1}_\lambda D^{j_1}_x R^k_\lambda\cdots D^{i_b}_\lambda D^{j_b}_x R^k_\lambda$$
with $j_1+\cdots+j_b=j$, $i_1+\cdots+i_b=m-a$ and $(i_l,j_l)\neq(0,0)$ for $1\le l\le b$.

Proof. We start by computing the m^{th} derivative of $H_\lambda\circ R^k_\lambda$ with respect to λ.

Let $T_m=D^m_\lambda[H_\lambda\circ R^k_\lambda]$. It is easily checked by induction that

$$T_m=\sum_{(a,b)\in\Sigma^*_{m,0}}\ \sum_{i_1,\dots,i_b} C D^a_\lambda D^b_x H_\lambda\circ R^k_\lambda\, D^{i_1}_\lambda R^k_\lambda\cdots D^{i_b}_\lambda R^k_\lambda\ , \tag{77}$$

where $i_1+\cdots+i_b=m-a$ and C is a coefficient which depends on $a,b,i_1,\dots,i_b$.

Then, we differentiate the expression (77) j times with respect to x. It is also easy to see by induction that

$$D^j_x T_m=\sum_{(a,b)\in\Sigma^*_{m,j}}\ \sum_{i_1,j_1,\dots,i_b,j_b} C D^a_\lambda D^b_x H_\lambda\circ R^k_\lambda\, D^{i_1}_\lambda D^{j_1}_x R^k_\lambda\cdots D^{i_b}_\lambda D^{j_b}_x R^k_\lambda\ , \tag{78}$$

where $i_1+\cdots+i_b=m-a$, $j_1+\cdots+j_b=j$ and C is a combinatorial coefficient depending on the indices.

Finally,

$$D^i_\lambda D^j_x[A_\lambda^{-(k+1)}H_\lambda\circ R^k_\lambda]=\sum_{m=0}^{i} C[D^{i-m}_\lambda A_\lambda^{-(k+1)}]D^j_x T_m\ . \tag{79}$$

Substituting (78) into (79), the resulting expression contains derivatives $D^a_\lambda D^b_x H_\lambda$ of orders $(a,b)\in\Sigma^*_{0,j}\cup\Sigma^*_{1,j}\cup\cdots\cup\Sigma^*_{i,j}=\Sigma^*_{i,j}=\tilde\Sigma_{i,j}$ if $j>0$, or $\{(a,b)\mid a+b\le i\}=\tilde\Sigma_{i,0}$ if $j=0$. $\square$

Lemma 5.4. *Let R be the map given in Lemma 3.1. Assuming that $R \in C^{i,\infty}$, $A_\lambda \in C^{i,\infty}$ and that $j \geq L$, then $\mathcal{S} : \Gamma_{i,j,L} \longrightarrow \Gamma_{i,j,L}$ defined by $(\mathcal{S}H)_\lambda = A_\lambda H_\lambda - H_\lambda \circ R_\lambda$ is an invertible operator, and $\|\mathcal{S}^{-1}\|$ can be bounded by a constant independent of the scaling parameter.*

Proof. The formal expression for $\mathcal{S}^{-1}$ is the same to the one we have obtained in Lemma 3.3:

$$(\mathcal{S}^{-1}\eta)_\lambda = \sum_{k=0}^{\infty} A_\lambda^{-(k+1)} \eta_\lambda \circ R_\lambda^k \ . \tag{80}$$

Now we have to see that the series (80) converges in the norm of $\Gamma_{i,j,L}$ and that

$$\| \sum_{k=0}^{\infty} A_\lambda^{-(k+1)} \eta_\lambda \circ R_\lambda^k \|_{\Gamma_{i,j,L}} \leq C \|\eta\|_{\Gamma_{i,j,L}} \ .$$

We begin by showing that if $\eta \in \Gamma_{i,j,L}$ and $(a,b) \in \Sigma_{i,j}$ then

$$\sum_{k=0}^{\infty} D_\lambda^a D_x^b \Big[A_\lambda^{-(k+1)} \eta_\lambda \circ R_\lambda^k \Big](x) \tag{81}$$

converges absolutely.

It is not hard to prove that

$$\|D_\lambda^m A_\lambda^{-(k+1)}\| \leq C_m (k+1)^m \|A_\lambda^{-1}\|^{k+1} \tag{82}$$

and that

$$|D_\lambda^m D_x^l R_\lambda^k(x)| \leq C_{m,l} (\|A_1\| + \varepsilon)^k |x|^{(1-l)_+} \ . \tag{83}$$

By Lemma 5.3 every term in the sum (81) is bounded by a sum of terms of the form

$$C \|D_\lambda^{a-m} A_\lambda^{-(k+1)}\| \, |D_\lambda^\alpha D_x^\beta \eta_\lambda (R_\lambda^k(x))| \, |D_\lambda^{i_1} D_x^{j_1} R_\lambda^k(x)| \cdots |D_\lambda^{i_\beta} D_x^{j_\beta} R_\lambda^k(x)| \tag{84}$$

with $0 \leq m \leq a$, $(\alpha, \beta) \in \tilde{\Sigma}_{m,b}$, $i_1 + \cdots + i_\beta = m - \alpha$ and $j_1 + \cdots + j_\beta = b$. Hence, every term in the sum (81) is bounded by

$$\begin{aligned}
C(k+1)^{a-m} &\|A_\lambda^{-1}\|^{k+1} \|\eta\|_{\Gamma_{i,j,L}} |R_\lambda^k(x)|^{(L-\beta+1)_+} (\|A_1\| + \varepsilon)^{k\beta} |x|^{(\beta-b)_+} \\
&\leq C \|\eta\|_{\Gamma_{i,j,L}} (k+1)^a \|A_\lambda^{-1}\|^{k+1} (\|A_1\| + \varepsilon)^{k(L-\beta+1)_+} |x|^{(L-\beta+1)_+} \\
&\quad \cdot (\|A_1\| + \varepsilon)^{k\beta} |x|^{(\beta-b)_+} \\
&\leq C \|\eta\|_{\Gamma_{i,j,L}} (k+1)^a \Big(\|A_\lambda^{-1}\| (\|A_1\| + \varepsilon)^{L+1} \Big)^k |x|^{(L-b+1)_+}
\end{aligned}$$

with C independent of k, where we have used

$$\begin{gathered}
(1-j_1)_+ + (1-j_2)_+ + \cdots + (1-j_\beta)_+ \geq (\beta - (j_1 + \cdots + j_\beta))_+ = (\beta - b)_+ \ , \\
(L - \beta + 1)_+ + \beta \geq L + 1, \\
(L - \beta + 1)_+ + (\beta - b)_+ \geq (L - b + 1)_+
\end{gathered}$$

which hold because the function $(\cdot)_+$ is subadditive.

Since $\sum_{k=1}^{\infty}(k+1)^a\Big(\|A_\lambda^{-1}\|(\|A_1\|+\varepsilon)^{L+1}\Big)^k$ converges, we deduce that (81) converges absolutely in $\Lambda\times B_1$.

Hence, in the same way as in Lemma 3.3, we conclude that the expression (80) defines a function H in $C^{\Sigma_{i,j}}$ which satisfies $\mathcal{S}H=\eta$. Moreover

$$|D_\lambda^a D_x^b(\mathcal{S}^{-1}\eta)_\lambda(x)|\le C\|\eta\|_{\Gamma_{i,j,L}}|x|^{(L-b+1)_+}\ ,$$

and then, clearly, $\|\mathcal{S}^{-1}\eta\|_{\Gamma_{i,j,L}}\le C\|\eta\|_{\Gamma_{i,j,L}}$. □

Lemma 5.5. *Under the hypotheses of Theorem 5.2 and assuming that the norm* $\|N\|_{C^{\Sigma_{i,j}}}$ *is sufficiently small in the ball of radius* 3*, we have that* $\mathcal{T}:\Gamma_{i,j-1,L}\longrightarrow\Gamma_{i,j-1,L}$ *sends the closed unit ball* $\overline{B}_1^{j-1}$ *of* $\Gamma_{i,j-1,L}$ *into itself, and it is a contraction there.*

Proof. As we remarked before, we can assume that $K^{\le}$ is arbitrarily close to the identity and R is arbitrarily close to A_1, uniformly in $\lambda\in\Lambda$.

In the same way as we did in the proof of Lemma 3.4, one checks that $\mathcal{T}$ maps the unit ball of $\Gamma_{i,j-1,L}$ into itself. Indeed, let $K^{>}\in\overline{B}_1^{j-1}$. We claim that $\|H\|_{\Gamma_{i,j-1,L}}$ can be made arbitrarily small, taking $\|N\|_{C^{\Sigma_{i,j}}}$ sufficiently small and therefore $\|\mathcal{T}(K^{>})\|_{\Gamma_{i,j-1,L}}\le 1$. Indeed, we already know that for all $\lambda\in\Lambda$, $D_x^bH_\lambda(0)=0$ for $0\le b\le L$ and then the $D_\lambda^aD_x^b$ derivative of H at $x=0$ will be zero for $0\le a\le i$, $0\le b\le L$. We are going to bound $D_\lambda^aD_x^bH_\lambda(x)$ for $(a,b)\in\Sigma_{i,j}$. In the case $b\le L$, by Taylor's theorem we have

$$|D_\lambda^aD_x^bH_\lambda(x)|\le C\|D_\lambda^aD_x^LH\|_{C^0}\,|x|^{L-b}\ .$$

We decompose H_λ as $H_\lambda=H_\lambda^1+H_\lambda^2$ with

$$H_\lambda^1=-N_\lambda\circ K_\lambda^{\le}-A_\lambda K_\lambda^{\le}+K_\lambda^{\le}\circ R_\lambda$$

and

$$H_\lambda^2=-N_\lambda\circ(K_\lambda^{\le}+K_\lambda^{>})+N_\lambda\circ K_\lambda^{\le}\ .$$

Using the previous expressions and that R is a polynomial of degree L in x, we have

$$|D_\lambda^aD_x^LH_\lambda^1(x)|\le\|D_\lambda^aD_x^{L+1}[-N_\lambda\circ K_\lambda^{\le}+(K_\lambda^{\le}-\mathrm{Id})\circ R_\lambda]\|_{C^0}\,|x| \tag{85}$$

and

$$\begin{aligned} D_\lambda^aD_x^LH_\lambda^2 &= -\int_0^1\frac{d}{ds}\Big\{D_\lambda^aD_x^L[N_\lambda\circ(K_\lambda^{\le}+sK_\lambda^{>})]\Big\}\,ds\\ &= -\int_0^1 D_\lambda^aD_x^L[D_xN_\lambda\circ(K_\lambda^{\le}+sK_\lambda^{>})\,K_\lambda^{>})]\,ds\ . \end{aligned}\tag{86}$$

In the case $b>L$ we write

$$D_\lambda^aD_x^bH_\lambda=D_\lambda^aD_x^b[N_\lambda\circ(K_\lambda^{\le}+K_\lambda^{>})+(K_\lambda^{\le}-\mathrm{Id})\circ R_\lambda]\ . \tag{87}$$

The smallness assumptions on N and $K^{\leq} - \mathrm{Id}$ together with the fact that $K^{>} \in \Gamma_{i,j-1,L}$ give that (85), (86) and (87) are small and hence $|D_\lambda^a D_x^b H_\lambda(x)|$ and $\sup_{(\lambda,x)} |D_\lambda^a D_x^L H_\lambda(x)|/|x|$ are also small.

As in Lemma 3.4, to compute the Lipschitz constant of $\mathcal{T}$ we write

$$\begin{aligned}(\mathcal{T}(K^{>} + \Delta) - \mathcal{T}(K^{>}))_\lambda \\ = - \int_0^1 \mathcal{S}^{-1} DN_\lambda \circ (K_\lambda^{\leq} + K_\lambda^{>} + s\Delta_\lambda)\Delta_\lambda \, ds \; .\end{aligned} \tag{88}$$

Taking the $D_\lambda^a D_x^b$ derivative of the previous expression with indices (a, b) in $\Sigma_{i,j-1}$, we easily see that

$$\|D_\lambda^a D_x^b [\mathcal{T}(K^{>} + \Delta) - \mathcal{T}(K^{>})]\|_{C^0} \leq C\|\mathcal{S}^{-1}\| \|DN\|_{C^{\Sigma_{i,j-1}}} \cdot \|\Delta\|_{C^{\Sigma_{i,j-1}}} \; .$$

When $b = L$, note that the $D_\lambda^a D_x^b$ derivative of $DN_\lambda \circ (K_\lambda^{\leq} + K_\lambda^{>} + s\Delta_\lambda)\Delta_\lambda$ is a sum of terms, each one having a derivative of Δ with respect to x of order $\beta \leq L$, which is bounded by $C|x|^{L-\beta+1}$. Then the supremum of the $D_\lambda^a D_x^b$ derivative divided by $|x|$ is less than $C\|D_x N\|_{C^{\Sigma_{i,j-1}}}$.

By the smallness requirement on N the map $\mathcal{T}$ can be made a contraction in $\overline{B}_1^{j-1}$. □

As a consequence of Lemma 5.5, $\mathcal{T}$ has a unique fixed point $K^{>}$ in the unit ball of $\Gamma_{i,j-1,L}$. Since, for any $\lambda \in \Lambda$ fixed, the $K_\lambda^{>}$ obtained in Lemma 5.5 belongs to the closed unit ball of $\Gamma_{i+j-1,L}$, by uniqueness it must coincide with the $K^{>}$ obtained in Section 3.2.

Now we are going to improve the differentiability conclusions from $C^{\Sigma_{i,j-1}}$ to $C^{\Sigma_{i,j}}$.

We follow essentially the same arguments that we used in Section 3.5. We have that $D_x K^{>}$ satisfies the following equation, which is the parameter version of (45):

$$\begin{aligned}A_\lambda D_x K_\lambda^{>} - D_x K_\lambda^{>} \circ R_\lambda \; D_x R_\lambda \\ = - D_x N_\lambda \circ (K_\lambda^{\leq} + K_\lambda^{>})(D_x K_\lambda^{\leq} + D_x K_\lambda^{>}) \\ - A_\lambda D_x K_\lambda^{\leq} + D_x K_\lambda^{\leq} \circ R_\lambda \; D_x R_\lambda \; .\end{aligned} \tag{89}$$

We also need the parameter version of the operators $\tilde{\mathcal{S}}$, $\tilde{\mathcal{T}}$ and U:

$$(\tilde{\mathcal{S}}G)_\lambda = A_\lambda G_\lambda - G_\lambda \circ R_\lambda \; D_x R_\lambda \; ,$$

$$(\tilde{\mathcal{T}}G)_\lambda = -D_x N_\lambda \circ (K_\lambda^{\leq} + K_\lambda^{>})G_\lambda \; ,$$

and

$$U_\lambda = -D_x N_\lambda \circ (K_\lambda^{\leq} + K_\lambda^{>})D_x K_\lambda^{\leq} - A_\lambda D_x K_\lambda^{\leq} + D_x K_\lambda^{\leq} \circ R_\lambda \; D_x R_\lambda \; ,$$

with $K^{>}$ being the function obtained in Lemma 5.5, which belongs to $\Gamma_{i,j-1,L}$. We consider $\tilde{\mathcal{S}}$ and $\tilde{\mathcal{T}}$ acting on the space $\Gamma_{i,s,L-1}$ defined in (5.1) with $Y = \mathcal{L}(X_1, X)$.

Lemma 5.6. *Under the assumptions of Theorem 5.2 and under the standing assumptions arranged by scaling at the beginning of Section 3.2, if* $s \in \mathbb{N}$ *and* $L-1 \le s \le j-1$ *then* $\tilde{\mathcal{S}}$ *and* $\tilde{\mathcal{T}}$ *are bounded linear operators from* $\Gamma_{i,s,L-1}$ *into itself. Moreover, taking* $\|N\|_{C^{\Sigma_{i,j}}}$ *sufficiently small,* $\tilde{\mathcal{S}}$ *is invertible and* $\|\tilde{\mathcal{S}}^{-1}\|\|\tilde{\mathcal{T}}\| < 1$.

Proof. Let $G \in \Gamma_{i,s,L-1}$ and $(a,b) \in \Sigma_{i,s}$. We have that

$$
\begin{aligned}
D^a_\lambda D^b_x[N_\lambda \circ (K^{\le}_\lambda + K^{>}_\lambda)G_\lambda] \\
= \sum_{\substack{0\le l\le a,\\ 0\le m\le b}} C D^l_\lambda D^m_x[N_\lambda \circ (K^{\le}_\lambda + K^{>}_\lambda)] D^{a-l}_\lambda D^{b-m}_x G_\lambda
\end{aligned}
\tag{90}
$$

and

$$
\begin{aligned}
& D^l_\lambda D^m_x[N_\lambda \circ (K^{\le}_\lambda + K^{>}_\lambda)] \\
= & \sum_{(\alpha,\beta)\in\Sigma^*_{l,m}} \sum_{\substack{i_1+\cdots+i_\beta = l-\alpha\\ j_1+\cdots+j_\beta=m}} C D^\alpha_\lambda D^\beta_x N_\lambda \circ (K^{\le}_\lambda + K^{>}_\lambda) \\
& \cdot \quad D^{i_1}_\lambda D^{j_1}_x (K^{\le}_\lambda + K^{>}_\lambda) \cdots \quad D^{i_\beta}_\lambda D^{j_\beta}_x (K^{\le}_\lambda + K^{>}_\lambda)
\end{aligned}
$$

as a consequence of Faa di Bruno's formula. Therefore, each term in (90) has a derivative $D^\alpha_\lambda D^\beta_x N$ with $(\alpha,\beta) \in \Sigma_{a,b} \subset \Sigma_{i,s}$ and a derivative $D^\alpha_\lambda D^\beta_x G$ with $0 \le \alpha \le a$, $0 \le \beta \le b$. Note that $|D^\alpha_\lambda D^\beta_x G_\lambda(x)| \le \frac{1}{(L-\beta-1)_+!}\|G\|_{\Gamma_{i,s,L-1}} |x|^{(L-\beta)_+}$. Then

$$
|[D^a_\lambda D^b_x[N_\lambda \circ (K^{\le}_\lambda + K^{>}_\lambda)G_\lambda](x)| \le C\|N\|_{C^{\Sigma_{a,b}}}\|G\|_{\Gamma_{i,s,L-1}}|x|^{(L-b+m)_+} \ .
$$

Hence

$$
\|D^a_\lambda D^L_x[N_\lambda \circ (K^{\le}_\lambda + K^{>}_\lambda)G_\lambda]\|_{\Gamma_{i,s,L-1}} \le C\|N\|_{C^{\Sigma_{i,j}}}\|G\|_{\Gamma_{i,s,L-1}}
$$

and $\tilde{\mathcal{T}}$ is bounded and $\|\tilde{\mathcal{T}}\|$ is as small as we need.

To bound $\tilde{\mathcal{S}}$, we argue as in the proof of Lemma 3.6. It is clear that the series

$$
\sum_{k=0}^\infty A_\lambda^{-(k+1)} \eta_\lambda \circ R^k_\lambda \ DR^k_\lambda
$$

provide the formal solution of $\tilde{\mathcal{S}}H = \eta$. To finish the proof, it is enough to check that

$$
\sum_{k=0}^\infty \|A_\lambda^{-(k+1)} \eta_\lambda \circ R^k_\lambda \ D_x R^k_\lambda\|_{\Gamma_{i,s,L-1}} \le C\|\eta\|_{\Gamma_{i,s,L-1}}
\tag{91}
$$

for some constant C independent of the scaling.

For $(a, b) \in \Sigma_{i,s}$, we have

$$
\begin{aligned}
&D^a_\lambda D^b_x[A_\lambda^{-(k+1)} \eta_\lambda \circ R^k_\lambda \; D_x R^k_\lambda] \\
&\qquad = \sum_{m=0}^{a} C[D_\lambda^{a-m} A_\lambda^{-(k+1)}]\, D^m_\lambda D^b_x[\eta_\lambda \circ R^k_\lambda \; D_x R^k_\lambda] \; .
\end{aligned}
\tag{92}
$$

Using the Faa-di-Bruno formula, (82), (83) and

$$|D^\alpha_\lambda D^\beta_x \eta_\lambda(R^k_\lambda(x))| \le C\|\eta\|_{\Gamma_{i,s,L-1}} |R^k_\lambda(x)|^{(L-\beta)_+} \; ,$$

we deduce that the sum in (92) is bounded by

$$C\|\eta\|_{\Gamma_{i,s,L-1}} \Big(\|A_\lambda^{-1}\|(\|A_1\| + \varepsilon)^{L+1}\Big)^k \; .$$

Moreover, when $b = L-1$, either $\beta \le L-1$ and then $|R^k_\lambda(x)|^{(L-\beta)_+}$ provides a factor $|x|^{(L-\beta)_+}$, or else $\beta > L-1$ and then, since $j_1 + \cdots + j_\beta = b = L-1$, at least one index j of $\{j_1, \ldots, j_\beta\}$ must be zero and then $D^{i_1}_\lambda D^{j_1}_x R^k_\lambda \cdots D^{i_\beta}_\lambda D^{j_\beta}_x R^k_\lambda$ provides a factor $|x|$. This means that

$$
\begin{aligned}
&|D^a_\lambda D^{L-1}_x [A_\lambda^{-(k+1)} \eta \circ R^k_\lambda \; DR^k_\lambda](x)|/|x| \\
&\qquad \le C\|\eta\|_{\Gamma_{i,s,L-1}} \Big(\|A_\lambda^{-1}\|(\|A_1\| + \varepsilon)^{L+1}\Big)^k \; ,
\end{aligned}
$$

which proves Lemma 5.6. □

To finish the proof of Theorem 5.2, that is, to prove that $K^> \in C^{\Sigma_{i,j}}$, we use Lemma 5.6 exactly in the same way as we have used Lemma 3.5 to prove of Proposition 3.5.

5.2. **Dependence on parameters for Theorem 1.2 .**

The following result implies the result on differentiability with respect to parameters claimed in Theorem 1.2.

Theorem 5.7. *Let F be a family of maps. Assume that:*

1) *For some $(i, j) \in (\mathbb{Z}^+)^2$,*

$$F \in C^{\Sigma_{i,j}}$$

where $\Sigma_{i,j}$ is the set introduced in Definition 5.1 and $C^{\Sigma_{i,j}}$ denotes the set of functions which have all the derivatives of indices in $\Sigma_{i,j}$ defined, continuous and bounded.

2) 2.1) *For all $\lambda \in \Lambda$, $D_x F_\lambda(0)$ preserves the space X_1, which is independent of λ.*

   2.2) *For all values $\lambda \in \Lambda$, conditions 0)–4) of Theorem 1.2 are satisfied with uniform constants.*

3) $j \ge L+1$.

Then, the family of maps K_λ given by Theorem 1.2 belongs to $C^{\Sigma_{i,j}}$.

We call attention to Remark 16 after Theorem 5.2 about the meaning of 2.1).

The proof of the regularity with respect to parameters claimed in Theorem 1.2 using Theorem 5.7 is the same as the one corresponding to Theorem 1.1 using Theorem 5.2 and that was given in the previous subsection.

We consider the parameter version of the operators $\mathcal{N}$ and $\mathcal{M}$ introduced in Section 4,

$$(\mathcal{N}(w))_\lambda = A_2^{-1} w_\lambda \circ \psi_{w_\lambda} - A_2^{-1} N_2 \circ (\mathrm{Id}, w_\lambda)$$

where

$$\psi_{w_\lambda} = A_1 + N_1 \circ (\mathrm{Id}, w_\lambda) + B_\lambda w_\lambda \ ,$$

$$(\mathcal{M}(w^>))_\lambda = (\mathcal{N}(w^\leq + w^>))_\lambda - w_\lambda^\leq \tag{93}$$

and $w^\leq$ is the polynomial (in x) given by Lemma 4.1, which belongs to $C^{i,\infty}$. We will write $w_\lambda = w_\lambda^\leq + w_\lambda^>$, and ψ_λ for ψ_{w_λ}.

Proposition 5.8. *Assume that $w^\leq$ is C^i with respect to λ, and that F and $w^>$ belong to $C^{\Sigma_{i,j}}$. Then, the following formula holds*:

$$\begin{aligned}
&[D_\lambda^i D_x^j \mathcal{M}(w^>)]_\lambda \\
&= \ A_2^{-1} D_\lambda^i D_x^j w_\lambda^> \circ \psi_\lambda (D_x \psi_\lambda)^{\otimes j} \\
&\quad + A_2^{-1} \Big[D_x w_\lambda \circ \psi_\lambda (D_2 N_1 \circ (\mathrm{Id}, w_\lambda) + B_\lambda) - D_2 N_2 \circ (\mathrm{Id}, w_\lambda) \Big] D_\lambda^i D_x^j w_\lambda^> \\
&\quad + V_{i,j}(w_\lambda^>) \ ,
\end{aligned} \tag{94}$$

where $V_{i,j}$ is a multilinear operator in $D_\lambda^a D_x^b w^>$ with $(a,b) \in \Sigma_{i,j} - \{(i,j)\}$, whose coefficients are derivatives of N of orders in $\Sigma_{i,j}$, or derivatives of A_1 or B with respect to λ of orders in $\{0, \ldots, i\}$.

Proof. The proof is an immediate consequence of the following lemmas, which provide a quite detailed structure of the derivatives of the composition. □

As in the previous subsection, here we see that the rules for computing higher order derivatives with respect to parameters of a composition suggest the introduction of special sets of indices.

We introduce two more auxiliary sets of indices for $i \in \mathbb{N}, j \geq 1$:

$$\Sigma_{i,0}^{**} = \{(a,b,0) \in (\mathbb{Z}^+)^3 \mid \ a+b \leq i, \ b \geq 1\} \cup \{(i,0,0)\},$$

$$\Sigma_{i,j}^{**} = \{(a,b,c) \in (\mathbb{Z}^+)^3 \mid \ a+b+c \leq i+j, \ a \leq i, \ c \leq j, \ b \geq 1\} \ .$$

In the rest of the section, C means a generic constant which depends only on the map F. Hence, the same letter may denote different constants in different places.

Lemma 5.9. *Under the hypotheses of Proposition 5.8, with the notation above, we have that $\mathcal{M}(w^>) \in C^{\Sigma_{i,j}}$ and that*

$$D_\lambda^i D_x^j [\mathcal{M}(w^>)]$$

is an expression of the form

$$
\begin{aligned}
&A_2^{-1} \sum_{(a,b)\in\Sigma^*_{i,j}} \sum_{\substack{i_1+\cdots+i_b=i-a,\\ j_1+\cdots+j_b=j}} C D_\lambda^a D_x^b w_\lambda \circ \psi_\lambda D_\lambda^{i_1} D_x^{j_1} \psi_\lambda \cdots D_\lambda^{i_b} D_x^{j_b} \psi_\lambda \\
&+ (D_\lambda A_2^{-1}) D_\lambda^{i-1} D_x^j [w_\lambda \circ \psi_\lambda] + \cdots + (D_\lambda^i A_2^{-1}) D_x^j [w_\lambda \circ \psi_\lambda] \\
&- A_2^{-1} \sum_{(a,b,c)\in\Sigma^{**}_{i,j}} \sum_{\substack{i_1+\cdots+i_b=i-a,\\ j_1+\cdots+j_b=j-c}} C D_\lambda^a D_1^c D_2^b N_2 \circ (\mathrm{Id}, w_\lambda)\, D_\lambda^{i_1} D_x^{j_1} w_\lambda \cdots D_\lambda^{i_b} D_x^{j_b} w_\lambda \\
&- (D_\lambda A_2^{-1}) D_\lambda^{i-1} D_x^j [N_2 \circ (\mathrm{Id}, w_\lambda)] - \cdots - (D_\lambda^i A_2^{-1}) D_x^j [N_2 \circ (\mathrm{Id}, w_\lambda)] \\
&- D_\lambda^i D_x^j w_\lambda^{\le} \ .
\end{aligned}
$$

Proof. The proof of Lemma 5.9 follows directly from the next two lemmas. □

Lemma 5.10. *Let $w \in C^{\Sigma_{i,j}}$ and $\psi \in C^{i,j}$. Then the derivatives of $w_\lambda \circ \psi_\lambda(x) = w(\lambda, \psi(\lambda, x))$ have the form*

$$
\begin{aligned}
D_x^j D_\lambda^i [w_\lambda \circ \psi_\lambda] = &\sum_{(a,b)\in\Sigma^*_{i,j}} \sum_{\substack{i_1+\cdots+i_b=i-a,\\ j_1+\cdots+j_b=j}} C D_\lambda^a D_x^b w_\lambda \circ \psi_\lambda \\
&\cdot D_\lambda^{i_1} D_x^{j_1} \psi_\lambda \cdots D_\lambda^{i_b} D_x^{j_b} \psi_\lambda \ ,
\end{aligned}
$$

where C is a combinatorial coefficient which depends on $a, b, i_1, \dots, i_b, j_1, \dots, j_b$, and $\Sigma^*_{i,j}$ are defined in (74) *and* (75).

Proof. The proof of this lemma is contained in the proof of Lemma 5.3 when one identifies w and ψ with H and R^k of that lemma respectively. □

Lemma 5.11. *Let $N \in C^{\Sigma_{i_0,j_0}}$ and $w \in C^{\Sigma_{i_0,j_0}}$. Then we have that $N_\lambda \circ (\mathrm{Id}, w_\lambda) \in C^{\Sigma_{i_0,j_0}}$, and if $(i,j) \in \Sigma_{i_0,j_0}$*

$$
\begin{aligned}
&D_\lambda^i D_x^j [N_\lambda \circ (\mathrm{Id}, w_\lambda)] \\
&\quad = \sum_{(a,b,c)\in\Sigma^{**}_{i,j}} \sum_{i_1,j_1,\dots,i_b,j_b} C D_\lambda^a D_1^c D_2^b N_\lambda \circ (\mathrm{Id}, w_\lambda)\, D_\lambda^{i_1} D_x^{j_1} w_\lambda \cdots D_\lambda^{i_b} D_x^{j_b} w_\lambda
\end{aligned} \tag{95}
$$

where $i_1 + \cdots + i_b = i - a$, $j_1 + \cdots + j_b = j - c$ and C is a coefficient depending on the indices. The only term in the sum (95) *that contains $D_\lambda^i D_x^j w_\lambda$ is $D_2 N_\lambda \circ (\mathrm{Id}, w_\lambda)\, D_\lambda^i D_x^j w_\lambda$.*

Proof. We begin by computing the i^{th} derivative of $N_\lambda \circ (\mathrm{Id}, w_\lambda)$ with respect to λ, which we denote by $T_i = D_\lambda^i [N_\lambda \circ (\mathrm{Id}, w_\lambda)]$. We have

$$
D_\lambda^i \Big[N_\lambda \circ (\mathrm{Id}, w_\lambda) \Big] = \sum_{(a,b)\in\Sigma^*_{i,0}} \sum C D_\lambda^a D_2^b N_\lambda \circ (\mathrm{Id}, w_\lambda) D_\lambda^{i_1} w_\lambda \cdots D_\lambda^{i_b} w_\lambda \ ,
$$

where the second sum is taken for $i_1 + \cdots + i_b = i - a$ and C is a coefficient which depends on $a, b, i_1, \dots, i_b$. This follows from Lemma 5.10 identifying $N_\lambda(x, .)$ with w_λ, and w_λ with ψ_λ.

Note that $D_x^j T_i = D_\lambda^i D_x^j [N_\lambda \circ (\mathrm{Id}, w_\lambda)]$. The rest of the proof is by induction in j. When $j = 0$ the formula (95) holds. The induction step incrementing j corresponds to taking one more derivative with respect to x. Assuming that (95) is true for j,

$$\begin{aligned} D_x^{j+1} T_i = & \sum_{(a,b,c)\in\Sigma_{i,j}^{**}} \sum_{i_1,j_1,\dots,i_b,j_b} C\Big[D_\lambda^a D_1^{c+1} D_2^b N_\lambda \circ (\mathrm{Id}, w_\lambda) \\ & \qquad + D_\lambda^a D_1^c D_2^{b+1} N_\lambda \circ (\mathrm{Id}, w_\lambda)\, D_x w_\lambda\Big] \cdot [D_\lambda^{i_1} D_x^{j_1} w_\lambda \cdots D_\lambda^{i_b} D_x^{j_b} w_\lambda] \\ & + \sum_{(a,b,c)\in\Sigma_{i,j}^{**}} \sum_{i_1,j_1,\dots,i_b,j_b} C D_\lambda^a D_1^c D_2^b N_\lambda \circ (\mathrm{Id}, w_\lambda) \cdot \\ & \qquad D_x [D_\lambda^{i_1} D_x^{j_1} w_\lambda \cdots D_\lambda^{i_b} D_x^{j_b} w_\lambda] \,. \end{aligned}$$

From the previous formula we see that taking one more derivative has the effect that each term labeled with indices (a, b, c) generates three new terms labeled with indices (a, b, c), $(a, b+1, c)$ and $(a, b, c+1)$, except for the term with $b = 0$ which only appears when $j = 0$ and it must have indices $(i, 0, 0)$. Such term only generates two terms with indices $(i + 1, 0, 0)$ and $(i, 1, 0)$.

For $j > 0$ we decompose $\Sigma_{i,j}^{**} = \tilde\Sigma_{i,j}^0 \cup \cdots \cup \tilde\Sigma_{i,j}^m \cup \cdots \cup \tilde\Sigma_{i,j}^i$, where

$$\tilde\Sigma_{i,j}^m = \{(m, b, c) \mid \ b + c \le i + j - m,\ c \le j,\ b \ge 1\} \,.$$

In this way, it is easy to see that the symbolic process

$$(a, b, c) \mapsto (a, b, c) + (a, b + 1, c) + (a, b, c + 1)$$

is exhaustive from $\tilde\Sigma_{i,j}^m$ to $\tilde\Sigma_{i,j+1}^m$. For the case $j = 0$ the same argument works, but one has to consider the term $(i, 0, 0)$ separately.

We note that the highest derivative of w appears accompanying $D_2 N$ (this means $a = c = 0$, $b = 1$) and it is $D_\lambda^i D_x^j w$. □

Given a subset $S \subset \Lambda \times B_1$, we introduce the space

$$\begin{aligned} &\Gamma_{i,j}^*(S) \\ &= \{v \in C^0(S, L^i(\Lambda_0; L^j(X_1; X_2))) \mid \sup_{(\lambda,x)\in S,\, x\ne 0} |v_\lambda(x)|/|x|^{(L-j+1)_+} < \infty\} \end{aligned}$$

with $\|v\|_{\Gamma_{i,j}^*(S)} = \sup_{(\lambda,x)\in S,\, x\ne 0} |v_\lambda(x)|/|x|^{(L-j+1)_+}$ and, given a set of indices Σ, the space

$$\begin{aligned} &\Gamma(\Sigma, S) \\ &= \{w \in C^\Sigma(S) \mid \sup_{(\lambda,x)\in S,\, x\ne 0} |D_\lambda^a D_x^b w_\lambda(x)|/|x|^{(L-b+1)_+} < \infty, \forall (a, b) \in \Sigma\} \end{aligned}$$

with $\|w\|_{\Gamma(\Sigma,S)} = \max_{(a,b)\in\Sigma} \sup_{(\lambda,x)\in S,\, x\ne 0} |D_\lambda^a D_x^b w_\lambda(x)|/|x|^{(L-b+1)_+}$.

Proposition 5.12. *Assume that $F \in C^{\Sigma_{i_0,j_0}}(\Lambda \times B_1)$, with $j_0 \geq L+1$, and that N has small enough $C^{\Sigma_{i_0,j_0}}$ norm. Assume that $w^{\leq} \in C^{i,\infty}$. Let $(i,j) \in \Sigma_{i_0,j_0}$. Let $w_n^{>}$ be the sequence of families of functions defined by*

$$w_n^{>} = \mathcal{M}^n(0). \tag{96}$$

Assume that for every $(a,b) \in \Sigma'_{i,j} \equiv \Sigma_{i,j} - \{(i,j)\}$,

1) *$w_n^{>}$ converges to $w_\infty^{>}$ in $\Gamma(\Sigma_{0,1}, \Lambda \times B_1)$ and $w_\infty^{>} \in C^{\Sigma'_{i,j}}$.*
2) *For each given compact set $G \subset \Lambda \times B_1$, $w_n^{>}$ converges to $w_\infty^{>}$ in $\Gamma(\Sigma'_{i,j}, G)$.*

Then,

a) *$w_\infty^{>} \in C^{\Sigma_{i,j}}(\Lambda \times B_1)$.*
b) *For each given compact set $G \subset \Lambda \times B_1$, $D_\lambda^i D_x^j w_n^{>}$ converges to $D_\lambda^i D_\lambda^j w_\infty^{>}$ in $\Gamma^*_{i,j}(G)$, and therefore $w_n^{>}$ converges to $w_\infty^{>}$ in $\Gamma(\Sigma_{i,j}, G)$.*

Before proving Proposition 5.12 we will introduce some notation and we will establish some preliminary lemmas.

First we remark that, since $j_0 \geq 1$, if $F \in C^{\Sigma_{i_0,j_0}}$ then $N = F - D_x F(0) \in C^{\Sigma_{i_0,j_0}}$.

We write (96) in inductive form,

$$w_{n+1}^{>} = \mathcal{M}(w_n^{>}), \quad n \in \mathbb{Z}^+, \qquad w_0^{>} = 0 \ . \tag{97}$$

We introduce

$$w_n = w^{\leq} + w_n^{>}$$

and

$$\psi_{\lambda,n}(x) = A_1 x + N_1(x, w_{\lambda,n}(x)) + B_\lambda w_{\lambda,n}(x), \qquad n \in \mathbb{Z}^+ \cup \{\infty\}.$$

Differentiating in (97) we have

$$D_\lambda^i D_x^j w_{\lambda,n+1}^{>} = D_\lambda^i D_x^j [\mathcal{M}(w_{\lambda,n}^{>})] \ . \tag{98}$$

By Proposition 5.8 and Lemmas 5.9 to 5.11 we can expand the right hand side of (98).

We define the operator $\mathcal{A}_n$ in such a way that $\mathcal{A}_n D_\lambda^i D_x^j w_n^{>}$ contains the only terms in such expansion with $D_\lambda^i D_x^j w_n^{>}$. We take $\mathcal{B}_n$ as the sum of all the remaining terms of the expansion. We allow n to be ∞. More explicitly, we define

$$\begin{aligned} [\mathcal{A}_n v]_\lambda &= A_2^{-1} v_\lambda \circ \psi_{\lambda,n} (D_x \psi_{\lambda,n})^{\otimes j} \\ &\quad + A_2^{-1} \Big[D_x w_{\lambda,n} \circ \psi_{\lambda,n} \ (D_2 N_1 \circ (\mathrm{Id}, w_{\lambda,n}) + B_\lambda) - D_2 N_2 \circ (\mathrm{Id}, w_{\lambda,n}) \Big] v_\lambda \end{aligned} \tag{99}$$

for $n \in \mathbb{Z}^+ \cup \{\infty\}$ and

$$\mathcal{B}_n = D_\lambda^i D_x^j [\mathcal{M}(w_n^{>})] - \mathcal{A}_n D_\lambda^i D_x^j w_n^{>}, \qquad n \in \mathbb{Z}^+ \cup \{\infty\},$$

that is, $\mathcal{B}_n$ is the multilinear expression in the derivatives $D_\lambda^a D_x^b w_n^{>}$ for $(a,b) \in \Sigma'_{i,j}$, $(a,b) \neq (0,0)$, which contains all terms in the expansion of the right hand side of (5.9) except the only two ones which contain $D^i_\lambda D^j_x w_n^>$, which are included in $\mathcal{A}_n$.

We remark that $\mathcal{A}_n$ and $\mathcal{B}_n$ actually depend on i and j, but we suppress them from the notation.

To study the convergence of the sequence $D^i_\lambda D^j_x w_n^>$ on compact sets, to each compact set $G \subset \Lambda \times B_1$ we associate to it the larger compact set

$$G^* = \bigcup_{n \in \mathbb{N} \cup \{\infty\},\, l \in \mathbb{Z}^+} \{(\lambda, \psi^l_{\lambda,n}(x)) \mid (\lambda, x) \in G\} ,$$

which has the important property of being invariant by $(\pi_\lambda, \psi_{\lambda,n})$ for every $n \in \mathbb{N} \cup \{\infty\}$, where $\pi_\lambda(\lambda, x) = \lambda$.

Lemma 5.13. *Assume the hypotheses and notation of Proposition 5.12. Then, for $n \in \mathbb{Z}^+ \cup \{\infty\}$,*

a) Given a compact set $G \subset \Lambda \times B_1$, $\mathcal{A}_n : \Gamma^*_{i,j}(G^*) \longrightarrow \Gamma^*_{i,j}(G^*)$ is a well defined operator and, denoting $Y = L(\Gamma^*_{i,j}(G^*), \Gamma^*_{i,j}(G^*))$, its norm satisfies $\|\mathcal{A}_n\|_Y \le \gamma < 1$, for some γ independent on n.
b) $\mathcal{B}_n \in \Gamma^*_{i,j}(G^*)$ and $\|\mathcal{B}_n\|_{\Gamma^*_{i,j}(G^*)} \le M$ for some $M > 0$ independent of n.
c) For every $n \in \mathbb{N}$ and $(\lambda, x) \in \Lambda \times B_1$, $|D^i_\lambda D^j_x w^>_{\lambda,n}(x)| \le M_{i,j} |x|^{(L-j+1)_+}$ for some $M_{i,j}$ independent of n.
d) $\mathcal{B}_n$ converges uniformly on compact sets to $\mathcal{B}_\infty$.

Proof. We begin by noting that from hypothesis 1) of Proposition 5.12 and (95), we have that for $n \in \mathbb{Z}^+ \cup \{\infty\}$

$$|\psi_{\lambda,n}(x)| \le (\|A_1\| + \varepsilon)|x|, \tag{100}$$
$$|D_x \psi_{\lambda,n}(x)| \le \|A_1\| + \varepsilon, \tag{101}$$

for $(\lambda, x) \in \Lambda \times B_1$, with ε as small as we need, and

$$|D^a_\lambda \psi_{\lambda,n}(x)| \le C|x|, \quad a \le i \text{ if } j \ge 1, \text{ or } a \le i-1 \text{ if } j = 0, \tag{102}$$

in a compact subset of $\Lambda \times B_1$ with C being a constant independent of n.

a) $\mathcal{A}_n$ is a linear operator. We estimate its norm:

$$\begin{aligned}
&\|\mathcal{A}_n v\|_{\Gamma^*_{i,j}(G^*)} \\
&= \sup_{(\lambda,x)\in G^*} \frac{1}{|x|^{(L-j+1)_+}} \|A_2^{-1}\| \Big\{ |v_\lambda(\psi_{\lambda,n}(x))|\, |D_x\psi_{\lambda,n}(x)|^j \\
&\quad + \Big[|D_x w_{\lambda,n}(\psi_{\lambda,n}(x))|\, |D_2N_1(x,w_{\lambda,n}(x)) + B_\lambda| \\
&\qquad + |D_2N_2(x,w_{\lambda,n}(x))| \Big] |v_\lambda(x)| \Big\} \\
&\le \sup_{(\lambda,x)\in G^*} \|A_2^{-1}\| \Big\{ \|v\|_{\Gamma^*_{i,j}(G^*)} \frac{|\psi_{\lambda,n}(x)|^{(L-j+1)_+}}{|x|^{(L-j+1)_+}} |D_x\psi_{\lambda,n}(x)|^j \\
&\quad + \Big[|D_x w_{\lambda,n}(\psi_{\lambda,n}(x))|\, |D_2N_1(x,w_{\lambda,n}(x)) + B_\lambda| \\
&\qquad + |D_2N_2(x,w_{\lambda,n}(x))| \Big] \|v\|_{\Gamma^*_{i,j}(G^*)} \Big\} \\
&\le \|A_2^{-1}\| \Big[\Big(\|A_1\| + \varepsilon \Big)^{L+1} + \varepsilon \Big] \|v\|_{\Gamma^*_{i,j}(G^*)}
\end{aligned}$$

by (100), (101) and the fact that $\|N\|_{C^1} + \|B_\lambda\|$ is as small as we want.

b) Note that if $F \in C^{\Sigma_{i_0,j_0}}$ then $N \in C^{\Sigma_{i_0,j_0}}$, if $w \in C^{\Sigma_{i,j}}$ then $N_\lambda \circ (\mathrm{Id}, w_\lambda) \in C^{\Sigma_{i,j}}$, and if $w^> \in C^{\Sigma_{i,j}}$ then $\mathcal{M}(w^>) \in C^{\Sigma_{i,j}}$.

By the way that $\mathcal{N}$ was constructed (see Section 3.1),

$$[D_x^\ell \mathcal{M}(0)]_\lambda(0) = D_x^\ell[(\mathcal{N}(w^\le)_\lambda - w_\lambda^\le](0) = 0, \qquad 0 \le \ell \le L .$$

Then

$$[D_\lambda^a D_x^b \mathcal{M}(0)]_\lambda(0) = 0, \qquad 0 \le a \le i, \quad 0 \le b \le L ,$$

and hence

$$\mathcal{M}(0) \in \Gamma(\Sigma_{i,j}, \Lambda \times B_1) . \tag{103}$$

To prove that $|\mathcal{B}_n(\lambda,x)| \le M|x|^{(L-j+1)_+}$ for $(\lambda,x) \in G^*$, we write $\mathcal{B}_n$ as:

$$D_\lambda^i D_x^j \mathcal{M}(0) + D_\lambda^i D_x^j \Big[\mathcal{M}(w_n^>) - \mathcal{M}(0) \Big] - \mathcal{A}_n D_\lambda^i D_x^j w_n^> .$$

We are going to decompose the terms in the expansion of

$$D_\lambda^i D_x^j \Big[\mathcal{M}(w_n^>) - \tilde{\mathcal{M}}(0) \Big]$$

given by (5.9) in telescopic form, and see that all of them have the factor $|x|^{(L-j+1)_+}$ and are bounded uniformly in n. Since $\mathcal{M}(0) = \mathcal{M}(w_0^>)$, we study the more general expression

$$D_\lambda^i D_x^j \Big[\mathcal{M}(w_n^>) - \tilde{\mathcal{M}}(w_m^>) \Big] - \mathcal{A}_n D_\lambda^i D_x^j w_n^> + \mathcal{A}_m D_\lambda^i D_x^j w_m^> \tag{104}$$

to be able to use the conclusions in the case $m = \infty$.

The following Lemma 5.14 estimates the different kinds of terms that will appear in the telescopic decomposition of (104). These terms are of the following forms:

$$\begin{aligned} T_{1,\lambda} = \Big[& D_\lambda^a D_1^c D_2^b N_1 \circ (\mathrm{Id}, w_{\lambda,n}) \\ & - D_\lambda^a D_1^c D_2^b N_1 \circ (\mathrm{Id}, w_{\lambda,m}) \Big] D_\lambda^{i_1} D_x^{j_1} w_{\lambda,m} \cdots D_\lambda^{i_b} D_x^{j_b} w_{\lambda,m} \end{aligned} \tag{105}$$

with $(a,b,c) \in \Sigma_{i,j}^{**}$ and $i_1+\cdots+i_b = i-m-a$, $0 \le m \le i-a$, $j_1+\cdots+j_b = j-c$,

$$\begin{aligned} T_{2,\lambda} = D_\lambda^a D_1^c D_2^b N_1 \circ (\mathrm{Id}, w_{\lambda,n}) & \Big[D_\lambda^{i_1} D_x^{j_1} w_{\lambda,n} \cdots D_\lambda^{i_b} D_x^{j_b} w_{\lambda,n} \\ & - D_\lambda^{i_1} D_x^{j_1} w_{\lambda,m} \cdots D_\lambda^{i_b} D_x^{j_b} w_{\lambda,m} \Big] \end{aligned} \tag{106}$$

with $(a,b,c) \in \Sigma_{i,j}^{**}$ and $i_1+\cdots+i_b = i-m-a$, $0 \le m \le i-a$, $j_1+\cdots+j_b = j-c$,

$$T_{3,\lambda}(x) = D_\lambda^{\tilde\imath} D_x^{\tilde\jmath} \psi_{\lambda,n}(x) - D_\lambda^{\tilde\imath} D_x^{\tilde\jmath} \psi_{\lambda,m}(x) \tag{107}$$

with $\tilde\imath \le i$, $\tilde\jmath \le j$, $(\tilde\imath, \tilde\jmath) \ne (i,j)$,

$$T_{4,\lambda} = \Big[D_\lambda^a D_x^b w_{\lambda,n} \circ \psi_{\lambda,n} - D_\lambda^a D_x^b w_{\lambda,m} \circ \psi_{\lambda,m} \Big] D_\lambda^{i_1} D_x^{j_1} \psi_{\lambda,m} \cdots D_\lambda^{i_b} D_x^{j_b} \psi_{\lambda,m} \tag{108}$$

with $(a,b) \in \Sigma'_{i,j}$, $i_1 + \cdots + i_b = i-m-a$, $0 \le m \le i-a$, and $j_1 + \cdots + j_b = j$,

$$\begin{aligned} T_{5,\lambda} = {} & D_\lambda^a D_x^b w_{\lambda,n}^{>} \circ \psi_{\lambda,n} \\ & \Big[D_\lambda^{i_1} D_x^{j_1} \psi_{\lambda,n} \cdots D_\lambda^{i_b} D_x^{j_b} \psi_{\lambda,n} - D_\lambda^{i_1} D_x^{j_1} \psi_{\lambda,m} \cdots D_\lambda^{i_b} D_x^{j_b} \psi_{\lambda,m} \Big] \end{aligned} \tag{109}$$

with $(a,b) \in \Sigma'_{i,j}$, $i_1 + \cdots + i_b = i-m-a$, $0 \le m \le i-a$, and $j_1 + \cdots + j_b = j$.

Before finishing the proof of Lemma 5.13, we state and prove the following lemma:

Lemma 5.14. *Assume the hypotheses of Lemma 5.13. Then, there exists a constant C independent on n such that*

$$\begin{aligned} |T_{1,\lambda}(x)| &\le C|x|^{L+1}, \\ |T_{2,\lambda}(x)| &\le C|x|^{(L-j+1)_+}, \\ |T_{3,\lambda}(x)| &\le C|x|^{(L-\tilde\jmath+1)_+}, \\ |T_{4,\lambda}(x)| &\le C|x|^{(L-j+1)_+} \quad \textit{in case that } m=0, \\ |T_{5,\lambda}(x)| &\le C|x|^{(L-j+1)_+} . \end{aligned}$$

Proof. For (105) we distinguish two cases. When $a+b+c < i_0+j_0$, then we use the mean value theorem and we bound T_1 by

$$\|D_\lambda^a D_1^c D_2^{b+1} N_1\|_{C^0} |w_{\lambda,n}(x) - w_{\lambda,m}(x)| \, |D_\lambda^{i_1} D_x^{j_1} w_{\lambda,m}(x) \cdots D_\lambda^{i_b} D_x^{j_b} w_{\lambda,m}(x)| \ ,$$

and we recall that

$$|w_{\lambda,n}(x) - w_{\lambda,m}(x)| = |w^>_{\lambda,n}(x) - w^>_{\lambda,m}(x)| \le \|w^>_n - w^>_m\|_{\Gamma(\Sigma'_{i,j},G^*)}|x|^{L+1}.$$

In the second case when $a+b+c = i_0 + j_0$, since $a+b+c \le i+j$ then $j \ge i_0 + j_0 - i \ge j_0 \ge L+1$ and thus $(L-j+1)_+ = 0$. Hence it is enough to see that T_1 is bounded, which follows immediately form 2) of Proposition 5.12.

Decomposing the differences of (106) in telescopic form, each difference will have a factor of the form

$$|D^{i_l}_\lambda D^{j_l}_x w_{\lambda,n}(x) - D^{i_l}_\lambda D^{j_l}_x w_{\lambda,m}(x)| \le \|w_n - w_m\|_{\Gamma(\Sigma'_{i,j},G^*)}|x|^{(L-j_l+1)_+}, \tag{110}$$

with $j_l \le j$. The factor (110) is bounded by $C|x|^{(L-j+1)_+}$ because $j_l \le j$.

For T_3 we write

$$\begin{aligned}
&D^{\tilde{\imath}}_\lambda D^{\tilde{\jmath}}_x \psi_{\lambda,n} - D^{\tilde{\imath}}_\lambda D^{\tilde{\jmath}}_x \psi_{\lambda,m}\\
&\quad = \sum_{(a,b,c)\in\Sigma^{**}_{\tilde{\imath},\tilde{\jmath}}} \sum_{\substack{i_1+\cdots+i_b=\tilde{\imath}-a,\\ j_1+\cdots+j_b=\tilde{\jmath}-c}}\\
&\quad C\Big[D^a_\lambda D^c_1 D^b_2 N_1 \circ (\mathrm{Id}, w_{\lambda,n})\, D^{i_1}_\lambda D^{j_1}_x w_{\lambda,n} \cdots D^{i_b}_\lambda D^{j_b}_x w_{\lambda,n}\\
&\qquad - D^a_\lambda D^c_1 D^b_2 N_1 \circ (\mathrm{Id}, w_{\lambda,m})\, D^{i_1}_\lambda D^{j_1}_x w_{\lambda,m} \cdots D^{i_b}_\lambda D^{j_b}_x w_{\lambda,m}\Big]\\
&\quad + \sum_{l=0}^{\tilde{\imath}} C D^l_\lambda B\lambda \left[D^{\tilde{\imath}-l}_\lambda D^{\tilde{\jmath}}_x w_{\lambda,n} - D^{\tilde{\imath}-l}_\lambda D^{\tilde{\jmath}}_x w_{\lambda,m}\right].
\end{aligned}$$

We deduce the bound for T_3 claimed in Lemma 5.14 by applying the bounds for T_1 and T_2 with $i = \tilde{\imath}$ and $j = \tilde{\jmath}$.

To establishe the bound for T_4 we take into account that $m = 0$ and hence $w_0 = w^{\le}_0$. Noting that w_0 is a polynomial and, therefore, has derivatives, we bound the first factor in the definition of T_4 by

$$\begin{aligned}
&|D^a_\lambda D^b_x w^>_{\lambda,n} \circ \psi_{\lambda,n}(x)| + |D^a_\lambda D^b_x w_{\lambda,0} \circ \psi_{\lambda,n}(x) - D^a_\lambda D^b_x w_{\lambda,0} \circ \psi_{\lambda,m}(x)|\\
&\quad \le \|D^a_\lambda D^b_x w^>_n\|_{\Gamma(\Sigma'_{i,j},G^*)}|\psi_{\lambda,n}(x)|^{(L-b+1)_+}\\
&\qquad + \|D^a_\lambda D^{b+1}_x w_0\|_{C^0}|\psi_{\lambda,n}(x) - \psi_{\lambda,m}(x)|\\
&\quad (\|A_1\| + \varepsilon)^{(L-b+1)_+}\|D^a_\lambda D^b_x w^>_n\|_{\Gamma(\Sigma'_{i,j},G^*)}|x|^{(L-b+1)_+} + C|x|^{L+1}
\end{aligned}$$

where we have used the bound for T_3.

Moreover the second factor of T_4, $D^{i_1}_\lambda D^{j_1}_x \psi_m \cdots D^{i_b}_\lambda D^{j_b}_x \psi_m$, is such that $j_1 + \cdots + j_b = j$ and hence it must have at least $(j-b)_+$ derivatives with index $j_l = 0$. Therefore by (101) this factor is bounded by $C|x|^{(j-b)_+}$.

For T_5, decomposing the difference

$$D^{i_1}_\lambda D^{j_1}_x \psi_n \cdots D^{i_b}_\lambda D^{j_b}_x \psi_n - D^{i_1}_\lambda D^{j_1}_x \psi_m \cdots D^{i_b}_\lambda D^{j_b}_x \psi_m$$

in a telescopic sum, each term will have a factor

$$|D^{i_l}_\lambda D^{j_l}_x \psi_{\lambda,n}(x) - D^{i_l}_\lambda D^{j_l}_x \psi_{\lambda,m}(x)|$$

which is bounded by $C|x|^{(L-j_l+1)_+} \leq C|x|^{(L-j+1)_+}$ by the bounds of T_3. $\square$

Continuation of the proof of Lemma 5.13. With the estimates of Lemma 5.14 and (103) one easily checks that $\mathcal{B}_n \in \Gamma^*_{i,j}(G^*)$ and that there exists a constant M independent on n such that

$$|\mathcal{B}_n(\lambda, x)| \leq M|x|^{(L-j+1)_+}.$$

The same estimates work in the case $n = \infty$.

To establish c) of Lemma 5.13, from the definition of $\mathcal{A}_n$ and $\mathcal{B}_n$ we write

$$D^i_\lambda D^j_x w^>_{n+1} = \mathcal{A}_n D^i_\lambda D^j_x w^>_n + \mathcal{B}_n\ .$$

Then

$$\|D^i_\lambda D^j_x w^>_{n+1}\|_{\Gamma^*_{i,j}(G^*)} \leq \|\mathcal{A}_n\|_Y\, \|D^i_\lambda D^j_x w^>_n\|_{\Gamma^*_{i,j}(G^*)} + \|\mathcal{B}_n\|_{\Gamma^*_{i,j}(G^*)}$$

and hence c) follows taking $M_{i,j} = \frac{M}{1-\gamma}$.

Now we will prove d) of Lemma 5.13. Given a compact set G of $\Lambda \times B_1$, we have to prove that $\mathcal{B}_n \to \mathcal{B}_\infty$ converges uniformly on G. For that we have to consider the differences (104) with $m = \infty$ in the larger compact set G^*. Examining the terms T_1 to T_5 in the analogous way as in Lemma 5.14, we see that all terms are bounded by bounded factors times one of the following terms:

$$\|w^>_n - w^>_m\|_{\Gamma(\Sigma'_{i,j}, G^*)} \tag{111}$$

$$|D^a_\lambda D^c_1 D^b_2 N_j \circ (x, w_{\lambda,n}(x)) - D^a_\lambda D^c_1 D^b_2 N_j \circ (x, w_{\lambda,\infty}(x))| \tag{112}$$

$$\|D^a_\lambda D^b_x w_{\lambda,\infty} \circ \psi_{\lambda,n} - D^a_\lambda D^b_x w_{\lambda,\infty} \circ \psi_{\lambda,\infty}\| \tag{113}$$

and

$$\|D^a_\lambda D^b_x w_{\lambda,n} \circ \psi_{\lambda,n} - D^a_\lambda D^b_x w_{\lambda,\infty} \circ \psi_{\lambda,n}\|\ . \tag{114}$$

The terms (112), (113) and (114) go to zero because $D^a_\lambda D^c_1 D^b_2 N_j$, $D^a_\lambda D^b_x w_\infty$ are continuous and $\{(\lambda, x, w_{\lambda,\infty}(x)) \mid\ (\lambda, x) \in G^*\}$ and

$$\{(\lambda, \psi_{\lambda,m}(x)) \mid\ (\lambda, x) \in G^*,\ m \in \mathbb{Z}^+ \cup \{\infty\}\}$$

are compact. $\square$

Proof of Proposition 5.12. First we establish part b). If we call $v_n = D^i_\lambda D^j_x w^>_n$, from the definition of $w^>_n$, (98) and the definitions of $\mathcal{A}_n$ and $\mathcal{B}_n$ we have

$$\begin{aligned} v_0 &= 0, \\ v_{n+1} &= \mathcal{A}_n v_n + \mathcal{B}_n, \qquad n \geq 0. \end{aligned}$$

Hence we can write

$$v_{n+1} = \mathcal{B}_n + \mathcal{A}_n \mathcal{B}_{n-1} + \cdots + \mathcal{A}_n \mathcal{A}_{n-1} \cdots \mathcal{A}_1 \mathcal{B}_0\ .$$

We claim that v_n converges to

$$\mathcal{B}_\infty + \mathcal{A}_\infty \mathcal{B}_\infty + \mathcal{A}_\infty^2 \mathcal{B}_\infty + \dots$$

uniformly on compact sets. Indeed

$$\begin{aligned}
&\mathcal{A}_n \mathcal{A}_{n-1} \cdots \mathcal{A}_{n-k} \mathcal{B}_{n-k-1} - \mathcal{A}_\infty^{k+1} \mathcal{B}_\infty \\
&\quad = \mathcal{A}_n \mathcal{A}_{n-1} \cdots \mathcal{A}_{n-k} (\mathcal{B}_{n-k-1} - \mathcal{B}_\infty) + \mathcal{A}_n \mathcal{A}_{n-1} \cdots \mathcal{A}_{n-k} \mathcal{B}_\infty - \mathcal{A}_\infty^{k+1} \mathcal{B}_\infty \\
&\quad = \mathcal{A}_n \mathcal{A}_{n-1} \cdots \mathcal{A}_{n-k} (\mathcal{B}_{n-k-1} - \mathcal{B}_\infty) \\
&\qquad + \sum_{i=1}^{k} \mathcal{A}_n \cdots \mathcal{A}_{n-i+1} (\mathcal{A}_{n-i} - \mathcal{A}_\infty) \mathcal{A}_\infty^{k-i} \mathcal{B}_\infty \\
&\qquad + (\mathcal{A}_n - \mathcal{A}_\infty) \mathcal{A}_\infty^k \mathcal{B}_\infty
\end{aligned}$$

and since given g defined on $\Lambda \times B_1$, $\mathcal{A}_n g \to \mathcal{A}_\infty g$ uniformly on compact sets G^* of $\Lambda \times B_1$ then

$$\mathcal{A}_n \mathcal{A}_{n-1} \cdots \mathcal{A}_{n-k} \mathcal{B}_{n-k-1} \to \mathcal{A}_\infty^{k+1} \mathcal{B}_\infty$$

uniformly on compact sets.

For $k < n$ we have

$$\begin{aligned}
&v_{n+1}(x) - \Big(\mathcal{B}_n(x) + \mathcal{A}_n \mathcal{B}_{n-1}(x) + \cdots + \mathcal{A}_n \mathcal{A}_{n-1} \cdots \mathcal{A}_{n-k} \mathcal{B}_{n-k-1}(x)\Big) \\
&\qquad = \mathcal{A}_n \mathcal{A}_{n-1} \cdots \mathcal{A}_{n-k-1} \mathcal{B}_{n-k-2}(x) + \cdots + \mathcal{A}_n \mathcal{A}_{n-1} \cdots \mathcal{A}_1 \mathcal{B}_0(x)
\end{aligned}$$

and we can bound

$$\begin{aligned}
&|v_{n+1}(x) - \Big(\mathcal{B}_n(x) + \mathcal{A}_n \mathcal{B}_{n-1}(x) + \cdots + \mathcal{A}_n \mathcal{A}_{n-1} \cdots \mathcal{A}_{n-k} \mathcal{B}_{n-k-1}(x)\Big)| \\
&\quad \le \sum_{l=2}^{n-k} \nu^{k+l} \|\mathcal{B}_{n-k-l}\|_{\Gamma^*_{i,j}(G^*)} |x|^{(L-j+1)_+} \\
&\quad \le \frac{\nu^{k+2}}{1-\nu} |x|^{(L-j+1)_+} \sup_{0 \le l \le n-k} \|\mathcal{B}_l\|_{\Gamma^*_{i,j}(G^*)} .
\end{aligned}$$

Taking the limit $n \to \infty$, we obtain

$$\begin{aligned}
&|v_\infty(x) - (\mathcal{B}_\infty(x) + \mathcal{A}_\infty \mathcal{B}_\infty(x) + \cdots + \mathcal{A}_\infty^{k+1} \mathcal{B}_\infty(x))| \\
&\qquad \le \frac{\nu^{k+2}}{1-\nu} M |x|^{(L-j+1)_+} \sup_{l \ge 0} \|\mathcal{B}_l\|_{\Gamma^*_{i,j}(G^*)} .
\end{aligned}$$

Then $D^i_\lambda D^j_x w_n^>$ converges to

$$v_\infty = \mathcal{B}_\infty + \mathcal{A}_\infty \mathcal{B}_\infty + \mathcal{A}_\infty^2 \mathcal{B}_\infty + \dots \tag{115}$$

uniformly on compact sets G^* as well as with the norm $\|\cdot\|_{\Gamma^*_{i,j}(G^*)}$.

Note that, since $\mathcal{A}_\infty$ is a contraction in $\Gamma^*_{i,j}(G^*)$ and $\mathcal{B}_\infty$ is in $\Gamma^*_{i,j}(G^*)$, then the resulting function v_∞ is in $\Gamma^*_{i,j}(G^*)$. Moreover $\mathcal{B}_\infty$ is continuous. Then v_∞ is the only solution of

$$u = \mathcal{A}_\infty u + \mathcal{B}_\infty$$

and hence it is continuous.

If $1 \le j \le i_0 - i + j_0$ we will show that the D_x derivative of $D_\lambda^i D_x^{j-1} w_n^>$ converges uniformly to v_∞ on compact sets of $\Lambda \times B_1$.

Since for any n we have

$$D_\lambda^i D_x^{j-1} w_{\lambda,n}^>(y) - D_\lambda^i D_x^{j-1} w_{\lambda,n}^>(x) = \int_0^1 dt\, D_\lambda^i D_x^j w_{\lambda,n}^>(x + t(y-x))(y-x),$$

passing to the limit, using that $D_\lambda^i D_x^j w_n^>$ converges to v_∞ in the compact set G^* with $G = \{x + t(y-x) \mid\ t \in [0,1]\}$

$$D_\lambda^i D_x^{j-1} w_{\mu,\infty}^>(x) - D_\lambda^i D_x^{j-1} w_{\lambda,\infty}^>(x) = \int_0^1 dt\, v_{\lambda,\infty}(x + t(y-x))(y-x),$$

and using that v_∞ is continuous, we obtain that v_∞ is indeed the $D_\lambda^i D_x^j$ derivative of $w^>$.

If $j = 0$ and $i \ge 1$ we show that the D_λ derivative of $D_\lambda^{i-1} D_x^j w_n^>$ converges uniformly to v_∞ on compact sets of $\Lambda \times B_1$.

From the identity

$$D_\lambda^{i-1} D_x^j w_{\mu,n}^>(x) - D_\lambda^{i-1} D_x^j w_{\lambda,n}^>(x) = \int_0^1 dt\, D_\lambda^i D_x^j w_{\lambda+t(\mu-\lambda),n}^>(x)(\mu-\lambda),$$

passing to the limit

$$D_\lambda^{i-1} D_x^j w_{\mu,\infty}^>(x) - D_\lambda^{i-1} D_x^j w_{\lambda,\infty}^>(x) = \int_0^1 dt\, v_{\lambda+t(\mu-\lambda),\infty}(x)(\mu-\lambda)$$

and the argument finishes as before. This ends the proof of Proposition 5.12. $\square$

Now, we are in a position to complete the proof of Theorem 5.7.

Proof of Theorem 5.7. Since $F \in C^{\Sigma_{i,j}}$ for fixed λ, as a function of x we have that $F_\lambda(\cdot) \in C^{i+j}$.

From Proposition 4.3 we know that $w_\lambda^> = \lim_n \mathcal{M}^n(0)$, where the limit is obtained with λ fixed and in the topology given in the space $\Gamma_{r-1,L}$, with $r = i+j$, and that this limit is C^{i+j}.

This means that $\|D_x^{r-1} w_{\lambda,n}^> - D_x^{r-1} w_{\lambda,\infty}^>\|_{C^0} \to 0$ and therefore, for $0 \le j \le r-1$,

$$\sup \frac{|D_x^j w_{\lambda,n}^>(x) - D_x^j w_{\lambda,\infty}^>(x)|}{|x|^{(L-j+1)_+}} \le \frac{1}{(L-j)_+!} \|w_{\lambda,n}^> - w_{\lambda,\infty}^>\|_{C^{r-1}} \ .$$

Then, we apply the induction Proposition 5.12 to $i_0 = 0$ starting with $j_0 = i + j - 1$, and then we apply Proposition 5.12 repeatedly from $i_0 = 1$ to $i_0 = i$ and for each i_0 from $j_0 = 0$ to $j_0 = i + j - i_0$. $\square$

6. Acknowledgments

This work has been supported by the *Comisión Conjunta Hispano Norte-americana de Cooperación Científica y Tecnológica.* The final version was prepared while Rafael de la Llave was enjoying a *Cátedra de la Fundación FBBV.* X.C. was partially supported by CICYT (Spain), grant PB98-0932-C02-01. E.F.

acknowledges the partial support of the the Spanish Grant DGICYT BFM2000-0805, the Catalan grant CIRIT 2000SGR–00027 and the INTAS project 00-221. The work of R.L. has been supported by NSF grants. R.L. thanks Universitat de Barcelona and Universitat Politècnica de Catalunya for hospitality during several visits.

Appendix A. Some results in spectral theory

The results of this Appendix are well known in finite dimensional spaces (see, for instance, [Nel69]). Here we present proofs which are also valid in infinite dimensions and, at the same time, we simplify some of the proofs in the finite dimensional case.

We recall that given a bounded linear operator A on a Banach space X, we say that $\lambda \in \mathrm{Res}(A)$ (the resolvent set of A) if and only if $A - \lambda := A - \lambda\,\mathrm{Id}$ is invertible, i.e., $A - \lambda$ is one to one and onto. By the open mapping theorem, the inverse $(A-\lambda)^{-1}$ is automatically a bounded operator. We also define the spectrum of A by $\mathrm{Spec}(A) = \mathbb{C} - \mathrm{Res}(A)$, a compact subset of $\mathbb{C}$.

An important subset of the spectrum of A is given by the approximate point spectrum of A, denoted by $\mathrm{Spec}_{ap}(A)$. By definition,

$$\lambda \in \mathrm{Spec}_{ap}(A) \Leftrightarrow \exists \{x_n\} \text{ with } \|x_n\| \geq \alpha > 0 \text{ and } \|(A-\lambda)x_n\| \to 0 \ . \tag{116}$$

The sequence $\{x_n\}$ is called an approximate eigenvector of A for λ.

A well known result states that the boundary of the spectrum is contained in the approximate point spectrum, i.e.,

$$\partial(\mathrm{Spec}(A)) \subset \mathrm{Spec}_{ap}(A). \tag{117}$$

This is easily verified as follows. Let $\lambda_n \to \lambda$, with $\lambda \in \mathrm{Spec}(A)$ and $\lambda_n \in \mathrm{Res}(A)$ for all n. Recall that, for every operator B with $\|B\| < 1$, $\mathrm{Id} + B$ is invertible (the inverse of $\mathrm{Id} + B$ is just given by the standard Neumann power series). This fact and the identity $A - \lambda = (A - \lambda_n)\{\mathrm{Id} + (\lambda_n - \lambda)(A-\lambda_n)^{-1}\}$ lead to $|\lambda - \lambda_n| \|(A-\lambda_n)^{-1}\| \geq 1$. Hence, there exists y_n such that

$$\|(A-\lambda_n)^{-1} y_n\| \geq \frac{\|y_n\|}{2|\lambda - \lambda_n|}.$$

By scaling y_n, we may also assume that $x_n := (A-\lambda_n)^{-1} y_n$ satisfies $\|x_n\| = 1$. Since $\|(A-\lambda_n)x_n\| = \|y_n\| \leq 2|\lambda - \lambda_n|$, we have that $\|(A-\lambda)x_n\| \leq 3|\lambda-\lambda_n| \to 0$, and hence that $\lambda \in \mathrm{Spec}_{ap}(A)$.

A.1. **Adapted norms.** Let X be a Banach space and A a linear map on X. It is well known that given any $\varepsilon > 0$, we can find a norm in X equivalent to the original one and such that

$$\|A\| \leq \rho(A) + \varepsilon,$$

where $\rho(A) = \sup_{z \in \mathrm{Spec}(A)} |z|$ is the spectral radius of A. We shall say that such norm is ε-adapted to A. Moreover, if A is invertible we can find a norm in X

equivalent to the original one and ε-adapted to A and to A^{-1} simultaneously, i.e.,

$$\|A\| \le \rho(A) + \varepsilon, \qquad \|A^{-1}\| \le \rho(A^{-1}) + \varepsilon \ .$$

An example of norm ε-adapted to A and A^{-1} is given by

$$|||x||| = \sum_{i=0}^{\infty} (\rho(A) + \varepsilon)^{-i} \|A^i x\| + \sum_{i=1}^{\infty} (\rho(A^{-1}) + \varepsilon)^{-i} \|A^{-i} x\| \ .$$

To verify this, it suffices to use the well known fact that

$$\rho(A) = \lim_{i \to +\infty} \|A^i\|^{1/i}.$$

A more general result is the following:

Proposition A.1. *Let X be a Banach space, $X = X_1 \oplus X_2$ be a direct sum decomposition into closed subspaces, and let*

$$A = \begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix}$$

with respect to the above decomposition. Then,

a) $\rho(A) = \max\big(\rho(A_1), \rho(A_2)\big)$.

b) *Assume further that A_1 and A_2 are invertible. Then, for every $\varepsilon > 0$, there exists a norm in X which is equivalent to the original one and such that*

$$\begin{array}{rclrcl} \|A_1\| & \le & \rho(A_1) + \varepsilon \ , & \|A_1^{-1}\| & \le & \rho(A_1^{-1}) + \varepsilon \ , \\ \|A_2\| & \le & \rho(A_2) + \varepsilon \ , & \|A_2^{-1}\| & \le & \rho(A_2^{-1}) + \varepsilon \ , \\ \|A\| & \le & \rho(A) + \varepsilon \ , & \|A^{-1}\| & \le & \rho(A^{-1}) + \varepsilon \ , \qquad \textit{and} \\ \|B\| & \le & \varepsilon \ . & & & \end{array}$$

Proof. To prove part b) once we have established a), it suffices to construct norms $\|\cdot\|_{X_1}$, $\|\cdot\|_{X_2}$ in X_1 and X_2 which are $(\varepsilon/2)$-adapted to A_1 and A_1^{-1}, and to A_2 and A_2^{-1}, respectively. Note that

$$A^{-1} = \begin{pmatrix} A_1^{-1} & -A_1^{-1} B A_2^{-1} \\ 0 & A_2^{-1} \end{pmatrix} \ .$$

Hence, by part a) applied to A and to A^{-1}, we have

$$\rho(A) = \max\big(\rho(A_1), \rho(A_2)\big) \qquad \text{and} \qquad \rho(A^{-1}) = \max\big(\rho(A_1^{-1}), \rho(A_2^{-1})\big) \ .$$

Using these equalities and defining, for $\delta > 0$ sufficiently small,

$$\|(x^1, x^2)\|_X = \max(\delta \|x^1\|_{X_1}, \|x^2\|_{X_2}) \qquad \text{for } x = (x^1, x^2) \ ,$$

it is easy to verify all the statements of part b).

Next, we prove a). Since $(A - \lambda)x = y$ is equivalent to

$$\begin{aligned} &(A_1 - \lambda)x^1 + Bx^2 = y^1 \\ &(A_2 - \lambda)x^2 = y^2 \ , \end{aligned}$$

we see that

$$\lambda \in \mathrm{Res}(A_1) \cap \mathrm{Res}(A_2) \Rightarrow \lambda \in \mathrm{Res}(A) \ .$$

Hence

$$\mathrm{Spec}(A) \subset \mathrm{Spec}(A_1) \cup \mathrm{Spec}(A_2) \tag{118}$$

and, in particular, $\rho(A) \leq \max\big(\rho(A_1), \rho(A_2)\big)$. Therefore, we only need to show that

$$\rho(A) \geq \max\big(\rho(A_1), \rho(A_2)\big) \ . \tag{119}$$

To prove this, we first claim that

$$\mathrm{Spec}_{ap}(A) \supset \mathrm{Spec}_{ap}(A_1) \cup \Big(\mathrm{Spec}_{ap}(A_2) \cap \mathrm{Res}(A_1) \Big) \ . \tag{120}$$

Then, using that $\partial(\mathrm{Spec}(A_i)) \subset \mathrm{Spec}_{ap}(A_i)$ for $i = 1, 2$ (i.e., property (117) applied to A_1 and A_2), we conclude inequality (119) and the proposition.

To establish (120), let first $\lambda \in \mathrm{Spec}_{ap}(A_1)$ and $\{x_n^1\}$ be an approximate eigenvector of A_1 for λ. Then, $\{x_n\} = \{(x_n^1, 0)\}$ clearly is an approximate eigenvector of A for λ.

Finally, if $\lambda \in \mathrm{Spec}_{ap}(A_2) \cap \mathrm{Res}(A_1)$ and $\{x_n^2\}$ is an approximate eigenvector of A_2 for λ, then

$$\{x_n\} = \{(-(A_1 - \lambda)^{-1} B x_n^2 \ , \ x_n^2)\}$$

is an approximate eigenvector of A for λ. This establishes (120). □

Of course, in finite dimensions we have

$$\mathrm{Spec}(A) = \mathrm{Spec}(A_1) \cup \mathrm{Spec}(A_2) \ .$$

However, in infinite dimensions the inclusion (118) could be strict, as the following example shows.

Example A.2. Let $\Sigma = \{\sigma : \mathbb{N} \longrightarrow X\}$ be the Banach space of sequences into a Banach space X equipped with the sup norm (an ℓ^2 norm would also produce similar effects). Let $A : \Sigma \times \Sigma \longrightarrow \Sigma \times \Sigma$ be defined by

$$A(\sigma, \tau) = (\tilde{\sigma}, \tilde{\tau})$$

with $\tilde{\sigma}(1) = \tau(1)$, $\tilde{\sigma}(i) = \sigma(i-1)$ for $i \geq 2$, and $\tilde{\tau}(j) = \tau(j+1)$ for $j \geq 1$. Then, it is easy to verify that A is invertible, but A_1 and A_2 are not.

In [dlL98] one can find examples arising in dynamical systems where both (118) and (120) are strict. This is due to the presence of residual spectrum.

A.2. **Proof of Proposition 3.2 .** The proof of the equalities in (25) when X and Y are finite dimensional is very easy, and we do it first.

Since the spectrum depends continuously on the matrix, it suffices to establish the equalities when A and B are diagonalizable matrices (a dense subset in the space of matrices).

Let us first consider the operator $\mathcal{L}_{n,A,B}$. If $Au_j = \mu_j u_j$, $Bv_i = \lambda_i v_i$, consider the form $M_{i;j_1,\dots,j_n}$ defined by the conditions

$$\begin{aligned} M_{i;j_1,\dots,j_n}(u_{j_1},\dots,u_{j_n}) &= v_i, \\ M_{i;j_1,\dots,j_n}(u_{s_1},\dots,u_{s_n}) &= 0, \qquad \text{when } (s_1,\dots,s_n) \neq (j_1,\dots,j_n) \ . \end{aligned}$$

Clearly,

$$\mathcal{L}_{n,A,B} M_{i;j_1,\dots,j_n} = \lambda_i \mu_{j_1} \cdots \mu_{j_n} M_{i;j_1,\dots,j_n} \ .$$

Moreover, the set formed by the $M_{i;j_1,\dots,j_n}$ is linearly independent and, under the assumption that A,B are diagonalizable, its cardinal is equal to the dimension of $\mathbf{M}_n$. Hence, the spectrum of $\mathcal{L}_{n,A,B}$ is indeed the set of numbers $\lambda_i \mu_{j_1} \cdots \mu_{j_n}$, as claimed.

The equality for the operator $\mathcal{L}_B$ is a particular case of the previous one, since $\mathcal{L}_B = \mathcal{L}_{n,\mathrm{Id},B}$. A similar argument also proves the equality for the operator $\mathcal{R}^k_A$.

For the case of symmetric forms, a very similar argument works. We can consider the form $S_{i;j_1,\dots,j_n} = \sum_\pi M_{i;j_{\pi(1)},\dots,j_{\pi(n)}}$ where the variable π in the sum runs over the permutations of $\{1,\dots,n\}$. Each form $S_{i;j_1,\dots,j_n}$ is an eigenvector of $\mathcal{L}_{n,A,B}$ with eigenvalue $\lambda_i \mu_{j_1} \cdots \mu_{j_n}$. Moreover, they are linearly independent, and there are as many of them as the dimension of the space $\mathbf{S}_n$. This proves the equality for the spectrum of $\mathcal{L}_{n,A,B}$ in $\mathbf{S}_n$.

Now, we turn to the proof of Proposition 3.2 for general Banach spaces.

First note that the operators $\mathcal{L}_B, \mathcal{R}^1_A, \dots, \mathcal{R}^n_A$ commute, and that

$$\mathcal{L}_{n,A,B} = \mathcal{L}_B \mathcal{R}^1_A \cdots \mathcal{R}^n_A \ . \tag{121}$$

Moreover, we have:

Proposition A.3. *With the notations of Proposition 3.2, one has*

$$\mathrm{Spec}(\mathcal{L}_B, \mathbf{M}_n) \subset \mathrm{Spec}(B, Y) \ , \tag{122}$$

$$\mathrm{Spec}(\mathcal{R}^k_A, \mathbf{M}_n) \subset \mathrm{Spec}(A, X) \ . \tag{123}$$

Proof. Note that $\mathcal{L}_{BB'} = \mathcal{L}_B \mathcal{L}_{B'}$ and $\mathcal{L}_{B+B'} = \mathcal{L}_B + \mathcal{L}_{B'}$, so that $\mathcal{L}$ can be considered as a representation of the Banach algebra of bounded operators in Y into the Banach algebra of operators in $\mathbf{M}_n$. Therefore, the spectrum is smaller. Similarly for all the $\mathcal{R}^k_A$. □

We also recall the following well known result in Banach algebras (see e.g., Theorem 11.23 in [Rud91]).

Theorem A.4. *Let $\mathcal{A}$ and $\mathcal{B}$ be two commuting elements in a Banach algebra. Then*

$$\mathrm{Spec}(\mathcal{A}\mathcal{B}) \subset \mathrm{Spec}(\mathcal{A})\,\mathrm{Spec}(\mathcal{B}) \ .$$

The result of Proposition 3.2 for $\mathbf{M}_n$ follows immediately from the fact that the operators $\mathcal{L}_B, \mathcal{R}^1_A, \dots, \mathcal{R}^n_A$ commute, Theorem A.4, (121), (122) and (123).

The result for $\mathbf{S}_n$ follows from the following:

Proposition A.5. *With the notations of Proposition 3.2, we have*

$$\operatorname{Spec}(\mathcal{L}_{n,A,B}, \mathbf{S}_n) \subset \operatorname{Spec}(\mathcal{L}_{n,A,B}, \mathbf{M}_n) \ . \tag{124}$$

Proof. If $\lambda \notin \operatorname{Spec}(\mathcal{L}_{n,A,B}, \mathbf{M}_n)$, given any $\eta \in \mathbf{M}_n$ there is a unique $\Gamma \in \mathbf{M}_n$ such that

$$(\mathcal{L}_{n,A,B} - \lambda)\Gamma = \eta \ . \tag{125}$$

Such Γ satisfies $\|\Gamma\| \le K\|\eta\|$.

If $\Pi(x_1, \ldots, x_n) = (x_{\pi(1)}, \ldots, x_{\pi(n)})$ is a permutation of the n vectors, it is easy to check that

$$(\mathcal{L}_{n,A,B} - \lambda)(\Gamma\Pi) = [(\mathcal{L}_{n,A,B} - \lambda)\Gamma]\Pi \ . \tag{126}$$

Therefore, if $\eta\Pi = \eta$ we obtain, using the uniqueness of Γ and (126), that $\Gamma\Pi = \Gamma$.

Hence, if equation (125) can be solved uniquely and boundedly in $\mathbf{M}_n$, it can be similarly solved in $\mathbf{S}_n$. □

Finally, we note that the construction of the forms $M_{i;j_1,\ldots,j_n}$ can be carried out in general Banach spaces –by using the Hahn-Banach theorem to extend from the space spanned by eigenvectors to the whole space. This construction shows that

$$\begin{aligned} \operatorname{Spec}_{pp}(\mathcal{L}_{n,A,B}, \mathbf{M}_n) &\supset \operatorname{Spec}_{pp}(B)\big(\operatorname{Spec}_{pp}(A)\big)^n \ , \\ \operatorname{Spec}_{pp}(\mathcal{L}_{n,A,B}, \mathbf{S}_n) &\supset \operatorname{Spec}_{pp}(B)\big(\operatorname{Spec}_{pp}(A)\big)^n \ , \end{aligned} \tag{127}$$

where Spec_{pp} denotes the closure of the set of eigenvalues.

Therefore, if we assume $\operatorname{Spec}_{pp}(B) = \operatorname{Spec}(B)$ and $\operatorname{Spec}_{pp}(A) = \operatorname{Spec}(A)$ then, combining (127) and (25), we obtain equalities in (25).

Appendix B. Historical remarks and information on the literature

Invariant manifold theory seems to have a not very well known history that abounds with re-discoveries –this has been eloquently mentioned in an often repeated quote in the fourth paragraph of [Ano69], page 23.

We have made an effort in finding works were results related to ours appeared. This cannot be considered a definitive effort (for instance, we have not been able to trace the work of Darboux, which is mentioned by Poincaré and Lyapunov). We can only hope that our modest search can inspire others.

B.1. **Early history.**

It seems to us that one-dimensional invariant submanifolds were more or less known in the analytic case, and with resonance conditions somewhat stronger than those considered in the present paper.

It seems well accepted that some versions of invariant manifold theory, at least for the analytic case, were known to Darboux, Poincaré and Lyapunov. Unfortunately, we have not been able to locate the works of Darboux, but we will comment on some works of Poincaré and Lyapunov.

B.2. **Two results of Poincaré.** One of us (R.L.) learned about the existence of [Poi90] (reproduced in [Poi50]) from conversations with D. Ruelle in the early 80's.

The motivation for [Poi90] was the theory of special functions.

When F is a polynomial and $X_1 = \mathbb{C}$, the equation

$$F \circ K(t) = K(\lambda t) \tag{128}$$

can be interpreted as saying that the system of functions given by the components of K admits a multiplication rule (*théoreme de multiplication*). Examples of such systems of functions (or systems satisfying the closely related addition rules) are the trigonometric functions and the elliptic functions.

For instance,$K(\theta) = (\sin\theta, \cos\theta)$ satisfies $K(2\theta) = F(K(\theta))$, where $F(x, y) = (2xy, y^2 - x^2)$. Note that $F(0, 1) = (0, 1)$. Similar formulas for the duplication of the argument are known for elliptic integrals. The fact that there are duplication formulas is related to the solvability of the quintic using elliptic functions and their inverses.

The paper [Poi90] shows that, given a map F and provided that λ, $|\lambda| > 1$, is a simple eigenvalue of $DF(0)$ and that there are no eigenvalues of $DF(0)$ which are powers of λ, one can find a formal series for K. Moreover, using the majorant method, one can show that the formal series for K converges.

This paper also contains the interesting observation –see page 541 in [Poi50]– that when F is a polynomial, every function K satisfying (128) is entire. The reason is that, when F is a polynomial, the functional equation (128) forces the domain of definition of K to be invariant under multiplication by λ. Hence, if it contains a ball, it is the whole complex plane. We note that this observation generalizes without difficulty to the situation when F is an entire function and we are working on a Banach space.

In [Poi90], Poincaré also studies the case when F^{-1} is a rational transformation, that he calls *Cremona*. In this case, he makes some dynamical observations. For instance, in the bottom half of page 561 of [Poi50], he relates the question of existence of solution to whether the iterates of the transformation converges to a fixed point –this is indeed the dynamical characterization of invariant manifold.

From a more dynamical point of view, similar series were considered in [Poi87], where all chapter VII is devoted to asymptotic expansions around periodic solutions of periodic vector fields. Taking time-T maps, this problem reduces to the setting about maps that we have considered in this paper. The logarithms of the eigenvalues of the time-T map are called *exposants charactéristiques*. In modern language, they are the Floquet exponents. Note that what we would call today Lyapunov exponents (which can be considered in more general settings than periodic systems) are, in the case of periodic systems, the real part of Poincaré's *exposants charactéristiques*. More confusingly, in the translation of Lyapunov that we have used, the name *characteristic exponent* refers to the negative of what we call now Lyapunov exponent. This would be, of course, the negative of the real part of the *exposant charactéristique* for the particular case of periodic systems.

In [Poi87] the crucial chapters dealing with stable and unstable manifolds are 104 and 105.

In paragraph 104, under the assumption that there are no resonances (the non-resonance condition is the last formula of paragraph 104), it is shown that one can obtain a formal power series expansion of exponentials with arbitrary constants.

The convergence of the series is studied in 105. The first paragraph asserts the convergence of the series of expansions in powers of the exponential under the assumption that the eigenvalues belong to what we now call the Poincaré domain (i.e., when the convex hull of the eigenvalues does not include zero). Of course, the reason why this condition enters is that, for eigenvalues satisfying these conditions, the small divisors that appear are bounded away from zero.

We note that, even if it is not said explicitly, the condition that the eigenvalues are different is indeed assumed. The proof of convergence is rather succinct. Nevertheless, it should have been quite clear to Poincaré and his contemporaries since it is very similar to arguments that had been done in detail in his thesis [Poi79] (reproduced in [Poi16]).

From the point of view of invariant manifold theory, the last paragraph of page 339 is quite interesting. Here, Poincaré discusses the case when there are stable and unstable characteristic exponents at the same time. He observes that the series for K remains convergent if one sets to zero the constants corresponding to coordinates along the expanding or neutral eigendirections. The arguments here are somewhat skimpy, but a modern mathematician can supply them without too much trouble. One is left with a set of solutions which tend to zero parameterized by as many constants as stable directions. This is, of course, our modern stable manifold. A similar construction works for the unstable solutions. Poincaré called these solutions *solutions asymptotiques.*

The rest of chapter VII contains a variety of expansions of these sets of solutions. It includes, quite notably, the expansions in terms of a slow parameter, which are then shown to be divergent. Of course, much modern work is still being done in these slow perturbations and related areas.

B.3. **The work of Lyapunov.** In chapters 11-33 [Lya92] (see also the summary in chapter 3 and the proofs of convergence in chapter 23), Lyapunov introduces the method of arbitrary constants, which consists in finding exponential solutions with arbitrary constants. Since the constants do not evolve in time, this is closely related to the problem of linearization. (Compare the expansions of the system studied and those of the linear systems.) Invariant manifolds can be obtained by setting some of the constants to zero.

One important difference between [Lya92] and [Poi90] [Poi87] is that [Lya92] considers systems which are *regular* (roughly, the definition is that the forward and backward Lyapunov exponents agree). This is a more general setting than that of periodic systems. In the case of regular systems, [Lya92] contains expansions of the solutions in terms of arbitrary constants. The derivation of the formal expansions in [Lya92] does not need non-resonance conditions.

In Chapter 23 of [Lya92], the question of convergence of these formal expansions is studied. This is done under the condition that there are no resonances and no repeated eigenvalues, and that all the eigenvalues are stable or unstable. Here one can find a note giving credit to [Poi79] for dealing with the more general case of the Poincaré domain.

In particular, we call attention to Theorem II of Section 24, which is a complete statement of the strong stable manifold theorem for analytic systems (see also Theorem II of chapter 13).

One interesting remark of Lyapunov in chapter 11 is that one can consider families that correspond to any subset of eigenvalues. This amounts to setting to zero a subset of the arbitrary constants used in the expansion. This is hard to interpret from the dynamical point of view since the arbitrary constants do not have a dynamical interpretation. In particular, the set obtained setting them to zero does not need to be invariant. Of course, setting to zero all the non-decreasing modes is an invariant set, as pointed out by Poincaré. With modern hindsight, setting to zero all the modes that are non-decreasing or decreasing more slowly than a certain rate is invariant. Indeed, it is the strongly stable manifold. As it was shown in examples in [dlL97], in general, one cannot get invariant manifolds tangent to a subspace if there are resonances of the type we have excluded in the present paper.

Overall, one cannot be but surprised by the enormous similarities in the problems and in the results between the contents of these chapters and the corresponding ones of the book by Poincaré, which appeared in the same year. Of course, there are big differences in style and in the methods as well as in the way that proofs are presented.

A modern exposition of some of the convergence results of Lyapunov can be found in [Lef77] V.4. It contains a statement and a proof of the expansion in arbitrary constants under non-resonance assumptions and provided that all eigenvalues are stable, and that the linearization is a constant (we remark that using Floquet theory, one can reduce the periodic case to the constant case). We have not been able to locate in any of these classical works the consideration of resonant terms.

B.4. **Modern work.** It seems that the particular case of one dimensional stable invariant manifolds (when there are no resonances) has appeared several times in the modern literature.

The papers [FR81] and [FG92] use the parameterization method for one dimensional manifolds, specially in conjunction with numerical analysis. They establish not only convergence of the series involved, but they also estimate the errors incurred when using a numerical approximation. Indeed, both papers have taken care of estimating actually the roundoff error so that a finite calculation can establish facts about transversality of intersections, etc.

It seems to us that similar results could be obtained using the functional equations (4) and the theory developed in the present article.

Numerical work for higher dimensional maps has been studied in [BK98], which undertook the task of systematically computing Taylor expansions of invariant manifolds. This could be considered one implementation of our result in Lemma 3.1 for finite dimensional systems. The authors of [BK98] indeed made the observation that the calculations can be carried out to any order provided that there are no resonances.

We note that our formalism could be used to provide an *a posteriori* estimate of the error of these numerical calculations. Once a polynomial satisfies (4) quite accurately, then it is close to being a fixed point of $\mathcal{T}$. Since $\mathcal{T}$ is a contraction, there is a fixed point at a distance that can be estimated by the error of the numerical approximation. This is the usual *a posteriori* estimates of numerical analysis.

The work [Pös86] considers invariant manifolds associated to non-resonant eigenspaces. It also studies the equation of semi-conjugacy of the motion on the manifold to a linear motion. In contrast to the classical works, it can deal with situations when the eigenvalues are not stable and all of the same sign. Small divisors appear, but they can be overcome using the majorant method and improvements of estimates in [Sie42]. It is interesting to note that the non-resonance conditions obtained in [Pös86] are more general than those required by the straightforward application of usual KAM method. For more modern developments, see [Sto94].

The paper [CF94] took up the task of making sense of the one-dimensional manifolds of Poincaré and Lyapunov, in the case of finite differentiability. It reduced the problem to a fixed point equation that was solved by the contraction method. In this paper there was some consideration given to the resonant case, and the result proved was that the equation (4) could be solved if and only if our equations for K and R provide a formal solution.

We note that all papers mentioned above seem to require the dynamics in the invariant manifold considered to be linearizable.

The work whose results are most closely related to those of the present paper is [dlL97]. It contains a study of invariant manifolds which does not require the motion on the invariant manifold to be linearizable. In the present paper, we improve the results of [dlL97] by not requiring the splitting to be invariant. Hence, we can associate invariant manifolds to eigenspaces in a non-trivial Jordan form. Also, we can improve the regularity obtained from $C^{r-1+\mathrm{Lip}}$ to C^r (this later improvement in the regularity was obtained also in [ElB98]). The interest of this sharp regularity is that the same method that we use to optimize the regularity also allows us to obtain rather sharp results on the dependence on parameters. The methods of the present article and of [dlL97] are quite different.

There are some generalizations of the non-resonant manifolds considered in the present paper to some rather general semigroups –which include many parabolic equations– in [dlLW97]. The main difference is that in [dlLW97], the assumption on invertibility of the linearized operator used in the present paper is dropped, since the linearization of the evolution of an elliptic equation is a compact operator, hence, not invertible.

B.5. **Slow manifolds in applications.** One of the cases which is most useful in applications is when the invariant manifolds correspond to the spaces that have eigenvalues closest to the unit circle. These are called *slow manifolds*. The reason why they are useful is because they correspond to the modes that converge the slowest and, hence, they can be used to control the long term behavior.

It should be noted that, for slow manifolds, one could also apply the Irwin theory of pseudo-stable manifolds. These two methods, as it can be seen in examples in [dlL97], do not lead to the same results.

In [LMS95], it is argued that beta function calculations of renormalization groups correspond to computation of the jets of smooth invariant manifolds as the ones we consider in Lemma 3.1 –see specially Remark 10. Nevertheless, it is also argued that the ones which should be considered in several problems of phase transitions are other slow manifolds.

This is a very nice idea, and it would be quite interesting to make it precise. Of course, the mathematically precise definition of renormalization group as a well defined differentiable operator in a Banach space, has only been achieved in a few cases (mainly dynamical systems and hierarchical models). Even in these cases where the renormalization operator is a bona fide differentiable operator in Banach spaces, its linearization is a compact operator, hence not invertible, and the theorems of this paper do not apply. One has to use those of [dlLW97].

In [dlL97] it is shown that the non-coincidence of these manifolds can be used to prove that the invariant circles predicted by the Hopf bifurcation for maps are often not C^∞.

A field where slow manifolds have played an important role is chemical kinetics. Here, slow manifolds correspond to the so called *quasi-steady state approximations*. There is a very extensive literature on these approximations. We quote [Fra88], [MP92] for a geometric interpretation of them. We believe that the methods suggested in these two papers lead, respectively to the smooth non-resonance manifolds discussed here and to the Irwin manifolds, which are, in general very different. Therefore, speaking of *the* slow invariant manifold may bee misleading since there are several different objects that can be given that name and should not be identified. This, we think, would explain some of the discrepancies reported in [Dav97]. Some more recent references in the chemical literature are [Fra98], [Smo91], [PR93], [WMD96].

For some specialized semigroups –including the very important Navier-Stokes equations– some slow invariant manifolds have been constructed in the remarkable articles [FS84b], [FS84a],[FS87] (see also the surveys [FS86a], [FS86b]). In these works, slow manifolds are constructed under non-resonance conditions. We note in particular that [FS86b] establishes that these invariant manifolds are analytic. This is in contrast with examples in [dlLW97] which show that, for semigroups, the slow manifolds may not be analytic. We thank E. Titi for bringing these papers to our attention.

References

[Ano69] D. V. Anosov. *Geodesic flows on closed Riemann manifolds with negative curvature.* American Mathematical Society, Providence, R.I., 1969.

[Bar88] Dwight Barkley. Slow manifolds and mixed-mode oscillations in the Belousov-Zhabotinskiĭ reaction. *J. Chem. Phys.*, 89(9):5547–5559, 1988.

[BdlLW96] Augustin Banyaga, Rafael de la Llave, and C. Eugene Wayne. Cohomology equations near hyperbolic points and geometric versions of Sternberg linearization theorem. *J. Geom. Anal.*, 6(4):613–649, 1996.

[BK98] Wolf-Jürgen Beyn and Winfried Kless. Numerical Taylor expansions of invariant manifolds in large dynamical systems. *Num. Math.*, 80(1):1–38, 1998.

[CF94] Xavier Cabré and Ernest Fontich. Regularity and uniqueness of one dimensional invariant manifolds. *U. of Barcelona Preprint*, 1994.

[CFdlL00] Xavier Cabré, Ernest Fontich, and Rafael de la Llave. The parameterization method for invariant manifolds II: overview and applications. *In preparation*, 2000.

[CL99] Carmen Chicone and Yuri Latushkin. *Evolution semigroups in dynamical systems and differential equations.* American Mathematical Society, Providence, RI, 1999.

[Dav97] Michael J. Davis. Some multiple scale problems in chemical dynamics, spectroscopy, and kinetic modelling. Notes of a talk given at IMA 10-28-1997, 1997.

[dlL97] Rafael de la Llave. Invariant manifolds associated to nonresonant spectral subspaces. *J. Statist. Phys.*, 87(1-2):211–249, 1997.

[dlL98] Rafael de la Llave. The spectrum of transfer operators acting on smooth bundles. *Preprint*, 1998.

[dlLMM86] Rafael de la Llave, Jose Manuel Marco, and Roberto Moriyón. Canonical perturbation theory of Anosov systems and regularity results for the Livšic cohomology equation. *Ann. of Math. (2)*, 123(3):537–611, 1986.

[dlLO99] Rafael de la Llave and Rafael Obaya. Regularity of the composition operator in spaces of Hölder functions. *Discrete Contin. Dynam. Systems*, 5(1):157–184, 1999.

[dlLW95] Rafael de la Llave and C. Eugene Wayne. On Irwin's proof of the pseudostable manifold theorem. *Math. Z.*, 219(2):301–321, 1995.

[dlLW97] Rafael de la Llave and C. Eugene Wayne. Non-resonant invariant manifolds for some semigroups. *Preprint*, 1997.

[ElB98] Mohamed Sami ElBialy. Sub-stable and weak-stable manifolds associated with finitely non-resonant spectral decompositions. *MP_ARC*, 1998.

[FG92] John Erik Fornæss and Estela A. Gavosto. Existence of generic homoclinic tangencies for Hénon mappings. *J. Geom. Anal.*, 2(5):429–444, 1992.

[FR81] Valter Franceschini and Lucio Russo. Stable and unstable manifolds of the Hénon mapping. *J. Statist. Phys.*, 25(4):757–769, 1981.

[Fra88] Simon J. Fraser. The stady state and equilibrium approximations: A geometrical picture. *Jour. Chem. Phys.*, 88(8):4732–4738, 1988.

[Fra98] Simon J. Fraser. Double perturbation series in the differential equations of enzyme kinetics. *Jour. Chem. Phys.*, 109(2):411–423, 1998.

[FS84a] C. Foiaş and J.-C. Saut. Asymptotic behavior, as $t \to +\infty$, of solutions of Navier–Stokes equations and nonlinear spectral manifolds. *Indiana Univ. Math. J.*, 33(3):459–477, 1984.

[FS84b] C. Foiaş and J.-C. Saut. On the smoothness of the nonlinear spectral manifolds associated to the Navier–Stokes equations. *Indiana Univ. Math. J.*, 33(6):911–926, 1984.

[FS86a] C. Foiaş and J.-C. Saut. Nonlinear spectral manifolds for the Navier–Stokes equations. In *Nonlinear functional analysis and its applications, Part 1 (Berkeley, Calif., 1983)*, pages 439–448. Amer. Math. Soc., Providence, R.I., 1986.

[FS86b] Ciprian Foiaş and Jean-Claude Saut. Variétés invariantes à décroissance exponentielle lente pour les équations de Navier–Stokes avec forces potentielles. *C. R. Acad. Sci. Paris Sér. I Math.*, 302(16):563–566, 1986.

[FS87] C. Foiaş and J.-C. Saut. Linearization and normal form of the Navier–Stokes equations with potential forces. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 4(1):1–47, 1987.

[HP70] Morris W. Hirsch and Charles C. Pugh. Stable manifolds and hyperbolic sets. In *Global Analysis (Proc. Sympos. Pure Math., Vol. XIV, Berkeley, Calif., 1968)*, pages 133–163. Amer. Math. Soc., Providence, R.I., 1970.

[JPdlL95] M. Jiang, Ya. B. Pesin, and R. de la Llave. On the integrability of intermediate distributions for Anosov diffeomorphisms. *Ergodic Theory Dynam. Systems*, 15(2):317–331, 1995.

[Lef77] Solomon Lefschetz. *Differential equations: geometric theory.* DoverPublications, Inc., New York,, second edition, 1977.

[LI73] Oscar E. Lanford III. Bifurcation of periodic solutions into invariant tori: the work of Ruelle and Takens. In Ivar Stakgold, Daniel D. Joseph, and David H. Sattinger, editors, *Nonlinear problems in the Physical Sciences and biology: Proceedings of a Battelle summer institute*, pages 159–192, Berlin, 1973. Springer-Verlag. Lecture Notes in Mathematics, Vol. 322.

[LI83] Oscar E. Lanford III. Introduction to the mathematical theory of dynamical systems. In *Chaotic behavior of deterministic systems (Les Houches, 1981)*, pages 3–51. North-Holland, Amsterdam, 1983.

[LMS95] Bin Li, Neal Madras, and Alan D. Sokal. Critical exponents, hyperscaling, and universal amplitude ratios for two- and three-dimensional self-avoiding walks. *J. Statist. Phys.*, 80(3-4):661–754, 1995.

[Lor92] Edward N. Lorenz. The slow manifold—what is it? *J. Atmospheric Sci.*, 49(24):2449–2451, 1992.

[Lya92] Aleksander Mikhailovich Lyapunov. *The general problem of the stability of motion.* Taylor and Francis, Bristol, 1992. Translated and edited from the French by A. T. Fuller, with Foreword by I. Stewart, Introduction by A. T. Fuller, Biography of A. M. Lyapunov by V.I. Smirnov, Bibliography of A. M. Lyapunov by J.F. Barrett.

[Mat68] John N. Mather. Characterization of Anosov diffeomorphisms. *Nederl. Akad. Wetensch. Proc. Ser. A 71 = Indag. Math.*, 30:479–483, 1968.

[MP92] U. Maas and S. B. Pope. Simplifying chemical kinetics:intrinsic low-dimensional manifolds in composition space. *Combust. Flame*, pages 239–264, 1992.

[Nel69] Edward Nelson. *Topics in dynamics. I: Flows.* Princeton University Press, Princeton, N.J., 1969. Mathematical Notes.

[Pes73] Ja. B. Pesin. The existence of invariant foliations for a diffeomorphism of a smooth manifold. *Mat. Sb. (N.S.)*, 91(133):202–210, 287, 1973.

[Poi79] Henri Poincaré. Sur les propriétés des fonctions définies par les équations aux différences partielles. *Théses présentées a la Faculté des scinces de Paris*, 1879.

[Poi90] Henri Poincaré. Sur une classe nouvelle de transcendentes uniformes. *Jour. de Math.*, 6:313–365, 1890.

[Poi16] Henri Poincaré. *Œuvres de Henri Poincaré, vol I.* Gauthier Villars, Paris, 1916.

[Poi50] Henri Poincaré. *Œuvres de Henri Poincaré, vol IV.* Gauthier Villars, Paris, 1950.

[Poi87] H. Poincaré. *Les méthodes nouvelles de la mécanique céleste. Tome I.* Librairie Scientifique et Technique Albert Blanchard, Paris, 1987. Solutions périodiques. Non-existence des intégrales uniformes. Solutions asymptotiques. [Periodic solutions. Nonexistence of uniform integrals. Asymptotic solutions], Reprint of the 1892 original, With a foreword by J. Kovalevsky, Bibliothèque Scientifique Albert Blanchard. [Albert Blanchard Scientific Library].

[Pös86] Jürgen Pöschel. On invariant manifolds of complex analytic mappings near fixed points. *Exposition. Math.*, 4(2):97–109, 1986.

[PR93] Norbert Peters and Bernd Rogg, editors. *Reduced kinetic mechanisms for applications in combustion systems.* Springer-Verlag, Berlin, New-York, 1993.

[Rud91] Walter Rudin. *Functional analysis.* McGraw-Hill Inc., New York, second edition, 1991.

[Sie42] Carl Ludwig Siegel. Iteration of analytic functions. *Ann. of Math. (2)*, 43:607–612, 1942.

[Smo91] Mitchell D. Smooke, editor. *Reduced kinetic mechanisms and asymptotic approximations for methane-air flames : a topical volume.* Springer-Verlag, Berlin,New-York, 1991. Lecture Notes in Physics, Vol. 384.

[Sto94] Laurent Stolovitch. Sur un théorème de Dulac. *Ann. Inst. Fourier (Grenoble)*, 44(5):1397–1433, 1994.

[WMD96] Jürgen Warnatz, Ulrich Maas, and Robert W. Dibble. *Combustion: physical and chemical fundamentals, modelling and simulation, experiments, pollutant formation.* Springer-Verlag, Berlin, New-York, 1996.

Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya, Diagonal 647, 08028 Barcelona, Spain,

E-mail address, X. Cabré: `cabre@ma1.upc.es`

Departament de Matemàtica Aplicada i Anàlisi, Universitat de Barcelona, Gran Via, 585, 08007 Barcelona, Spain

E-mail address, E. Fontich: `fontich@mat.ub.es`

Department of Mathematics, The University of Texas at Austin, Austin, TX 78712-1082, U.S.A,

E-mail address, R. de la Llave: `llave@math.utexas.edu`