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 to find simple polynomial maps which are conjugated to the dynamics on them. 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.

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 a parameterization of the manifold in such a way that it satisfies a functional equation that expresses that its range is invariant and that it semi-conjugates the dynamics to the dynamics of a simpler map (in some cases linear). More precisely, we look

Date: Nov 27, 2002.

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. Writing $X = X_1 \oplus X_2$ and $\pi_i : X \to X_i$ the corresponding projections, the conditions involve non-resonances between the spectrum of $\pi_1 DF(0)|_{X_1}$ and the spectrum of $\pi_2 DF(0)|_{X_2}$.

The study of the regularity with respect to parameters is deferred to the follow-up paper [CFdlL00].

Invariant manifolds theory has a long history. Among many relevant references about invariant manifolds associated to fixed points we quote [HP70], [Irw70], [Irw80], and we refer to them for the original literature. We point out that most of the literature deals with invariant manifolds associated to spectral spaces corresponding to intersections of the spectrum with discs and complements of discs, in which case there is no need to make non-resonance assumptions.

Regarding the more general case of invariant manifolds associated to non-resonant subspaces, [Poi90] studies one dimensional invariant manifolds associated to a simple eigenvalue satisfying non-resonance conditions, in the analytic case. Also for analytic systems, some results for the stable and unstable manifolds (and submanifolds of them) corresponding to periodic orbits of differential equations are found in [Poi87] and [Lya92] (see also [Lef77]). The work [Pös86] considers invariant manifolds associated to non-resonant eigenspaces. In contrast to the classical works, it can deal with eigenvalues of modulus one and hence small divisors appear. 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. A study of non-resonant manifolds in infinite dimensional spaces was undertaken in [dlL97] (see also [ElB98]). We refer to [CFdlL02] for more historical references and comments.

We point out that [dlL97] contains examples which show that the non-resonance conditions are necessary for existence of smooth invariant manifolds.

The results of the present paper extend the results of [dlL97]. 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].

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. 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.

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.

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]. We refer to [CFdlL02] for a longer discussion on the applied literature.

It is interesting to note that there is another possible construction of slow manifolds, namely Irwin's pseudostable manifolds [Irw70, Irw80]. These papers associate manifolds $W^{p(a),f}$ to spectral subsets $\{|z| \le a\}$, with $a > 1$. By considering $W^{\mathrm{Irwin}} = W^{p(a),f^{-1}} \cap W^{s,f}$, one obtains an invariant manifold for f associated to the spectral projection on $\{|z| \in [a^{-1}, 1)\}$.

Perhaps surprisingly (see [dlL97]), it turns out that our smooth manifolds and the Irwin's ones may not coincide even in systems in which both of them can be defined. Irwin's manifolds are unique under conditions of global behavior of orbits. Hence, the pseudostable manifolds in an arbitrarily small neighborhood of the origin can be affected by changing the map arbitrarily far away. On the other hand, the manifolds considered in the present paper are unique under smoothness conditions at the origin and, therefore, their restrictions to a small neighborhood are not affected by changes of the map outside this neighborhood. 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. We think that this fact explains certain discrepancies in the numerical literature in chemical kinetics computing slow invariant manifolds. Methods based in computing jets –hence assuming implicitly differentiability– compute the manifolds in this paper, whereas methods based on iteration compute the Irwin ones. Methods based on a combination of the two compute yet other objects.

Roughly, the situation of slow manifolds (with respect to their local regularity) can be summarized as saying that we can find numbers $r_0 \le r_1$ depending only on the spectrum of the linearization such that:

- There are infinitely many invariant manifolds which are $C^{r_0-\delta}$ (r_0 is the regularity claimed in Irwin's work; see [dlLW95]).
- There is at most one $C^{r_1+\delta}$ at the origin (this is the manifold singled out in the present paper).
- 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 exists– depends smoothly on parameters.

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.

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.

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 $\tilde{X}$.

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*

$$\big(\mathrm{Spec}(A_1)\big)^i \cap \mathrm{Spec}(A_1) = \emptyset \qquad \textit{for every integer } i \textit{ with } 2 \le i \le L \ , \tag{7}$$

then we can choose R *in a) above to be linear. More generally, if*

$$\big(\mathrm{Spec}(A_1)\big)^i \cap \mathrm{Spec}(A_1) = \emptyset \qquad \textit{for every integer } i \textit{ with } M \le i \le L \ , \tag{8}$$

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

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).*

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 $L \ge 1$ for which (2) holds. We also point out that, by hypothesis 5), we always have $r \ge 2$ (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, which is just a polynomial. Moreover, under the additional hypothesis (7), R 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 4).

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 $K(U_1)$, 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.

These situations happen naturally when the map preserves a geometric structure (e.g. volume, contact or symplectic) or a symmetry. Such geometric features often produce resonances, but at the same time give cancellations that allow the equations to be solved.

Remark 2. The parameterization method of this paper has been implemented numerically.

The equations to compute the jet of K and R (that is, the polynomials $K^{\leq}$ and R) alluded to in Remark 1 (and in Theorem 1.2 below) reduce to equations that are called Sylvester equations in [BK98].

One of the advantages of the parameterization method is that, since invariant manifolds often bend and fold along themselves, a parametric representation is less affected by foldings than a representation as a graph.

In particular, in the case that F has an entire inverse and that we are in the case b) of Theorem 1.1 with $M = 2$ (hence R is linear), the parameterization K is entire.

Indeed, note that equation (4) reads $K = F^{-1} \circ K \circ R$. Hence, if a set S is contained in the domain of K, then $R^{-1}(S)$ is also contained in the domain. When R is a linear contraction, we see that this implies that the domain of K is the whole space.

Remark 3. 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 4. 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

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

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. However, 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.

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:

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

Note that, in finite dimensions, (11) is obvious.

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

$$\big(\mathrm{Spec}(A_1)\big)^{L+1}\,\mathrm{Spec}(A_2^{-1}) \subset \{z \in \mathbb{C} \mid |z| < 1\} \ . \tag{12}$$

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 the polynomial 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.

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

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

Then,

a) *We can find a C^r map K and a C^r map R satisfying* (4),

$$K(0) = 0, \qquad \pi_1 K = \mathrm{Id}, \qquad \pi_2 DK(0) = 0$$

and

$$R(0) = 0, \qquad DR(0) = A_1.$$

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 $\sup_{x \in U_1}(|h(x)|/|x|^{L+1}) < \infty$.*

The polynomial w_L can be explicitly computed out of the L-jet of the map F.

Remark 5. Examining the proof of Theorem 1.2 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

$$\|A_2^{-1}\|\,\|A_1\|^{L+1} < 1. \tag{14}$$

This condition (14) is what is actually used through the proof. Hence, Theorem 1.2 also holds if we do not assume that A is invertible but, instead of (13), we assume (14).

In finite dimensional spaces, the spectral condition (13) and the verification of (14) for some equivalent norm are equivalent. However, this equivalence is not true in infinite dimensions (see Example A.2). We prefer to use the spectral condition on the statement because it is intrinsic, even if the result is less general than possible.

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. This generalizes 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

$$[\mathcal{L}_f v](x) = \exp_x^{-1} f(\exp_{f^{-1}(x)} v(f^{-1}(x))), \tag{15}$$

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 (15) as

$$[\mathcal{L}_f v](f(y)) = f(y + v(y)) - f(y) \tag{16}$$

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, (16) 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 a study of differentiability 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

$$\mathcal{L}_f(0) = 0 \qquad \text{and} \qquad D\mathcal{L}_f(0) = f_* \ ,$$

where $(f_* v)(x) = Df(f^{-1}(x)) \cdot v(f^{-1}(x))$ is the usual pushforward of differential geometry.

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*

$$\operatorname{Spec}(f_*, C^0(N, TM)) \subset \bigcup_{i=1}^{n} \{z \in \mathbb{C} \mid \lambda_i^- \leq |z| \leq \lambda_i^+\} \tag{17}$$

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

Moreover, we can find a continuous decomposition

$$T_x M = \bigoplus_{i=1}^{n} E_x^i \tag{18}$$

such that

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

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 (17). 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 $\Sigma \subset \{1, \dots, n\}$ 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)
$$\Big(\bigcup_{i\in\Sigma} [\lambda_i^-, \lambda_i^+]\Big)^j \cap \Big(\bigcup_{i\notin\Sigma} [\lambda_i^-, \lambda_i^+]\Big) = \emptyset \qquad \text{for every } 2 \le j \le L\ .$$
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

$$f(x + v(x)) = f(x) + \mathcal{L}_f v(f(x)), \tag{19}$$

imply that

$$f(W_x^\Sigma) \subset W_{f(x)}^\Sigma. \tag{20}$$

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 (20).

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

$$c(\lambda_\Sigma^- - \varepsilon)^n \le d(f^n(x), f^n(y)) \le c(\lambda_\Sigma^+ + \varepsilon)^n, \qquad \text{for all } n \ge 0, y \in W_x^\Sigma. \quad (21)$$

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, (21) is not only consequence of $y \in W_x^\Sigma$, but is also equivalent to it. Since (21) 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 result similar to Theorem 2.2, proved by a different method, can be found in [Pes73].

When Σ contains intermediate components, then (21) may not be equivalent to $x \in W_y^\Sigma$. Indeed, in [JPdlL95] it is shown that the W_x^Σ may fail very strongly 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.

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 $K^{\le}$ and R so obtained and reduce the equation for the higher order terms to a fixed point problem for $K^{>}$ (where $K = K^{\le} + K^{>}$), 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.

It will 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.

Indeed, if $K = K^{\leq} + K^{>}$ and $\mathcal{T}$ is an operator such that $\mathcal{T}K^{>} = K^{>}$ with

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

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

$$\|K^{>} - K_0^{>}\| \leq \delta/(1-\kappa) \; .$$

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

We emphasize the fact 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.

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).

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.

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 homogeneous polynomials 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*

$$\left(\mathrm{Spec}(A_1)\right)^j \cap \mathrm{Spec}(A_1) = \emptyset \qquad \text{for some } j \text{ with } 2 \leq j \leq L \ , \qquad (22)$$

then we can choose $R_j = 0$.

Remark 10. 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 $\mathcal{L}_B$, $\mathcal{R}_A^k$ and $\mathcal{L}_{n,A,B}$ acting on $\mathbf{M}_n$ by

$$\begin{aligned} (\mathcal{L}_B\, M)(x_1, \ldots, x_n) &= B\, M(x_1, \ldots, x_n), \\ (\mathcal{R}_A^k\, M)(x_1, \ldots, x_n) &= M(x_1, \ldots x_{k-1}, Ax_k, x_{k+1}, \ldots, x_n), \\ (\mathcal{L}_{n,A,B} M)(x_1, \ldots, x_n) &= BM(Ax_1, \ldots, Ax_n) \ . \end{aligned} \qquad (23)$$

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

Then,

$$
\begin{aligned}
\mathrm{Spec}(\mathcal{L}_B, \mathbf{M}_n) &\subset \mathrm{Spec}(B, Y)\ , \\
\mathrm{Spec}(\mathcal{R}^k_A, \mathbf{M}_n) &\subset \mathrm{Spec}(A, X)\ , \\
\mathrm{Spec}(\mathcal{L}_{n,A,B}, \mathbf{S}_n) \subset \mathrm{Spec}(\mathcal{L}_{n,A,B}, \mathbf{M}_n) &\subset \mathrm{Spec}(B, Y)\big(\mathrm{Spec}(A, X)\big)^n\ .
\end{aligned}
\tag{24}
$$

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 (24) *are equalities.*

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

The operators in (23) 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 (24) 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

$$AK_1 = K_1 R_1\ .$$

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

$$AK_i + \Gamma_i = K_i A_1^{\otimes i} + K_1 R_i\ , \tag{25}$$

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

We study the system of equations (25) by induction on i, by considering (25) as an equation to be solved for K_i and R_i once $K_1, \ldots, K_{i-1}$, $R_1, \ldots, R_{i-1}$ –and therefore Γ_i– are known. So, we turn our efforts into studying the solvability of (25) 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 (25) is equivalent to

$$A_1 K_i^1 - K_i^1 A_1^{\otimes i} - R_i = -\Gamma_i^1 - BK_i^2, \tag{26}$$

$$A_2 K_i^2 - K_i^2 A_1^{\otimes i} = -\Gamma_i^2\ . \tag{27}$$

With the notation (23), equations (26) and (27) can be written as

$$\Big(\mathcal{L}_{A_1} - \mathcal{L}_{i,A_1,\mathrm{Id}}\Big) K_i^1 - R_i = -\Gamma_i^1 - BK_i^2\ , \tag{28}$$

$$\Big(\mathcal{L}_{A_2} - \mathcal{L}_{i,A_1,\mathrm{Id}}\Big) K_i^2 = -\Gamma_i^2\ . \tag{29}$$

The crux of the problem is the second equation (29). Once equation (29) 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 (27), that is (29), 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

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

at the level $i = j$, then we can choose $R_j = 0$ and solve uniquely K_j^1 in the first equation, (28), 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).

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. $\square$

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 (26) and (27) 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.

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 $i \leq L$, and such that $K = K^{\leq} + K^{>}$ satisfies (4), which can be written as

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

or in the equivalent form

$$AK^{>} - K^{>} \circ R = -N \circ (K^{\leq} + K^{>}) - AK^{\leq} + K^{\leq} \circ R . \qquad (30)$$

Note that the way that $K^{\leq}$ 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 (30) also vanish at the origin. Hence, we will consider (30) 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 10 we know that we can assume $K^{\leq}$ 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 (30), 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

$$\|DR\|_{C^0(B_1)} \leq \|A_1\| + \varepsilon < 1 , \qquad (31)$$

with ε small enough such that

$$\|DR\|^{L+1}_{C^0(B_1)} \|A^{-1}\| \leq (\|A_1\| + \varepsilon)^{L+1} \|A^{-1}\| < 1 .$$

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

$$|D^k R^j(x)| \leq C_k(\|A_1\| + \varepsilon)^j \qquad (32)$$

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

3.3. **Study of the linearized problem.** Since N may be assumed to be small, to solve equation (30) we first study the linear operator

$$\mathcal{S}H = AH - H \circ R \tag{33}$$

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

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

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 functions in $\Gamma_{\omega,l}$ have all derivatives at 0 up to order l equal to zero. With this norm, $\Gamma_{\omega,l}$ is a Banach space.

When $s \in \mathbb{N}$, we also 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 the norm $\|\cdot\|_{\Gamma_{s,l}}$ is relevant only when $s = l$, in the sense that we could omit this term when $s > l$ and still get an equivalent norm (since $D^l H(0) = 0$).

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 $r \in \mathbb{N} \cup \{\omega\}$ then $\mathcal{S} : \Gamma_{r-1,L} \longrightarrow \Gamma_{r-1,L}$ is a bounded invertible operator. Moreover, $\|\mathcal{S}^{-1}\|$ 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

$$\mathcal{S}H = \eta. \tag{34}$$

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

The equation (34) for H is equivalent to

$$H = A^{-1}H \circ R + A^{-1}\eta \ . \tag{35}$$

We claim that the solution of (35) is given by

$$H = \sum_{j=0}^{\infty} A^{-(j+1)}\eta \circ R^j \ . \tag{36}$$

To establish the claim, we will show that the series in (36) converges absolutely in $\Gamma_{r-1,L}$ and that

$$\sum_{j=0}^{\infty} \|A^{-(j+1)}\eta \circ R^j\|_{\Gamma_{r-1,L}} \le C\|\eta\|_{\Gamma_{r-1,L}}, \qquad \forall \eta \in \Gamma_{r-1,L} \ . \tag{37}$$

In particular, when substituted in (35), we can rearrange the terms and show that H is indeed a solution.

We also claim that the solution is unique. Indeed, if $\eta = 0$, by (35) we have $H = A^{-1}H \circ R$, and hence $H = A^{-j}H \circ R^j$ for every $j \ge 1$. But $\|A^{-j}H \circ R^j\|_{\Gamma_{r-1,L}} \to 0$ as $j \to \infty$, by (37) applied with $\eta = H$. We conclude that $H = 0$.

We now establish (37) when $r \in \mathbb{N}$. Since $\eta \in \Gamma_{r-1,L}$, we have $|D^L\eta(\xi)| \le \|\eta\|_{\Gamma_{r-1,L}}|\xi|$ for $\xi \in B_1$, and $\eta(0) = 0, \dots, D^L\eta(0) = 0$. Hence, by Taylor's formula, we have that $|\eta(y)| \le C\|\eta\|_{\Gamma_{r-1,L}}|y|^{L+1}$, for $y \in B_1$. Moreover, $\mathrm{Lip}(R) \le \|A_1\| + \varepsilon$ where $\varepsilon > 0$ is small by the scaling argument. We conclude that

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

Since $\|A^{-1}\|(\|A_1\| + \varepsilon)^{L+1} < 1$, we deduce that the right hand side of (36) 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

$$D^k(\eta \circ R^j) = \sum_{i=0}^{k} \sum_{\substack{1\le k_1,\dots,k_i \le k \\ k_1+\cdots+k_i=k}} \sigma^{i,k}_{k_1,\dots,k_i}([D^i\eta] \circ R^j)\, D^{k_1}R^j \cdots D^{k_i}R^j, \tag{38}$$

where $\sigma^{i,k}_{k_1,\dots,k_i}$ is an explicit combinatorial coefficient. We recall that, by (32),

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

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

$$\begin{aligned} |[D^i\eta] \circ R^j(x)| &\le C\|\eta\|_{\Gamma_{r-1,L}}|R^j(x)|^{(L+1-i)_+} \\ &\le C\|\eta\|_{\Gamma_{r-1,L}}(\|A_1\| + \varepsilon)^{j(L+1-i)_+}|x|^{(L+1-i)_+} \end{aligned} \tag{39}$$

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 (38), we deduce that

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

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 (37). 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 (38) 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 (37). □

3.4. **Solution of the fixed point problem.** We want to solve equation (30), that can be rewritten using the operator $\mathcal{S}$ introduced in (33) 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

$$\mathcal{T}(K^{>})=\mathcal{S}^{-1}\big[-N\circ(K^{\le}+K^{>})-AK^{\le}+K^{\le}\circ R\ \big]\ . \tag{40}$$

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 $r\in\mathbb{N}\cup\{\omega\}$ 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 [CFdlL02]) it gives less differentiability for the solution. Of course, the standard proof of the implicit function theorem reduces to a contraction mapping theorem.

Proof of Lemma 3.4. By Remark 10 we can assume that $K^\leq$ 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 $K^\leq + K^>$ is contained in the ball of radius 3. Hence, the composition $N \circ (K^\leq + K^>)$ 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

$$D\mathcal{T}(K^>)\Delta = -\mathcal{S}^{-1} DN \circ (K^\leq + K^>)\Delta \ , \tag{41}$$

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 11 below). Nevertheless, for $K^>$ and $K^> + \Delta$ in the closed unit ball of $\Gamma_{r-1,L}$, we have the finite increments formula

$$\begin{aligned} \mathcal{T}(K^> + \Delta) - \mathcal{T}(K^>) &= \int_0^1 \frac{d}{ds}\Big[\mathcal{T}(K^> + s\Delta)\Big]\, ds \\ &= -\int_0^1 \mathcal{S}^{-1} DN \circ (K^\leq + K^> + s\Delta)\Delta\, ds \ . \end{aligned} \tag{42}$$

Taking derivatives of this expression up to order $r-1$ 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 (42) 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 the $(L+1)$-derivative of (42) on B_1. In both cases, we deduce

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

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 $\|K^>\|_{\Gamma_{r-1,L}} \leq 1$ then

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

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 $\|\mathcal{T}(K^>) - \mathcal{T}(0)\|_{\Gamma_{r-1,L}} \leq \nu\|K^>\|_{\Gamma_{r-1,L}} \leq \nu < 1$. We conclude that $\mathcal{T}$ maps the closed unit ball of $\Gamma_{r-1,L}$ into itself. $\square$

Remark 11. The reason why (42) does not imply that $\mathcal{T}$ is differentiable is because the map $K^> \mapsto DN \circ (K^\leq + K^>)$ in (41) 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 $D^r N$ 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 (40) 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 $r \in \mathbb{N}$.

Finally, we deal with the case $F \in C^\infty$. According to Lemma 3.4, for any $r \in \mathbb{N}$ with $r \geq L+1$, 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

$$K = F^{-j} \circ K \circ R^j \qquad \text{for every } j \geq 1 \ . \tag{43}$$

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.** In this section 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 proposition is based on differentiating the equation

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

satisfied by $K^>$, to obtain:

$$\begin{aligned} ADK^> - DK^> \circ R\, DR = {} & -DN \circ (K^\leq + K^>)(DK^\leq + DK^>) \\ & - ADK^\leq + DK^\leq \circ R\, DR \ , \end{aligned} \tag{44}$$

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 (44) for $G := DK^> \in \Gamma_{r-2,L-1}$, that we can rewrite as

$$\tilde{\mathcal{S}} DK^> = \tilde{\mathcal{T}} DK^> + U \ , \tag{45}$$

where

$$\tilde{\mathcal{S}}G = AG - G \circ R\ DR\ ,$$
$$\tilde{\mathcal{T}}G = -DN \circ (K^{\leq} + K^{>})\ G$$

and

$$U = -DN \circ (K^{\leq} + K^{>})DK^{\leq} - ADK^{\leq} + DK^{\leq} \circ R\ DR\ . \tag{46}$$

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 (45), 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 (45), 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 $\|\tilde{\mathcal{S}}^{-1}\tilde{\mathcal{T}}\| \leq \|\tilde{\mathcal{S}}^{-1}\|\,\|\tilde{\mathcal{T}}\| < 1$ by the lemma. Hence

$$DK^{>} = \sum_{j=0}^{\infty} (\tilde{\mathcal{S}}^{-1}\tilde{\mathcal{T}})^j \tilde{\mathcal{S}}^{-1}U \qquad \text{in } \Gamma_{r-2,L-1}\ . \tag{47}$$

Now, by the definition (46) of U we have that $U \in C^{r-1}$ and, since we already know that $U \in \Gamma_{r-2,L-1}$, we conclude 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 (47) 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 $\tilde{\mathcal{T}}G = -DN \circ (K^{\leq} + K^{>})\ G$ are easily proved. Indeed, if $G \in \Gamma_{s,L-1}$ with $L-1 \leq s \leq r-1$, 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 beginning of the proof of Lemma 3.4, $K^{\leq} + K^{>}$ remains in the ball of radius 3 of X 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

$$AG - G \circ R\ DR = \eta\ . \tag{48}$$

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 (48). To finish the proof we only need to show that

$$\sum_{j=0}^{\infty} \|A^{-(j+1)} \eta \circ R^j\ DR^j\|_{\Gamma_{s,L-1}} \le C \|\eta\|_{\Gamma_{s,L-1}} \tag{49}$$

for some constant C independent of the scaling.

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

$$\begin{aligned} |D^l(A^{-(j+1)} \eta \circ R^j\ DR^j)(x)| &\le C\|A^{-1}\|^j \sum_{k=0}^{l} |D^k(\eta \circ R^j)(x)||D^{(l-k)} DR^j(x)| \\ &\le C\|A^{-1}\|^j (\|A_1\| + \varepsilon)^j \sum_{k=0}^{l} |D^k(\eta \circ R^j)(x)|\ , \end{aligned} \tag{50}$$

by (32). 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 (38) and the bound (32) , 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 (50), 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 (49). $\square$

4. Proof of Theorem 1.2

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

$$\begin{aligned} A_1 K^1 + N_1 \circ K + BK^2 &= K^1 \circ R\ , \\ A_2 K^2 + N_2 \circ K &= K^2 \circ R\ . \end{aligned}$$

If we now decide to solve this equation by setting $K^1 = \mathrm{Id}$ and by determining 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

$$K^2(x) = A_2^{-1}\Big[K^2\Big(A_1x + N_1(x, K^2(x)) + BK^2(x)\Big) - N_2(x, K^2(x))\Big] \ . \tag{51}$$

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 $B \neq 0$, 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 (51) (see the uniqueness argument at the end of Section 4.1 below).

The study of the equation under non-resonance conditions was undertaken for $B = 0$ in [dlL97] by performing some preliminary changes of variables which reduce 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 the proof of 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 both sides of (51) match.

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

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

has a unique solution in the class of C^{r-1} functions. We point out that, unfortunately, the operator $\mathcal{N}$ defined in (52) 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 third step will improve the regularity of the function to conclude that it is C^r.

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 (52) and that, with this notation, becomes

$$\mathcal{N}(w) = A_2^{-1}\Big[w \circ \psi_w - N_2 \circ (\mathrm{Id}, w)\Big] \tag{53}$$

with

$$\psi_w = A_1 + N_1 \circ (\mathrm{Id}, w) + Bw \ . \tag{54}$$

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 $K^{\leq} = \sum_{i=1}^{L} K_i$ of degree L such that $K_1 = (\mathrm{Id}, 0)$, $\pi_1 K^{\leq} = K^{\leq,1} = \mathrm{Id}$ (that is, $K_i^1 = 0$ for $2 \leq i \leq L$) and such that $F \circ K^{\leq}(x) = K^{\leq} \circ R(x) + o(|x|^L)$ for some polynomial R of degree L with $R_1 = A_1$. We know that, having set $K^{\leq,1} = \mathrm{Id}$, $F \circ K^{\leq}(x) = K^{\leq} \circ R(x) + o(|x|^L)$ 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*

$$w^{\leq}(x) = \mathcal{N}(w^{\leq})(x) + o(|x|^L) \ , \tag{55}$$

$$w^{\leq}(0) = 0 \ , \qquad Dw^{\leq}(0) = 0 \ . \tag{56}$$

Moreover, with a suitable scaling of F we can get $w^{\leq}$ as small as we want.

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

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

as well as (56) 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

$$w^{>} = \mathcal{M}(w^{>}) \ , \tag{57}$$

where

$$\mathcal{M}(w^{>}) = -w^{\leq} + \mathcal{N}(w^{\leq} + w^{>}) \ . \tag{58}$$

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 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 $w_i = w^{\le} + w_i^{>}$, $i = 1, 2$. 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 seminorm

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

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

$$|\psi_{w_2}(x)| \le (\|A_1\| + \varepsilon)|x|, \tag{59}$$

with ε small such that $\|A_2^{-1}\|(\|A_1\| + \varepsilon)^{L+1} < 1$. Using the bound (59), and expressions (53) and (54), we have:

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

for some constant $\nu < 1$, under the smallness assumptions (that here may depend on $\operatorname{Lip} w_1^{>}$ 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 $\sup_{x\in B_1}(|h(x)|/|x|^{L+1}) < \infty$. $\square$

4.2. **Existence of solution of the fixed point equation.** 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 by* (58) *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 $\nu < 1$,

$$\begin{aligned} \|\mathcal{M}(w_2^>) - &\mathcal{M}(w_1^>)\|_{\Gamma_{r-1,L}} \\ &\leq \nu \|w_2^> - w_1^>\|_{\Gamma_{r-1,L}} \quad \textit{if } \|w_2^>\|_{\Gamma_{r,L}} \leq 1 \textit{ and } \|w_1^>\|_{\Gamma_{r,L}} \leq 1\ . \end{aligned} \tag{60}$$

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 (60), 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).

To prove the bounds in the proposition, note that

$$\begin{aligned} D[\mathcal{N}(w)] = &A_2^{-1}(Dw \circ \psi_w) D\psi_w - A_2^{-1} D_1 N_2 \circ (\mathrm{Id}, w) \\ &- A_2^{-1}[D_2 N_2 \circ (\mathrm{Id}, w)] Dw \end{aligned} \tag{61}$$

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, we have, for $2 \le i \le r$,

$$\begin{aligned} D^i[\mathcal{N}(w)] \quad = \quad & A_2^{-1}(D^i w \circ \psi_w) D\psi_w^{\otimes i} \\ & + A_2^{-1}(Dw \circ \psi_w)[D_2N_1 \circ (\mathrm{Id}, w) + B]D^i w \\ & - A_2^{-1}[D_2N_2 \circ (\mathrm{Id}, w)]D^i w + V_i \end{aligned} \tag{62}$$

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 (60). 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

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

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

$$D^i[\mathcal{M}(w^> + \Delta) - \mathcal{M}(w^>)] = \int_0^1 ds \left\{ A_2^{-1}[(D^i\Delta) \circ \psi_{w+s\Delta}] D\psi_{w+s\Delta}^{\otimes i} + W_i \right\} \ , \tag{63}$$

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 (63) and $|D^L\Delta(y)| \le \|\Delta\|_{\Gamma_{r-1,L}}|y|$ if $y \in B_1$, we deduce for $x \in B_1$

$$|D^L\mathcal{I}(x)|/|x| \le [\|A_2^{-1}\|(\|A_1\| + \varepsilon)^{L+1} + \varepsilon]\|\Delta\|_{\Gamma_{r-1,L}} \le \nu\|\Delta\|_{\Gamma_{r-1,L}} \tag{64}$$

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

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

Note also that for $i \le L$, (63) 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 (64) 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 (60).

In case $r = \omega$, we use (63) 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 (60).

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 $\|\mathcal{M}(w^>)\|_{\Gamma_{r-1,L}} \le \varepsilon + \nu$. Finally, when $r \in \mathbb{N}$, (62) leads to

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

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

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 (61), 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

$$\psi_w(x) = A_1x + N_1(x, w(x)) + Bw(x) \tag{65}$$

and

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

Therefore, $H_0 := Dw$ is a solution of

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

where we define, with ψ_w given by (65),

$$\begin{aligned}\tilde{\mathcal{N}}(H)(x) =& A_2^{-1}\Big\{H(\psi_w(x))\Big[A_1 + D_1N_1(x, w(x)) \\ &+ \Big(D_2N_1(x, w(x)) + B\Big)H(x)\Big] \\ &- D_1N_2(x, w(x)) - D_2N_2(x, w(x))H(x)\Big\} \ .\end{aligned} \tag{66}$$

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

$$\tilde{\mathcal{M}}(H^>) = -H^{\leq} + \tilde{\mathcal{N}}(H^{\leq} + H^>) \ . \tag{67}$$

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

$$\begin{aligned}
&[\tilde{\mathcal{M}}(H^> + \Delta) - \tilde{\mathcal{M}}(H^>)](x) = [\tilde{\mathcal{N}}(H + \Delta) - \tilde{\mathcal{N}}(H)](x) \\
&\quad = \int_0^1 \frac{d}{ds}[\tilde{\mathcal{N}}(H + s\Delta)(x)]\, ds \\
&\quad = \int_0^1 ds\ A_2^{-1}\Big\{\Delta(\psi_w(x)) \\
&\qquad \cdot \Big[A_1 + D_1N_1(x, w(x)) + \Big(D_2N_1(x, w(x)) + B\Big)(H + s\Delta)(x)\Big] \\
&\qquad + (H + s\Delta)(\psi_w(x))\Big(D_2N_1(x, w(x)) + B\Big)\Delta(x) \\
&\qquad - D_2N_2(x, w(x))\Delta(x)\Big\}\ .
\end{aligned} \tag{68}$$

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

$$D^i[\tilde{\mathcal{M}}(H^> + \Delta) - \tilde{\mathcal{M}}(H^>)] = \int_0^1 ds\, \big\{A_2^{-1}[(D^i\Delta)\circ\psi_w]D\psi_w^{\otimes i} \otimes A_1 + U_i\big\}\ , \tag{69}$$

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 $i + 1 \leq r$.

Using the notation $\mathcal{J} = \tilde{\mathcal{M}}(H^> + \Delta) - \tilde{\mathcal{M}}(H^>)$, (69) and $|D^{L-1}\Delta(y)| \leq \|\Delta\|_{\Gamma_{s,L-1}}|y|$ for $y \in B_1$, we deduce, for some constant $\nu < 1$,

$$\begin{aligned}
|D^{L-1}\mathcal{J}(x)|/|x| &\leq \|A_2^{-1}\|(\|A_1\| + \varepsilon)^L|D^{L-1}\Delta(\psi_w(x))| + \varepsilon\|\Delta\|_{\Gamma_{s,L-1}} \\
&\leq [\|A_2^{-1}\|(\|A_1\| + \varepsilon)^{L+1} + \varepsilon]\|\Delta\|_{\Gamma_{s,L-1}} \leq \nu\|\Delta\|_{\Gamma_{s,L-1}} \quad \text{for } x \in B_1
\end{aligned} \tag{70}$$

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 $i \le L-1$, (69) 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 (70) 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

$$\tilde{\mathcal{M}}(H^>) = Dw^> + [\tilde{\mathcal{M}}(0) - \tilde{\mathcal{M}}(Dw^>)] + [\tilde{\mathcal{M}}(H^>) - \tilde{\mathcal{M}}(0)] \ , \tag{71}$$

since, by definition (67), $Dw^> - \tilde{\mathcal{M}}(Dw^>) = Dw - \tilde{\mathcal{N}}(Dw) = 0$.

Expression (71) 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 (71) belongs to $\Gamma_{r-2,L-1}$ and has small norm after scaling.

On the other hand, the previous proof of the contraction property gives that the two last terms of (71) 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 $\|\tilde{\mathcal{M}}(H^>)\|_{\Gamma_{r-2,L-1}} \le 2\varepsilon + \nu$. 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 (67) 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 (66) $r-1$ times and using the smallness of $\|B\|$ and $\|N\|_{C^r}$, we conclude

$$\begin{aligned} \|D^{r-1}[\tilde{\mathcal{M}}(H^>)]\|_{C^0} &= \|D^{r-1}[\tilde{\mathcal{N}}(H)]\|_{C^0} \\ &\le \Big(\|A_2^{-1}\|(\|A_1\| + \varepsilon)^r + \varepsilon\Big)\|H\|_{C^{r-1}} + \varepsilon \\ &\le \Big(\|A_2^{-1}\|(\|A_1\| + \varepsilon)^r + \varepsilon\Big)(\varepsilon + 1) + \varepsilon \\ &\le 1 \end{aligned} \tag{72}$$

if we take ε small enough. □

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,

$$\lambda \in \mathrm{Spec}_{ap}(A) \Leftrightarrow \exists \{x_n\} \text{ with } \|x_n\| \geq \alpha > 0 \text{ and } \|(A - \lambda)x_n\| \to 0 \;. \tag{73}$$

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.,

$$\partial(\mathrm{Spec}(A)) \subset \mathrm{Spec}_{ap}(A). \tag{74}$$

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 let A be a bounded 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.,

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

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

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

together with the inequality $(\rho(A^{-1})+\varepsilon)^{-1} \le \rho(A)+\varepsilon$, which follows from the fact that $\rho(A^{-1})\rho(A) \ge 1$ (a consequence of the first statement in Remark 3).

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 be a bounded linear map on X such that*

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

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*

$$\begin{array}{rclcrcll} \|A_1\| & \le & \rho(A_1)+\varepsilon\ , & & \|A_1^{-1}\| & \le & \rho(A_1^{-1})+\varepsilon\ , & \\ \|A_2\| & \le & \rho(A_2)+\varepsilon\ , & & \|A_2^{-1}\| & \le & \rho(A_2^{-1})+\varepsilon\ , & \\ \|A\| & \le & \rho(A)+\varepsilon\ , & & \|A^{-1}\| & \le & \rho(A^{-1})+\varepsilon\ , & \textit{and} \\ \|B\| & \le & \varepsilon\ . & & & & & \end{array}$$

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

$$A^{-1} = \begin{pmatrix} A_1^{-1} & -A_1^{-1}BA_2^{-1} \\ 0 & A_2^{-1} \end{pmatrix} .$$

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

$$\mathrm{Spec}(A) \subset \mathrm{Spec}(A_1) \cup \mathrm{Spec}(A_2) \tag{75}$$

and, in particular, $\rho(A) \le \max\big(\rho(A_1), \rho(A_2)\big)$. Therefore, we only need to show that

$$\rho(A) \ge \max\big(\rho(A_1), \rho(A_2)\big)\ . \tag{76}$$

To prove this, we first claim that

$$\mathrm{Spec}_{ap}(A) \supset \mathrm{Spec}_{ap}(A_1) \cup \Big(\mathrm{Spec}_{ap}(A_2) \cap \mathrm{Res}(A_1) \Big) \ . \tag{77}$$

Then, using that $\partial(\mathrm{Spec}(A_i)) \subset \mathrm{Spec}_{ap}(A_i)$ for $i = 1, 2$ (i.e., property (74) applied to A_1 and A_2), we conclude inequality (76) and the proposition.

To establish (77), 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 (77). $\square$

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 (75) 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 (75) and (77) are strict. This is due to the presence of residual spectrum.

A.2. **Proof of Proposition 3.2.** The proof of the equalities in (24) 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

$$\begin{array}{rcll} M_{i;j_1,\dots,j_n}(u_{j_1}, \dots, u_{j_n}) & = & v_i, & \\ M_{i;j_1,\dots,j_n}(u_{s_1}, \dots, u_{s_n}) & = & 0, & \qquad \text{when } (s_1, \dots, s_n) \neq (j_1, \dots, j_n) \ . \end{array}$$

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

$$\mathcal{L}_{n,A,B} = \mathcal{L}_B\mathcal{R}^1_A\cdots\mathcal{R}^n_A \ . \tag{78}$$

Moreover, we have:

Proposition A.3. *With the notations of Proposition 3.2, one has*

$$\mathrm{Spec}(\mathcal{L}_B, \mathbf{M}_n) \subset \mathrm{Spec}(B, Y) \ , \tag{79}$$

$$\mathrm{Spec}(\mathcal{R}^k_A, \mathbf{M}_n) \subset \mathrm{Spec}(A, X) \ . \tag{80}$$

Proof. The proof is immediate. If $B-\lambda$ is invertible, then $\mathcal{L}_B-\lambda$ is also invertible since its inverse is given by $(\mathcal{L}_B-\lambda)^{-1} = \mathcal{L}_{(B-\lambda)^{-1}}$. The same proof applies to $\mathcal{R}^k_A$.

In an alternative way, 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, (78), (79) and (80).

The result for $\mathbf{S}_n$ follows from the following:

Proposition A.5. *With the notations of Proposition 3.2, we have*

$$\mathrm{Spec}(\mathcal{L}_{n,A,B}, \mathbf{S}_n) \subset \mathrm{Spec}(\mathcal{L}_{n,A,B}, \mathbf{M}_n) \ . \tag{81}$$

Proof. If $\lambda \notin \mathrm{Spec}(\mathcal{L}_{n,A,B}, \mathbf{M}_n)$, given any $\eta \in \mathbf{M}_n$ there is a unique $\Gamma \in \mathbf{M}_n$ such that

$$(\mathcal{L}_{n,A,B} - \lambda)\Gamma = \eta \ . \tag{82}$$

Such Γ satisfies $\|\Gamma\| \le K\|\eta\|$.

If $\Pi(x_1, \dots, x_n) = (x_{\pi(1)}, \dots, x_{\pi(n)})$ is a permutation of the n vectors, it is easy to check that

$$(\mathcal{L}_{n,A,B} - \lambda)(\Gamma\Pi) = [(\mathcal{L}_{n,A,B} - \lambda)\Gamma]\Pi \ . \tag{83}$$

Therefore, if $\eta\Pi = \eta$ we obtain, using the uniqueness of Γ and (83), that $\Gamma\Pi = \Gamma$.

Hence, if equation (82) can be solved uniquely and boundedly in $\mathbf{M}_n$, it can be similarly solved in $\mathbf{S}_n$. $\square$

Finally, we note that the construction of the forms $M_{i;j_1,\dots,j_n}$ 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

$$\begin{aligned} \mathrm{Spec}_{pp}(\mathcal{L}_{n,A,B}, \mathbf{M}_n) &\supset \mathrm{Spec}_{pp}(B)\big(\mathrm{Spec}_{pp}(A)\big)^n \ , \\ \mathrm{Spec}_{pp}(\mathcal{L}_{n,A,B}, \mathbf{S}_n) &\supset \mathrm{Spec}_{pp}(B)\big(\mathrm{Spec}_{pp}(A)\big)^n \ , \end{aligned} \tag{84}$$

where Spec_{pp} denotes the closure of the set of eigenvalues.

Therefore, if we assume $\mathrm{Spec}_{pp}(B) = \mathrm{Spec}(B)$ and $\mathrm{Spec}_{pp}(A) = \mathrm{Spec}(A)$ then, combining (84) and (24), we obtain equalities in (24).

Acknowledgments

This work has been supported by the *Comisión Conjunta Hispano Norteamericana 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 the Spanish CICYT grant PB98-0932-C02-01, and by a Harrington Faculty Fellowship at the University of Texas-Austin. 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.

References

[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: regularity with respect to parameters. *Preprint*, 2000.

[CFdlL02] Xavier Cabré, Ernest Fontich, and Rafael de la Llave. The parameterization method for invariant manifolds III: overview and applications. *In preparation*, 2002.

[CL99] Carmen Chicone and Yuri Latushkin. *Evolution semigroups in dynamical systems and differential equations.* American Mathematical Society, Providence, RI, 1999.

[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.

[Fra98] Simon J. Fraser. Double perturbation series in the differential equations of enzyme kinetics. *Jour. Chem. Phys.*, 109(2):411–423, 1998.

[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.

[Irw70] M.C. Irwin. On the stable manifold theorem. *Bull. London Math. Soc.*, 2:196–198, 1970.

[Irw80] M.C. Irwin. A new proof of the pseudostable manifold theorem. *J. London Math. Soc.*, 21:557–566, 1980.

[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.

[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.

[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.

[Poi90] Henri Poincaré. Sur une classe nouvelle de transcendentes uniformes. *Jour. de Math.*, 6:313–365, 1890.

[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.

[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.

[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.

Department of Mathematics, RLM 8.100, The University of Texas at Austin, Austin, TX 78712-1082, U.S.A,

E-mail address, X. Cabré: `cabre@math.utexas.edu`

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, RLM 8.100, The University of Texas at Austin, Austin, TX 78712-1082, U.S.A,

E-mail address, R. de la Llave: `llave@math.utexas.edu`