J. Eur. Math. Soc. 23, 3323–3349 (2021) DOI 10.4171/JEMS/1086

Ivar Ekeland · Éric Séré

A surjection theorem for maps with singular perturbation and loss of derivatives

Received November 15, 2018

Abstract. In this paper we introduce a new algorithm for solving perturbed nonlinear functional equations which admit a right-invertible linearization, but with an inverse that loses derivatives and may blow up when the perturbation parameter ε goes to zero. These equations are of the form $F_\varepsilon(u) = v$ with $F_\varepsilon(0) = 0$, v small and given, u small and unknown. The main difference from the by now classical Nash–Moser algorithm is that, instead of using a regularized Newton scheme, we solve a sequence of Galerkin problems thanks to a topological argument. As a consequence, in our estimates there are *no quadratic terms*. For problems without perturbation parameter, our results require weaker regularity assumptions on F and v than earlier ones, such as those of Hörmander [17]. For singularly perturbed functionals F_ε, we allow v to be larger than in previous works. To illustrate this, we apply our method to a nonlinear Schrödinger Cauchy problem with concentrated initial data studied by Texier–Zumbrun [26], and we show that our result improves significantly on theirs.

Keywords. Inverse function theorem, loss of derivatives, singular perturbations, Nash–Moser theorem, Cauchy problem, nonlinear Schrödinger system, Ekeland's variational principle

1. Introduction

The basic idea of the inverse function theorem (henceforth IFT) is that, if a map F is differentiable at a point u_0 and the derivative $DF(u_0)$ is invertible, then the map itself is invertible in some neighbourhood of u_0. It has a long and distinguished history (see [20] for instance), going back to the inversion of power series in the seventeenth century, and has been extended since to maps between infinite-dimensional spaces. If the underlying space is Banach, and if one is only interested in the local surjectivity of F, that is, the existence, near u_0, of a solution u to the equation $F(u) = v$ for v close to $F(u_0)$, one just needs to assume that F is of class C^1 and that $DF(u_0)$ has a right-inverse $L(u_0)$. The

Ivar Ekeland: Université Paris-Dauphine, PSL University, CNRS, CEREMADE, Place de Lattre de Tassigny, 75016 Paris, France; ekeland@ceremade.dauphine.fr

Éric Séré: Université Paris-Dauphine, PSL University, CNRS, CEREMADE, Place de Lattre de Tassigny, 75016 Paris, France; sere@ceremade.dauphine.fr

Mathematics Subject Classification (2020): Primary 47J07; Secondary 35B25, 35G25, 35Q55, 47J25, 58C15

standard proof is based on the Picard scheme

$$u_n = u_{n-1} - L(u_0)(F(u_{n-1}) - v)$$

which converges geometrically to a solution of $F(u) = v$ provided $\|F(u_0) - v\|$ is small enough. In the C^2 case, the Newton algorithm

$$u_n = u_{n-1} - L(u_{n-1})(F(u_{n-1}) - v)$$

uses the right-invertibility of $DF(u)$ for u close to u_0, and provides local quadratic convergence.

In functional analysis, u will typically be a function. In many situations the IFT on Banach spaces will be enough, but in the study of Hamiltonian systems and PDEs, one encounters cases when the right-inverse $L(u)$ of $DF(u)$ loses derivatives, i.e. when $L(u)F(w)$ has fewer derivatives than u and w. In such a case, the Picard and Newton schemes lose derivatives at each step. The first solutions to this problem are due, on the one hand, to Kolmogorov [19] and Arnol'd [3], [2], [4] who investigated perturbations of completely integrable Hamiltonian systems in the analytic class, and showed that invariant tori persist under small perturbations, and, on the other hand, to Nash [23], who showed that any smooth compact Riemannian manifold can be isometrically imbedded into a Euclidean space of sufficiently high dimension.[1]

In both cases, the fast convergence of Newton's scheme was used to overcome the loss of regularity. Since Nash was considering functions with finitely many derivatives, he had to introduce a sequence of smoothing operators $\mathcal{S}_n$, in order to regularize $L(u_{n-1})(F(u_{n-1}) - v)$, and the new scheme was

$$u_n = u_{n-1} - \mathcal{S}_n L(u_{n-1})(F(u_{n-1}) - v).$$

An early presentation of Nash's method can be found in Schwartz' notes [24]. It was further improved by Moser [22], who used it to extend the Kolmogorov–Arnol'd results to C^k Hamiltonians. The Nash–Moser method has been the source of a considerable amount of work in many different situations, giving rise in each case to a so-called "hard" IFT. We will not attempt to review this line of work in the present paper. A survey up to 1982 is found in [15]. In [17], Hörmander introduced a refined version of the Nash–Moser scheme which provides the best estimates to date on the regularity loss. We refer to [1] for a pedagogical account of this work, and to [5] for recent improvements. We also gained much insight into the Nash–Moser scheme from the papers [7], [8], [9], [10], [26].

The question we want to address here is the following. The IFT implies that the range of F contains a neighbourhood $\mathcal{V}$ of $v_0 = F(u_0)$. What is the size of $\mathcal{V}$?

In general, when one tries to apply the abstract Nash–Moser theorem directly, the estimates which can be derived from its proof are unreasonably small, many orders of

[1] Nash's theorem on isometric embeddings was later re-proved by Gunther [14], who found a different formulation of the problem and was able to use the classical IFT in Banach spaces.

magnitude away from what can be observed in numerical simulations or physical experiments. Moreover, precise estimates for the Nash–Moser method are difficult to compute, and most theoretical papers simply do not address the question.

So we shall address instead a “hard” singular perturbation problem with loss of derivatives. The same issue appears in such problems, as we shall explain in a moment, but it takes a simpler form: one tries to find a good estimate on the size of $\mathcal{V}$ as a power of the perturbation parameter ε. Such an asymptotic analysis has been carefully done in the paper of Texier and Zumbrun [26] which has been an important source of inspiration to us, and we will be able to compare our results with theirs. As noted by those authors, the use of Newton’s scheme implies an intrinsic limit to the size of $\mathcal{V}$.

Let us explain this in the “soft” case, without loss of derivatives. Suppose that for every $0 < \varepsilon \leq 1$ we have a C^2 map F_ε between two Banach spaces X and Y, such that $F_\varepsilon(0) = 0$, and, for all $\|u\| \leq R$,

$$|||D_u F_\varepsilon(u)^{-1}||| \leq \varepsilon^{-1} M, \qquad |||D^2_{uu} F_\varepsilon(u)||| \leq K.$$

Then the Newton–Kantorovich theorem (see [11, Section 7.7] for a comprehensive discussion) tells us that the solution u_ε of $F_\varepsilon(u) = v$ exists for $\|v\| < \varepsilon^2/(2KM^2)$, and this is essentially the best result one can hope for when using Newton’s algorithm, as mentioned by Texier and Zumbrun in [26, Remark 2.22]. Note that the use of a Picard iteration would give a similar condition.

However, in this simple situation where no derivatives are lost, it is possible, using topological arguments instead of Newton’s method, to find a solution u provided $\|v\| \leq \varepsilon R/M$: one order of magnitude in ε has been gained. The first result of this kind, when F is C^1 and $\dim X = \dim Y < \infty$, is due to Ważewski [27] who used a continuation method. See also [18] and [25] and the references in these papers, for more general results in this direction. In [12, Theorem 2], using Ekeland’s variational principle, Ważewski’s result is proved in Banach spaces, assuming only that F is continuous and Gâteaux differentiable, the differential having a uniformly bounded right-inverse (in §2 below, we recall this result as Theorem 5).

Our goal is to extend such a topological approach to “hard” problems with loss of derivatives, which up to now have been tackled by the Nash–Moser algorithm. A first attempt in this direction was made in [12, Theorem 1], in the case when the estimates on the right-inverse do not depend on the base point, but it is very hard to find examples of such situations. The present paper fulfills the program in the general case, where estimates on the inverse depend on the base point.

In [10], Berti, Bolle and Procesi prove a new version of the Nash–Moser theorem by solving a sequence of Galerkin problems $\Pi'_n F(u_n) = \Pi'_n v$, $u_n \in E_n$, where Π_n and Π'_n are projectors and E_n is the range of Π_n. They find the solution of each projected equation thanks to a Picard iteration:

$$u_n = \lim_{k\to\infty} w^k \quad \text{with} \quad w^0 = u_{n-1} \text{ and } w^{k+1} = w^k - L_n(u_{n-1})(F(w^k) - v),$$

where $L_n(u_{n-1})$ is a right-inverse of $D(\Pi'_n F_{|E_n})(u_{n-1})$. So, in [10] the regularized Newton step is not really absent: it is essentially the first step in each Picard iteration. As a consequence, the proof in [10] involves quadratic estimates similar to the ones of more standard Nash–Moser schemes. Moreover, Berti, Bolle and Procesi assume the right-invertibility of $D(\Pi'_n F_{|E_n})(u_{n-1})$. This assumption is perfectly suitable for the applications they consider (periodic solutions of a nonlinear wave equation), but in general it is not a consequence of the right-invertibility of $DF(u_{n-1})$, and this restricts the generality of their method as compared with the standard Nash–Moser scheme.

As in [10], we work with projectors and solve a sequence of Galerkin problems. But in contrast with [10], the Newton steps are completely absent in our new algorithm, they are replaced by the topological argument from [12, Theorem 2], ensuring the solvability of each projected equation. Incidentally, this allows us to work with functionals F that are only continuous and Gâteaux-differentiable, while the standard Nash–Moser scheme requires twice-differentiable functionals. Our regularity assumption on v also seems to be optimal, and even weaker than in [17]. Moreover, our method works assuming either the right-invertibility of $D(\Pi'_n F_{|E_n})(u)$ as in [10], or the right-invertibility of $DF(u)$ (in the second case, our proof is more complicated). But in our opinion, the main advantage of our approach is the following: there are *no more quadratic terms* in our estimates, as a consequence we can deal with larger v's, and this advantage is particularly obvious in the case of singular perturbations.

To illustrate this, we will give an abstract existence theorem with a precise estimate of the range of F for a singular perturbation problem: this is Theorem 3 below. Comparing our result with the abstract theorem of [26], one can see that we have weaker assumptions and a stronger conclusion. Then we will apply Theorem 3 to an example given in [26], namely a Cauchy problem for a quasilinear Schrödinger system first studied by Métivier and Rauch [21]. Texier and Zumbrun use their abstract Nash–Moser theorem to prove the existence of solutions of this system on a fixed time interval, for concentrated initial data. Our abstract theorem allows us to increase the order of magnitude of the oscillation in the initial data. After reading our paper, Baldi and Haus [6] have been able to increase this order of magnitude even more, using their own version [5] of the Newton scheme for Nash–Moser, combined with a clever modification of the norms considered in [26] and an improved estimate on the second derivative of the functional. In contrast, our proof follows directly from our abstract theorem, taking exactly the same norms and estimates as in [26], and without even considering the second derivative of the functional.

The paper is constructed as follows. In Section 2, we present the general framework: we are trying to solve the equation $F_\varepsilon(u) = v$ near $F_\varepsilon(0) = 0$, when F_ε maps a scale of Banach spaces of functions into another and admits a right-invertible Gâteaux differential with "tame estimates" involving losses of derivatives and negative powers of ε. After giving our precise assumptions, we state our main theorem. Section 3 is devoted to its proof. In Section 4, we apply it to the example taken from Texier and Zumbrun [26], and we compare our results with theirs.

2. Main assumptions and results

2.1. Two tame scales of Banach spaces

Let $(V_s, \|\cdot\|_s)_{0\le s\le S}$ be a scale of Banach spaces, that is,

$$0 \le s_1 \le s_2 \le S \implies [V_{s_2} \subset V_{s_1} \text{ and } \|\cdot\|_{s_1} \le \|\cdot\|_{s_2}].$$

We shall assume that to each $\Lambda \in [1, \infty)$ is associated a continuous linear projection $\Pi(\Lambda)$ on V_0, with range $E(\Lambda) \subset V_S$. We shall also assume that the spaces $E(\Lambda)$ form a nondecreasing family of sets indexed by $[1, \infty)$, while the spaces $\operatorname{Ker}\Pi(\Lambda)$ form a nonincreasing family. In other words,

$$1 \le \Lambda \le \Lambda' \implies \Pi(\Lambda)\Pi(\Lambda') = \Pi(\Lambda')\Pi(\Lambda) = \Pi(\Lambda).$$

Finally, we assume that the projections $\Pi(\Lambda)$ are "smoothing operators" satisfying the following estimates:

Polynomial growth and approximation. *There are constants $A_1, A_2 \ge 1$ such that, for all $0 \le s \le S$, all $\Lambda \in [1, \infty)$ and all $u \in V_s$, we have*

$$\forall t \in [0, S], \quad \|\Pi(\Lambda)u\|_t \le A_1\, \Lambda^{(t-s)^+}\|u\|_s, \tag{2.1}$$

$$\forall t \in [0, s], \quad \|(1 - \Pi(\Lambda))u\|_t \le A_2\, \Lambda^{-(s-t)}\|u\|_s. \tag{2.2}$$

When the above properties are met, we shall say that $(V_s, \|\cdot\|_s)_{0\le s\le S}$ endowed with the family of projectors $\{\, \Pi(\Lambda) : \Lambda \in [1, \infty)\,\}$, is a *tame* Banach scale.

It is well-known (see e.g. [10]) that (2.1,2.2) imply:

Interpolation inequality. *For $0 \le t_1 \le s \le t_2 \le S$,*

$$\|u\|_s \le A_3\|u\|_{t_1}^{\frac{t_2-s}{t_2-t_1}}\,\|u\|_{t_2}^{\frac{s-t_1}{t_2-t_1}}. \tag{2.3}$$

Let $(W_s, \|\cdot\|'_s)_{0\le s\le S}$ be another tame scale of Banach spaces. We shall denote by $\Pi'(\Lambda)$ the corresponding projections defined on W_0 with ranges $E'(\Lambda) \subset W_S$, and by A'_i $(i = 1, 2, 3)$ the corresponding constants in (2.1), (2.2) and (2.3).

Remark. In many practical situations, the projectors form a discrete family, for instance, $\{\Pi(N) : N \in \mathbb{N}^*\}$, or $\{\Pi(2^j) : j \in \mathbb{N}\}$. The first case occurs when $\Pi(N)$ acts on periodic functions by truncating their Fourier series, keeping only frequencies of size less than or equal to N, as in [10]. The second case occurs when truncating orthogonal wavelet expansions as in an earlier version of the present work [13]. Our choice of notation and assumptions covers these cases, taking $\Pi(\Lambda) = \Pi(\lfloor\Lambda\rfloor)$ or $\Pi(\Lambda) = \Pi(2^{\lfloor\log_2(\Lambda)\rfloor})$, where $\lfloor\cdot\rfloor$ denotes the integer part.

2.2. Main theorem

We state our result in the framework of singular perturbations, in the spirit of Texier and Zumbrun [26]. The norms $\|\cdot\|_s$, $\|\cdot\|'_s$ on the tame scales (V_s), (W_s) may depend on the perturbation parameter $\varepsilon \in (0,1]$, as well as the projectors $\Pi(\Lambda)$, $\Pi'(\Lambda)$ and their ranges $E(\Lambda)$, $E'(\Lambda)$. But we require that S and the constants A_i, A'_i appearing in estimates (2.1)–(2.3) be independent of ε. In order to avoid burdensome notations, the dependence of the norms, projectors and subspaces on ε will not be explicit in what follows.

Denote by B_s the unit ball in V_s:

$$B_s = \{u : \|u\|_s \leq 1\}$$

We fix nonnegative constants s_0, m, ℓ, ℓ' and g, independent of ε. We will assume that S is large enough.

We first recall the definition of Gâteaux-differentiability, in a form adapted to our framework:

Definition 1. We shall say that a function $F : B_{s_0+m} \to W_{s_0}$ is *Gâteaux-differentiable* (henceforth *G-differentiable*) if for every $u \in B_{s_0+m}$, there exists a linear map $DF(u) : V_{s_0+m} \to W_{s_0}$ such that for every $s \in [s_0, S-m]$, if $u \in B_{s_0+m} \cap V_{s+m}$, then $DF(u)$ maps continuously V_{s+m} into W_s, and

$$\forall h \in V_{s+m}, \quad \lim_{t \to 0} \left\| \frac{1}{t}[F(u+th) - F(u)] - DF(u)h \right\|'_s = 0.$$

Note that, even in finite dimension, a G-differentiable map need not be C^1, or even continuous. However, if $DF : V_{s+m} \to \mathcal{L}(V_{s+m}, W_s)$ is locally bounded, then $F : V_{s+m} \to W_s$ is locally Lipschitz, hence continuous. In the present paper, we will always be in such a situation.

We now consider a family $(F_\varepsilon)_{0<\varepsilon\leq 1}$ of maps with $F_\varepsilon : B_{s_0+m} \to W_{s_0}$. We are ready to state our assumptions on this family:

Definition 2.

- We shall say that the maps $F_\varepsilon : B_{s_0+m} \to W_{s_0}$ $(0 < \varepsilon \leq 1)$ form an *S-tame differentiable family* if they are G-differentiable with respect to u, and, for some positive constant a, for all $\varepsilon \in (0,1]$ and all $s \in [s_0, S-m]$, if $u \in B_{s_0+m} \cap V_{s+m}$ and $h \in V_{s+m}$, then $DF_\varepsilon(u)h \in W_s$ with the tame direct estimate

$$\|DF_\varepsilon(u)h\|'_s \leq a(\|h\|_{s+m} + \|u\|_{s+m}\|h\|_{s_0+m}). \tag{2.4}$$

- Further, we shall say that $(DF_\varepsilon)_{0<\varepsilon\leq 1}$ is *tame right-invertible* if there are $b > 0$ and $g, \ell, \ell' \geq 0$ such that for all $0 < \varepsilon \leq 1$ and $u \in B_{s_0+\max\{m,\ell\}}$, there is a linear map $L_\varepsilon(u) : W_{s_0+\ell'} \to V_{s_0}$ satisfying

$$\forall k \in W_{s_0+\ell'}, \quad DF_\varepsilon(u)L_\varepsilon(u)k = k, \tag{2.5}$$

and for all $s_0 \le s \le S - \max\{\ell, \ell'\}$, if $u \in B_{s_0+\max\{m,\ell\}} \cap V_{s+\ell}$ and $k \in W_{s+\ell'}$, then $L_\varepsilon(u)k \in V_s$, with the tame inverse estimate

$$\|L_\varepsilon(u)k\|_s \le b\varepsilon^{-g}(\|k\|'_{s+\ell'} + \|k\|'_{s_0+\ell'}\|u\|_{s+\ell}). \tag{2.6}$$

- Alternatively, we shall say that $(DF_\varepsilon)_{0<\varepsilon\le 1}$ is *tame Galerkin right-invertible* if there are $\underline{\Lambda} \ge 1$, $b > 0$ and $g, \ell, \ell' \ge 0$ such that for all $\Lambda \ge \underline{\Lambda}$, $0 < \varepsilon \le 1$ and any $u \in B_{s_0+\max\{m,\ell\}} \cap E(\Lambda)$, there is a linear map $L_{\Lambda,\varepsilon}(u) : E'(\Lambda) \to E(\Lambda)$ satisfying

$$\forall k \in E'(\Lambda), \qquad \Pi'(\Lambda)DF_\varepsilon(u)L_{\Lambda,\varepsilon}(u)k = k, \tag{2.7}$$

and for all $s_0 \le s \le S - \max\{\ell, \ell'\}$, we have the tame inverse estimate

$$\forall k \in E'(\Lambda), \qquad \|L_{\Lambda,\varepsilon}(u)k\|_s \le b\varepsilon^{-g}(\|k\|'_{s+\ell'} + \|k\|'_{s_0+\ell'}\|u\|_{s+\ell}). \tag{2.8}$$

In this definition, the integers m, ℓ, ℓ' denote the loss of derivatives for DF_ε and its right-inverse, and $g > 0$ denotes the strength of the singularity at $\varepsilon = 0$. The unperturbed case of a fixed map can be recovered by setting $\varepsilon = 1$.

We want to solve the equation $F_\varepsilon(u) = v$. There are three things of importance. How regular is v? How regular is u, or, equivalently, how small is the loss of derivatives between v and u? How does the existence domain depend on ε? The following result answers them in a near-optimal way.

Theorem 3. *Assume that the maps F_ε $(0 < \varepsilon \le 1)$ form an S-tame differentiable family between the tame scales $(V_s)_{0\le s\le S}$ and $(W_s)_{0\le s\le S}$, with $F_\varepsilon(0) = 0$ for all $0 < \varepsilon \le 1$. Assume, in addition, that $(DF_\varepsilon)_{0<\varepsilon\le 1}$ is either tame right-invertible or tame Galerkin right-invertible. Let s_0, m, g, ℓ, ℓ' be the associated parameters. Let $s_1 \ge s_0 + \max\{m, \ell\}$, $\delta > s_1 + \ell'$ and $g' > g$. Then, for S large enough, there is $r > 0$ such that, whenever $0 < \varepsilon \le 1$ and $\|v\|'_\delta \le r\varepsilon^{g'}$, there exists some $u_\varepsilon \in B_{s_1}$ satisfying*

$$F_\varepsilon(u_\varepsilon) = v, \qquad \|u_\varepsilon\|_{s_1} \le r^{-1}\,\varepsilon^{-g'}\|v\|'_\delta.$$

As we will see, the proof of Theorem 3 is much shorter under the assumptions that DF_ε is Galerkin right-invertible. But in many applications, it is easier to check that DF_ε is tame right-invertible than tame Galerkin right-invertible. See [10], however, where an assumption similar to (2.7), (2.8) is used.

All other "hard" surjection theorems that we know of require some additional conditions on the second derivative of F_ε. Here we do not need such assumptions, in fact we only assume that F_ε is G-differentiable, not C^2.

As for the three questions we raised, let us explain in which sense the answers are almost optimal in Theorem 3. For the tame estimates (2.4), (2.6) to hold, one needs $u \in B_{s_1}$ with $s_1 \ge s_0 + \max\{m, \ell\}$. When solving the linearized equation $DF_\varepsilon(u)h = k$ in V_{s_1} by $h = L_\varepsilon(u)k$, one needs $k \in W_{s_1+\ell'}$, so it seems necessary to assume $\delta \ge s_1 + \ell'$, and we find that the strict inequality is sufficient. Replacing s_1 with its minimal value, our condition on δ becomes

$$\delta > s_0 + \max\{m, \ell\} + \ell'.$$

We have not found this condition in the literature: in [17], for instance, a stronger assumption is made, namely $\delta > s_0 + \max\{2m + \ell', \ell\} + \ell'$.

For the dependence of $\|v\|'_\delta$ on ε, the constraint $g' > g$ also seems to be nearly optimal. Indeed, the solution u_ε has to be in B_{s_1}, but the right-inverse L_ε of DF_ε has a norm of order ε^{-g}, so the condition $\|v\|'_\delta \lesssim \varepsilon^g$ seems necessary. We find that the condition $\|v\|'_\delta \lesssim \varepsilon^{g'}$ is sufficient.

Our condition on S is of the form $S \geq S_0$ where S_0 depends only on the parameters s_0, m, g, ℓ, ℓ' and g', s_1, δ. Then r depends only on these parameters and the constants A_i, A'_i associated with the tame scales. In principle, all these constants could be made explicit, but we will not do it here. Let us just mention that one can take $S_0 = \mathcal{O}\big(\frac{1}{g'-g}\big)$ as $g' \to g$, all other parameters remaining fixed. This follows from the inequality $\sigma < \zeta g/\eta$ in Lemma 1.

In the case of a tame right-invertible differential, we can restate our theorem in a form that allows direct comparison with [26, Theorem 2.19 and Remarks 2.9, 2.14]. For this purpose, we consider two tame Banach scales $(V_s, \|\cdot\|_s)$ and $(W_s, \|\cdot\|'_s)$ with associated projectors Π_Λ, Π'_Λ, we take $\gamma > 0$ and we introduce the norms $|\cdot|_s := \varepsilon^\gamma\|\cdot\|_s$ and $|\cdot|'_s := \varepsilon^\gamma\|\cdot\|'_s$. We then denote $\mathfrak{B}_s(\rho) = \{u : |u|_s \leq \rho\}$ and we consider functions F_ε of the form $F_\varepsilon(u) = \Phi_\varepsilon(\mathfrak{a}_\varepsilon + u) - \Phi_\varepsilon(\mathfrak{a}_\varepsilon)$, where Φ_ε is defined on $\mathfrak{B}_{s_0+m}(2\varepsilon^\gamma)$ and $\mathfrak{a}_\varepsilon \in \mathfrak{B}_S(\varepsilon^\gamma)$ is chosen such that $v_\varepsilon := -\Phi_\varepsilon(\mathfrak{a}_\varepsilon)$ is very small. A point u in B_{s_0+m} satisfies $F_\varepsilon(u) = v_\varepsilon$ if and only if it solves the equation $\Phi_\varepsilon(\mathfrak{a}_\varepsilon + u) = 0$ in $\mathfrak{B}_{s_0+m}(\varepsilon^\gamma)$. We make the following assumptions on Φ_ε:

For some $\gamma > 0$ and any $0 < \varepsilon \leq 1$, the map $\Phi_\varepsilon : \mathfrak{B}_{s_0+m}(2\varepsilon^\gamma) \to W_{s_0}$ is G-differentiable with respect to u, and there are constants a, b and $g > 0$ such that:

- *for all $0 < \varepsilon \leq 1$ and $s_0 \leq s \leq S - m$, if $u \in \mathfrak{B}_{s_0+m}(2\varepsilon^\gamma) \cap V_{s+m}$ and $h \in V_{s+m}$, then $D\Phi_\varepsilon(u)h \in W_s$, with the tame direct estimate*
$$|D\Phi_\varepsilon(u)h|'_s \leq a(|h|_{s+m} + \varepsilon^{-\gamma}|u|_{s+m}|h|_{s_0+m}); \tag{2.9}$$
- *for all $0 < \varepsilon \leq 1$ and $u \in \mathfrak{B}_{s_0+\max\{m,\ell\}}(2\varepsilon^\gamma)$, there is $L_\varepsilon(u) : W_{s_0+\ell'} \to V_{s_0}$ linear, satisfying*
$$\forall k \in W_{s_0+\ell'}, \qquad D\Phi_\varepsilon(u)L_\varepsilon(u)k = k, \tag{2.10}$$
and for all $s_0 \leq s \leq S - \max\{\ell, \ell'\}$, if $u \in \mathfrak{B}_{s_0+\max\{m,\ell\}}(2\varepsilon^\gamma) \cap V_{s+\ell}$ and $k \in W_{s+\ell'}$, then $L_\varepsilon(u)k \in V_s$, with the tame inverse estimate
$$|L_\varepsilon(u)k|_s \leq b\varepsilon^{-g}(|k|'_{s+\ell'} + \varepsilon^{-\gamma}|k|'_{s_0+\ell'}|u|_{s+\ell}). \tag{2.11}$$

Under these assumptions, the maps $F_\varepsilon : B_{s_0+m} \to W_{s_0}$ form an S-tame differentiable family for the "old" norms $\|\cdot\|_s$, $\|\cdot\|'_s$. So the following result holds, as a direct consequence of our main theorem:

Corollary 4. *Consider two tame Banach scales $(V_s, |\cdot|_s)_{0\leq s\leq S}$ and $(W_s, |\cdot|'_s)_{0\leq s\leq S}$, nonnegative constants s_0, m, ℓ, ℓ', g, γ, and positive constants a, b. Take any $g' > g$, $s_1 \geq s_0 + \max\{m, \ell\}$ and $\delta > s_1 + \ell'$. For S large enough and $r > 0$ small, if a family* *of G-differentiable maps* $\Phi_\varepsilon : \mathfrak{B}_{s_0+m}(2\varepsilon^\gamma) \to W_{s_0}$ $(0 < \varepsilon \le 1)$ *satisfies* (2.9)–(2.11) *and, in addition, for some* $\mathfrak{a}_\varepsilon \in \mathfrak{B}_S(\varepsilon^\gamma)$, $|\Phi_\varepsilon(\mathfrak{a}_\varepsilon)|'_\delta \le r\varepsilon^{\gamma+g'}$, *then there exists some* $u_\varepsilon \in \mathfrak{B}_{s_1,\varepsilon}(\varepsilon^\gamma)$ *such that*

$$\Phi_\varepsilon(\mathfrak{a}_\varepsilon + u_\varepsilon) = 0, \quad |u_\varepsilon|_{s_1} \le r^{-1}\varepsilon^{-g'}|\Phi_\varepsilon(\mathfrak{a}_\varepsilon)|'_\delta.$$

In [26, Theorem 2.19 and Remarks 2.9, 2.14], the assumptions are stronger, since they involve the second derivative of Φ_ε. More importantly, we only need the norm of $\Phi_\varepsilon(\mathfrak{a}_\varepsilon)$ to be controlled by $\varepsilon^{\gamma+g'}$ with $g' > g$, provided $S \ge S_0$ with $S_0 = \mathcal{O}\big(\frac{1}{g'-g}\big)$, while in [26, Assumption 2.15 and Remark 2.23], due to quadratic estimates, one needs $g' > 2g$ with the faster growth $S_0 = \mathcal{O}\big(\frac{1}{(g'-2g)^2}\big)$.

3. Proof of Theorem 3

The proof consists in constructing a sequence $(u_n)_{n\ge 1}$ which converges to a solution u of $F(u) = v$. At each step, in order to find u_n, we solve a nonlinear equation in a Banach space, using [12, Theorem 2], which we restate below for the reader's convenience (the notation $|||L|||$ stands for the operator norm of any continuous linear map L between two Banach spaces):

Theorem 5. *Let X and Y be Banach spaces. Let $f : B_X(0, R) \to Y$ be continuous and Gâteaux-differentiable, with $f(0) = 0$. Assume that the derivative $Df(u)$ has a right-inverse $L(u)$, uniformly bounded on the ball $B_X(0, R)$:*

$$\forall (u, k) \in B_X(0, R) \times Y, \quad Df(u)L(u)k = k,$$
$$\sup\{|||L(u)||| : \|u\|_X < R\} < M.$$

Then, for every $v \in Y$ with $\|v\|_Y < RM^{-1}$ there is some $u \in X$ satisfying

$$f(u) = v \quad \text{and} \quad \|u\|_X \le M\|v\|_Y < R.$$

Note first that this is a local surjection theorem, not an inverse function theorem: compared to the IFT, we lose uniqueness. On the other hand, the regularity requirement on f and the smallness condition on v are much weaker. As mentioned in the Introduction, for a C^1 functional in finite dimensions, this theorem has been proved a long time ago by Ważewski [27] by a continuation argument (we thank Sotomayor for drawing our attention to this result). For a comparison of the existence and uniqueness domains in the C^2 case with $\dim X = \dim Y$, see [16, Chapter II, Exercise 2.3].

It turns out that the proof of Theorem 3 is much easier if one assumes that the family (DF_ε) is tame *Galerkin* right-invertible. But most applications require that (DF_ε) be tame right-invertible. Let us explain why the proof is longer in this case. In our algorithm, we will use two sequences of projectors $\Pi_n := \Pi(\Lambda_n)$ and $\Pi'_n := \Pi'(M_n)$ with associated ranges $E_n = E(\Lambda_n)$ and $E'_n = E'(M_n)$, where $\Lambda^0 \approx \varepsilon^{-\eta}$ for some small $\eta > 0$, $\Lambda_n = \Lambda_0^{\alpha^n}$ for some $\alpha > 1$ close to 1, and $M_n = \Lambda_n^\vartheta$ for some $\vartheta \le 1$ such that $\vartheta\alpha > 1$. The algorithm consists in finding, by induction on n and using Theorem 5 at each step, a solution $u_n \in E_n$ of the problem $\Pi'_n F_\varepsilon(u_n) = \Pi'_{n-1} v$. For this, we need $\Pi'_n DF_\varepsilon(u)_{|E_n}$ to be invertible for u in a certain ball $\mathcal{B}_n$, with estimates on the right-inverse for a certain norm $\|\cdot\|_{\mathcal{N}_n}$.

When the family (DF_ε) is tame Galerkin right-invertible, we can take $\vartheta = 1$ so that $M_n = \Lambda_n$, instead of assuming $\vartheta < 1$. Then the right-invertibility of $\Pi'_n DF_\varepsilon(u)_{|E_n}$ follows immediately from the definition.

But when (DF_ε) is only tame right-invertible, it is crucial to take $\vartheta < 1$. The intuitive idea is the following. One can think of $DF_\varepsilon(u)$ as a very large right-invertible matrix. The topological argument we use requires $\Pi'_n DF_\varepsilon(u)_{|E_n}$ to have a right-inverse for u in a suitable ball. If we take $M_n = \Lambda_n$, this is like asking that a square submatrix of a right-invertible matrix be invertible. In general this is not true. But a rectangular submatrix, with more columns than lines, will be right-invertible if the full matrix is and if there are enough columns in the submatrix. This is why we impose $M_n < \Lambda_n$ when we do not assume the tame Galerkin right-invertibility.

In what follows, we assume that the family (DF_ε) is tame right-invertible, so we take $\vartheta < 1$, and we point out the specific places where the arguments would be easier assuming, instead, that (DF_ε) is tame *Galerkin* right-invertible.

The sequence u_n depends on a number of parameters $\eta, \alpha, \beta, \vartheta$ and σ satisfying various conditions; in the first subsection we prove that these conditions are compatible. In the next one, we construct an initial point u_1 depending on η, α and ϑ. In the third one we construct, by induction, the remaining points u_n which also depend on β and σ. Finally, we prove that the sequence (u_n) converges to a solution u of the problem, satisfying the desired estimates.

3.1. *Choosing the values of the parameters*

We are given $s_1 \geq s_0 + \max\{m, \ell\}$, $\delta > s_1 + \ell'$ and $g' > g$. These are fixed throughout the proof.

We introduce positive parameters $\eta, \alpha, \beta, \vartheta$ and σ satisfying the following conditions:

$$\eta < \frac{g' - g}{\max\{\vartheta \ell', \ell\}}, \tag{3.1}$$

$$1/\alpha < \vartheta < 1, \tag{3.2}$$

$$(1 - \vartheta)(\sigma - \delta) > \vartheta m + \max\{\ell, \vartheta \ell'\} + g/\eta, \tag{3.3}$$

$$\sigma > \alpha\beta + s_1, \tag{3.4}$$

$$(1 + \alpha - \vartheta\alpha)(\sigma - s_0) > \alpha\beta + \alpha(m + \ell) + \ell' + g/\eta, \tag{3.5}$$

$$(1 - \vartheta)(\sigma - s_0) > m + \vartheta \ell' + \frac{g}{\alpha\eta}, \tag{3.6}$$

$$\delta > s_0 + \frac{\alpha}{\vartheta}(\sigma - s_0 - \alpha\beta + \ell''), \tag{3.7}$$

$$(\alpha - 1)\beta > (1 - \vartheta)(\sigma - s_0) + \vartheta m + \ell'' + g/\eta, \tag{3.8}$$

$$\ell'' = \max\{(\alpha - 1)\ell + \ell', \alpha\vartheta\ell'\}. \tag{3.9}$$

Note that condition (3.3) implies that $\delta < \sigma$. Note also that condition (3.7) may be rewritten as

$$\beta > \frac{1}{\alpha}(\sigma - \delta) + \left(1 - \frac{\vartheta}{\alpha}\right)\frac{\delta - s_0}{\alpha} + \frac{\ell''}{\alpha},$$

which implies the simpler inequality

$$\beta > \frac{1}{\alpha}(\sigma - \delta). \tag{3.10}$$

Inequality (3.10) will also be used in the proof.

If we assume tame Galerkin right-invertibility instead of tame right-invertibility, we can replace condition (3.3) by the weaker condition $\delta < \sigma$, we do not need conditions (3.5), (3.6) any more, and we can take $\vartheta = 1$ instead of $\vartheta < 1$.

Lemma 1. *The set of parameters* $(\eta, \alpha, \beta, \vartheta, \sigma)$ *satisfying the above conditions is non-empty. More precisely, there are some* $\alpha > 1$ *and* $\zeta > 0$ *depending only on* $(s_0, m, \ell, \ell', s_1, \delta)$ *such that, for* $\vartheta = \alpha^{-1/2}$ *and for every* $0 < \eta < 1$*, there exist* (β, σ) *with* $\sigma < \zeta g/\eta$ *such that the constraints* (3.3) *to* (3.9) *are satisfied.*

Proof. Since $\delta > s_1 + \ell'$, and $\ell'' \to \ell'$ when both α and ϑ tend to 1, it is possible to choose ϑ and $\alpha = \vartheta^{-2}$ close enough to 1 so that $\delta > s_0 + \frac{\alpha}{\vartheta}(s_1 - s_0 + \ell'')$. Take some τ with $0 < \tau < \frac{\vartheta}{\alpha}(\delta - s_0) - s_1 + s_0 - \ell''$, and set

$$\beta = \frac{\sigma}{\alpha} - \frac{s_1 + \tau}{\alpha}. \tag{3.11}$$

Then conditions (3.2), (3.4) and (3.7) are satisfied.

The remaining inequalities are constraints on β and σ. They can be rewritten as follows:

$$\sigma > \delta + \frac{1}{1-\vartheta}\left[\vartheta m + \max\{\ell, \vartheta\ell'\} + \frac{g}{\eta}\right], \tag{3.12}$$

$$\beta < \left(\frac{1}{\alpha} + 1 - \vartheta\right)\sigma - m - \ell - \frac{\ell'}{\alpha} - \left(\frac{1}{\alpha} + 1 - \vartheta\right)s_0 - \frac{g}{\alpha\eta}, \tag{3.13}$$

$$\sigma > s_0 + \frac{1}{1-\vartheta}\left(m + \vartheta\ell' + \frac{g}{\alpha\eta}\right), \tag{3.14}$$

$$\beta > \frac{1-\vartheta}{\alpha-1}\sigma + \frac{1}{\alpha-1}\left(\vartheta m + \ell'' + \frac{g}{\eta} - (1-\vartheta)s_0\right). \tag{3.15}$$

These inequalities define half-planes in the (σ, β)-plane. Since $\alpha\vartheta > 1$, the slopes in (3.11), (3.13) and (3.15) are ordered as follows:

$$0 < \frac{1-\vartheta}{\alpha-1} < \frac{1}{\alpha} < \frac{1}{\alpha} + 1 - \vartheta < 1.$$

As a consequence, for the chosen values of α, ϑ and τ, the domain defined by these three conditions in the (σ, β)-plane is an infinite half-line stretching to the North-East.

The remaining two, (3.12) and (3.14), just tell us that σ should be large enough. So the set of solutions is of the form $\sigma > \bar{\sigma}$, $\beta = \frac{\sigma}{\alpha} - \frac{s_1+\tau}{\alpha}$ and $\bar{\sigma}$ is clearly a piecewise affine function of g/η. We may thus choose $\sigma < \zeta g/\eta$ for some constant ζ. ∎

Remark. As already mentioned, if we assume that (DF_ε) is tame Galerkin right-invertible, (3.3) can be replaced by the condition $\delta < \sigma$, and (3.5) and (3.6) are not needed. The remaining conditions can be satisfied by taking $\vartheta = 1$ and for a larger set of the other parameters. The corresponding variant of Lemma 1 has a simpler proof. We can choose $\alpha > 1$ such that $\delta > s_0 + \alpha(s_1 - s_0 + \ell'')$ and τ such that $0 < \tau < \frac{1}{\alpha}(\delta - s_0) - s_1 + s_0 - \ell''$, and we may impose condition (3.11). Then conditions (3.12)–(3.14) are no longer required, and the last conditions $\delta < \sigma$ and (3.15) are easily satisfied by taking σ large enough.

The values $(\eta, \alpha, \beta, \vartheta, \sigma)$ are now fixed. For the remainder of the proof we introduce an important notation. By writing

$$x \lesssim y$$

we mean that there is some constant C such that $x \le Cy$. This constant depends on A_i, A'_i, a, b, s_0, m, ℓ, ℓ', g, g', s_1, δ and our additional parameters $(\eta, \alpha, \beta, \vartheta, \sigma)$, but NOT on ε, nor on the regularity index $s \in [0, S]$ or the rank n in any of the sequences which will be introduced below. For instance, the tame inequalities become

$$\|DF_\varepsilon(u)h\|_s \lesssim \|u\|_{s+m}\|h\|_{s_0+m} + \|h\|_{s+m},$$
$$\|L_\varepsilon(u)k\|_s \lesssim \varepsilon^{-g}(\|u\|_{s+l}\|k\|_{s_0+l'} + \|k\|_{s+l'}).$$

In the iteration process, we will need the following result:

Lemma 2. *If the maps F_ε form an S-tame differentiable family and $F_\varepsilon(0) = 0$, then, for $u \in B_{s_0+m} \cap V_{s+m}$ and $s_0 \le s \le S - m$, we have*

$$\|F_\varepsilon(u)\|_s \lesssim \|u\|_{s+m}.$$

Proof. Consider the function $\varphi(t) = \|F_\varepsilon(tu)\|_s$. Since F_ε is G-differentiable, we have

$$\varphi'(t) = \left\langle DF_\varepsilon(tu)u, \frac{F_\varepsilon(tu)}{\|F_\varepsilon(tu)\|_s} \right\rangle_s \le a(t\|u\|_{s_0+m}\|u\|_{s+m} + \|u\|_{s+m}),$$

and since $\varphi(0) = 0$, we get the result. ∎

3.2. Initialization

3.2.1. Defining appropriate norms. This subsection uses condition (3.2) and the inequalities $s_1 + \ell' < \delta < \sigma$, which, as already noted, follow from (3.3).

We are given $(\eta, \alpha, \vartheta, \delta, \sigma)$. We fix a large constant $K > 1$, to be chosen later independently of $0 < \varepsilon \le 1$.

We set $\Lambda_0 = (K\varepsilon^{-\eta})^{1/\alpha}$, $\Lambda_1 := (\Lambda_0)^\alpha = K\varepsilon^{-\eta}$, $M_0 := (\Lambda_0)^\vartheta = (K\varepsilon^{-\eta})^{\vartheta/\alpha}$ and $M_1 := (\Lambda_1)^\vartheta = (K\varepsilon^{-\eta})^\vartheta$. We then have the inequalities $M_0 < \Lambda_0 < M_1 < \Lambda_1$.

Let $E_1 := E(\Lambda_1)$, $\Pi_1 := \Pi(\Lambda_1)$, $E_1' = E(M_1)$ and $\Pi_i' := \Pi'(M_i)$ for $i = 0, 1$.

We choose the following norms on E_1 , E_1':

$$\|h\|_{\mathcal{N}_1} := \|h\|_\delta + \Lambda_1^{-\frac{\vartheta}{\alpha}(\sigma-\delta)}\|h\|_\sigma, \qquad \|k\|'_{\mathcal{N}_1} := \|k\|'_\delta + \Lambda_1^{-\frac{\vartheta}{\alpha}(\sigma-\delta)}\|k\|_\sigma.$$

Endowed with these norms, E_1 and E_1' are Banach spaces. We shall use the notation $|||L|||_{\mathcal{N}_1}$ for the operator norm of any continuous linear map L from the Banach space E_1' to either E_1 or E_1'.

The map F_ε induces a map $f_1 : B_{s_0+m} \cap E_1 \to E_1'$ defined by

$$f_1(u) := \Pi_1' F_\varepsilon(u)$$

for $u \in B_{s_0+m} \cap E_1$. Note that $f_1(0) = 0$. We will use the local surjection theorem to show that the range of f_1 covers a neighbourhood of 0 in E_1'. We begin by showing that Df_1 has a right-inverse.

Note that if we assume that DF is tame Galerkin right-invertible, we can take $M_1 = \Lambda_1 \geq \underline{\Lambda}$, and Df_1 is automatically right-invertible, with the tame estimate (2.8). *So the next subsection is only necessary if we assume that DF is tame right-invertible.*

3.2.2. $Df_1(u)$ has a right-inverse for $\|u\|_{\mathcal{N}_1} \leq 1$. This subsection uses condition (3.3). We recall it here for the reader's convenience:

$$(1-\vartheta)(\sigma-\delta) > \vartheta m + \max\{\ell, \vartheta\ell'\} + g/\eta.$$

Lemma 3. *For K large enough and for all $u \in E_1$ with $\|u\|_{\mathcal{N}_1} \leq 1$,*

$$|||\Pi_1' DF_\varepsilon(u)(1-\Pi_1)L_\varepsilon(u)|||_{\mathcal{N}_1} \leq 1/2.$$

Proof. From $\|u\|_{\mathcal{N}_1} \leq 1$, it follows that $\|u\|_\delta \leq 1$, and since $\delta > s_0 + \max\{\ell, m\} + \ell'$, the tame estimates hold at u.

Take any $k \in E_1'$ and set $h = (1-\Pi_1)L_\varepsilon(u)k$.

We have $\|h\|_\delta \lesssim \Lambda_1^{\delta-\sigma}\|L_\varepsilon(u)k\|_\sigma$, and

$$\|\Pi_1' DF_\varepsilon(u)h\|'_{\delta-m} \lesssim \|h\|_{s_0+m}\|u\|_\delta + \|h\|_\delta \lesssim \|h\|_\delta,$$
$$\|\Pi_1' DF_\varepsilon(u)h\|'_\delta \lesssim M_1^m \|\Pi_1' DF_\varepsilon(u)h\|_{\delta-m} \lesssim M_1^m\|h\|_\delta.$$

Hence

$$\|\Pi_1' DF_\varepsilon(u)h\|'_\delta \lesssim M_1^m \Lambda_1^{\delta-\sigma}\|L_\varepsilon(u)k\|_\sigma.$$

Writing $\|\Pi_1' DF_\varepsilon(u)h\|'_\sigma \lesssim M_1^{\sigma-\delta}\|\Pi_1' DF_\varepsilon(u)h\|_\delta$ we finally get

$$\|\Pi_1' DF_\varepsilon(u)h\|'_{\mathcal{N}_1} \lesssim M_1^m \Lambda_1^{\delta-\sigma}(1 + \Lambda_1^{-\frac{\vartheta}{\alpha}(\sigma-\delta)} M_1^{\sigma-\delta})\|L_\varepsilon(u)k\|_\sigma. \tag{3.16}$$

We now have to estimate $\|L_\varepsilon(u)k\|_\sigma$. By the tame estimates, we have

$$\|L_\varepsilon(u)k\|_\sigma \lesssim \varepsilon^{-g}(\|k\|'_{\sigma+\ell'} + \|u\|_{\sigma+\ell}\|k\|'_{s_0+\ell'}) \lesssim \varepsilon^{-g}(M_1^{\ell'}\|k\|'_\sigma + \Lambda_1^\ell\|u\|_\sigma\|k\|'_\delta).$$

Since $\|u\|_{\mathcal{N}_1} \le 1$, we have $\|u\|_\sigma \le \Lambda_1^{\frac{\vartheta}{\alpha}(\sigma-\delta)}$. Substituting, we get

$$\begin{aligned}\|L_\varepsilon(u)k\|_\sigma &\lesssim \varepsilon^{-g}(M_1^{\ell'}\|k\|'_\sigma + \Lambda_1^{\frac{\vartheta}{\alpha}(\sigma-\delta)+\ell}\|k\|'_\delta)\\ &\lesssim \varepsilon^{-g}\Lambda_1^{\frac{\vartheta}{\alpha}(\sigma-\delta)}(M_1^{\ell'} + \Lambda_1^\ell)\|k\|'_{\mathcal{N}_1}.\end{aligned} \tag{3.17}$$

Putting (3.16) and (3.17) together, we get

$$\|\Pi'_1 DF_\varepsilon(u)h\|'_{\mathcal{N}_1} \lesssim \varepsilon^{-g}M_1^m\Lambda_1^{\delta-\sigma}(\Lambda_1^{\frac{\vartheta}{\alpha}(\sigma-\delta)} + M_1^{\sigma-\delta})(M_1^{\ell'} + \Lambda_1^\ell)\|k\|'_{\mathcal{N}_1}$$

Since $\alpha > 1$, we have $\Lambda_1^{\frac{\vartheta}{\alpha}(\sigma-\delta)} \le \Lambda_1^{\vartheta(\sigma-\delta)} = M_1^{\sigma-\delta}$, so that

$$\begin{aligned}\|\Pi'_1 DF_\varepsilon(u)h\|'_{\mathcal{N}_1} &\lesssim \varepsilon^{-g}M_1^m\Lambda_1^{\delta-\sigma}M_1^{\sigma-\delta}(M_1^{\ell'} + \Lambda_1^\ell)\|k\|'_{\mathcal{N}_1}\\ &\lesssim \varepsilon^{-g}\Lambda_1^{\vartheta m-(1-\vartheta)(\sigma-\delta)+\max\{\ell,\vartheta\ell'\}}\|k\|'_{\mathcal{N}_1}.\end{aligned}$$

Since $\Lambda_1 = K\varepsilon^{-\eta}$, the inequality becomes

$$\|\Pi'_1 DF_\varepsilon(u)h\|'_{\mathcal{N}_1} \lesssim K^{-C_0}\varepsilon^{-g+\eta C_0}\|k\|'_{\mathcal{N}_1}$$

with $C_0 := (1-\vartheta)(\sigma-\delta) - \vartheta m - \max\{\ell, \vartheta\ell'\}$.

By condition (3.3), the exponent C_0 is larger than g/η, and the conclusion follows by choosing K large enough independently of $0 < \varepsilon \le 1$. ■

Introduce the map $\mathcal{L}_1(u) = \Pi_1 L_\varepsilon(u)_{|E'_1}$. Since $DF_\varepsilon(u)L_\varepsilon(u) = 1$, it follows from Lemma 3 that, for $k \in E'_1$, $u \in E_1$ and $\|u\|_{\mathcal{N}_1} \le 1$, we have

$$\|k - Df_1(u)\mathcal{L}_1(u)k\|'_{\mathcal{N}_1} \le \tfrac{1}{2}\|k\|'_{\mathcal{N}_1}$$

This implies that the Neumann series $\sum_{i\ge 0}(I_{E'_1} - Df_1(u)\mathcal{L}_1(u))^i$ converges in operator norm. Its sum is $S_1(u) = (Df_1(u)\mathcal{L}_1(u))^{-1}$ and it has operator norm at most 2.

Then $T_1(u) := \mathcal{L}_1(u)S_1(u)$ is a right-inverse of $Df_1(u)$ and $|||T_1(u)|||_{\mathcal{N}_1} \le 2|||\mathcal{L}_1(u)|||_{\mathcal{N}_1}$. By the tame estimates, if $u \in E_1$, $\|u\|_{\mathcal{N}_1} \le 1$ and $k \in E'_1$, we have

$$\begin{aligned}\|\mathcal{L}_1(u)k\|_\delta \lesssim \|L_\varepsilon(u)k\|_\delta &\lesssim \varepsilon^{-g}(\|k\|'_{\delta+\ell'} + \|u\|_{\delta+\ell}\|k\|'_{s_0+\ell'})\\ &\lesssim \varepsilon^{-g}(M_1^{\ell'} + \Lambda_1^\ell)\|k\|_\delta.\end{aligned}$$

Combining this with (3.17), we find

$$\sup_{\|u\|_{\mathcal{N}_1}\le 1} |||T_1(u)|||_{\mathcal{N}_1} \lesssim \varepsilon^{-g}(M_1^{\ell'} + \Lambda_1^\ell) = m_1.$$

3.2.3. Local inversion of f_1. Applying Theorem 5, we find that if $\|\Pi_0' v\|_{\mathcal{N}_1}' < 1/m_1$, then the equation $f_1(u) = \Pi_0' v$ has a solution $u_1 \in E_1$ with $\|u_1\|_{\mathcal{N}_1} \le 1$ and $\|u_1\|_{\mathcal{N}_1} \le m_1 \|\Pi_0' v\|_{\mathcal{N}_1}'$.

Note that $\|\Pi_0' v\|_\sigma' \lesssim M_0^{\sigma-\delta} \|\Pi_0' v\|_\delta' \lesssim \Lambda_1^{\frac{\vartheta}{\alpha}(\sigma-\delta)} \|\Pi_0' v\|_\delta'$. It follows that

$$\|\Pi_0' v\|_{\mathcal{N}_1}' = \|\Pi_0' v\|_\delta' + \Lambda_1^{-\frac{\vartheta}{\alpha}(\sigma-\delta)} \|\Pi_0' v\|_\sigma' \lesssim \|\Pi_0' v\|_\delta.$$

Assume from now on that

$$\|v\|_\delta' \lesssim \varepsilon^g (M_1^{\ell'} + \Lambda_1^{\ell})^{-1}. \tag{3.18}$$

Then $\|\Pi_0' v\|_{\mathcal{N}_1}' \lesssim m_1^{-1}$, and Theorem 5 applies. The estimate on u_1 implies

$$\|u_1\|_\delta \lesssim \varepsilon^{-g} (M_1^{\ell'} + \Lambda_1^{\ell}) \|v\|_\delta' \le 1. \tag{3.19}$$

It also implies an estimate in higher norm:

$$\|u_1\|_\sigma \lesssim \varepsilon^{-g} \Lambda_1^{\frac{\vartheta}{\alpha}(\sigma-\delta)} (M_1^{\ell'} + \Lambda_1^{\ell}) \|v\|_\delta' \lesssim \Lambda_1^{\frac{\vartheta}{\alpha}(\sigma-\delta)}. \tag{3.20}$$

3.3. *Induction*

3.3.1. Finding uniform bounds. In addition to $(\alpha, \vartheta, \delta, \varepsilon, \eta)$ we are given β satisfying relations (3.4) and (3.10) . We recall them here for the reader's convenience. With $s_1 \ge s_0 + \max\{m, \ell\}$ and $\delta > s_1 + \ell'$,

$$\sigma > \alpha\beta + s_1, \qquad \beta > \frac{1}{\alpha}(\sigma - \delta).$$

We also inherit $\Lambda_1 = K\varepsilon^{-\eta}$ and u_1 from the preceding section. Combining (3.10) and (3.20), we immediately obtain the estimate

$$\|u_1\|_\sigma \lesssim \varepsilon^{-g} \Lambda_1^{\beta} (M_1^{\ell'} + \Lambda_1^{\ell}) \|v\|_\delta' \lesssim \Lambda_1^{\beta}. \tag{3.21}$$

Consider the sequences of integers M_n and Λ_n, $n \ge 1$, defined by $\Lambda_n := \Lambda_1^{\alpha^{n-1}}$ and $M_n := \Lambda_n^{\vartheta}$. Let $\Pi_n := \Pi(\Lambda_n)$, $\Pi_n' := \Pi'(M_n)$, $E_n := E(\Lambda_n)$, $E_n' := E'(M_n)$.

We will construct a sequence $u_n \in E_n$, $n \ge 1$, starting from the initial point u_1 we found in the preceding section. For all $n \ge 2$ the remaining points should satisfy the following conditions:

$$\Pi_n' F_\varepsilon(u_n) = \Pi_{n-1}' v, \tag{3.22}$$

$$\|u_n - u_{n-1}\|_{s_0} \le \varepsilon^{-g} \Lambda_{n-1}^{\alpha\beta-\sigma+s_0} (M_1^{\ell'} + \Lambda_1^{\ell}) \|v\|_\delta', \tag{3.23}$$

$$\|u_n - u_{n-1}\|_\sigma \le \varepsilon^{-g} \Lambda_{n-1}^{\alpha\beta} (M_1^{\ell'} + \Lambda_1^{\ell}) \|v\|_\delta'. \tag{3.24}$$

We proceed by induction. Suppose we have found $u_2, \ldots, u_{n-1}$ satisfying these conditions. We want to construct u_n.

Lemma 4. *Suppose $K \geq 2$. For all t with $s_0 \leq t < \sigma - \alpha\beta$, and all i with $2 \leq i \leq n-1$, we have*

$$\sum_{i=2}^{n-1} \|u_i - u_{i-1}\|_t \leq \varepsilon^{-g} (M_1^{\ell'} + \Lambda_1^{\ell}) \Sigma(t) \|v\|'_\delta$$

where $\Sigma(t)$ is finite and independent of n, ε.

Proof. By the interpolation formula,

$$\|u_i - u_{i-1}\|_t \leq \varepsilon^{-g} \Lambda_{i-1}^{\alpha\beta-\sigma+t} (M_1^{\ell'} + \Lambda_1^{\ell}) \|v\|'_\delta$$

for all $2 \leq i \leq n$. Since $\Lambda_1 = K\varepsilon^{-\eta} \geq 2$, we have

$$\begin{aligned}\sum_{i=2}^{n-1} \|u_i - u_{i-1}\|_t &\leq \varepsilon^{-g} (M_1^{\ell'} + \Lambda_1^{\ell}) \sum_{i=2}^{\infty} \Lambda_{i-1}^{\alpha\beta-\sigma+t} \|v\|'_\delta \\ &\leq \varepsilon^{-g} (M_1^{\ell'} + \Lambda_1^{\ell}) \sum_{j=0}^{\infty} 2^{\alpha^{j(\alpha\beta-\sigma+t)}} \|v\|'_\delta. \end{aligned}$$ ∎

By (3.4) we can take $t = s_1$, and we find a uniform bound for u_{n-1} in the s_1-norm:

$$\begin{aligned}\|u_{n-1}\|_{s_1} &\leq \|u_1\|_\delta + \sum_{i=2}^{n-1} \|u_i - u_{i-1}\|_{s_1} \\ &\lesssim \varepsilon^{-g} (M_1^{\ell'} + \Lambda_1^{\ell})(1 + \Sigma(s_1)) \|v\|'_\delta \\ &\lesssim \varepsilon^{-g} (M_1^{\ell'} + \Lambda_1^{\ell}) \|v\|'_\delta.\end{aligned}$$

In particular, we will have $\|u_{n-1}\|_{s_1} \leq 1$ if $\|v\|'_\delta \lesssim \varepsilon^g (M_1^{\ell'} + \Lambda_1^{\ell})^{-1}$, so the tame estimates hold at u_{n-1}.

Similarly, if $\|v\|'_\delta \lesssim \varepsilon^g (M_1^{\ell'} + \Lambda_1^{\ell})^{-1}$ we find a uniform bound in the σ-norm. We have

$$\|u_{n-1}\|_\sigma \leq \|u_1\|_\sigma + \sum_{i=2}^{n-1} \|u_i - u_{i-1}\|_\sigma$$

and

$$\begin{aligned}\sum_{i=2}^{n-1} \|u_i - u_{i-1}\|_\sigma &\lesssim \varepsilon^{-g} (M_1^{\ell'} + \Lambda_1^{\ell}) \sum_{i=1}^{n-1} \Lambda_i^{\beta} \|v\|'_\delta \\ &\lesssim \varepsilon^{-g} (M_1^{\ell'} + \Lambda_1^{\ell}) \Lambda_{n-1}^{\beta} \|v\|'_\delta,\end{aligned}$$

so combining this with (3.21), we get

$$\|u_{n-1}\|_\sigma \lesssim \varepsilon^{-g} \Lambda_{n-1}^{\beta} (M_1^{\ell'} + \Lambda_1^{\ell}) \|v\|'_\delta \lesssim \Lambda_{n-1}^{\beta}. \tag{3.25}$$

3.3.2. Setting up the induction step. Suppose, as above, that $\|v\|'_\delta \lesssim \varepsilon^g (M_1^{\ell'} + \Lambda_1^\ell)^{-1}$and that $u_2, \dots, u_{n-1}$ have been found. We have seen that $\|u_{n-1}\|_{s_1} \leq 1$, so that the tame estimates hold at u_{n-1}, and we also have $\|u_{n-1}\|_\sigma \lesssim \Lambda_{n-1}^\beta$. We want to find u_n satisfying (3.22)–(3.24). Since $\Pi'_{n-1} F_\varepsilon(u_n) = \Pi'_{n-2} v$, we can write

$$\Pi'_n(F_\varepsilon(u_n) - F_\varepsilon(u_{n-1})) + (\Pi'_n - \Pi'_{n-1}) F_\varepsilon(u_{n-1}) = (\Pi'_{n-1} - \Pi'_{n-2}) v. \tag{3.26}$$

Define a map $f_n : E_n \to E'_n$ with $f_n(0) = 0$ by

$$f_n(z) = \Pi'_n(F_\varepsilon(u_{n-1} + z) - F_\varepsilon(u_{n-1})).$$

Equation (3.26) can be rewritten as follows:

$$f_n(z) = \Delta_n v + e_n, \tag{3.27}$$

$$\Delta_n v = \Pi'_{n-1}(1 - \Pi'_{n-2}) v, \tag{3.28}$$

$$e_n = -\Pi'_n(1 - \Pi'_{n-1}) F_\varepsilon(u_{n-1}). \tag{3.29}$$

We choose the following norms on E_n and E'_n:

$$\|x\|_{\mathcal{N}_n} = \|x\|_{s_0} + \Lambda_{n-1}^{-\sigma+s_0} \|x\|_\sigma, \quad \|y\|'_{\mathcal{N}_n} = \|y\|'_{s_0} + \Lambda_{n-1}^{-\sigma+s_0} \|y\|'_\sigma.$$

Endowed with these norms, E_n and E'_n are Banach spaces. We shall use $|||L|||_{\mathcal{N}_n}$ for the operator norm of any continuous linear map L from E'_n to either E_n or E'_n.

Lemma 5. *If $0 \leq t \leq \sigma - s_0$, then*

$$\|x\|_{s_0+t} \lesssim \Lambda_{n-1}^t \|x\|_{\mathcal{N}_n}, \quad \|y\|'_{s_0+t} \lesssim \Lambda_{n-1}^t \|y\|'_{\mathcal{N}_n}.$$

Proof. Use the interpolation inequality. ∎

We will solve the system (3.27)–(3.29) by applying the local surjection theorem to f_n on the ball $B_{\mathcal{N}_n}(0, r_n) \subset E_n$ where

$$r_n = \varepsilon^{-g} \Lambda_{n-1}^{\alpha\beta-\sigma+s_0} (M_1^{\ell'} + \Lambda_1^\ell) \|v\|'_\delta. \tag{3.30}$$

Note that if the solution z belongs to $B_{\mathcal{N}_n}(0, r_n)$, then

$$\|z\|_{s_0} \leq \varepsilon^{-g} \Lambda_{n-1}^{\alpha\beta-\sigma+s_0} (M_1^{\ell'} + \Lambda_1^\ell) \|v\|'_\delta \quad \text{and} \quad \|z\|_\sigma \leq \varepsilon^{-g} \Lambda_{n-1}^{\alpha\beta} (M_1^{\ell'} + \Lambda_1^\ell) \|v\|'_\delta.$$

In other words, $u_n = u_{n-1} + z$ satisfies (3.23) and (3.24), so that the induction step is proved.

We begin by showing that $Df_n(z)$ has a right-inverse.

Note that if we assume that DF_ε is tame Galerkin right-invertible, we can take $M_n = \Lambda_n$, and the result of the next subsection is obvious. This subsection is only useful if we assume that DF is tame right-invertible but not tame Galerkin right-invertible.

3.3.3. $Df_n(z)$ has a right-inverse for $\|z\|_{\mathcal{N}_n} \le r_n$. In this subsection, we use conditions (3.5) and (3.6). We recall them for the reader's convenience:

$$(1+\alpha-\vartheta\alpha)(\sigma-s_0) > \alpha\beta+\alpha(m+\ell)+\ell'+g/\eta,$$
$$(1-\vartheta)(\sigma-s_0) > m+\vartheta\ell'+\frac{g}{\alpha\eta}.$$

Take now any $z \in B_{\mathcal{N}_n}(0, r_n)$. Arguing as above, we find that

$$\|u_{n-1}+z\|_{s_1} \lesssim 1, \tag{3.31}$$
$$\|u_{n-1}+z\|_{\sigma} \lesssim \Lambda_n^{\beta}. \tag{3.32}$$

By (3.31) the tame estimates hold on $z \in B_{\mathcal{N}_n}(0, r_n)$.

Lemma 6. *Take $\Lambda_1 = K\varepsilon^{-\eta}$ with $K > 1$ chosen large enough, independently of n and $\varepsilon \in (0, 1]$. Then, for all $z \in B_{\mathcal{N}_n}(0, r_n)$,*

$$|||\Pi_n' DF_\varepsilon(u_{n-1}+z)(1-\Pi_n)L_\varepsilon(u_{n-1}+z)|||_{\mathcal{N}_n} \le 1/2.$$

Proof. We proceed as in the proof of Lemma 3. For $k \in E_n'$, we set

$$h = (1-\Pi_n)L_\varepsilon(u_{n-1}+z)k.$$

We have

$$\|h\|_{s_0+m} \lesssim \Lambda_n^{-\sigma+s_0+m}\|L_\varepsilon(u_{n-1}+z)k\|_\sigma.$$

By (3.32) and the tame estimates for L_ε, we get

$$\begin{aligned}\|L_\varepsilon(u_{n-1}+z)k\|_\sigma &\lesssim \varepsilon^{-g}(\|u_{n-1}+z\|_{\sigma+\ell}\|k\|'_{s_0+\ell'}+\|k\|'_{\sigma+\ell'})\\ &\lesssim \varepsilon^{-g}(\Lambda_n^{\beta+\ell}\Lambda_{n-1}^{\ell'}+M_n^{l'}\Lambda_{n-1}^{\sigma-s_0})\|k\|'_{\mathcal{N}_n},\end{aligned} \tag{3.33}$$

where we have used Lemma 5. Substituting this in the preceding formula, we get

$$\|h\|_{s_0+m} \lesssim \varepsilon^{-g}(\Lambda_n^{\beta+\ell-\sigma+s_0+m}\Lambda_{n-1}^{l'}+M_n^{\ell'}\Lambda_{n-1}^{-(\alpha-1)(\sigma-s_0)+\alpha m})\|k\|'_{\mathcal{N}_n}.$$

By the tame estimate (2.4), we have

$$\|\Pi_n' DF_\varepsilon(u_{n-1}+z)h\|'_{s_0} \lesssim \|h\|_{s_0+m}.$$

From this it follows that

$$\|\Pi_n' DF_\varepsilon(u_{n-1}+z)h\|'_{\sigma} \lesssim M_n^{\sigma-s_0}\|h\|_{s_0+m}.$$

Hence

$$\|\Pi_n' DF_\varepsilon(u_{n-1}+z)h\|'_{\mathcal{N}_n} \lesssim (1+\Lambda_{n-1}^{-\sigma+s_0}M_n^{\sigma-s_0})\|h\|_{s_0+m}.$$

We have $\Lambda_{n-1}^{-\sigma+s_0} M_n^{\sigma-s_0} \lesssim \Lambda_{n-1}^{(\alpha\vartheta-1)(\sigma-s_0)}$. Since $\alpha\vartheta > 1$, the dominant term in the parenthesis is the second one, and

$$\begin{aligned}\|\Pi_n' DF_\varepsilon(u_{n-1}+z)h\|_{\mathcal{N}_n} &\lesssim \Lambda_{n-1}^{-\sigma+s_0} M_n^{\sigma-s_0}\|h\|_{s_0+m} \\ &\lesssim \varepsilon^{-g} M_n^{\sigma-s_0}(\Lambda_{n-1}^{\alpha(\beta+\ell-\sigma+s_0+m)+\ell'-\sigma+s_0} + M_n^{\ell'}\Lambda_{n-1}^{-\alpha(\sigma-s_0)+\alpha m})\|k\|'_{\mathcal{N}_n}.\end{aligned}$$

From (3.5) and (3.6), it follows that the right-hand side is a decreasing function of n. To check that it is less than $1/2$ for all $n \geq 2$, it is enough to check it for $n = 2$. Since $\Lambda_1 = K\varepsilon^{-\eta}$, substituting this in the right-hand side we get

$$\|\Pi_n' DF_\varepsilon(u_{n-1}+z)h\|_{\mathcal{N}_n} \lesssim (K^{-\min\{C_1,C_2\}})^{\alpha^{n-2}}(\varepsilon^{\min\{C_1,C_2\}-\alpha^{2-n}g/\eta})^{\eta\alpha^{n-2}}\|k\|'_{\mathcal{N}_n}$$

with

$$\begin{aligned}C_1 &= -\alpha(\beta+\ell+m) - \ell' + (1+\alpha-\alpha\vartheta)(\sigma-s_0),\\ C_2 &= \alpha((1-\vartheta)(\sigma-s_0) - \vartheta\ell' - m).\end{aligned}$$

By (3.5) and (3.6), both exponents C_1 and C_2 are larger than g/η. As a consequence, $\|\Pi_n' DF_\varepsilon(u_{n-1}+z)h\|_{\mathcal{N}_n} \leq \frac{1}{2}\|k\|'_{\mathcal{N}_n}$ for K chosen large enough, independently of n and $0 < \varepsilon \leq 1$. ■

Define $\mathcal{L}_n(z) = \Pi_n L_\varepsilon(u_{n-1}+z)_{|E_n'}$. Arguing as in Subsection 3.2.2, we find that the Neumann series $\sum_{i\geq 0}(I_{E_n'} - Df_n(u)\mathcal{L}_n(u))^i$ converges in operator norm. Its sum is $S_n(u) = (Df_n(u)\mathcal{L}_n(u))^{-1}$ and it has operator norm at most 2. Then $T_n(u) := \mathcal{L}_n(u)S_n(u)$ is a right-inverse of $Df_n(u)$, with $|||T_n(u)|||_{\mathcal{N}_n} \leq 2|||\mathcal{L}_n(u)|||_{\mathcal{N}_n}$.

We have already derived estimate (3.33), which immediately implies

$$\|\mathcal{L}_n(z)k\|_\sigma \lesssim \varepsilon^{-g}\Lambda_{n-1}^{\sigma-s_0}(\Lambda_n^{\beta+\ell}\Lambda_{n-1}^{-\sigma+s_0+\ell'} + M_n^{\ell'})\|k\|'_{\mathcal{N}_n}.$$

From the tame estimates and Lemma 5, we also have

$$\|\mathcal{L}_n(z)k\|_{s_0} \lesssim \varepsilon^{-g}\|k\|'_{s_0+\ell'} \lesssim \varepsilon^{-g}\Lambda_{n-1}^{\ell'}\|k\|'_{\mathcal{N}_n}.$$

Since $\alpha\vartheta > 1$, we have $\Lambda_{n-1}^{\ell'} \lesssim M_n^{\ell'}$. So the preceding two estimates can be combined, and we get the final estimate for the right-inverse in operator norm:

$$|||T_n(z)|||_{\mathcal{N}_n} \lesssim \varepsilon^{-g}(\Lambda_n^{\beta+\ell}\Lambda_{n-1}^{-\sigma+s_0+\ell'} + M_n^{\ell'}). \tag{3.34}$$

3.3.4. Finding u_n. In this subsection, we use relations (3.4), (3.7), (3.8) and (3.9). We recall them for the reader's convenience:

$$\begin{aligned}&\sigma > \alpha\beta + s_1,\\ &\delta > s_0 + \frac{\alpha}{\vartheta}(\sigma - s_0 - \alpha\beta + \ell''),\\ &(\alpha-1)\beta > (1-\vartheta)(\sigma-s_0) + \vartheta m + \ell'' + g/\eta,\\ &\ell'' = \max\{(\alpha-1)\ell + \ell', \alpha\vartheta\ell'\}.\end{aligned}$$

Let us go back to (3.27). By Theorem 5, to solve $\Pi'_n f_n(z) = \Delta_n v + e_n$ with z in $B_{\mathcal{N}_n}(0, r_n)$ it is enough that

$$|||T_n(z)|||_{\mathcal{N}_n} (\|\Delta_n v\|_{\mathcal{N}_n} + \|e_n\|_{\mathcal{N}_n}) \le r_n. \tag{3.35}$$

Here r_n is given by (3.30). We can estimate $|||T_n(z)|||_{\mathcal{N}_n}$ using (3.34). We need to estimate $\|\Delta_n v\|_{\mathcal{N}_n}$ and $\|e_n\|_{\mathcal{N}_n}$.

From (3.28) we have

$$\begin{aligned} \|\Delta v\|'_{s_0} &\lesssim M_{n-2}^{s_0-\delta} \|v\|'_\delta, \\ \|\Delta v\|'_\sigma &\lesssim M_{n-1}^{\sigma-\delta} \|v\|'_\delta, \\ \|\Delta v\|'_{\mathcal{N}_n} &\lesssim \max\{M_{n-2}^{s_0-\delta}, \Lambda_{n-1}^{-\sigma+s_0} M_{n-1}^{\sigma-\delta}\} \|v\|'_\delta. \end{aligned}$$

An easy calculation yields

$$\sigma - s_0 - \vartheta(\sigma - \delta) + \frac{\vartheta}{\alpha}(s_0 - \delta) = (1 - \vartheta)(\sigma - \delta) + \left(1 - \frac{\vartheta}{\alpha}\right)(\delta - s_0).$$

Since $s_0 < \delta < \sigma$ and $\vartheta < 1 < \alpha$, the two terms on the right-hand side are positive, so $\Lambda_{n-1}^{-\sigma+s_0} M_{n-1}^{\sigma-\delta} \lesssim M_{n-2}^{s_0-\delta}$. It follows that

$$\|\Delta v\|'_{\mathcal{N}_n} \lesssim M_{n-2}^{s_0-\delta} \|v\|'_\delta. \tag{3.36}$$

From (3.29), we derive

$$\|e_n\|'_{s_0} \lesssim M_{n-1}^{-\sigma+m+s_0} \|e_n\|'_{\sigma-m}.$$

By Lemma 2, $\|F_\varepsilon(u_{n-1})\|_{\sigma-m} \lesssim \|u_{n-1}\|_\sigma$. So, remembering (3.29) and (3.25), we get

$$\|e_n\|'_{s_0} \lesssim M_{n-1}^{-\sigma+m+s_0} \|u_{n-1}\|_\sigma \lesssim \varepsilon^{-g} M_{n-1}^{-\sigma+m+s_0} \Lambda_{n-1}^\beta (M_1^{\ell'} + \Lambda_1^\ell) \|v\|'_\delta.$$

Similarly,

$$\|e_n\|'_\sigma \lesssim \|u_{n-1}\|_{\sigma+m} \lesssim \Lambda_{n-1}^m \|u_{n-1}\|_\sigma \lesssim \varepsilon^{-g} \Lambda_{n-1}^{\beta+m} (M_1^{\ell'} + \Lambda_1^\ell) \|v\|'_\delta.$$

Finally, since $M_{n-1} < \Lambda_{n-1}$ and $\sigma > m + s_0$, we get

$$\|e_n\|'_{\mathcal{N}_n} \lesssim \varepsilon^{-g} \Lambda_{n-1}^\beta M_{n-1}^{-\sigma+m+s_0} (M_1^{\ell'} + \Lambda_1^\ell) \|v\|'_\delta. \tag{3.37}$$

Substituting (3.34), (3.30), (3.36), (3.37) in (3.35), we get the following sufficient condition:

$$(\Lambda_n^{\beta+\ell} \Lambda_{n-1}^{-\sigma+s_0+\ell'} + M_n^{\ell'})(M_{n-2}^{s_0-\delta} + \varepsilon^{-g} \Lambda_{n-1}^\beta M_{n-1}^{-\sigma+m+s_0}) \lesssim \Lambda_{n-1}^{\alpha\beta-\sigma+s_0}. \tag{3.38}$$

We estimate both sides separately. Remembering that $M_{n-i} = (\Lambda_{n-1})^{\alpha^{1-i}\vartheta}$ and $\Lambda_{n-1} = (K\varepsilon^{-\eta})^{\alpha^{n-2}}$, we find

$$\begin{aligned} &(\Lambda_n^{\beta+\ell} \Lambda_{n-1}^{-\sigma+s_0+\ell'} + M_n^{\ell'})(M_{n-2}^{s_0-\delta} + \varepsilon^{-g} \Lambda_{n-1}^\beta M_{n-1}^{-\sigma+m+s_0}) \\ &\qquad \lesssim \left(\varepsilon^{-\eta\alpha^{n-2}}\right)^{\max\{C_3, C_4\} + \max\{C_5, C_6\}} \end{aligned}$$

and

$$\Lambda_{n-1}^{\alpha\beta-\sigma+s_0} \gtrsim \left(\varepsilon^{-\eta\alpha^{n-2}}\right)^{C_7}$$

with

$$\begin{aligned} C_3 &:= \alpha(\beta+\ell) - \sigma + s_0 + \ell', \\ C_4 &:= \alpha\vartheta\ell', \\ C_5 &:= \vartheta\alpha^{-1}(s_0 - \delta), \\ C_6 &:= g/\eta + \beta + \vartheta(-\sigma + m + s_0), \\ C_7 &:= \alpha\beta - \sigma + s_0. \end{aligned}$$

By (3.4), we have $\sigma - \alpha\beta > s_1 > s_0 + \max\{m, \ell\}$. It follows that

$$C_3 < (\alpha - 1)\ell + \ell'.$$

So, defining $\ell'' = \max\{(\alpha-1)\ell + \ell', \alpha\vartheta\ell'\}$ as in (3.9), we see that

$$\max\{C_3, C_4\} + \max\{C_5, C_6\} \le \max\{\ell'' + C_5, \ell'' + C_6\}.$$

So condition (3.38) is implied by the inequalities $\ell'' + C_5 < C_7$ and $\ell'' + C_6 < C_7$, which are the same as conditions (3.7) and (3.8). So inequality (3.35) holds, and the induction holds by Theorem 5.

3.4. End of proof

First of all, for the above construction to work, the only constraint on S is $S > \sigma$, and Lemma 1 gives us the estimate $\sigma < \zeta g/\eta$. The constant η is only constrained by condition (3.1), and we can choose, for instance, $\eta = \frac{g'-g}{2\max\{\vartheta\ell',\ell\}}$. So we only need a condition on S of the form $S \ge S_0$ with $S_0 = O\left(\frac{1}{g'-g}\right)$ as $g' \to g$, all the other parameters being fixed.

Let us now check that the estimate $\|v\|'_\delta \lesssim \varepsilon^{g'}$ is sufficient for the above construction. In (3.18) we made the assumption $\|v\|'_\delta \lesssim \varepsilon^g(\Lambda_1^\ell + M_1^{\ell'})^{-1}$ on v, and we have $M_1 \lesssim \varepsilon^{-\vartheta\eta}$, $\Lambda_1 \lesssim \varepsilon^{-\eta}$, hence $\Lambda_1^\ell + M_1^{\ell'} \lesssim \varepsilon^{-\eta\max\{\vartheta\ell',\ell\}}$. So the condition $\|v\|'_\delta \lesssim \varepsilon^{g+\eta\max\{\vartheta\ell',\ell\}}$ guarantees the existence of the sequence (u_n). But (3.1) may be rewritten in the form

$$g + \eta\max\{\vartheta\ell',\ell\} < g',$$

so the preceding condition is implied by the estimate $\|v\|'_\delta \lesssim \varepsilon^{g'}$, which is thus sufficient, as desired.

Now we can translate the symbol $\lesssim$ into more explicit estimates. Choosing $r > 0$ small enough, our construction gives, for every $v \in W_\delta$ with $\|v\|'_\delta \le r\varepsilon^{g'}$, a sequence u_n, $n \ge 1$, such that $u_n \in E_n$, $\|u_n\|_{s_1} \le r^{-1}\varepsilon^{-g'}\|v\|'_\delta \le 1$, and

$$\Pi'_n F_\varepsilon(u_n) = \Pi'_{n-1} v.$$

It follows from Lemma 4 that for any $t < \sigma - \alpha\beta$, (u_n) is a Cauchy sequence for $\|\cdot\|_t$. We recall that, by condition (3.4), $s_1 < \sigma - \alpha\beta$. So we can choose $t_1 \in (s_1, \sigma - \alpha\beta)$. Then (u_n) converges to some u_ε in V_{t_1} with $\|u_\varepsilon\|_{s_1} \le r^{-1}\varepsilon^{-g'}\|v\|'_\delta \le 1$.

Since $t_1 \ge s_0 + m$, the map F_ε is continuous from the t_1-norm to the $(t_1 - m)$-norm, so $F_\varepsilon(u_n)$ converges to $F_\varepsilon(u_\varepsilon)$ in W_{t_1-m}. Then $F_\varepsilon(u_n)$ is a bounded sequence in W_{t_1-m}, and $t_1 - m > s_0$. So, using the approximation estimate (2.2), we find that $\|(1 - \Pi'_n)F_\varepsilon(u_n)\|_{s_0} \to 0$, and finally $\|\Pi'_n F_\varepsilon(u_n) - F_\varepsilon(u_\varepsilon)\|_{s_0} \to 0$ as $n \to \infty$.

On the right-hand side, using (2.2) again, we find that $\Pi'_{n-1}v$ converges to v in W_{s_0}, since $\delta > s_0$.

We conclude that $F_\varepsilon(u_\varepsilon) = v$, as desired, and this ends the proof of Theorem 3.

4. An application of the singular perturbation theorem

4.1. The result

In this section, we consider a Cauchy problem for nonlinear Schrödinger systems arising in nonlinear optics, a question recently studied by Métivier–Rauch [21] and Texier–Zumbrun [26]. Métivier–Rauch proved the existence of local-in-time solutions, with an existence time T converging to 0 when the H^s norm of the initial datum goes to infinity. Texier–Zumbrun, thanks to their version of the Nash–Moser theorem adapted to singular perturbation problems, were able to find a uniform lower bound on T for certain highly concentrated initial data. The H^s norm of these initial data could go to infinity. By applying our "semiglobal" version of the Nash–Moser theorem, we are able to extend Texier–Zumbrun's result to even larger initial data. Below we follow closely their exposition, but some parameters are named differently to avoid confusion with other notations.

The problem takes the following form:

$$\begin{cases} \partial_t u + iA(\partial_x)u = B(u, \partial_x)u, \\ u(0, x) = \varepsilon^\kappa (a_\varepsilon(x), \bar{a}_\varepsilon(x)) \end{cases} \tag{4.1}$$

with $u(t, x) = (\psi(t, x), \bar{\psi}(t, x)) \in \mathbb{C}^{2n}$, $(t, x) \in [0, T] \times \mathbb{R}^d$,

$$A(\partial_x) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n, -\lambda_1, \ldots, -\lambda_n)\Delta_x$$

and

$$B = \begin{pmatrix} \mathcal{B} & \mathcal{C} \\ \bar{\mathcal{C}} & \bar{\mathcal{B}} \end{pmatrix}.$$

The coefficients $b_{jj'}$, $c_{jj'}$ of the $n \times n$ matrices $\mathcal{B}$, $\mathcal{C}$ are first-order operators with smooth coefficients: $b_{jj'} = \sum_{k=1}^d b_{kjj'}(u)\partial_{x_k}$, $c_{jj'} = \sum_{k=1}^d c_{kjj'}(u)\partial_{x_k}$, with $b_{kjj'}$ and $c_{kjj'}$ smooth complex-valued functions of u satisfying, for some integer $p \ge 2$, some $C > 0$, all $0 \le |\alpha| \le p$ and all $u = (\psi, \bar{\psi}) \in \mathbb{C}^{2n}$,

$$|\partial^\alpha b_{kjj'}(u)| + |\partial^\alpha c_{kjj'}(u)| \le C|u|^{p-|\alpha|}.$$

Moreover, we assume that the following “transparency” conditions hold: the functions b_{kjj} are real-valued, the coefficients λ_j are real and pairwise distinct, and $c_{jj'} = c_{j'j}$ for any $j,\ j'$ such that $\lambda_j + \lambda_{j'} = 0$.

We consider initial data of the form $\varepsilon^\kappa(a_\varepsilon(x), \bar{a}_\varepsilon(x))$ with $a_\varepsilon(x) = a_1(x/\varepsilon)$ where $0 < \varepsilon \leq 1, a_1 \in H^S(\mathbb{R}^d)$ for some S large enough and $\|a_1\|_{H^S}$ small enough.

Our goal is to prove that the Cauchy problem has a solution on $[0, T] \times \mathbb{R}^d$ for all $0 < \varepsilon \leq 1$, with $T > 0$ independent of ε. Texier–Zumbrun obtain existence and uniqueness of the solution, under some conditions on κ, which should be large enough. This corresponds to a smallness condition on the initial datum when ε approaches zero. Our local surjection theorem only provides existence, but our condition on κ is less restrictive, so our initial datum is allowed to be larger. Note that once existence is proved, uniqueness is easily obtained for this Cauchy problem; indeed, local-in-time uniqueness implies global-in-time uniqueness. Our result is the following:

Theorem 6. *Under the above assumptions and notations, suppose additionally that*

$$\kappa > \frac{d}{2(p-1)}. \tag{4.2}$$

Let $s_1 > d/2 + 4$. If $0 < \varepsilon \leq 1$, $a_1 \in H^S(\mathbb{R}^d)$ for S large enough, and $\|a_1\|_{H^S}$ is small enough, then the Cauchy problem (4.1) *has a unique solution in the function space $C^1([0, T],\ H^{s_1-2}(\mathbb{R}^d)) \cap C^0([0, T],\ H^{s_1}(\mathbb{R}^d))$.*

Métivier–Rauch already provide existence for a fixed positive T when $\kappa \geq 1$. So we obtain something new in comparison with them when $\frac{d}{2}\frac{1}{p-1} < 1$, that is, when

$$p > 1 + d/2.$$

Let us now compare our results with those of Texier–Zumbrun [26]. In order to do so, we consider the same particular values as in their Remark 4.7 and Examples 4.8, 4.9 pages 517–518. Let us illustrate this in two and three space dimensions.

In two space dimensions, $d = 2$ [26, Example 4.8], our condition becomes $\frac{1}{p-1} < \kappa$, while Texier and Zumbrun need the stronger condition $\frac{9}{2(p+1)} < \kappa$.

In three space dimensions, $d = 3$ [26, Example 4.9], our condition becomes $\frac{3}{2(p-1)} < \kappa$, while they need the stronger condition $\frac{4}{p+1} < \kappa$.

In both cases, we improve over Métivier–Rauch when $p \geq 3$, while Texier–Zumbrun need $p \geq 4$.

Remark. After reading our paper, Baldi and Haus [6] have been able to relax the condition on κ even further, based on their version [5] of the classical Newton scheme in the spirit of Hörmander. A key point in their proof is a clever modification of the norms considered by Texier–Zumbrun, allowing better C^2 estimates on the functional. They also explain that their approach can be extended to other C^2 functionals consisting of a linear term perturbed by a nonlinear term of homogeneity at least $p + 1$. Our abstract theorem, however, seems more general since we do not need such a structure.

4.2. Proof of Theorem 6

We have to show that our Corollary 4 applies. Our functional setting is the same as in [26], with slightly different notations.

We introduce the norm $\|f\|_{H^s_\varepsilon(\mathbb{R}^d)} = \|(-\varepsilon^2\Delta + 1)^{s/2} f\|_{L^2(\mathbb{R}^d)}$, and we take

$$V_s = \mathcal{C}^1([0,T], H^{s-2}(\mathbb{R}^d)) \cap \mathcal{C}^0([0,T], H^s(\mathbb{R}^d)),$$
$$|u|_s = \sup_{0\le t\le T} \{\|\varepsilon^2 \partial_t u(t,\cdot)\|_{H^{s-2}_\varepsilon(\mathbb{R}^d)} + \|u(t,\cdot)\|_{H^s_\varepsilon(\mathbb{R}^d)}\},$$

and

$$W_s = \mathcal{C}^0([0,T], H^s(\mathbb{R}^d)) \times H^{s+2}(\mathbb{R}^d),$$
$$|(v_1,v_2)|'_s = \sup_{0\le t\le T} \{\|v_1(t,\cdot)\|_{H^s_\varepsilon(\mathbb{R}^d)}\} + \|v_2\|_{H^{s+2}_\varepsilon(\mathbb{R}^d)}.$$

Our projectors are

$$\Pi_\Lambda u = \mathcal{F}_x^{-1}(1_{|\varepsilon\xi|\le\Lambda} \mathcal{F}_x u(t,\xi)),$$
$$\Pi'_\Lambda(v_1,v_2) = \big(\mathcal{F}_x^{-1}(1_{|\varepsilon\xi|\le\Lambda} \mathcal{F}_x v_1(t,\xi)), \mathcal{F}^{-1}(1_{|\varepsilon\xi|\le\Lambda} \mathcal{F} v_2(\xi))\big).$$

We take

$$\Phi_\varepsilon(u) = \big(\varepsilon^2\partial_t u + iA(\varepsilon\partial_x)u - \varepsilon B(u,\varepsilon\partial_x)u,\ u(0,\cdot) - \varepsilon^\kappa(a_\varepsilon, \bar{a}_\varepsilon)\big)$$

and

$$\mathfrak{a}_\varepsilon(t,x) = \varepsilon^\kappa\big(\exp(-itA(\partial_x))a_\varepsilon, \exp(itA(\partial_x))\bar{a}_\varepsilon\big).$$

We have $\Phi_\varepsilon(\mathfrak{a}_\varepsilon) = (-\varepsilon B(\mathfrak{a}_\varepsilon, \varepsilon\partial_x)\mathfrak{a}_\varepsilon, 0)$. A solution of the functional equation $\Phi_\varepsilon(u) = 0$ is a solution on $[0,T]\times\mathbb{R}^d$ of the Cauchy problem (4.1).

Our Corollary 4 requires a direct estimate (2.9) on $D\Phi_\varepsilon$ and an estimate (2.11) on the right-inverse L_ε.

Take $s_0 > d/2 + 2$, $m = 2$, $\gamma = \frac{dp}{2(p-1)}$ and S large. Since $\kappa > \frac{d}{2(p-1)}$, we have an estimate of the form $|\mathfrak{a}_\varepsilon|_S \lesssim \varepsilon^\gamma \|a_1\|_{H^S}$, so, taking $\|a_1\|_{H^S}$ small, we can ensure that $\mathfrak{a}_\varepsilon \in \mathfrak{B}_S(\varepsilon^\gamma)$. Moreover the inequality $\kappa > \frac{d}{2(p-1)}$ implies the condition

$$1 - dp/2 + p\gamma \ge 0.$$

So we see that the assumptions of [26, Lemma 4.4] are satisfied by the parameters $\gamma_0 = \gamma_1 = \gamma$ (note that our exponent p is denoted ℓ in [26]). The direct estimate (2.9) thus follows from [26, Lemma 4.4]. Note that [26, Lemma 4.4] also gives an estimate on the second derivative of $\Phi_\varepsilon(\cdot)$, but we do not need such an estimate.

Choosing, in addition, $\ell = 2$, $\ell' = 0$, $g = 2$, our inverse estimate (2.11) follows from from [26, Lemma 4.5].

To summarize, the assumptions (2.9)–(2.11) of our Corollary 4 are satisfied for $s_0 > d/2$, $m = 2$, $\gamma = \frac{p}{p-1}\frac{d}{2}$, $g = 2$, $\ell = 2$, $\ell' = 0$.

Moreover, in [26, proof of Theorem 4.6] one finds an estimate which can be written in the form

$$|\Phi_\varepsilon(\mathfrak{a}_\varepsilon)|'_{s_1-1} \le r\varepsilon^{1+\kappa(p+1)+d/2}$$

where r is small when $\|a_1\|_{H^{s_1}}$ is small.

So, using our Corollary 4, taking S large enough we can solve the equation $\Phi_\varepsilon(u) = 0$ in X_{s_1} under the additional condition $1 + \kappa(p+1) + d/2 > \gamma + g$, which can be rewritten as follows:

$$\kappa > \frac{1}{p+1} + \frac{d}{2(p+1)(p-1)}.$$

Since $d \ge 2$, this inequality is a consequence of our assumption $\kappa \ge \frac{d}{2(p-1)}$.

So our Corollary 4 implies the existence of a solution to the Cauchy problem (4.1). The uniqueness of this solution comes from the local-in-time uniqueness of solutions to the Cauchy problem. This proves Theorem 6 as a consequence of Corollary 4.

Remark. In [26, Examples 4.8 and 4.9], Texier and Zumbrun also study the case of oscillating initial data, i.e. $a_\varepsilon = a(x)e^{ix\cdot\xi_0/\varepsilon}$, and in the first submitted version of this paper we considered it as well. However, a referee pointed out to us that the corresponding statements were not fully justified in [26]. Indeed, in the proof of their Theorem 4.6, Texier and Zumbrun have to invert the linearized functional $D\Phi_\varepsilon(u)$ for u in a neighbourhood of the function $\mathfrak{a}_\varepsilon$, denoted a_f in their paper. For this purpose, it seems that they need the norm of their function a_f to be controlled by ε^γ. This condition appears in their Remark 2.14 and their Lemma 4.5, but not in the statement of their Theorem 4.6. This additional constraint does not affect their results for concentrating initial data in Examples 4.8, 4.9. But in the oscillating case, their statements seem overly optimistic. We did not want to investigate that issue, and that is why we only deal with the concentrating case. Note, however, that this difficulty with the oscillating case is overcome in the recent work [6], thanks to improved norms and estimates.

5. Conclusion

The purpose of this paper has been to introduce a new algorithm into the “hard” inverse function theorem, where both $DF(u)$ and its right-inverse $L(u)$ lose derivatives, in order to improve its range of validity. To highlight this improvement, we have considered singular perturbation problems with loss of derivatives. We have shown that, on the specific example of a Schrödinger-type system of PDEs arising from nonlinear optics, our method leads to substantial improvements of known results. We believe that our approach has the potential of improving the known estimates in many other “hard” inversion problems.

In the statement and proof of our abstract theorem, our main focus has been the existence of u solving $F(u) = v$ in the case when S is large and the regularity of v is as small as possible. We have not tried to give an explicit bound on S, but with some additional work, it can be done. In an earlier version [13] of this paper, the reader will find a study of the intermediate case of a tame Galerkin right-invertible differential DF, with precise estimates on the parameter S depending on the loss of regularity of the right-inverse, in the special case $s_0 = m = 0$ and $\ell = \ell'$.

Acknowledgements. We are grateful to Massimiliano Berti, Philippe Bolle, Jacques Fejoz and Louis Nirenberg for their interest in our work and their encouragements. It is a pleasure to thank Pietro Baldi for stimulating discussions in Naples and Paris, for a careful reading of the present paper and for a number of suggestions. We also thank the referees, whose remarks have helped us to improve this manuscript.

References

[1] Alinhac, S., Gérard, P.: Opérateurs pseudo-différentiels et théorème de Nash–Moser. InterEditions, Paris; Éditions du CNRS, Meudon (1991) Zbl 0791.47044 MR 1172111

[2] Arnol'd, V. I.: Small denominators. I. Mapping the circle onto itself. Izv. Akad. Nauk SSSR **25**, 21–86 (1961) (in Russian) Zbl 0152.41905 MR 0140699

[3] Arnol'd, V. I.: Some remarks on flows of line elements and frames. Dokl. Akad. Nauk SSSR **138**, 255–257 (1961) (in Russian) Zbl 0124.14601 MR 0158330

[4] Arnol'd, V. I.: Small denominators and problems of stability of motion in classical and celestial mechanics. Uspekhi Mat. Nauk **18**, no. 6, 91–192 (1963) (in Russian) Zbl 0135.42701 MR 0170705

[5] Baldi, P., Haus, E.: A Nash–Moser–Hörmander implicit function theorem with applications to control and Cauchy problems for PDEs. J. Funct. Anal. **273**, 3875–3900 (2017) Zbl 06792303 MR 3711883

[6] Baldi, P., Haus, E.: Size of data in implicit function problems and singular perturbations for nonlinear Schrödinger systems, arXiv:1906.12290 (2019)

[7] Berti, M., Bolle, P.: Cantor families of periodic solutions for completely resonant nonlinear wave equations. Duke Math. J. **134**, 359–419 (2006) Zbl 1103.35077 MR 2248834

[8] Berti, M., Bolle, P.: Cantor families of periodic solutions of wave equations with C^k nonlinearities. Nonlinear Differential Equations Appl. **15**, 247–276 (2008) Zbl 1155.35004 MR 2408354

[9] Berti, M., Bolle, P.: Sobolev periodic solutions of nonlinear wave equations in higher spatial dimensions. Arch. Ration. Mech. Anal. **195**, 609–642 (2010) Zbl 1186.35113 MR 2592290

[10] Berti, M., Bolle, P., Procesi, M.: An abstract Nash–Moser theorem with parameters and applications to PDEs. Ann. Inst. H. Poincaré Anal. Non Linéaire **27**, 377–399 (2010) Zbl 1203.47038 MR 2580515

[11] Ciarlet, P. G.: Linear and Nonlinear Functional Analysis with Applications. SIAM, Philadelphia, PA (2013) Zbl 1293.46001 MR 3136903

[12] Ekeland, I.: An inverse function theorem in Fréchet spaces. Ann. Inst. H. Poincaré Anal. Non Linéaire **28**, 91–105 (2011) Zbl 1256.47037 MR 2765512

[13] Ekeland, I., Séré, E.: An implicit function theorem for non-smooth maps between Fréchet spaces. arXiv:1502.01561 (2015)

[14] Günther, M.: Isometric embeddings of Riemannian manifolds. In: Proc. International Congress of Mathematicians (Kyoto, 1990), Vol. II, Math. Soc. Japan, Tokyo, 1137–1143 (1991) Zbl 0745.53031 MR 1159298

[15] Hamilton, R. S.: The inverse function theorem of Nash and Moser. Bull. Amer. Math. Soc. (N.S.) **7**, 65–222 (1982) Zbl 0499.58003 MR 656198

[16] Hartman, P.: Ordinary Differential Equations. Wiley, New York (1964) Zbl 0125.32102 MR 0171038

[17] Hörmander, L.: The boundary problems of physical geodesy. Arch. Ration. Mech. Anal. **62**, 1–52 (1976) Zbl 0331.35020 MR 602181

[18] John, F.: On quasi-isometric mappings. I. Comm. Pure Appl. Math. **21**, 77–110 (1968) Zbl 0157.45803 MR 222666

[19] Kolmogorov, A. N.: On conservation of conditionally periodic motions for a small change in Hamilton's function. Dokl. Akad. Nauk SSSR (N.S.) **98**, 527–530 (1954) (in Russian) Zbl 0056.31502 MR 0068687

[20] Krantz, S. G., Parks, H. R.: The Implicit Function Theorem: History, Theory, and Applications. Birkhäuser Boston, Boston, MA (2002) Zbl 1269.58003 MR 2977424

[21] Métivier, G., Rauch, J.: Dispersive stabilization. Bull. London Math. Soc. **42**, 250–262 (2010) Zbl 1192.35165 MR 2601551

[22] Moser, J.: A new technique for the construction of solutions of nonlinear differential equations. Proc. Nat. Acad. Sci. U.S.A. **47**, 1824–1831 (1961) Zbl 0104.30503 MR 132859

[23] Nash, J.: The imbedding problem for Riemannian manifolds. Ann. of Math. (2) **63**, 20–63 (1956) Zbl 0217.47504 MR 75639

[24] Schwartz, J.: Nonlinear functional analysis. Courant Institute lecture notes

[25] Sotomayor, J.: Inversion of smooth mappings. Z. Angew. Math. Phys. **41**, 306–310 (1990) Zbl 0706.58005 MR 1045818

[26] Texier, B., Zumbrun, K.: Nash–Moser iteration and singular perturbations. Ann. Inst. H. Poincaré Anal. Non Linéaire **28**, 499–527 (2011) Zbl 1237.47066 MR 2823882

[27] Ważewski, T.: Sur l'évaluation du domaine d'existence des fonctions implicites réelles ou complexes. Ann. Soc. Polon. Math. **20**, 81–120 (1947) Zbl 0032.05601 MR 0026100