Ann. Inst. H. Poincaré
Anal. Non Linéaire 40 (2023), 1051–1091
DOI 10.4171/AIHPC/56

Size of data in implicit function problems and singular perturbations for nonlinear Schrödinger systems

Pietro Baldi and Emanuele Haus

Abstract. We investigate a general question about the size and regularity of the data and the solutions in implicit function problems with loss of regularity. First, we give a heuristic explanation of the fact that the optimal data size found by Ekeland and Séré with their recent nonquadratic version of the Nash–Moser theorem can also be recovered, for a large class of nonlinear problems, with quadratic schemes. Then we prove that this heuristic observation applies to the singular perturbation Cauchy problem for the nonlinear Schrödinger system studied by Métivier, Rauch, Texier, Zumbrun, Ekeland, and Séré. Using a "free flow component" decomposition and applying an abstract Nash–Moser–Hörmander theorem, we improve the existing results regarding both the size of the data and the regularity of the solutions.

1. Introduction

This paper is motivated by a general question concerning the size and regularity of the data and the solutions in implicit function problems with loss of regularity. In the recent work [4], Ekeland and Séré introduce a new iteration scheme in Banach spaces for solving nonlinear functional equations of the form

$$F(u) = v,$$

where the linearized operator $F'(u)$ admits a right inverse that loses derivatives. In such situations, a well-established strategy for constructing a solution u consists in applying a Nash–Moser iteration, essentially based on a quadratic Newton scheme combined with smoothing operators. The scheme in [4] differs from the standard Nash–Moser approach in that it is not quadratic, and it consists in solving a sequence of Galerkin problems by a topological argument (Ekeland's variational principle). This gives two main improvements with respect to the standard quadratic approach: the map F need not be twice differentiable, and a larger ball for the datum v is covered.

The first point of the present paper is the observation that, for operators of the form

$$F(u) = Lu + \mathcal{N}(u),$$

2020 Mathematics Subject Classification. Primary 47J07; Secondary 35B25, 35Q55, 35L52.
Keywords. Nash–Moser theorem, nonlinear Schrödinger equation, nonlinear Schrödinger system, Cauchy problem, singular perturbation, size of data.

where L is linear and $\mathcal{N}(u) = O(\|u\|^\alpha)$ for some $\alpha > 1$ in a ball $\|u\| \leq R$, the same size of ball for the datum v as in [4] can also be obtained by quadratic Nash–Moser schemes. In Section 2 we explain the heuristics behind this simple, general observation.

In Sections 3–6 we consider the singular perturbation Cauchy problem for the nonlinear Schrödinger system studied by Métivier and Rauch [10], Texier and Zumbrun [11], and Ekeland and Séré [4], and we rigorously prove that the observation of Section 2 applies to this PDE problem. The result of Sections 3–6 is stated in Theorem 3.4, which improves the results in [4, 11] regarding the size of the data and also the regularity of the solution: for initial data in a Sobolev space $H^s(\mathbb{R}^d)$ we prove that the solution of the Cauchy problem belongs to $C([0,T], H^s(\mathbb{R}^d))$ with the *same* regularity s, as is expected, and we give the corresponding estimate for the solution in terms of its initial datum. For initial data of a special "concentrating" form, see (3.5), Theorem 3.4 also improves the size of the ball for the data with respect to [4, 11]; see Remark 3.7.

For initial data of the other special form considered in [11] ("fast oscillating" data; see (3.5)), we improve the size of initial data in Theorem 3.5, which is proved in Sections 7–8. With respect to Theorem 3.4, the new ingredient is a "free flow decomposition" of the unknown, which is a natural way of exploiting the interplay between the linear and nonlinear parts of the system and the better L^∞ embedding properties of concentrating or highly oscillating free flows (see Lemma 7.2), inspired by the "shifted map" trick of [11]. The price to pay for this improvement in the size of data is a loss of *one* derivative: for data in $H^s(\mathbb{R}^d)$, the solution belongs to $C([0,T], H^{s-1}(\mathbb{R}^d))$. Theorem 3.5 improves the results of [4, 11] regarding both the regularity of the solution and the size of the data; see Remark 3.7.

We point out that the loss of regularity in Theorem 3.5 is *not* due to the Nash–Moser iteration: the loss of one derivative is introduced when solving the linearized Cauchy problem as a triangular system (see (7.14)) in two components, which are the "free flow" component of the unknown and its correction – the Nash–Moser–Hörmander Theorem A.1 just replicates the loss of one derivative for the nonlinear problem, without introducing additional losses. The loss of regularity in Theorem 3.5 equals exactly the number of derivatives in the nonlinearity, which is 1 in system (3.1).

The main difference between our "free flow decomposition" and the "shifted map" trick of [11] is that we treat the free flow as an unknown, although it is already completely determined by the initial datum of the problem. In this way, Theorem A.1 regularizes the free flow, introducing just one new dyadic Fourier packet at each step of the iteration. This is the key ingredient for preserving the regularity of the linearized problem in the nonlinear one, and it is somewhat reminiscent of a similar idea in Hörmander [5].

Technical details of the fact that the heuristic observation of Section 2 rigorously applies to Theorems 3.4 and 3.5 are contained in Remarks 6.1 and 8.1. Other general observations about the optimization of the data size in Nash–Moser schemes are in Remarks 7.3 and 7.4.

2. Large radius with quadratic schemes: an informal explanation

Consider a nonlinear problem of the kind

$$F(u) = v,$$

where v is given, u is the unknown, and F is a twice differentiable nonlinear operator in some Banach spaces satisfying $F(0) = 0$. Assume that for all u in a ball $\|u\| \leq R$ the linearized operator $F'(u)$ admits a right inverse $\Psi(u)$ satisfying

$$\|\Psi(u)h\| \leq A\|h\| \quad \text{for all } \|u\| \leq R, \tag{2.1}$$

and the second derivative $F''(u)$ satisfies

$$\|F''(u)[h,w]\| \leq B\|h\|\|w\| \quad \text{for all } \|u\| \leq R \tag{2.2}$$

(in this discussion we ignore completely the questions about loss of derivatives, and we only care about size). As explained in [4], the quadratic Newton scheme gives a solution u of the equation $F(u) = v$ for all v of size

$$\|v\| \lesssim \min\left\{\frac{1}{A^2 B}, \frac{R}{A}\right\},$$

while, with topological arguments, one can prove the existence of a solution u for all v in the larger ball

$$\|v\| \lesssim \frac{R}{A}.$$

Our observation is that, for operators F in some large class, the two radii are of the same order.

Indeed, assume that F is given by the sum of a linear part $\mathscr{L}$ and a nonlinear one $\mathscr{N}$,

$$F(u) = \mathscr{L}u + \mathscr{N}(u).$$

Assume that $\mathscr{N}$ satisfies

$$\|\mathscr{N}(u)\| \lesssim \|u\|^{p+1},$$

$$\|\mathscr{N}'(u)h\| \lesssim \|u\|^{p}\|h\|, \tag{2.3}$$

$$\|\mathscr{N}''(u)[h,w]\| \lesssim \|u\|^{p-1}\|h\|\,\|w\| \tag{2.4}$$

for some $p \geq 1$, for all u in the ball $\|u\| \leq 1$, so that

$$\|F''(u)[h,w]\| \lesssim \|u\|^{p-1}\,\|h\|\,\|w\|.$$

Suppose that $\mathscr{L}$ has a right inverse $\mathscr{L}_r^{-1}$ (namely $\mathscr{L}\mathscr{L}_r^{-1} = I$) and that

$$\|\mathscr{L}_r^{-1}\mathscr{N}'(u)\| \leq \frac{1}{2} \tag{2.5}$$

for u sufficiently small, say $\|u\| \le R$, so that, by Neumann series, the linearized operator

$$F'(u) = \mathcal{L} + \mathcal{N}'(u) = \mathcal{L}(I + \mathcal{L}_r^{-1}\mathcal{N}'(u))$$

has the right inverse

$$\Psi(u) = (I + \mathcal{L}_r^{-1}\mathcal{N}'(u))^{-1}\mathcal{L}_r^{-1},$$

with

$$\|\Psi(u)\| \le 2\|\mathcal{L}_r^{-1}\|.$$

Hence (2.1) holds with

$$A := 2\|\mathcal{L}_r^{-1}\|.$$

What is the "intrinsic" size of R? By (2.3), condition (2.5) holds for

$$\|\mathcal{L}_r^{-1}\|\,\|u\|^p \le \frac{1}{2}, \quad \text{i.e. } \|u\| \le \Big(\frac{1}{2\|\mathcal{L}_r^{-1}\|}\Big)^{\frac{1}{p}},$$

and therefore we fix

$$R := \Big(\frac{1}{2\|\mathcal{L}_r^{-1}\|}\Big)^{\frac{1}{p}} = A^{-\frac{1}{p}}. \tag{2.6}$$

Moreover, by (2.4), condition (2.2) holds with

$$B := R^{p-1} = A^{-1+\frac{1}{p}}.$$

Thus

$$\frac{1}{A^2 B} = A^{-1-\frac{1}{p}}, \qquad \frac{R}{A} = A^{-1-\frac{1}{p}},$$

namely the two balls have the same size.

Remark 2.1. Even when $\mathcal{L}_r^{-1}\mathcal{N}'(u)$ is an unbounded operator, so that the right invertibility of $F'(u)$ cannot be directly obtained by Neumann series, the heuristic argument above still catches the right size of R, provided that the invertibility of $F'(u)$ is obtained by a perturbative procedure.

3. Application to a singular perturbation problem

Like Ekeland and Séré in [4], we consider the Cauchy problem studied by Métivier and Rauch [10], and Texier and Zumbrun [11], which is a nonlinear system of Schrödinger equations arising in nonlinear optics. In [10], Métivier and Rauch prove the existence of local solutions of the Cauchy problem, with existence time T converging to 0 when the Sobolev $H^s(\mathbb{R}^d)$ norm of the initial datum goes to infinity. In [11], Texier and Zumbrun use a Nash–Moser scheme to improve this result, giving a uniform lower bound for T for two classes of initial data (concentrating and highly oscillating) whose $H^s(\mathbb{R}^d)$ norm goes to infinity. In [4], Ekeland and Séré apply their nonquadratic version of the Nash–Moser theorem, extending the result in [11] to even larger initial data.

Like in the aforementioned papers, we consider the system

$$\partial_t v_j + i\lambda_j \Delta v_j = \sum_{k=1}^{N} \big(b_{jk}(v, \partial_x) v_k + c_{jk}(v, \partial_x) \overline{v_k} \big), \quad j = 1, \dots, N, \tag{3.1}$$

where $v = v(t,x) = (v_1, \dots, v_N) \in \mathbb{C}^N$ is the unknown, $(t,x) \in [0,T] \times \mathbb{R}^d$, $\lambda_1, \dots, \lambda_N$ are constants, and $b_{jk}(v, \partial_x)$, $c_{jk}(v, \partial_x)$ are first-order differential operators

$$b_{jk}(v, \partial_x) = \sum_{\ell=1}^{d} b_{\ell jk}(v) \partial_{x_\ell}, \quad c_{jk}(v, \partial_x) = \sum_{\ell=1}^{d} c_{\ell jk}(v) \partial_{x_\ell}, \tag{3.2}$$

with $b_{\ell jk}$, $c_{\ell jk}$ complex-valued C^∞ functions of $\mathrm{Re}(v_1), \dots, \mathrm{Re}(v_N)$, $\mathrm{Im}(v_1), \dots, \mathrm{Im}(v_N)$ of order

$$b_{\ell jk}(v) = O(|v|^p), \quad c_{\ell jk}(v) = O(|v|^p) \tag{3.3}$$

in a ball around the origin, for some integer $p \geq 1$.

Following [4, 10, 11], we assume these "transparency conditions":

Assumption 3.1. We assume that

(i) $\lambda_1, \dots, \lambda_N$ are real and pairwise distinct;

(ii) for all j, k such that $\lambda_j + \lambda_k = 0$ there holds $c_{jk} = c_{kj}$;

(iii) for all j, b_{jj} is real.

Under these assumptions, the Cauchy problem for (3.1) is locally well posed in the Sobolev space $H^s(\mathbb{R}^d)$ for $s > 1 + d/2$ ([10, Theorem 1.5]). As is natural in the case of general initial data, the result in [10] gives an existence time T going to 0 as the initial datum goes to ∞ in $H^s(\mathbb{R}^d)$. In [4, 11] it is assumed that $p \geq 2$, and special initial data

$$v(0,x) = \varepsilon^\sigma \mathrm{a}_\varepsilon(x) \tag{3.4}$$

are considered, either concentrating or fast oscillating,

$$\mathrm{a}_\varepsilon(x) = \mathrm{a}_0(x/\varepsilon) \text{ (concentrating)}, \quad \mathrm{a}_\varepsilon(x) = \mathrm{a}_0(x) e^{ix\cdot\xi_0/\varepsilon} \text{ (oscillating)}, \tag{3.5}$$

with $\xi_0 \in \mathbb{R}^d$, and in both cases $0 < \varepsilon \leq 1$, $\sigma > 0$, $\mathrm{a}_0 \in H^{s_1}(\mathbb{R}^d)$ for some large s_1.

In [4, 11] the following results are proved.

Theorem 3.2 ([11, Theorem 4.6]). *Under the assumptions above, let $d, p \geq 2$ and*

$$\sigma > \frac{k_c - \sigma_{\mathrm{a}} - 1}{p+1}, \quad \sigma > \frac{d}{2}\frac{p}{p-1} - \sigma_{\mathrm{a}}, \tag{3.6}$$

where $\sigma_{\mathrm{a}} = d/2$ in the concentrating case, $\sigma_{\mathrm{a}} = 0$ in the oscillating case, and k_c is a constant depending on (d, p). Let s_1 be large enough, and let $T > 0$. If $\mathrm{a}_0 \in H^{\bar{s}}(\mathbb{R}^d)$

for $\bar{s}$ large enough, and $\|a_0\|_{H^{\bar{s}}}$ is small enough, then, for all $\varepsilon \in (0,1]$, the Cauchy problem (3.1)–(3.4)–(3.5) *has a unique solution in the space $C^1([0,T], H^{s_1-2}(\mathbb{R}^d)) \cap C^0([0,T], H^{s_1}(\mathbb{R}^d))$.*

The second condition in (3.6) is not written explicitly in the statement of [11, Theorem 4.6], but it is used in its proof. The constant k_c in (3.6) satisfies $k_c \geq \max\{6, 3 + \frac{dp}{2(p-1)}\}$; see Remark 3.7.

Theorem 3.3 ([4, Theorem 6]). *Under the assumptions above, let $d, p \geq 2$, let*

$$\sigma > \frac{d}{2(p-1)}, \tag{3.7}$$

and consider the concentrating case. Let $s_1 > d/2 + 4$ and $T > 0$. If $a_0 \in H^{\bar{s}}(\mathbb{R}^d)$ for $\bar{s}$ large enough, and $\|a_0\|_{H^{\bar{s}}}$ is small enough, then, for all $\varepsilon \in (0,1]$, the Cauchy problem (3.1)–(3.4)–(3.5) *has a unique solution in the space $C^1([0,T], H^{s_1-2}(\mathbb{R}^d)) \cap C^0([0,T], H^{s_1}(\mathbb{R}^d))$.*

Following [11], we introduce the "semiclassical" Sobolev norms

$$\begin{aligned} \|f\|_{H^s_\varepsilon} &:= \|(-\varepsilon^2 \Delta + 1)^{s/2} f\|_{L^2(\mathbb{R}^d)} \\ &= \|(1 + |\varepsilon\xi|^2)^{s/2} (\mathcal{F} f)(\xi)\|_{L^2(\mathbb{R}^d_\xi)}, \quad s \in \mathbb{R}, \end{aligned} \tag{3.8}$$

where $\mathcal{F}$ is the Fourier transform on $\mathbb{R}^d$, and $0 < \varepsilon \leq 1$. The first theorem we prove in this paper is the following.

Theorem 3.4. (i) (Existence) *In the assumptions above, let $T > 0$, $p \geq 1$, $d \geq 1$, and $s_1 > d/2 + 4$. Then there exist constants $C, C' > 0$, $\varepsilon_0 \in (0,1]$, depending on T, p, d, s_1, and on λ_j, b_{jk}, c_{jk} in system* (3.1)*, such that for all $\varepsilon \in (0, \varepsilon_0]$, for all initial data $v_0 \in H^{s_1}(\mathbb{R}^d)$ in the ball*

$$\|v_0\|_{H^{s_1}_\varepsilon} \leq C\varepsilon^q, \quad q := \frac{1}{p} + \frac{d}{2}, \tag{3.9}$$

the Cauchy problem for system (3.1) *with initial data $v(0,x) = v_0(x)$ has a solution*

$$v \in C^0([0,T], H^{s_1}(\mathbb{R}^d)) \cap C^1([0,T], H^{s_1-2}(\mathbb{R}^d)),$$

which satisfies

$$\sup_{t\in[0,T]} \|v(t)\|_{H^{s_1}_\varepsilon} + \varepsilon^2 \sup_{t\in[0,T]} \|\partial_t v(t)\|_{H^{s_1-2}_\varepsilon} \leq C' \|v_0\|_{H^{s_1}_\varepsilon}.$$

(ii) (Higher regularity) *If, in addition, $v_0 \in H^s(\mathbb{R}^d)$ for $s > s_1$, then*

$$\sup_{t\in[0,T]} \|v(t)\|_{H^s_\varepsilon} + \varepsilon^2 \sup_{t\in[0,T]} \|\partial_t v(t)\|_{H^{s-2}_\varepsilon} \leq C_s \|v_0\|_{H^s_\varepsilon}$$

where C_s depends on s (and it is independent of ε, v_0, v).

(iii) (Initial data of special form) *In particular, initial data* v_0 *of the form* (3.4)–(3.5), *with* $\|\mathrm{a}_0\|_{H^{s_1}(\mathbb{R}^d)} \leq 1$, *belong to the ball* (3.9) *for all* ε *sufficiently small if* $\sigma + \sigma_{\mathrm{a}} > q$, *namely*

$$\sigma > \frac{1}{p} + \frac{d}{2} - \sigma_{\mathrm{a}}, \tag{3.10}$$

where $\sigma_{\mathrm{a}} = d/2$ *in the concentrating case and* $\sigma_{\mathrm{a}} = 0$ *in the oscillating case.*

In the next theorem we deal with the case $p \geq 2$, where the power p of the nonlinearity is used to improve the lower bound for σ, at the price of a loss of one derivative in the solution with respect to the regularity of the datum.

Theorem 3.5. (i) (Existence) *In the assumptions above, let* $T > 0$, $p \geq 2$, $d \geq 1$, $s_1 > \max\{d + 4, 6\}$, *and*

$$\sigma > \frac{1 + d/2 - \sigma_{\mathrm{a}}}{p}, \tag{3.11}$$

where $\sigma_{\mathrm{a}} = d/2$ *in the concentrating case and* $\sigma_{\mathrm{a}} = 0$ *in the oscillating case.*

Then there exist constants $C > 0$, $\varepsilon_0 \in (0, 1]$, *depending on* T, p, d, s_1, *on* λ_j, b_{jk}, c_{jk} *in system* (3.1), *and on the difference* $\sigma - (1 + d/2 - \sigma_{\mathrm{a}})/p$, *such that for all* $\varepsilon \in (0, \varepsilon_0]$, *for all functions* $\mathrm{a}_0 \in H^{s_1}(\mathbb{R}^d)$ *in the ball*

$$\|\mathrm{a}_0\|_{H^{s_1}} \leq 1, \tag{3.12}$$

the Cauchy problem for system (3.1) *with initial data of the form* (3.4)–(3.5) *has a solution*

$$v \in C^0([0, T], H^{s_1 - 1}(\mathbb{R}^d)) \cap C^1([0, T], H^{s_1 - 3}(\mathbb{R}^d))$$

on the time interval $[0, T]$. *Such a solution* v *is the sum*

$$v = y + \tilde{v}$$

of a "free flow" component $y(t, x)$, *which is the solution of the Cauchy problem for the free Schrödinger system*

$$\begin{cases} \partial_t y_j + i\lambda_j \Delta y_j = 0, \quad j = 1, \dots, N, \\ y(0, x) = \varepsilon^{\sigma} \mathrm{a}_{\varepsilon}(x), \end{cases}$$

and a "correction" term $\tilde{v}(t, x)$ *satisfying* $\tilde{v}(0, x) = 0$ *and*

$$\sup_{t \in [0,T]} \|\tilde{v}(t)\|_{H_{\varepsilon}^{s_1 - 1}} + \varepsilon^2 \sup_{t \in [0,T]} \|\partial_t \tilde{v}(t)\|_{H_{\varepsilon}^{s_1 - 3}} \leq C \varepsilon^{\sigma + d/2} \|\mathrm{a}_0\|_{H^{s_1}}.$$

(ii) (Higher regularity) *If, in addition,* $\mathrm{a}_0 \in H^s(\mathbb{R}^d)$ *for* $s > s_1$, *then*

$$\sup_{t \in [0,T]} \|\tilde{v}(t)\|_{H_{\varepsilon}^{s - 1}} + \varepsilon^2 \sup_{t \in [0,T]} \|\partial_t \tilde{v}(t)\|_{H_{\varepsilon}^{s - 3}} \leq C_s \varepsilon^{\sigma + d/2} \|\mathrm{a}_0\|_{H^{s}},$$

where C_s *depends on* s *(and is independent of* ε, a_0*).*

Remark 3.6 (Smallness in low norm). In the higher regularity case, the smallness assumptions (3.9) in Theorem 3.4 and (3.12) in Theorem 3.5 are only required in the low norm s_1, with radii independent of the high regularity s.

Remark 3.7 (Comparison with the results in [4, 10, 11]). As observed in [4, 11], Métivier and Rauch [10] already provide existence for a fixed positive T, uniformly in ε, when

$$\sigma \geq \sigma_{\mathrm{MR}} := 1 + d/2 - \sigma_{\mathrm{a}}.$$

Hence [4, 11] and Theorems 3.4–3.5 give something new only for $\sigma < \sigma_{\mathrm{MR}}$.

The result of Texier and Zumbrun holds for $d \geq 2$, $p \geq 2$, and σ above the threshold

$$\sigma_{\mathrm{TZ}} := \frac{k_c - \sigma_{\mathrm{a}} - 1}{p+1}$$

([11, Theorem 4.6]), where the constant k_c satisfies some conditions; in particular, $k_c \geq 6$ and

$$k_c \geq 3 + \frac{d}{2}\frac{p}{p-1},$$

whence

$$\sigma_{\mathrm{TZ}} \geq \frac{1}{p+1}\Big(2 + \frac{d}{2}\frac{p}{p-1} - \sigma_{\mathrm{a}}\Big) =: c.$$

The threshold for σ in our Theorem 3.5 is

$$\sigma_1^* := \frac{1 + d/2 - \sigma_{\mathrm{a}}}{p} = \frac{\sigma_{\mathrm{MR}}}{p}.$$

For all pairs (d, p) covered by [11] (namely $d, p \geq 2$), one has $\sigma_1^* < c \leq \sigma_{\mathrm{TZ}}$, therefore we get a larger ball for the initial data. More precisely, regarding the data size, the improvement of Theorem 3.5 with respect to [11] corresponds to the exponent σ in the interval $\sigma_1^* < \sigma \leq \min\{\sigma_{\mathrm{TZ}}, \sigma_{\mathrm{MR}}\}$. Note that for some pairs (d, p) one has $\sigma_{\mathrm{TZ}} \geq \sigma_{\mathrm{MR}}$ (see [11, Examples 4.8–4.9]), so that [11] gives no improvement with respect to [10]; our result improves [10] in those cases also.

The result of Ekeland and Séré holds for $d, p \geq 2$, and σ above the threshold

$$\sigma_{\mathrm{ES}} := \frac{d}{2}\frac{p}{p-1} - \sigma_{\mathrm{a}} = \frac{d}{2(p-1)}$$

in the concentrating case $\sigma_{\mathrm{a}} = d/2$ ([4, Theorem 6]). The threshold for σ in our Theorem 3.4 is

$$\sigma_0^* := \frac{1}{p} + \frac{d}{2} - \sigma_{\mathrm{a}};$$

in particular, $\sigma_0^* = 1/p$ in the concentrating case. Since $\sigma_0^* < \sigma_{\mathrm{ES}}$ for all $d, p \geq 2$, we get a larger ball for the initial data also with respect to [4].

With respect to [4, 11] we also improve the regularity of the solution with respect to that of the initial data: using Theorem 3.5, the solution is one derivative less regular than the data (the loss of regularity is 1), while with Theorem 3.4 the solution has the same regularity as the data (the loss is 0). In [4, 11], instead, the loss of regularity depends in a nontrivial way on several parameters of the iteration scheme: it blows up to $+\infty$ in certain parameter regimes, and, in particular, can never be 0.

Remark 3.8 (Rôle of dispersion). Like in the approach of Métivier and Rauch [10], Texier and Zumbrun [11], and Ekeland and Séré [4], smoothing effects, Strichartz estimates, and dispersive properties of the linear Schrödinger flow play no direct rôle in the present paper.

Another natural approach to the study of system (3.1) in the singular perturbation regime (3.4)–(3.5) would be along the lines of the works of Kenig, Ponce, Vega, Cazenave, Chihara, etc. (see e.g. [6, 7] and the references therein), adapting "dispersive techniques" to the present singular perturbation issue.

It would be interesting to understand (although outside the scope of the present paper) whether the "inhomogeneous smoothing effect" of [6, 7], which provides a gain of one derivative, could be used to prove a stronger version of Theorem 3.5 where the loss of one derivative is removed and, simultaneously, the existence time, uniform in ε, and the threshold (3.11) are not deteriorated. Note, on the other hand, that in Theorem 3.4 there is no loss (even if we use Nash–Moser).

In fact, our point of view in the study of (3.1) in the singular perturbation regime (3.4)–(3.5) is very similar to that in [4, 11], which is somewhat the one of considering that problem *also* as a "concrete test for abstract Nash–Moser theorems" outside the traditional field of Hamiltonian dynamics where the loss of derivatives is due to the presence of small denominators in Fourier series.

4. Functional setting

In this section we introduce weighted Sobolev norms and recall the basic inequalities that will be used in the rest of the paper.

For $s \in \mathbb{R}$, we define

$$\|u\|_{H^s(\mathbb{R}^d)} := \|\Lambda^s u\|_{L^2(\mathbb{R}^d)}, \quad \|u\|_{H^s_\varepsilon(\mathbb{R}^d)} := \|\Lambda^s_\varepsilon u\|_{L^2(\mathbb{R}^d)}, \tag{4.1}$$

where $\Lambda^s = (1-\Delta)^{s/2}$ is the Fourier multiplier of symbol $(1+|\xi|^2)^{s/2}$ and $\Lambda^s_\varepsilon = (1-\varepsilon^2\Delta)^{s/2}$ is that of symbol $(1+\varepsilon^2|\xi|^2)^{s/2}$, namely, following [11],

$$\|u\|_{H^s_\varepsilon(\mathbb{R}^d)} = \|(1-\varepsilon^2\Delta)^{s/2}u\|_{L^2(\mathbb{R}^d)} = \|(1+|\varepsilon\xi|^2)^{s/2}\hat{u}(\xi)\|_{L^2(\mathbb{R}^d_\xi)}, \quad s \in \mathbb{R}, \tag{4.2}$$

where $\hat{u}$ is the Fourier transform of u on $\mathbb{R}^d$, and $0 < \varepsilon \leq 1$. For all $u \in H^s(\mathbb{R}^d)$, one has

$$\widehat{(R_\varepsilon u)}(\xi) = \varepsilon^{-d}\hat{u}(\varepsilon^{-1}\xi), \quad (R_\varepsilon u)(x) := u(\varepsilon x), \tag{4.3}$$

whence

$$\Lambda^s R_\varepsilon = R_\varepsilon \Lambda^s_\varepsilon, \quad \|u\|_{H^s_\varepsilon(\mathbb{R}^d)} = \varepsilon^{d/2}\|R_\varepsilon u\|_{H^s(\mathbb{R}^d)}. \tag{4.4}$$

We define the scalar product

$$\langle u, v\rangle_{H^s_\varepsilon(\mathbb{R}^d)} := \langle \Lambda^s_\varepsilon u, \Lambda^s_\varepsilon v\rangle_{L^2(\mathbb{R}^d)}. \tag{4.5}$$

To shorten the notation, we write $\| \, \|_{H^s}$ instead of $\| \, \|_{H^s(\mathbb{R}^d)}$, and so on. Using (4.4), it is immediate to obtain the Sobolev embedding and the standard tame estimates for products and compositions of functions in terms of the rescaled norms (4.2): for the Sobolev embedding, one has

$$\|u\|_{L^\infty} = \|R_\varepsilon u\|_{L^\infty} \leq C_{s_0}\|R_\varepsilon u\|_{H^{s_0}} = C_{s_0}\varepsilon^{-d/2}\|u\|_{H^{s_0}_\varepsilon} \tag{4.6}$$

for all $s_0 > d/2$, all $u \in H^{s_0}(\mathbb{R}^d)$, for some constant C_{s_0} depending on s_0, d; for the product, one has

$$\|uv\|_{H^s_\varepsilon} \leq C_s(\|u\|_{L^\infty}\|v\|_{H^s_\varepsilon} + \|u\|_{H^s_\varepsilon}\|v\|_{L^\infty}) \tag{4.7}$$

for all $u, h \in H^s(\mathbb{R}^d)$, all $s \geq 0$, for some constant C_s depending only on s, d; for the composition, given any C^∞ function f such that $f(y) = O(y^p)$ around the origin for some integer $p \geq 1$, one has

$$\|f(u)\|_{H^s_\varepsilon} \leq C_{s,M}\|u\|^{p-1}_{L^\infty}\|u\|_{H^s_\varepsilon} \tag{4.8}$$

for all $M > 0$, all $u \in H^s(\mathbb{R}^d)$ in the ball $\|u\|_{L^\infty} \leq M$, all $s \geq 0$, for some constant $C_{s,M}$ depending only on s, M, d, f. Moreover,

$$\varepsilon^{|\alpha|}\|\partial^\alpha_x u\|_{H^s_\varepsilon} \leq \|u\|_{H^{s+|\alpha|}_\varepsilon} \tag{4.9}$$

for all multi-indices $\alpha \in \mathbb{N}^d$.

For $m \geq 0$ integer, we define

$$\|u\|_{W^{m,\infty}} := \sum_{\substack{\alpha\in\mathbb{N}^d \\ |\alpha|\leq m}} \|\partial^\alpha_x u\|_{L^\infty}, \quad \|u\|_{W^{m,\infty}_\varepsilon} := \sum_{\substack{\alpha\in\mathbb{N}^d \\ |\alpha|\leq m}} \varepsilon^{|\alpha|}\|\partial^\alpha_x u\|_{L^\infty}. \tag{4.10}$$

One has

$$\partial^\alpha_x R_\varepsilon = \varepsilon^{|\alpha|}R_\varepsilon\partial^\alpha_x, \qquad \|u\|_{W^{m,\infty}_\varepsilon} = \|R_\varepsilon u\|_{W^{m,\infty}}. \tag{4.11}$$

Similarly to (4.8), given any C^∞ function f such that $f(y) = O(y^p)$ around the origin for some positive integer p, one has

$$\|f(u)\|_{W^{m,\infty}_\varepsilon} \leq C_{m,M}\|u\|^{p-1}_{L^\infty}\|u\|_{W^{m,\infty}_\varepsilon} \tag{4.12}$$

for all $M > 0$, all $u \in W^{m,\infty}(\mathbb{R}^d)$ in the ball $\|u\|_{L^\infty} \leq M$, all integers $m \geq 0$, for some constant $C_{m,M}$ depending on m, M, d, f. For the product of two functions, we also have

$$\|uv\|_{H^s_\varepsilon} \leq \varepsilon^{-d/2}(C_{s_0}\|u\|_{H^{s_0}_\varepsilon}\|v\|_{H^s_\varepsilon} + C_s\|u\|_{H^s_\varepsilon}\|v\|_{H^{s_0}_\varepsilon}) \tag{4.13}$$

for all $s \geq 0$, $s_0 > d/2$, all $u, v \in H^s(\mathbb{R}^d) \cap H^{s_0}(\mathbb{R}^d)$, and

$$\|uv\|_{H^s_\varepsilon} \leq 2\|u\|_{L^\infty}\|v\|_{H^s_\varepsilon} + C_s\|u\|_{W^{m,\infty}_\varepsilon}\|v\|_{L^2} \tag{4.14}$$

for all $s \geq 0$, all $v \in H^s(\mathbb{R}^d)$, all $u \in W^{m,\infty}(\mathbb{R}^d)$, where m is the smallest positive integer such that $m \geq s$, and C_s depends on s, d. Estimate (4.14) is proved in the appendix (see (B.8) in Lemma B.2). We remark that the constants C_{s_0}, C_s, $C_{s,M}$, $C_{m,M}$ in (4.6), (4.7), (4.8), (4.12), (4.13), (4.14) are independent of ε, and C_{s_0} is also independent of s.

For time-dependent functions $u(t,x)$, $t \in [0,T]$, we denote, in short,

$$\|u\|_{C^0H^s_\varepsilon} := \|u\|_{C([0,T],H^s_\varepsilon)}, \qquad \|u\|_{C^1_\varepsilon H^s_\varepsilon} := \|u\|_{C^0H^s_\varepsilon} + \varepsilon^2\|\partial_t u\|_{C^0H^{s-2}_\varepsilon}, \tag{4.15}$$

$$\|u\|_{C^0W^m_\varepsilon} := \|u\|_{C([0,T],W^{m,\infty}_\varepsilon)}, \quad \|u\|_{C^1_\varepsilon W^m_\varepsilon} := \|u\|_{C^0W^m_\varepsilon} + \varepsilon^2\|\partial_t u\|_{C^0W^{m-2}_\varepsilon}. \tag{4.16}$$

The notation $a \lesssim_s b$ means $a \leq C_s b$ for some constant C_s, independent of ε, possibly depending on s; also, $a \lesssim b$ means $a \leq Cb$ for some constant C independent of ε and s.

5. Analysis of the singular perturbation problem

In [4, 11], system (3.1) is written as

$$\partial_t u + iA(\partial_x)u = B(u, \partial_x)u, \tag{5.1}$$

where $u = (v, \bar{v}) = (v_1, \ldots, v_N, \overline{v_1}, \ldots, \overline{v_N})$ is the unknown, $A(\partial_x)$ is the constant coefficients operator of second order

$$A(\partial_x) = \operatorname{diag}(\lambda_1, \ldots, \lambda_n, -\lambda_1, \ldots, -\lambda_n)\Delta,$$

$B(u, \partial_x)$ is the operator matrix

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

$\mathcal{B}$, $\mathcal{C}$ are the operator matrices with entries $b_{jk}(v, \partial_x)$, $c_{jk}(v, \partial_x)$ respectively, and $\bar{\mathcal{B}}$, $\bar{\mathcal{C}}$ have conjugate entry coefficients. To deal with concentrating or highly oscillating initial data (3.5), in [11] the weighted Sobolev norms (4.2) are introduced. Recalling (4.9), it is natural, as is done in [4, 11], to write the powers of ε as separate factors, writing (5.1) as

$$\partial_t u + i\varepsilon^{-2}A(\varepsilon\partial_x)u = \varepsilon^{-1}B(u, \varepsilon\partial_x)u, \tag{5.2}$$

where $A(\varepsilon\partial_x) := \varepsilon^2 A(\partial_x)$ and $B(u, \varepsilon\partial_x) := \varepsilon B(u, \partial_x)$. In this way $A(\varepsilon\partial_x)$ and $B(u, \varepsilon\partial_x)$ satisfy estimates that are uniform in ε:

$$\|A(\varepsilon\partial_x)u\|_{H^s_\varepsilon} \leq C_0\|u\|_{H^{s+2}_\varepsilon} \tag{5.3}$$

for all $s \in \mathbb{R}$, all $u \in H^s(\mathbb{R}^d)$, with $C_0 = \max\{|\lambda_1|, \ldots, |\lambda_N|\}$;

$$\|B(u, \varepsilon\partial_x)h\|_{H^s_\varepsilon} \leq C_s(\|u\|^p_{L^\infty}\|h\|_{H^{s+1}_\varepsilon} + \|u\|^{p-1}_{L^\infty}\|u\|_{H^s_\varepsilon}\|\varepsilon\partial_x h\|_{L^\infty}) \tag{5.4}$$

for all $s \geq 0$, all $h \in H^{s+1}(\mathbb{R}^d)$, all $u \in H^s(\mathbb{R}^d)$ in the ball $\|u\|_{L^\infty} \leq 1$; also, by (4.14) and (4.12),

$$\|B(u, \varepsilon\partial_x)h\|_{H^s_\varepsilon} \leq C\|u\|^p_{L^\infty}\|h\|_{H^{s+1}_\varepsilon} + C_s\|u\|^{p-1}_{L^\infty}\|u\|_{W^{[s]+1,\infty}_\varepsilon}\|h\|_{H^1_\varepsilon} \tag{5.5}$$

for all $s \geq 0$, all $h \in H^{s+1}(\mathbb{R}^d)$, all $u \in W^{[s]+1,\infty}(\mathbb{R}^d)$ in the ball $\|u\|_{L^\infty} \leq 1$, where $[s]$ is the integer part of s; and, by (B.13) and (4.12),

$$\|B(u, \varepsilon\partial_x)h\|_{H^s_\varepsilon} \leq C\|u\|^{p-1}_{L^\infty}\|u\|_{W^{1,\infty}_\varepsilon}\|h\|_{H^{s+1}_\varepsilon} \tag{5.6}$$

for all $-1 \leq s \leq 0$, all $h \in H^{s+1}(\mathbb{R}^d)$, all $u \in W^{1,\infty}(\mathbb{R}^d)$ in the ball $\|u\|_{L^\infty} \leq 1$. The constants in (5.3), (5.4), (5.5), (5.6) do not depend on $\varepsilon \in (0, 1]$; C_0, C in (5.3), (5.5), and (5.6) are also independent of s.

We consider the Cauchy problem for (5.2) with initial data (3.4), namely

$$\begin{cases} \partial_t u + P(u) = 0, \\ u(0) = u_0, \end{cases} \tag{5.7}$$

where

$$P(u) := i\varepsilon^{-2}A(\varepsilon\partial_x)u - \varepsilon^{-1}B(u, \varepsilon\partial_x)u, \quad u_0(x) := \varepsilon^\sigma(\mathrm{a}_\varepsilon(x), \overline{\mathrm{a}_\varepsilon(x)}). \tag{5.8}$$

To apply our Nash–Moser theorem, we need to construct a right inverse for the linearized problem and to estimate the second derivative of the nonlinear operator. Let us begin with the linear inversion problem.

Analysis of the linearized problem. Given $u(t, x)$, $f_1(t, x)$, and $f_2(x)$, consider the linear Cauchy problem for the unknown $h(t, x)$,

$$\begin{cases} \partial_t h + P'(u)h = f_1, \\ h(0) = f_2, \end{cases} \tag{5.9}$$

where

$$P'(u)h = i\varepsilon^{-2}A(\varepsilon\partial_x)h - \varepsilon^{-1}B(u, \varepsilon\partial_x)h + R_0(u)h, \tag{5.10}$$

$$R_0(u)h := -\varepsilon^{-1}(\partial_u B)(u, \varepsilon\partial_x)[h]u. \tag{5.11}$$

Following [11], let

$$J := \{(j, k) : \lambda_j + \lambda_k = 0\},$$

and let $\chi \in C^\infty_c(\mathbb{R}^d, \mathbb{R})$ be a frequency truncation such that $0 \leq \chi(\xi) \leq 1$, $\chi(\xi) = 1$ for $|\xi| \leq 1/2$, and $\chi(\xi) = 0$ for $|\xi| \geq 1$. Like in [11], we decompose B into the sum of a resonant term, a nonresonant term, and a low-frequency term: $B = B_{\mathrm{r}} + B_{\mathrm{nr}} + B_{\mathrm{lf}}$, where the following hold:

- The resonant term is

$$B_{\mathrm{r}} := \begin{pmatrix} \mathcal{B}_d & \mathcal{C}_J \\ \overline{\mathcal{C}_J} & \mathcal{B}_d \end{pmatrix},$$

where $\mathcal{B}_d := \operatorname{diag}(b_{11}, \dots, b_{NN})$, $(\mathcal{C}_J)_{jk} := c_{jk}$ if $(j,k) \in J$, and $(\mathcal{C}_J)_{jk} := 0$ otherwise. By Assumption 3.1, the matrix $B_{\mathrm{r}}(v, \xi)$ is Hermitian.

- The nonresonant term is

$$B_{\mathrm{nr}} := \begin{pmatrix} \mathcal{B}^1 & \mathcal{C}^1 \\ \bar{\mathcal{C}}^1 & \bar{\mathcal{B}}^1 \end{pmatrix},$$

where $(\mathcal{B}^1)_{jk} := (1-\chi) b_{jk}$ if $j \neq k$, and $(\mathcal{B}^1)_{jk} := 0$ if $j = k$; $(\mathcal{C}^1)_{jk} := (1-\chi) c_{jk}$ if $(j,k) \notin J$, and $(\mathcal{C}^1)_{jk} := 0$ if $(j,k) \in J$.

- The low-frequency term is

$$B_{\mathrm{lf}} := \begin{pmatrix} \mathcal{B}^0 & \mathcal{C}^0 \\ \bar{\mathcal{C}}^0 & \bar{\mathcal{B}}^0 \end{pmatrix}$$

where $(\mathcal{B}^0)_{jk} := \chi b_{jk}$ if $j \neq k$, and $(\mathcal{B}^0)_{jk} := 0$ if $j = k$; $(\mathcal{C}^0)_{jk} := \chi c_{jk}$ if $(j,k) \notin J$, and $(\mathcal{C}^0)_{jk} := 0$ if $(j,k) \in J$.

We recall the normal form transformation of [11] (see [11, proof of Lemma 4.5]): define the pseudo-differential matrix symbol $M(u(t,x), \xi)$ as

$$M_{jk}(u(t,x), \xi) := \begin{cases} \dfrac{(B_{\mathrm{nr}})_{jk}(u(t,x), i\xi)}{i|\xi|^2(\omega_j - \omega_k)} & \text{if } \omega_j \neq \omega_k, \\ 0 & \text{if } \omega_j = \omega_k, \end{cases} \tag{5.12}$$

where

$$\omega_j := \begin{cases} -\lambda_j & \text{for } j = 1, \dots, N, \\ \lambda_{j-N} & \text{for } j = N+1, \dots, 2N. \end{cases}$$

Since the commutator of A and M is the matrix

$$[A(i\xi), M(u,\xi)] = \big(|\xi|^2(\omega_j - \omega_k) M_{jk}(u,\xi)\big)_{j,k=1,\dots,2N}, \tag{5.13}$$

one has

$$B_{\mathrm{nr}}(u(t,x), i\xi) - i[A(i\xi), M(u(t,x),\xi)] = 0.$$

Like in [11], we introduce the following semiclassical quantization of a symbol $\sigma(x,\xi)$:

$$\operatorname{op}_\varepsilon(\sigma) h(x) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} \sigma(x, \varepsilon\xi) \hat{h}(\xi) e^{i\xi\cdot x}\, d\xi.$$

By (5.12) and (5.5), one has

$$\|\operatorname{op}_\varepsilon(M) h\|_{H^s_\varepsilon} \le C \|u\|_{L^\infty}^p \|h\|_{H^{s-1}_\varepsilon} + C_s \|u\|_{L^\infty}^{p-1} \|u\|_{W_\varepsilon^{[s]+1,\infty}} \|h\|_{H_\varepsilon^{-1}}, \tag{5.14}$$

$$\|\operatorname{op}_\varepsilon(M) h\|_{L^2} \le C \|u\|_{L^\infty}^p \|h\|_{H_\varepsilon^{-1}} \tag{5.15}$$

for all $s \geq 0$, all $\|u\|_{L^\infty} \leq 1$, all h. Hence there exists $\rho_0 > 0$, independent of ε, such that, for u in the ball

$$\varepsilon \|u\|_{L^\infty}^p \leq \rho_0, \tag{5.16}$$

one has

$$\|\varepsilon \operatorname{op}_\varepsilon(M)h\|_{H_\varepsilon^{-1}} \leq \|\varepsilon \operatorname{op}_\varepsilon(M)h\|_{L^2} \leq C\varepsilon\|u\|_{L^\infty}^p \|h\|_{H_\varepsilon^{-1}} \leq \frac{1}{2}\|h\|_{H_\varepsilon^{-1}} \leq \frac{1}{2}\|h\|_{L^2}. \tag{5.17}$$

Therefore, by Neumann series, $I + \varepsilon \operatorname{op}_\varepsilon(M)$ is invertible in H_ε^{-1} and in L^2, and

$$\|(I + \varepsilon \operatorname{op}_\varepsilon(M))^{-1} h\|_{H_\varepsilon^s} \leq C\|h\|_{H_\varepsilon^s} + C_s \varepsilon \|u\|_{L^\infty}^{p-1} \|u\|_{W_\varepsilon^{[s]+1,\infty}} \|h\|_{H_\varepsilon^{-1}} \tag{5.18}$$

for all $s \geq 0$, for $u \in W_\varepsilon^{[s]+1,\infty}(\mathbb{R}^d)$ in the ball (5.16) (where ρ_0 is independent of s).

Under the change of variable

$$h = (I + \varepsilon \operatorname{op}_\varepsilon(M))\varphi, \tag{5.19}$$

the linear Cauchy problem (5.9) becomes

$$\begin{cases} \partial_t \varphi + Q(u)\varphi = g_1, \\ \varphi(0) = g_2, \end{cases} \tag{5.20}$$

where

$$g_1 := (I + \varepsilon \operatorname{op}_\varepsilon(M))^{-1} f_1, \quad g_2 := (I + \varepsilon \operatorname{op}_\varepsilon(M))^{-1}|_{t=0} f_2, \tag{5.21}$$

and, by (5.13),

$$\begin{aligned} \partial_t + Q(u) &:= (I + \varepsilon \operatorname{op}_\varepsilon(M))^{-1}(\partial_t + P'(u))(I + \varepsilon \operatorname{op}_\varepsilon(M)) \\ &= \partial_t + i\varepsilon^{-2} A(\varepsilon\partial_x) - \varepsilon^{-1} B_{\mathrm{r}}(u, \varepsilon\partial_x) + G(u), \end{aligned} \tag{5.22}$$

with

$$\begin{aligned} G(u) := (I + \varepsilon \operatorname{op}_\varepsilon(M))^{-1}\big(&\varepsilon \operatorname{op}_\varepsilon(M)\varepsilon^{-1} B_{\mathrm{r}}(u, \varepsilon\partial_x) - \varepsilon^{-1} B_{\mathrm{lf}}(u, \varepsilon\partial_x) \\ &+ \varepsilon \operatorname{op}_\varepsilon(\partial_t M) - \varepsilon^{-1} B(u, \varepsilon\partial_x)\varepsilon \operatorname{op}_\varepsilon(M) \\ &+ R_0(u)(I + \varepsilon \operatorname{op}_\varepsilon(M))\big) \end{aligned} \tag{5.23}$$

(we have used the trivial identity $I - (I + K)^{-1} = (I + K)^{-1} K$ for $K = \varepsilon \operatorname{op}_\varepsilon(M)$).

Now we prove an energy estimate for (5.22), and we start with the term $G(u)$. By (5.18), (5.14), (5.17), (5.5), (5.6), the first term in (5.23) satisfies, for $s \geq 0$,

$$\begin{aligned} &\|(I + \varepsilon \operatorname{op}_\varepsilon(M))^{-1} \varepsilon \operatorname{op}_\varepsilon(M)\varepsilon^{-1} B_{\mathrm{r}}(u, \varepsilon\partial_x)\varphi\|_{H_\varepsilon^s} \\ &\qquad \lesssim_s \|u\|_{L^\infty}^{2p} \|\varphi\|_{H_\varepsilon^s} + \|u\|_{L^\infty}^{2p-2} \|u\|_{W_\varepsilon^{[s]+1,\infty}} \|u\|_{W_\varepsilon^{1,\infty}} \|\varphi\|_{L^2} \end{aligned}$$

and

$$\|(I + \varepsilon \operatorname{op}_\varepsilon(M))^{-1} \varepsilon \operatorname{op}_\varepsilon(M)\varepsilon^{-1} B_{\mathrm{r}}(u, \varepsilon\partial_x)\varphi\|_{L^2} \lesssim \|u\|_{L^\infty}^{2p-1} \|u\|_{W_\varepsilon^{1,\infty}} \|\varphi\|_{L^2}.$$

The low-frequency term $B_{\rm lf}$ satisfies, for $s \geq 0$,

$$\|\varepsilon^{-1} B_{\rm lf}(u, \varepsilon\partial_x)\varphi\|_{H^s_\varepsilon} \lesssim_s \varepsilon^{-1}\|u\|^{p-1}_{L^\infty}\|u\|_{W^{[s]+1,\infty}_\varepsilon}\|\varphi\|_{L^2}.$$

The term containing the time derivative of the symbol M is estimated, for $s \geq 0$, by

$$\begin{aligned}\|\varepsilon\,\mathrm{op}_\varepsilon(\partial_t M)\varphi\|_{H^s_\varepsilon} \lesssim_s {}& \varepsilon\|u\|^{p-1}_{L^\infty}\|\partial_t u\|_{L^\infty}\|\varphi\|_{H^{s-1}_\varepsilon} \\ &+ \varepsilon(\|u\|^{p-1}_{L^\infty}\|\partial_t u\|_{W^{[s]+1,\infty}_\varepsilon} + \|u\|^{\nu}_{L^\infty}\|u\|_{W^{[s]+1,\infty}_\varepsilon}\|\partial_t u\|_{L^\infty})\|\varphi\|_{H^{-1}_\varepsilon},\end{aligned}$$

where

$$\nu := \max\{p-2, 0\}, \tag{5.24}$$

and, by (5.12),

$$\|\varepsilon\,\mathrm{op}_\varepsilon(\partial_t M)\varphi\|_{H^{-1}_\varepsilon} \leq \|\varepsilon\,\mathrm{op}_\varepsilon(\partial_t M)\varphi\|_{L^2} \lesssim \varepsilon\|u\|^{p-1}_{L^\infty}\|\partial_t u\|_{L^\infty}\|\varphi\|_{H^{-1}_\varepsilon}.$$

Next, R_0 defined in (5.11) satisfies, for $s \geq 0$,

$$\|R_0(u)\varphi\|_{H^s_\varepsilon} \lesssim_s \varepsilon^{-1}\|u\|^{p-1}_{L^\infty}(\|u\|_{W^{1,\infty}_\varepsilon}\|\varphi\|_{H^s_\varepsilon} + \|u\|_{W^{[s]+2,\infty}_\varepsilon}\|\varphi\|_{L^2}), \tag{5.25}$$

$$\|R_0(u)\varphi\|_{L^2} \lesssim \varepsilon^{-1}\|u\|^{p-1}_{L^\infty}\|u\|_{W^{1,\infty}_\varepsilon}\|\varphi\|_{L^2}. \tag{5.26}$$

Hence $G(u)$ in (5.23) satisfies, for all $s \geq 0$,

$$\begin{aligned}\|G(u)\varphi\|_{H^s_\varepsilon} \lesssim_s {}& \varepsilon^{-1}\|u\|^{p-1}_{L^\infty}(\|u\|_{W^{1,\infty}_\varepsilon} + \varepsilon^2\|\partial_t u\|_{L^\infty})\|\varphi\|_{H^s_\varepsilon} \\ &+ \varepsilon^{-1}\big\{\|u\|^{p-1}_{L^\infty}(\|u\|_{W^{[s]+2,\infty}_\varepsilon} + \varepsilon^2\|\partial_t u\|_{W^{[s]+1,\infty}_\varepsilon}) \\ &\quad + \varepsilon^2\|u\|^{\nu}_{L^\infty}\|u\|_{W^{[s]+1,\infty}_\varepsilon}\|\partial_t u\|_{L^\infty}\big\}\|\varphi\|_{L^2},\end{aligned} \tag{5.27}$$

$$\|G(u)\varphi\|_{L^2} \lesssim \varepsilon^{-1}\|u\|^{p-1}_{L^\infty}(\|u\|_{W^{1,\infty}_\varepsilon} + \varepsilon^2\|\partial_t u\|_{L^\infty})\|\varphi\|_{L^2}. \tag{5.28}$$

The constant coefficient operator $A(\varepsilon\partial_x)$ in (5.22) satisfies

$$\mathrm{Re}\langle i\varepsilon^{-2}A(\varepsilon\partial_x)\varphi, \varphi\rangle_{H^s_\varepsilon} = 0 \tag{5.29}$$

because λ_j are all real. To estimate the term with $B_{\rm r}(u, \varepsilon\partial_x)$ in (5.22), we recall that

$$2\,\mathrm{Re}\langle X\varphi, \varphi\rangle_{H^s_\varepsilon} = \langle (X + X^*)\Lambda^s_\varepsilon\varphi, \Lambda^s_\varepsilon\varphi\rangle_{L^2} + 2\,\mathrm{Re}\langle [\Lambda^s_\varepsilon, X]\varphi, \Lambda^s_\varepsilon\varphi\rangle_{L^2}$$

for any linear operator X, where X^* is the adjoint of X with respect to the L^2 scalar product and $[\,,\,]$ is the commutator, whence

$$2|\mathrm{Re}\langle X\varphi, \varphi\rangle_{H^s_\varepsilon}| \leq \|X + X^*\|_{\mathcal{L}(L^2,L^2)}\|\varphi\|^2_{H^s_\varepsilon} + 2\|[\Lambda^s_\varepsilon, X]\varphi\|_{L^2}\|\varphi\|_{H^s_\varepsilon}.$$

By the Hermitian structure of $B_{\rm r}(u, \varepsilon\partial_x)$,

$$\|X + X^*\|_{\mathcal{L}(L^2,L^2)} \lesssim \varepsilon^{-1}\|u\|^{p-1}_{L^\infty}\|u\|_{W^{1,\infty}_\varepsilon}, \quad X = \varepsilon^{-1}B_{\rm r}(u, \varepsilon\partial_x), \tag{5.30}$$

and, by (B.18), for $X = \varepsilon^{-1} B_{\rm r}(u, \varepsilon\partial_x)$ one has

$$\|[\Lambda_\varepsilon^s, X]\varphi\|_{L^2} \lesssim_s \varepsilon^{-1}\|u\|_{L^\infty}^{p-1}(\|u\|_{W_\varepsilon^{1,\infty}}\|\varphi\|_{H_\varepsilon^s} + \|u\|_{W_\varepsilon^{[s]+2,\infty}}\|\varphi\|_{L^2}).$$

Therefore,

$$\begin{aligned} |\mathrm{Re}\langle \varepsilon^{-1} B_{\rm r}(u,\varepsilon\partial_x)\varphi,\varphi\rangle_{H_\varepsilon^s}| &\lesssim_s \varepsilon^{-1}\|u\|_{L^\infty}^{p-1}(\|u\|_{W_\varepsilon^{1,\infty}}\|\varphi\|_{H_\varepsilon^s} \\ &\quad + \|u\|_{W_\varepsilon^{[s]+2,\infty}}\|\varphi\|_{L^2})\|\varphi\|_{H_\varepsilon^s}, \\ |\mathrm{Re}\langle \varepsilon^{-1} B_{\rm r}(u,\varepsilon\partial_x)\varphi,\varphi\rangle_{L^2}| &\lesssim \ \varepsilon^{-1}\|u\|_{L^\infty}^{p-1}\|u\|_{W_\varepsilon^{1,\infty}}\|\varphi\|_{L^2}^2. \end{aligned} \tag{5.31}$$

By (5.27)–(5.31), the solution φ of the linear equation $\partial_t\varphi + Q(u)\varphi = g_1$ (see (5.20) and (5.22)) satisfies

$$\begin{aligned} \partial_t(\|\varphi\|_{H_\varepsilon^s}^2) =& \ 2\,\mathrm{Re}\langle g_1 + \varepsilon^{-1}B_{\rm r}(u,\varepsilon\partial_x)\varphi - G(u)\varphi, \varphi\rangle_{H_\varepsilon^s} \\ \lesssim_s& \ \big\{\|g_1\|_{H_\varepsilon^s} + \varepsilon^{-1}\|u\|_{L^\infty}^{p-1}(\|u\|_{W_\varepsilon^{1,\infty}} + \varepsilon^2\|\partial_t u\|_{L^\infty})\|\varphi\|_{H_\varepsilon^s} \\ &+ \varepsilon^{-1}\big(\|u\|_{L^\infty}^{p-1}(\|u\|_{W_\varepsilon^{[s]+2,\infty}} + \varepsilon^2\|\partial_t u\|_{W_\varepsilon^{[s]+1,\infty}}) \\ &+ \varepsilon^2\|u\|_{L^\infty}^{\nu}\|u\|_{W_\varepsilon^{[s]+1,\infty}}\|\partial_t u\|_{L^\infty}\big)\|\varphi\|_{L^2}\big\}\|\varphi\|_{H_\varepsilon^s}, \end{aligned} \tag{5.32}$$

$$\partial_t(\|\varphi\|_{L^2}^2) \lesssim \ \varepsilon^{-1}\|u\|_{L^\infty}^{p-1}(\|u\|_{W_\varepsilon^{1,\infty}} + \varepsilon^2\|\partial_t u\|_{L^\infty})\|\varphi\|_{L^2}^2 + \|g_1\|_{L^2}\|\varphi\|_{L^2}. \tag{5.33}$$

If u satisfies

$$\varepsilon^{-1}\|u\|_{L^\infty}^{p-1}(\|u\|_{W_\varepsilon^{1,\infty}} + \varepsilon^2\|\partial_t u\|_{L^\infty}) \le 1 \tag{5.34}$$

on the time interval $[0,T]$, then for $s \ge 0$ the solution φ of (5.20) satisfies, with the notation introduced in (4.15), (4.16),

$$\|\varphi\|_{C^0L^2} \lesssim \ \|g_1\|_{C^0L^2} + \|g_2\|_{L^2}, \tag{5.35}$$

$$\begin{aligned} \|\varphi\|_{C^0H_\varepsilon^s} \lesssim_s& \ \|g_1\|_{C^0H_\varepsilon^s} + \|g_2\|_{H_\varepsilon^s} \\ &+ \varepsilon^{-1}(\|u\|_{C^0L^\infty}^{p-1}\|u\|_{C_\varepsilon^1W_\varepsilon^{[s]+3}} + \|u\|_{C^0L^\infty}^{\nu}\|u\|_{C^0W_\varepsilon^{[s]+1}}\|u\|_{C_\varepsilon^1W_\varepsilon^2}) \\ &\times(\|g_1\|_{C^0L^2} + \|g_2\|_{L^2}) \end{aligned} \tag{5.36}$$

(first use (5.33), (5.34), and Grönwall to get (5.35), then insert (5.35) into (5.32) and use Grönwall again).

By definitions (5.19), (5.21) and estimates (5.14), (5.15), (5.18), we deduce that the solution h of the linear Cauchy problem (5.9) satisfies the same estimates (5.35), (5.36) as φ with f_1, f_2 in place of g_1, g_2, namely, for all $s \ge 0$,

$$\|h\|_{C^0L^2} \lesssim \ \|f_1\|_{C^0L^2} + \|f_2\|_{L^2}, \tag{5.37}$$

$$\begin{aligned} \|h\|_{C^0H_\varepsilon^s} \lesssim_s& \ \|f_1\|_{C^0H_\varepsilon^s} + \|f_2\|_{H_\varepsilon^s} \\ &+ \varepsilon^{-1}(\|u\|_{C^0L^\infty}^{p-1}\|u\|_{C_\varepsilon^1W_\varepsilon^{[s]+3}} + \|u\|_{C^0L^\infty}^{\nu}\|u\|_{C^0W_\varepsilon^{[s]+1}}\|u\|_{C_\varepsilon^1W_\varepsilon^2}) \\ &\times(\|f_1\|_{C^0L^2} + \|f_2\|_{L^2}). \end{aligned} \tag{5.38}$$

From the equation $\partial_t h + P'(u)h = f_1$ one has, for all s real,

$$\|\partial_t h\|_{H^s_\varepsilon} \le \|f_1\|_{H^s_\varepsilon} + \|P'(u)h\|_{H^s_\varepsilon}. \tag{5.39}$$

By (5.6), (5.26), (5.34), for $-1 \le s \le 0$ one has

$$\|P'(u)h\|_{H^s_\varepsilon} \lesssim \varepsilon^{-2}\|h\|_{H^{s+2}_\varepsilon}, \tag{5.40}$$

and, by (5.5), (5.25), (5.34), for $s \ge 0$ one has

$$\|P'(u)h\|_{H^s_\varepsilon} \lesssim_s \varepsilon^{-2}\|h\|_{H^{s+2}_\varepsilon} + \varepsilon^{-1}\|u\|^{p-1}_{L^\infty}\|u\|_{W^{[s]+2,\infty}_\varepsilon}\|h\|_{L^2}. \tag{5.41}$$

Hence, by (5.37)–(5.41), for all $s \ge -1$ one has

$$\begin{aligned}\varepsilon^2\|\partial_t h\|_{H^s_\varepsilon} \lesssim_s\ & \|f_1\|_{C^0 H^{s+2}_\varepsilon} + \|f_2\|_{H^{s+2}_\varepsilon} \\ & + \varepsilon^{-1}(\|u\|^{p-1}_{C^0 L^\infty}\|u\|_{C^1_\varepsilon W^{[s]+5}_\varepsilon} + \|u\|^{\nu}_{C^0 L^\infty}\|u\|_{C^0 W^{[s]+3}_\varepsilon}\|u\|_{C^1_\varepsilon W^2_\varepsilon}) \\ & \times (\|f_1\|_{C^0 L^2} + \|f_2\|_{L^2}).\end{aligned}$$

Thus, recalling definition (4.15), h satisfies, for all $s \ge 1$,

$$\begin{aligned}\|h\|_{C^1_\varepsilon H^s_\varepsilon} \lesssim_s\ & \|f_1\|_{C^0 H^s_\varepsilon} + \|f_2\|_{H^s_\varepsilon} \\ & + \varepsilon^{-1}(\|u\|^{p-1}_{C^0 L^\infty}\|u\|_{C^1_\varepsilon W^{[s]+3}_\varepsilon} + \|u\|^{\nu}_{C^0 L^\infty}\|u\|_{C^0 W^{[s]+1}_\varepsilon}\|u\|_{C^1_\varepsilon W^2_\varepsilon}) \\ & \times (\|f_1\|_{C^0 L^2} + \|f_2\|_{L^2}).\end{aligned} \tag{5.42}$$

In conclusion, we have proved the following result.

Lemma 5.1 (Right inverse of the linearized problem). *Let $s \ge 1$ be real, and let u belong to $C([0,T], W^{[s]+3,\infty}(\mathbb{R}^d)) \cap C^1([0,T], W^{[s]+1,\infty}(\mathbb{R}^d))$, with (5.34) and (5.16). Then for all $f_1 \in C([0,T], H^s(\mathbb{R}^d))$, all $f_2 \in H^s(\mathbb{R}^d)$, the linear Cauchy problem (5.9) has a (unique) solution h, which satisfies (5.42).*

Estimate for the second derivative. By (3.2) and (5.8), the operator

$$\begin{aligned}P''(u)[h_1,h_2] = & -\varepsilon^{-1}(\partial_u B)(u,\varepsilon\partial_x)[h_1]h_2 - \varepsilon^{-1}(\partial_u B)(u,\varepsilon\partial_x)[h_2]h_1 \\ & - \varepsilon^{-1}(\partial_{uu} B)(u,\varepsilon\partial_x)[h_1,h_2]u\end{aligned}$$

is the sum of terms of the form

$$\varepsilon^{-1}g'(u)h_1\varepsilon\partial_x h_2 + \varepsilon^{-1}g'(u)h_2\varepsilon\partial_x h_1 + \varepsilon^{-1}g''(u)h_1h_2\varepsilon\partial_x u, \tag{5.43}$$

where $g(u)$ is a vector of components $b_{\ell jk}(u)$ or $c_{\ell jk}(u)$. By (3.3), $g(u) = O(|u|^p)$ with $p \ge 1$ integer. For $p \ge 3$, by (4.8) one has for all u in the ball $\|u\|_{L^\infty} \le 1$, for all $s \ge 0$,

$$\|g'(u)\|_{L^\infty} \lesssim \|u\|^{p-1}_{L^\infty}, \quad \|g'(u)\|_{H^s_\varepsilon} \lesssim_s \|u\|_{H^s_\varepsilon}\|u\|^{p-2}_{L^\infty}, \tag{5.44}$$

$$\|g''(u)\|_{L^\infty} \lesssim \|u\|^{p-2}_{L^\infty}, \quad \|g''(u)\|_{H^s_\varepsilon} \lesssim_s \|u\|_{H^s_\varepsilon}\|u\|^{p-3}_{L^\infty}. \tag{5.45}$$

For $p = 2$, $g(u) = g_2(u) + \tilde{g}(u)$, where $g_2(u)$ is homogeneous of degree 2 in u and $\tilde{g}(u) = O(|u|^3)$ (we do not distinguish whether $\tilde{g}$ is of order 3 or higher). Thus $\tilde{g}(u)$ satisfies (5.44)–(5.45) with 3 in place of p, and g_2 satisfies (5.44) with 2 in place of p, while $g_2''(u)$ is a constant, independent of u. For $p = 1$, one has $g(u) = g_1(u) + g_2(u) + \tilde{g}(u)$, where $g_1(u)$ is linear in u and $g_2, \tilde{g}$ are as above. Thus $g_1'(u)$ is a constant, independent of u, and $g_1''(u) = 0$.

By (5.43), (4.7), and (4.8), for all u in the ball $\|u\|_{L^\infty} \leq 1$, for all real $s \geq 0$, all integer $p \geq 1$, one has

$$\begin{aligned}
&\|P''(u)[h_1, h_2]\|_{H^s_\varepsilon} \\
&\quad\lesssim_s \varepsilon^{-1} \|u\|^{p-1}_{L^\infty} (\|h_1\|_{H^{s+1}_\varepsilon} \|h_2\|_{L^\infty} + \|h_1\|_{H^s_\varepsilon} \|h_2\|_{W^{1,\infty}_\varepsilon} \\
&\qquad\qquad + \|h_1\|_{W^{1,\infty}_\varepsilon} \|h_2\|_{H^s_\varepsilon} + \|h_1\|_{L^\infty} \|h_2\|_{H^{s+1}_\varepsilon}) \\
&\qquad + \varepsilon^{-1} \|u\|^{\nu}_{L^\infty} \|u\|_{W^{1,\infty}_\varepsilon} (\|h_1\|_{H^s_\varepsilon} \|h_2\|_{L^\infty} + \|h_1\|_{L^\infty} \|h_2\|_{H^s_\varepsilon}) \\
&\qquad + \varepsilon^{-1} \|u\|^{\nu}_{L^\infty} \|u\|_{H^s_\varepsilon} (\|h_1\|_{W^{1,\infty}_\varepsilon} \|h_2\|_{L^\infty} + \|h_1\|_{L^\infty} \|h_2\|_{W^{1,\infty}_\varepsilon}) \\
&\qquad + \varepsilon^{-1} (\|u\|^{\nu}_{L^\infty} \|u\|_{H^{s+1}_\varepsilon} + \|u\|^{\nu_3}_{L^\infty} \|u\|_{W^{1,\infty}_\varepsilon} \|u\|_{H^s_\varepsilon}) \|h_1\|_{L^\infty} \|h_2\|_{L^\infty}, \qquad (5.46)
\end{aligned}$$

where $\nu = \max\{p-2, 0\}$ has been defined in (5.24), and $\nu_3 := \max\{p-3, 0\}$.

Estimates in H^s_ε spaces only. For the result in the concentrating case, it is convenient to work directly in the H^s_ε class, avoiding the $W^{m,\infty}_\varepsilon$ spaces. Thus, by (5.4), one has

$$\|B(u, \varepsilon\partial_x)h\|_{H^s_\varepsilon} \leq \varepsilon^{-pd/2} (C_{s_0} \|u\|^p_{H^{s_0}_\varepsilon} \|h\|_{H^{s+1}_\varepsilon} + C_s \|u\|^{p-1}_{H^{s_0}_\varepsilon} \|u\|_{H^s_\varepsilon} \|h\|_{H^{s_0+1}_\varepsilon}) \quad (5.47)$$

for all $s \geq s_0 > d/2$, all u in the ball

$$C_{s_0} \varepsilon^{-d/2} \|u\|_{H^{s_0}_\varepsilon} \leq 1, \quad (5.48)$$

so that $\|u\|_{L^\infty} \leq 1$. By (5.47) and (5.12),

$$\|\operatorname{op}_\varepsilon(M)h\|_{H^s_\varepsilon} \leq \varepsilon^{-pd/2} (C_{s_0} \|u\|^p_{H^{s_0}_\varepsilon} \|h\|_{H^{s-1}_\varepsilon} + C_s \|u\|^{p-1}_{H^{s_0}_\varepsilon} \|u\|_{H^s_\varepsilon} \|h\|_{H^{s_0-1}_\varepsilon}) \quad (5.49)$$

for $s \geq s_0 > d/2$, u in the ball (5.48). Thus there exists $\rho_3 > 0$, independent of ε, such that for u in the ball

$$\varepsilon^{1-pd/2} \|u\|^p_{H^{s_0}_\varepsilon} \leq \rho_3, \quad (5.50)$$

one has

$$\|\varepsilon \operatorname{op}_\varepsilon(M)h\|_{H^{s_0}_\varepsilon} \leq C_{s_0} \varepsilon^{1-pd/2} \|u\|^p_{H^{s_0}_\varepsilon} \|h\|_{H^{s_0-1}_\varepsilon} \leq \frac{1}{2} \|h\|_{H^{s_0-1}_\varepsilon}. \quad (5.51)$$

Therefore, by Neumann series, $I + \varepsilon \operatorname{op}_\varepsilon(M)$ is invertible in $H^{s_0}(\mathbb{R}^d)$, and

$$\|(I + \varepsilon \operatorname{op}_\varepsilon(M))^{-1} h\|_{H^s_\varepsilon} \leq C_{s_0} \|h\|_{H^s_\varepsilon} + C_s \varepsilon^{1-pd/2} \|u\|^{p-1}_{H^{s_0}_\varepsilon} \|u\|_{H^s_\varepsilon} \|h\|_{H^{s_0-1}_\varepsilon} \quad (5.52)$$

for $s \geq s_0 > d/2$ and u in the ball (5.50). For u in the ball (5.50), for $s \geq s_0 > d/2$, we deduce the following estimates:

$$\begin{aligned} &\|(I + \varepsilon \operatorname{op}_\varepsilon(M))^{-1} \varepsilon \operatorname{op}_\varepsilon(M) \varepsilon^{-1} B_{\mathrm{r}}(u, \varepsilon\partial_x)\varphi\|_{H^s_\varepsilon} \\ &\qquad \lesssim_s \varepsilon^{-pd/2} \|u\|^{p-1}_{H^{s_0}_\varepsilon} (\|u\|_{H^{s_0+1}_\varepsilon} \|\varphi\|_{H^s_\varepsilon} + \|u\|_{H^s_\varepsilon} \|\varphi\|_{H^{s_0}_\varepsilon}) \end{aligned} \tag{5.53}$$

(to prove (5.53), we have used (B.22)),

$$\|\varepsilon^{-1} B_{\mathrm{lf}}(u, \varepsilon\partial_x)\varphi\|_{H^s_\varepsilon} \lesssim_s \varepsilon^{-1-pd/2} \|u\|^{p-1}_{H^{s_0}_\varepsilon} \|u\|_{H^s_\varepsilon} \|\varphi\|_{L^2}, \tag{5.54}$$

$$\begin{aligned} &\|\varepsilon \operatorname{op}_\varepsilon(\partial_t M)\varphi\|_{H^s_\varepsilon} \\ &\quad \lesssim_s \varepsilon^{1-pd/2} \big\{ \|u\|^{p-1}_{H^{s_0}_\varepsilon} \|\partial_t u\|_{H^{s_0}_\varepsilon} \|\varphi\|_{H^{s-1}_\varepsilon} \\ &\qquad\quad + (\|u\|^{p-1}_{H^{s_0}_\varepsilon} \|\partial_t u\|_{H^s_\varepsilon} + \|u\|^{\nu}_{H^{s_0}_\varepsilon} \|u\|_{H^s_\varepsilon} \|\partial_t u\|_{H^{s_0}_\varepsilon}) \|\varphi\|_{H^{s_0-1}_\varepsilon} \big\} \end{aligned} \tag{5.55}$$

with ν defined in (5.24), and

$$\|R_0(u)\varphi\|_{H^s_\varepsilon} \lesssim_s \varepsilon^{-1-pd/2} \|u\|^{p-1}_{H^{s_0}_\varepsilon} (\|u\|_{H^{s_0+1}_\varepsilon} \|\varphi\|_{H^s_\varepsilon} + \|u\|_{H^{s+1}_\varepsilon} \|\varphi\|_{H^{s_0}_\varepsilon}). \tag{5.56}$$

Hence

$$\begin{aligned} \|G(u)\varphi\|_{H^s_\varepsilon} \lesssim_s{}& \varepsilon^{-1-pd/2} \|u\|^{p-1}_{H^{s_0}_\varepsilon} (\|u\|_{H^{s_0+1}_\varepsilon} + \varepsilon^2 \|\partial_t u\|_{H^{s_0}_\varepsilon}) \|\varphi\|_{H^s_\varepsilon} \\ &+ \varepsilon^{-1-pd/2} \big\{ \|u\|^{p-1}_{H^{s_0}_\varepsilon} (\|u\|_{H^{s+1}_\varepsilon} + \varepsilon^2 \|\partial_t u\|_{H^s_\varepsilon}) \\ &\qquad + \varepsilon^2 \|u\|^{\nu}_{H^{s_0}_\varepsilon} \|u\|_{H^s_\varepsilon} \|\partial_t u\|_{H^{s_0}_\varepsilon} \big\} \|\varphi\|_{H^{s_0}_\varepsilon} \end{aligned} \tag{5.57}$$

for all $s \geq s_0$. By (5.30) and (B.27), for $X = \varepsilon^{-1} B_{\mathrm{r}}(u, \varepsilon\partial_x)$, for $s \geq s_0$, one has

$$\begin{aligned} &\|X + X^*\|_{\mathcal{L}(L^2, L^2)} \lesssim \varepsilon^{-1-pd/2} \|u\|^{p-1}_{H^{s_0}_\varepsilon} \|u\|_{H^{s_0+1}_\varepsilon}, \\ &\quad \|[\Lambda^s_\varepsilon, X]\varphi\|_{L^2} \lesssim_s \varepsilon^{-1-pd/2} \|u\|^{p-1}_{H^{s_0}_\varepsilon} (\|u\|_{H^{s_0+1}_\varepsilon} \|\varphi\|_{H^s_\varepsilon} + \|u\|_{H^{s+1}_\varepsilon} \|\varphi\|_{H^{s_0}_\varepsilon}), \\ &\quad |\mathrm{Re}\langle X\varphi, \varphi\rangle_{H^s_\varepsilon}| \lesssim_s \varepsilon^{-1-pd/2} \|u\|^{p-1}_{H^{s_0}_\varepsilon} (\|u\|_{H^{s_0+1}_\varepsilon} \|\varphi\|_{H^s_\varepsilon} \\ &\qquad\qquad + \|u\|_{H^{s+1}_\varepsilon} \|\varphi\|_{H^{s_0}_\varepsilon}) \|\varphi\|_{H^s_\varepsilon}. \end{aligned} \tag{5.58}$$

By (5.29), (5.57), and (5.58), we get energy estimates for φ: for u in the ball

$$\varepsilon^{-1-pd/2} \|u\|^p_{C^1_\varepsilon H^{s_0+2}_\varepsilon} \leq 1, \tag{5.59}$$

the solution φ of the linear Cauchy problem (5.20) satisfies

$$\|\varphi\|_{C^0 H^{s_0}_\varepsilon} \lesssim \|g_1\|_{C^0 H^{s_0}_\varepsilon} + \|g_2\|_{H^{s_0}_\varepsilon}, \tag{5.60}$$

$$\begin{aligned} \|\varphi\|_{C^0 H^s_\varepsilon} \lesssim_s{}& \|g_1\|_{C^0 H^s_\varepsilon} + \|g_2\|_{H^s_\varepsilon} \\ &+ \varepsilon^{-1-pd/2} \|u\|^{p-1}_{C^1_\varepsilon H^{s_0+2}_\varepsilon} \|u\|_{C^1_\varepsilon H^{s+2}_\varepsilon} (\|g_1\|_{C^0 H^{s_0}_\varepsilon} + \|g_2\|_{H^{s_0}_\varepsilon}) \end{aligned} \tag{5.61}$$

for all $s \ge s_0$. Hence, following the same argument as above, the solution h of the Cauchy problem (5.9) satisfies, for $s \ge s_0 + 2$,

$$\|h\|_{C^1_\varepsilon H^s_\varepsilon} \lesssim_s \|f_1\|_{C^0 H^s_\varepsilon} + \|f_2\|_{H^s_\varepsilon} + \varepsilon^{-1-pd/2} \|u\|^{p-1}_{C^1_\varepsilon H^{s_0+2}_\varepsilon} \|u\|_{C^1_\varepsilon H^{s+2}_\varepsilon} (\|f_1\|_{C^0 H^{s_0}_\varepsilon} + \|f_2\|_{H^{s_0}_\varepsilon}). \tag{5.62}$$

We have obtained the following inversion for the linear problem.

Lemma 5.2. *Let $s_0 > d/2$, $s \ge s_0 + 2$, and $u \in C([0,T], H^{s+2}(\mathbb{R}^d)) \cap C^1([0,T], H^s(\mathbb{R}^d))$, with (5.48), (5.50), and (5.59). Then for all $f_1 \in C([0,T], H^s(\mathbb{R}^d))$, all $f_2 \in H^s(\mathbb{R}^d)$, the linear Cauchy problem (5.9) has a (unique) solution h, which satisfies (5.62).*

Also, by (5.46) and (5.48), for $s \ge s_0$,

$$\|P''(u)[h_1, h_2]\|_{H^s_\varepsilon} \lesssim_s \varepsilon^{-1-pd/2} \|u\|^{p-1}_{H^{s_0}_\varepsilon} (\|h_1\|_{H^{s+1}_\varepsilon} \|h_2\|_{H^{s_0}_\varepsilon} + \|h_1\|_{H^{s_0}_\varepsilon} \|h_2\|_{H^{s+1}_\varepsilon}) + \varepsilon^{-1-(\nu+2)d/2} \|u\|^{\nu}_{H^{s_0}_\varepsilon} \|u\|_{H^{s+1}_\varepsilon} \|h_1\|_{H^{s_0}_\varepsilon} \|h_2\|_{H^{s_0}_\varepsilon}. \tag{5.63}$$

6. Proof of Theorem 3.4

For $a \ge 0$ real, let

$$E_a := C([0,T], H^{s_0+a}(\mathbb{R}^d)) \cap C^1([0,T], H^{s_0+a-2}(\mathbb{R}^d)), \tag{6.1}$$

$$F_a := C([0,T], H^{s_0+a}(\mathbb{R}^d)) \times H^{s_0+a}(\mathbb{R}^d), \tag{6.2}$$

and, recalling the notation in (4.15), define

$$\|u\|_{E_a} := \|u\|_{C^1_\varepsilon H^{s_0+a}_\varepsilon}, \quad \|f\|_{F_a} = \|(f_1, f_2)\|_{F_a} := \|f_1\|_{C^0 H^{s_0+a}_\varepsilon} + \|f_2\|_{H^{s_0+a}_\varepsilon}. \tag{6.3}$$

Define the smoothing operators S_j, $j \in \mathbb{N}$, as the "semiclassical" crude Fourier truncations

$$S_j u(x) := (2\pi)^{-d/2} \int_{\varepsilon|\xi| \le 2^j} \hat{u}(\xi) e^{i\xi\cdot x}\, d\xi, \tag{6.4}$$

which satisfy all (A.2)–(A.8) with constants independent of ε. Define

$$\Phi(u) := \begin{pmatrix} \partial_t u + P(u) \\ u(0) \end{pmatrix}, \tag{6.5}$$

where $P(u)$ is defined in (5.8). For $\|u\|_{E_2} \le 1$, the second derivative of Φ satisfies (5.63), which gives, for all $a \ge 0$,

$$\|\Phi''(u)[h_1, h_2]\|_{F_a} \lesssim_s \varepsilon^{-1-pd/2} \|u\|^{p-1}_{E_0} (\|h_1\|_{E_{a+1}} \|h_2\|_{E_0} + \|h_1\|_{E_0} \|h_2\|_{E_{a+1}}) + \varepsilon^{-1-(\nu+2)d/2} \|u\|^{\nu}_{E_0} \|u\|_{E_{a+1}} \|h_1\|_{E_0} \|h_2\|_{E_0}. \tag{6.6}$$

For u in the ball

$$\|u\|_{E_2} \leq \varepsilon^q, \quad q := \frac{1}{p} + \frac{d}{2}, \tag{6.7}$$

conditions (5.48), (5.50), (5.59) are all satisfied for ε sufficiently small – more precisely, for $\varepsilon \in (0, \varepsilon_0]$, where $\varepsilon_0 := \min\{1, C_{s_0}^{-p}, \rho_3^{1/2}\}$, and C_{s_0}, ρ_3 are the constants in (5.48), (5.50), independent of ε. Then, for u in the ball (6.7), Lemma 5.2 defines a right inverse $\Psi(u)$ of the linearized operator $\Phi'(u)$ (namely $h = \Psi(u) f$ solves the linear Cauchy problem $\Phi'(u)h = f$, which is (5.9)), with bound (5.62), which is

$$\|\Psi(u) f\|_{E_a} \lesssim_s \|f\|_{F_a} + \varepsilon^{-1-pd/2}\|u\|_{E_2}^{p-1}\|u\|_{E_{a+2}}\|f\|_{F_0}, \quad a \geq 2. \tag{6.8}$$

To reach the best radius for the initial data (see Remarks 7.3 and 7.4), we introduce the rescaled norm

$$\|u\|_{\mathcal{E}_a} := \varepsilon^{-q}\|u\|_{E_a}. \tag{6.9}$$

Thus (6.7) becomes

$$\|u\|_{\mathcal{E}_2} \leq 1. \tag{6.10}$$

By (6.6) and (6.8), for all u in the unit ball (6.10) one has

$$\begin{aligned} \|\Phi''(u)[h_1, h_2]\|_{F_a} \lesssim_s \varepsilon^q(&\|h_1\|_{\mathcal{E}_{a+1}}\|h_2\|_{\mathcal{E}_0} + \|h_1\|_{\mathcal{E}_0}\|h_2\|_{\mathcal{E}_{a+1}} \\ &+ \|u\|_{\mathcal{E}_{a+1}}\|h_1\|_{\mathcal{E}_0}\|h_2\|_{\mathcal{E}_0}) \end{aligned} \tag{6.11}$$

for $a \geq 0$, because $-1 - (\nu + 2)d/2 + q(\nu + 3) \geq q$ (recall that $\nu = \max\{p - 2, 0\}$), and

$$\|\Psi(u) f\|_{\mathcal{E}_a} \lesssim_s \varepsilon^{-q}(\|f\|_{F_a} + \|u\|_{\mathcal{E}_{a+2}}\|f\|_{F_0}) \tag{6.12}$$

for $a \geq 2$. Hence Φ satisfies the assumptions of Theorem A.1 with

$$\begin{gathered} a_0 = 0, \quad \mu = a_1 = 2, \quad \beta = \alpha > 4, \\ a_2 > 2\beta - 2, \quad U = \{u \in E_2 : \|u\|_{\mathcal{E}_2} \leq 1\}, \\ \delta_1 = 1, \quad M_1(a) = M_2(a) = C_a\varepsilon^q, \\ L_1(a) = L_2(a) = C_a\varepsilon^{-q}, \quad M_3(a) = L_3(a) = 0. \end{gathered} \tag{6.13}$$

For any function $u_0 = u_0(x) \in H^{s_0+\beta}(\mathbb{R}^d)$, the pair $g = (0, u_0) \in F_\beta$ trivially satisfies the first inequality in (A.12) with $A = 1$ (in fact, the inequality is an identity), because g does not depend on the time variable.

Hence, by Theorem A.1, if $\|g\|_{F_\beta} \leq \delta$, with $\delta = C\varepsilon^q$ given by (A.14), there exists $u \in E_\alpha$ such that $\Phi(u) = \Phi(0) + g = g$. This means that we have solved the nonlinear Cauchy problem (5.7), i.e. $\Phi(u) = (0, u_0)$, on the time interval $[0, T]$ for all initial data u_0 in the ball

$$\|u_0\|_{H_\varepsilon^{s_0+\beta}} \leq \delta = C\varepsilon^q, \tag{6.14}$$

for all $\varepsilon \in (0, \varepsilon_0]$. By (A.13), the solution u satisfies

$$\|u\|_{\mathcal{E}_\alpha} \leq C\varepsilon^{-q}\|g\|_{F_\beta}, \quad \text{i.e. } \|u\|_{C_\varepsilon^1 H_\varepsilon^{s_0+\beta}} \leq C\|u_0\|_{H_\varepsilon^{s_0+\beta}}.$$

The higher regularity part of Theorem 3.4 is also deduced from Theorem A.1.

For data u_0 of the form $u_0(x) = \varepsilon^\sigma(\mathrm{a}_\varepsilon(x), \overline{\mathrm{a}_\varepsilon(x)})$ (see (5.8)), where a_ε is defined in (3.5), one has

$$\|u_0\|_{H^s_\varepsilon} = \varepsilon^\sigma \|\mathrm{a}_\varepsilon\|_{H^s_\varepsilon} \lesssim_s \varepsilon^{\sigma+\sigma_{\mathrm{a}}} \|\mathrm{a}\|_{H^s}$$

(see (7.3), (7.1)), where $\sigma_{\mathrm{a}} = d/2$ in the concentrating case, and $\sigma_{\mathrm{a}} = 0$ in the fast oscillating case. Hence u_0 belongs to the ball (6.14) for all ε sufficiently small if

$$\|u_0\|_{H^{s_0+\beta}_\varepsilon} \leq C_{s_0+\beta}\,\varepsilon^{\sigma+\sigma_{\mathrm{a}}} \|\mathrm{a}\|_{H^{s_0+\beta}} \leq \delta = C\varepsilon^q.$$

For $\|\mathrm{a}\|_{H^{s_0+\beta}} \leq 1$, this holds for $\sigma + \sigma_{\mathrm{a}} > q$, namely

$$\sigma > \frac{1}{p} + \frac{d}{2} - \sigma_{\mathrm{a}}.$$

Finally, given $s_1 > d/2 + 4$, we define $\gamma := s_1 - (d/2 + 4)$, $s_0 := d/2 + \gamma/2$, $\beta := 4 + \gamma/2$, so that $s_0 > d/2$, $\beta > 4$, and $s_1 = s_0 + \beta$. This concludes the proof of Theorem 3.4.

Remark 6.1 (Confirmation of the heuristics discussion of Section 2 in Theorem 3.4). The radius δ given by the Nash–Moser Theorem A.1 is the minimum among $1/L$, δ_1/L, $1/(L^2M)$; here (see (6.13)) these three quantities are all of order ε^q. In particular, the "quadratic condition" $\delta \leq 1/(L^2M)$, coming from the use of the second derivative $\Phi''(u)$ in the Nash–Moser iteration, does not modify δ. This is a confirmation of the heuristic discussion of Section 2.

7. Free flow component decomposition

The "shifted map" trick used in [4, 11] consists in choosing the solution of the linear part of the PDE as a starting point for the Nash–Moser iteration. The reason the trick works is that the free flow of functions of special structure (3.5) satisfies better estimates in L^∞ norm than the free flow of general Sobolev functions. This, combined with the power p of the nonlinearity in the equation, makes it possible to obtain solutions of larger size, which are the sums of a free flow and a correction of smaller size.

Here we use this property in a different way, splitting the problem into components of special structure (3.5) and corrections, introducing nonisotropic norms to catch the different size effect.

For any function $\mathrm{a} \in H^s(\mathbb{R}^d)$ we define $\mathcal{T}_\varepsilon \mathrm{a}$, $0 < \varepsilon \leq 1$, as

$$(\mathcal{T}_\varepsilon \mathrm{a})(x) := \begin{cases} \mathrm{a}(x/\varepsilon) & \text{(concentrating case)}, \\ e^{ix\cdot\xi_0/\varepsilon}\mathrm{a}(x) & \text{(oscillating case)}, \end{cases} \tag{7.1}$$

so that, in both cases, (3.5) becomes $\mathrm{a}_\varepsilon = \mathcal{T}_\varepsilon \mathrm{a}_0$. To deal with conjugate pairs, define

$$\mathcal{T}_{\varepsilon,c}\mathrm{a} := (\mathcal{T}_\varepsilon \mathrm{a}, \overline{\mathcal{T}_\varepsilon \mathrm{a}}), \quad \mathcal{T}_{\varepsilon,c}^{-1}(\mathrm{b}, \bar{\mathrm{b}}) := \mathcal{T}_\varepsilon^{-1}\mathrm{b}.$$

Hence the initial datum u_0 defined in (5.8) can be written as $u_0 = \varepsilon^\sigma \mathcal{T}_{\varepsilon,c}\mathrm{a}_0$.

Lemma 7.1. *Let* $\mathrm{a} \in H^s(\mathbb{R}^d)$*,* $s \geq 0$*. Then the Fourier transform of* $\mathcal{T}_\varepsilon \mathrm{a}$ *is*

$$\widehat{(\mathcal{T}_\varepsilon \mathrm{a})}(\xi) = \varepsilon^d \hat{\mathrm{a}}(\varepsilon\xi) \ \ \textit{(concentrating)}, \quad \widehat{(\mathcal{T}_\varepsilon \mathrm{a})}(\xi) = \hat{\mathrm{a}}(\xi - \xi_0/\varepsilon) \ \ \textit{(oscillating)}, \tag{7.2}$$

and one has

$$\|\mathcal{T}_\varepsilon \mathrm{a}\|_{H^s_\varepsilon} \leq \varepsilon^{\sigma_\mathrm{a}}(2\|\mathrm{a}\|_{H^s} + C_s\|\mathrm{a}\|_{L^2}), \tag{7.3}$$

where

$$\sigma_\mathrm{a} = d/2 \ \ \textit{(concentrating)}, \quad \sigma_\mathrm{a} = 0 \ \ \textit{(oscillating)}. \tag{7.4}$$

Proof. Formula (7.2) is a direct calculation. Then, in the concentrating case, $\|\mathcal{T}_\varepsilon \mathrm{a}\|_{H^s_\varepsilon} = \varepsilon^{d/2}\|\mathrm{a}\|_{H^s}$. In the oscillating case, using the change of variable $\xi - \xi_0/\varepsilon = \eta$ and applying (B.10), one has $\|\mathcal{T}_\varepsilon \mathrm{a}\|_{H^s_\varepsilon} \leq 2\|\mathrm{a}\|_{H^s} + C_s|\xi_0|^s\|\mathrm{a}\|_{L^2}$. ■

Given any $y_0 \in H^s(\mathbb{R}^d)$, let $y = \mathcal{S} y_0$ denote the solution of the linear Cauchy problem

$$\begin{cases} \partial_t y + i\varepsilon^{-2} A(\varepsilon\partial_x) y = 0, \\ y(0,x) = y_0(x), \end{cases} \tag{7.5}$$

so that $\mathcal{S}$ is the free Schrödinger solution map. For initial data of type $\mathcal{T}_{\varepsilon,c}\mathrm{a}$, the flow $\mathcal{S}\mathcal{T}_{\varepsilon,c}\mathrm{a}$ has special properties, which are used in [11, proof of Theorem 4.6], which we recall in the following lemma.

Lemma 7.2. *For all real* $s \geq 0$*,* $s_0 > d/2$*, all multi-indices* $\alpha \in \mathbb{N}^d$*, for all* $t \in \mathbb{R}$ *the solution*

$$y = \mathcal{S}\mathcal{T}_{\varepsilon,c}\mathrm{a}$$

of the linear Cauchy problem (7.5) *with initial datum* $y_0 = \mathcal{T}_{\varepsilon,c}\mathrm{a}$ *satisfies*

$$\|y(t)\|_{L^\infty} \leq C_{s_0}\|\mathrm{a}\|_{H^{s_0}}, \tag{7.6}$$

$$\varepsilon^2\|\partial_t y(t)\|_{L^\infty} \leq C_{s_0}\|\mathrm{a}\|_{H^{s_0+2}}, \tag{7.7}$$

$$\varepsilon^{|\alpha|}\|\partial_x^\alpha y(t)\|_{L^\infty} \leq C_{|\alpha|,s_0}\|\mathrm{a}\|_{H^{s_0+|\alpha|}}, \tag{7.8}$$

$$\|y(t)\|_{H^s_\varepsilon} = \|\mathcal{T}_\varepsilon \mathrm{a}\|_{H^s_\varepsilon}. \tag{7.9}$$

Proof. At each t one has $|y(t,x)| \lesssim \|\hat{y}(t,\cdot)\|_{L^1}$ by the inverse Fourier formula, and $|\hat{y}(t,\xi)| = |\hat{y}(0,\xi)| = |\widehat{(\mathcal{T}_{\varepsilon,c}\mathrm{a})}|$ for all t, ξ because y solves (7.5). By (7.2), one has $\|\widehat{(\mathcal{T}_\varepsilon \mathrm{a})}\|_{L^1} = \|\hat{\mathrm{a}}\|_{L^1}$ in both cases. This proves (7.6) because, by Hölder's inequality, $\|\hat{\mathrm{a}}\|_{L^1} \lesssim_{s_0} \|\mathrm{a}\|_{H^{s_0}}$.

To prove (7.7) we use the equation in (7.5) recalling that $\varepsilon^{-2}A(\varepsilon\partial_x) = A(\partial_x)$. Proceeding as above, we get $|\partial_t y(t,x)| \lesssim \int |\xi|^2 |\widehat{(\mathcal{T}_\varepsilon \mathrm{a})}(\xi)|\, d\xi$, and then we use (7.2) to conclude. Similarly, (7.8) follows from $|\partial_x^\alpha y(t,x)| \lesssim \int |\xi|^{|\alpha|} |\widehat{(\mathcal{T}_\varepsilon \mathrm{a})}(\xi)|\, d\xi$. Finally, (7.9) is trivial. ■

We look for a solution of the Cauchy problem (5.7) by decomposing the unknown u into the sum of the solution of the free Schrödinger equation with initial datum u_0 of the form (5.8) and a "correction" $\tilde{u}(t,x)$ of smaller size.

For any pair $(\mathrm{a}, \tilde{u})$ where $\mathrm{a} = \mathrm{a}(x) \in H^s(\mathbb{R}^d)$ and $\tilde{u} = \tilde{u}(t,x) \in C^0([0,T], H^s(\mathbb{R}^d)) \cap C^1([0,T], H^{s-2}(\mathbb{R}^d))$ with $\tilde{u}(0,x) = 0$, we define

$$\widetilde{\Phi}(\mathrm{a}, \tilde{u}) := \begin{pmatrix} \partial_t u + P(u) \\ \mathrm{a} \end{pmatrix}, \quad \text{where } u = \varepsilon^\sigma \mathcal{S}\mathcal{T}_{\varepsilon,c}\mathrm{a} + \tilde{u}. \tag{7.10}$$

At time $t = 0$ the function u in (7.10) satisfies $u(0) = \varepsilon^\sigma \mathcal{T}_{\varepsilon,c}\mathrm{a}$. Hence the Cauchy problem (5.7) becomes

$$\widetilde{\Phi}(\mathrm{a}, \tilde{u}) = (0, \mathrm{a}_0). \tag{7.11}$$

We solve (7.11) by applying our Nash–Moser–Hörmander theorem; therefore we have to construct a right inverse for the linearized operator and to estimate the second derivative. We only have to adapt the general analysis of Section 5 to functions u of the form (7.10).

Right inverse of the linearized operator. The differential of $\widetilde{\Phi}$ at the point $(\mathrm{a}, \tilde{u})$ in the direction $(\mathrm{b}, \tilde{h})$ is

$$\widetilde{\Phi}'(\mathrm{a}, \tilde{u})(\mathrm{b}, \tilde{h}) = \begin{pmatrix} \partial_t h + P'(u)h \\ \mathrm{b} \end{pmatrix},$$
$$\text{where } u = \varepsilon^\sigma \mathcal{S}\mathcal{T}_{\varepsilon,c}\mathrm{a} + \tilde{u}, \quad h = \varepsilon^\sigma \mathcal{S}\mathcal{T}_{\varepsilon,c}\mathrm{b} + \tilde{h}, \tag{7.12}$$

and $\tilde{u}(0) = 0$, $\tilde{h}(0) = 0$. Given $(\mathrm{a}, \tilde{u})$ and $g = (g_1, g_2)$, with $g_1 = g_1(t,x)$ and $g_2 = g_2(x)$, the right inversion problem for the linearized operator $\widetilde{\Phi}'(\mathrm{a}, \tilde{u})$ consists in finding $(\mathrm{b}, \tilde{h})$ such that

$$\widetilde{\Phi}'(\mathrm{a}, \tilde{u})(\mathrm{b}, \tilde{h}) = g, \quad \text{i.e.} \begin{cases} \partial_t h + P'(u)h = g_1, \\ \mathrm{b} = g_2 \end{cases} \tag{7.13}$$

with u, h as in (7.12). Since the free flow $\varepsilon^\sigma \mathcal{S}\mathcal{T}_{\varepsilon,c}\mathrm{b} = \varepsilon^\sigma \mathcal{S}\mathcal{T}_{\varepsilon,c} g_2$ solves (7.5), and $\tilde{h}(0) = 0$ by construction, (7.13) is equivalent to the following problem for $\tilde{h}$:

$$\begin{cases} \partial_t \tilde{h} + P'(u)\tilde{h} = g_1 + \varepsilon^{-1} B(u, \varepsilon\partial_x)\varepsilon^\sigma \mathcal{S}\mathcal{T}_{\varepsilon,c} g_2 - R_0(u)\varepsilon^\sigma \mathcal{S}\mathcal{T}_{\varepsilon,c} g_2, \\ \tilde{h}(0) = 0, \end{cases} \tag{7.14}$$

namely $\tilde{h}$ has to solve the linear Cauchy problem (5.9) with

$$f_1 = g_1 + \varepsilon^{-1} B(u, \varepsilon\partial_x)\varepsilon^\sigma \mathcal{S}\mathcal{T}_{\varepsilon,c} g_2 - R_0(u)\varepsilon^\sigma \mathcal{S}\mathcal{T}_{\varepsilon,c} g_2, \quad f_2 = 0. \tag{7.15}$$

The solution of (5.9) is estimated in Lemma 5.1; to apply that lemma, now we check that u satisfies its hypotheses. By Lemma 7.2, (4.6), (4.9), (4.15), and (4.16), the function $u = \varepsilon^\sigma \mathcal{S}\mathcal{T}_{\varepsilon,c}\mathrm{a} + \tilde{u}$ satisfies

$$\|u\|_{C^1_\varepsilon W^m_\varepsilon} \lesssim_{s_0,m} \varepsilon^\sigma \|\mathrm{a}\|_{H^{s_0+m}} + \varepsilon^{-d/2} \|\tilde{u}\|_{C^1_\varepsilon H^{s_0+m}_\varepsilon}, \quad m \in \mathbb{N}. \tag{7.16}$$

For all s, let

$$\|(\mathrm{a}, \tilde{u})\|_{X^s} := \varepsilon^\sigma \|\mathrm{a}\|_{H^s} + \varepsilon^{-d/2} \|\tilde{u}\|_{C^1_\varepsilon H^s_\varepsilon}. \tag{7.17}$$

By (7.16), one has, in particular,

$$\|u\|_{W_\varepsilon^{2,\infty}} + \varepsilon^2 \|\partial_t u\|_{L^\infty} \lesssim_{s_0} \|(\mathrm{a}, \tilde{u})\|_{X^{s_0+2}}, \tag{7.18}$$

and therefore there exists $\rho_1 \in (0, 1]$, depending only on s_0 and on the nonlinearity of the problem, such that, for $(\mathrm{a}, \tilde{u})$ in the ball

$$\varepsilon^{-1} \|(\mathrm{a}, \tilde{u})\|^p_{X^{s_0+2}} \le \rho_1, \tag{7.19}$$

the function $u = \varepsilon^\sigma \mathcal{S} \mathcal{T}_{\varepsilon,c} \mathrm{a} + \tilde{u}$ satisfies (5.34) and (5.16). Hence Lemma 5.1 applies, and $\tilde{h}$ satisfies bound (5.42). Moreover, assuming (7.19), the factor in u appearing in (5.42) satisfies

$$\begin{aligned}
&(\|u\|^{p-1}_{C^0 L^\infty} \|u\|_{C^1_\varepsilon W_\varepsilon^{[s]+3}} + \|u\|^\nu_{C^0 L^\infty} \|u\|_{C^0 W_\varepsilon^{[s]+1}} \|u\|_{C^1_\varepsilon W_\varepsilon^2}) \\
&\quad \lesssim_s \|(\mathrm{a}, \tilde{u})\|^{p-1}_{X^{s_0+2}} \|(\mathrm{a}, \tilde{u})\|_{X^{[s]+s_0+3}} + \|(\mathrm{a}, \tilde{u})\|^{\nu+1}_{X^{s_0+2}} \|(\mathrm{a}, \tilde{u})\|_{X^{[s]+s_0+1}} \\
&\quad \lesssim_s \|(\mathrm{a}, \tilde{u})\|^{p-1}_{X^{s_0+2}} \|(\mathrm{a}, \tilde{u})\|_{X^{s+s_0+3}}
\end{aligned} \tag{7.20}$$

because $[s] \le s$, $\nu + 1 = \max\{p-2, 0\} + 1 \ge p - 1$ and $\|(\mathrm{a}, \tilde{u})\|_{X^{s_0+2}} \le 1$.

Thus we have to estimate f_1 in (7.15). By (5.5) and (5.25), using (7.18), (7.16), (7.17), (7.9), and Lemma 7.1, for all $s \ge 0$ one has

$$\begin{aligned}
&\|\varepsilon^{-1} B(u, \varepsilon\partial_x) \varepsilon^\sigma \mathcal{S} \mathcal{T}_{\varepsilon,c} g_2\|_{H^s_\varepsilon} \\
&\quad \lesssim_s \varepsilon^{\sigma+\sigma_{\mathrm{a}}-1} \|(\mathrm{a}, \tilde{u})\|^p_{X^{s_0+2}} \|g_2\|_{H^{s+1}} \\
&\qquad + \varepsilon^{\sigma+\sigma_{\mathrm{a}}-1} \|(\mathrm{a}, \tilde{u})\|^{p-1}_{X^{s_0+2}} \|(\mathrm{a}, \tilde{u})\|_{X^{[s]+s_0+1}} \|g_2\|_{H^1},
\end{aligned} \tag{7.21}$$

$$\begin{aligned}
&\|R_0(u) \varepsilon^\sigma \mathcal{S} \mathcal{T}_{\varepsilon,c} g_2\|_{H^s_\varepsilon} \\
&\quad \lesssim_s \varepsilon^{\sigma+\sigma_{\mathrm{a}}-1} \|(\mathrm{a}, \tilde{u})\|^p_{X^{s_0+2}} \|g_2\|_{H^s} \\
&\qquad + \varepsilon^{\sigma+\sigma_{\mathrm{a}}-1} \|(\mathrm{a}, \tilde{u})\|^{p-1}_{X^{s_0+2}} \|(\mathrm{a}, \tilde{u})\|_{X^{[s]+s_0+2}} \|g_2\|_{L^2}.
\end{aligned} \tag{7.22}$$

By (7.15), (7.21), (7.22), (7.20), and Lemma 5.1, for $(\mathrm{a}, \tilde{u})$ in the ball (7.19), for $s \ge 1$ we obtain

$$\begin{aligned}
\|\tilde{h}\|_{C^1_\varepsilon H^s_\varepsilon} &\lesssim_s \|g_1\|_{C^0 H^s_\varepsilon} + \varepsilon^{-1} \|(\mathrm{a}, \tilde{u})\|^{p-1}_{X^{s_0+2}} \|(\mathrm{a}, \tilde{u})\|_{X^{s+s_0+3}} \|g_1\|_{C^0 L^2} \\
&\quad + \varepsilon^{\sigma+\sigma_{\mathrm{a}}-1} \|(\mathrm{a}, \tilde{u})\|^p_{X^{s_0+2}} \|g_2\|_{H^{s+1}} \\
&\quad + \varepsilon^{\sigma+\sigma_{\mathrm{a}}-1} \|(\mathrm{a}, \tilde{u})\|^{p-1}_{X^{s_0+2}} \|(\mathrm{a}, \tilde{u})\|_{X^{s+s_0+3}} \|g_2\|_{H^1}.
\end{aligned} \tag{7.23}$$

Since $\mathrm{b} = g_2$, we get

$$\begin{aligned}
\|(\mathrm{b}, \tilde{h})\|_{X^s} &= \varepsilon^\sigma \|\mathrm{b}\|_{H^s} + \varepsilon^{-d/2} \|\tilde{h}\|_{C^1_\varepsilon H^s_\varepsilon} \\
&\lesssim_s \varepsilon^{-d/2} \|g_1\|_{C^0 H^s_\varepsilon} + \varepsilon^{-1-d/2} \|(\mathrm{a}, \tilde{u})\|^{p-1}_{X^{s_0+2}} \|(\mathrm{a}, \tilde{u})\|_{X^{s+s_0+3}} \|g_1\|_{C^0 L^2} \\
&\quad + \varepsilon^\sigma (1 + \varepsilon^{\sigma_{\mathrm{a}}-1-d/2} \|(\mathrm{a}, \tilde{u})\|^p_{X^{s_0+2}}) \|g_2\|_{H^{s+1}} \\
&\quad + \varepsilon^{\sigma+\sigma_{\mathrm{a}}-1-d/2} \|(\mathrm{a}, \tilde{u})\|^{p-1}_{X^{s_0+2}} \|(\mathrm{a}, \tilde{u})\|_{X^{s+s_0+3}} \|g_2\|_{H^1}
\end{aligned} \tag{7.24}$$

for all $(\mathrm{a}, \tilde{u})$ in the ball (7.19). As explained in Remark 7.3 in general, and in Remark (7.4) for our specific problem, for $p > 1$ it is convenient

(i) to consider $(\mathrm{a}, \tilde{u})$ in the ball

$$\|(\mathrm{a}, \tilde{u})\|_{X^{s_0+2}} \le \rho_2 \varepsilon^{(1+d/2-\sigma_{\mathrm{a}})/p},$$
$$\text{i.e. } \varepsilon^{\sigma_{\mathrm{a}}-1-d/2} \|(\mathrm{a}, \tilde{u})\|^p_{X^{s_0+2}} \le \rho_1, \ \rho_2 := \rho_1^{1/p}, \tag{7.25}$$

which is smaller than the ball (7.19) if $\sigma_{\mathrm{a}} = 0$, and it is the same ball if $\sigma_{\mathrm{a}} = d/2$;

(ii) to rescale $\| \ \|_{X^s}$ so that (7.25) becomes a ball with radius $O(1)$ (i.e. independent of ε) in the rescaled norm.

Thus we define

$$\|(\mathrm{a}, \tilde{u})\|_{\mathcal{Z}^s} := \varepsilon^{(\sigma_{\mathrm{a}}-1-d/2)/p} \|(\mathrm{a}, \tilde{u})\|_{X^s}, \tag{7.26}$$

and (7.24) becomes

$$\begin{aligned} \|(\mathrm{b}, \tilde{h})\|_{\mathcal{Z}^s} \lesssim_s \ & \varepsilon^{-d/2+(\sigma_{\mathrm{a}}-1-d/2)/p} \|g_1\|_{C^0 H^s_\varepsilon} \\ & + \varepsilon^{\sigma+(\sigma_{\mathrm{a}}-1-d/2)/p} (1 + \|(\mathrm{a}, \tilde{u})\|^p_{\mathcal{Z}^{s_0+2}}) \|g_2\|_{H^{s+1}} \\ & + \varepsilon^{\sigma+(\sigma_{\mathrm{a}}-1-d/2)/p} \|(\mathrm{a}, \tilde{u})\|^{p-1}_{\mathcal{Z}^{s_0+2}} \|(\mathrm{a}, \tilde{u})\|_{\mathcal{Z}^{s+s_0+3}} \\ & \times (\varepsilon^{-\sigma-\sigma_{\mathrm{a}}} \|g_1\|_{C^0 L^2} + \|g_2\|_{H^1}) \end{aligned} \tag{7.27}$$

for all $s \ge 1$, all $(\mathrm{a}, \tilde{u})$ in the ball

$$\|(\mathrm{a}, \tilde{u})\|_{\mathcal{Z}^{s_0+2}} \le \rho_2. \tag{7.28}$$

Therefore, in the case $p > 1$,

$$\begin{aligned} \|(\mathrm{b}, \tilde{h})\|_{\mathcal{Z}^s} \lesssim_s \ & \varepsilon^{\sigma+(\sigma_{\mathrm{a}}-1-d/2)/p} \big\{ (\varepsilon^{-\sigma-d/2} \|g_1\|_{C^0 H^s_\varepsilon} + \|g_2\|_{H^{s+1}}) \\ & + \|(\mathrm{a}, \tilde{u})\|_{\mathcal{Z}^{s+s_0+3}} (\varepsilon^{-\sigma-\sigma_{\mathrm{a}}} \|g_1\|_{C^0 L^2} + \|g_2\|_{H^1}) \big\} \end{aligned} \tag{7.29}$$

for all $s \ge 1$, all $(\mathrm{a}, \tilde{u})$ in the ball (7.28).

For $p = 1$, the restriction to the ball (7.25) is not convenient (see Remarks 7.3 and 7.4), and we take, instead, u in the entire ball (7.19). Hence, for $p = 1$, we define

$$\|(\mathrm{a}, \tilde{u})\|_{Z^s} := \varepsilon^{-1} \|(\mathrm{a}, \tilde{u})\|_{X^s}, \tag{7.30}$$

and (7.24) becomes

$$\begin{aligned} \|(\mathrm{b}, \tilde{h})\|_{Z^s} \lesssim_s \ & \varepsilon^{-1-d/2} \|g_1\|_{C^0 H^s_\varepsilon} + \varepsilon^{-1-d/2} \|(\mathrm{a}, \tilde{u})\|_{Z^{s+s_0+3}} \|g_1\|_{C^0 L^2} \\ & + \varepsilon^{\sigma-1} (1 + \varepsilon^{\sigma_{\mathrm{a}}-d/2} \|(\mathrm{a}, \tilde{u})\|_{Z^{s_0+2}}) \|g_2\|_{H^{s+1}} \\ & + \varepsilon^{\sigma+\sigma_{\mathrm{a}}-1-d/2} \|(\mathrm{a}, \tilde{u})\|_{Z^{s+s_0+3}} \|g_2\|_{H^1} \end{aligned} \tag{7.31}$$

for all $(\mathrm{a}, \tilde{u})$ in the ball

$$\|(\mathrm{a}, \tilde{u})\|_{Z^{s_0+2}} \le \rho_2. \tag{7.32}$$

Therefore, in the case $p = 1$,

$$\|(\mathrm{b}, \tilde{h})\|_{Z^s} \lesssim_s \varepsilon^{\sigma+\sigma_{\mathrm{a}}-1-d/2}\big\{(\varepsilon^{-\sigma-\sigma_{\mathrm{a}}}\|g_1\|_{C^0 H^s_\varepsilon} + \|g_2\|_{H^{s+1}}) + \|(\mathrm{a}, \tilde{u})\|_{Z^{s+s_0+3}}(\varepsilon^{-\sigma-\sigma_{\mathrm{a}}}\|g_1\|_{C^0 L^2} + \|g_2\|_{H^1})\big\} \tag{7.33}$$

for all $s \geq 1$, all $(\mathrm{a}, \tilde{u})$ in the ball (7.32).

Note that we have used norms $\| \ \|_{Z^s}$ for $p = 1$ and norms $\| \ \|_{\mathcal{Z}^s}$ for $p > 1$.

Estimate for the second derivative. By (7.17), (7.16), (7.9), and Lemma 7.1, any function $u = \varepsilon^\sigma \mathcal{S}\mathcal{T}_{\varepsilon,c}\mathrm{a} + \tilde{u}$ satisfies

$$\begin{aligned} \|u\|_{H^s_\varepsilon} &\lesssim_s \varepsilon^{\sigma+\sigma_{\mathrm{a}}}\|\mathrm{a}\|_{H^s} + \|\tilde{u}\|_{H^s_\varepsilon} \lesssim_s \varepsilon^{\sigma_{\mathrm{a}}}\|(\mathrm{a}, \tilde{u})\|_{X^s}, \\ \|u\|_{L^\infty} &\lesssim \|(\mathrm{a}, \tilde{u})\|_{X^{s_0}}, \\ \|u\|_{W^{1,\infty}_\varepsilon} &\lesssim \|(\mathrm{a}, \tilde{u})\|_{X^{s_0+1}}. \end{aligned}$$

From (5.46) we deduce that

$$\begin{aligned} &\|P''(u)[h_1, h_2]\|_{H^s_\varepsilon} \\ &\quad \lesssim_s \varepsilon^{\sigma_{\mathrm{a}}-1}\|(\mathrm{a}, \tilde{u})\|^{p-1}_{X^{s_0}}(\|(\mathrm{b}_1, \tilde{h}_1)\|_{X^{s+1}}\|(\mathrm{b}_2, \tilde{h}_2)\|_{X^{s_0}} + \|(\mathrm{b}_1, \tilde{h}_1)\|_{X^{s_0}}\|(\mathrm{b}_2, \tilde{h}_2)\|_{X^{s+1}}) \\ &\quad\quad + \varepsilon^{\sigma_{\mathrm{a}}-1}\|(\mathrm{a}, \tilde{u})\|^{\nu}_{X^{s_0}}\|(\mathrm{a}, \tilde{u})\|_{X^{s+1}}\|(\mathrm{b}_1, \tilde{h}_1)\|_{X^{s_0}}\|(\mathrm{b}_2, \tilde{h}_2)\|_{X^{s_0}} \end{aligned} \tag{7.34}$$

for $u = \varepsilon^\sigma \mathcal{S}\mathcal{T}_{\varepsilon,c}\mathrm{a} + \tilde{u}$, $h_i = \varepsilon^\sigma \mathcal{S}\mathcal{T}_{\varepsilon,c}\mathrm{b}_i + \tilde{h}_i$, $i = 1, 2$, and $s \geq 0$.

With the norms $\| \ \|_{\mathcal{Z}^s}$ defined in (7.26), which we use in the case $p > 1$, from (7.34) we get

$$\begin{aligned} &\|P''(u)[h_1, h_2]\|_{H^s_\varepsilon} \\ &\quad \lesssim_s \varepsilon^{d/2+(1+d/2-\sigma_{\mathrm{a}})/p}\|(\mathrm{a}, \tilde{u})\|^{p-1}_{\mathcal{Z}^{s_0}} \\ &\quad\quad \times (\|(\mathrm{b}_1, \tilde{h}_1)\|_{\mathcal{Z}^{s+1}}\|(\mathrm{b}_2, \tilde{h}_2)\|_{\mathcal{Z}^{s_0}} + \|(\mathrm{b}_1, \tilde{h}_1)\|_{\mathcal{Z}^{s_0}}\|(\mathrm{b}_2, \tilde{h}_2)\|_{\mathcal{Z}^{s+1}}) \\ &\quad\quad + \varepsilon^{d/2+(1+d/2-\sigma_{\mathrm{a}})/p}\|(\mathrm{a}, \tilde{u})\|^{\nu}_{\mathcal{Z}^{s_0}}\|(\mathrm{a}, \tilde{u})\|_{\mathcal{Z}^{s+1}}\|(\mathrm{b}_1, \tilde{h}_1)\|_{\mathcal{Z}^{s_0}}\|(\mathrm{b}_2, \tilde{h}_2)\|_{\mathcal{Z}^{s_0}}. \end{aligned}$$

Hence, for $(\mathrm{a}, \tilde{u})$ in the ball (7.28), for $s \geq 0$, in the case $p > 1$, one has

$$\begin{aligned} \|P''(u)[h_1, h_2]\|_{H^s_\varepsilon} \lesssim_s \varepsilon^{d/2+(1+d/2-\sigma_{\mathrm{a}})/p}\big\{&\|(\mathrm{b}_1, \tilde{h}_1)\|_{\mathcal{Z}^{s+1}}\|(\mathrm{b}_2, \tilde{h}_2)\|_{\mathcal{Z}^{s_0}} \\ &+ \|(\mathrm{b}_1, \tilde{h}_1)\|_{\mathcal{Z}^{s_0}}\|(\mathrm{b}_2, \tilde{h}_2)\|_{\mathcal{Z}^{s+1}} \\ &+ \|(\mathrm{a}, \tilde{u})\|_{\mathcal{Z}^{s+1}}\|(\mathrm{b}_1, \tilde{h}_1)\|_{\mathcal{Z}^{s_0}}\|(\mathrm{b}_2, \tilde{h}_2)\|_{\mathcal{Z}^{s_0}}\big\}. \end{aligned} \tag{7.35}$$

For $p = 1$, with the norms $\| \ \|_{Z^s}$ defined in (7.30), for $(\mathrm{a}, \tilde{u})$ in the ball (7.32), for $s \geq 0$, one has

$$\begin{aligned} \|P''(u)[h_1, h_2]\|_{H^s_\varepsilon} \lesssim_s \varepsilon^{\sigma_{\mathrm{a}}+1}\big\{&\|(\mathrm{b}_1, \tilde{h}_1)\|_{Z^{s+1}}\|(\mathrm{b}_2, \tilde{h}_2)\|_{Z^{s_0}} \\ &+ \|(\mathrm{b}_1, \tilde{h}_1)\|_{Z^{s_0}}\|(\mathrm{b}_2, \tilde{h}_2)\|_{Z^{s+1}} \\ &+ \|(\mathrm{a}, \tilde{u})\|_{Z^{s+1}}\|(\mathrm{b}_1, \tilde{h}_1)\|_{Z^{s_0}}\|(\mathrm{b}_2, \tilde{h}_2)\|_{Z^{s_0}}\big\}. \end{aligned} \tag{7.36}$$

Remark 7.3 (Best rescaling for Nash–Moser application). In this remark we discuss a general, simple way to choose the best rescaling to obtain the largest size ball for the solution when applying the Nash–Moser Theorem A.1 (or essentially any other Nash–Moser theorem).

Suppose we have a nonlinear operator Φ and a right inverse $\Psi(u)$ of its linearized operator $\Phi'(u)$, satisfying an estimate of the form

$$\|\Psi(u)g\|_{X^s} \leq (A + B\|u\|_{X^{s_0}}^p)\|g\|_{Y^s} + C\|u\|_{X^{s_0}}^{p-1}\|u\|_{X^s}\|g\|_{Y^{s_0}} \tag{7.37}$$

for all u in a low norm ball

$$\|u\|_{X^{s_0}} \leq R \tag{7.38}$$

for some positive constants A, B, C, R, where $\| \, \|_{X^s}$ are the norms on the domain of Φ, $\| \, \|_{Y^s}$ are those on its codomain, and s denotes high norms, while s_0 denotes low norms (we ignore any possible loss of regularity, which is not the point in this discussion). From (7.37), (7.38) we deduce the bound

$$\|\Psi(u)g\|_{X^s} \leq (A + BR^p)\|g\|_{Y^s} + CR^{p-1}\|u\|_{X^s}\|g\|_{Y^{s_0}} \tag{7.39}$$

for u in the ball (7.38). Then Theorem A.1 gives a solution of the problem $\Phi(u) = \Phi(0) + g$ for all data g in the ball

$$\|g\|_{Y^{s_0}} \leq \delta, \tag{7.40}$$

where (ignoring, at least for the moment, the contribution to δ coming from the second derivative $\Phi''(u)[h_1, h_2]$ of the operator Φ) the radius δ is essentially given by

$$\delta = \min\Big\{\frac{1}{L}, \frac{R}{L}\Big\}, \quad L = A + BR^p + CR^{p-1}. \tag{7.41}$$

Our goal is to find the best (i.e. the largest possible) radius δ that we can obtain in this situation.

First, we consider a rescaling of the norm $\| \, \|_{X^s}$: for any λ positive, let

$$\lambda\|u\|_{X^s} =: \|u\|_{Z^s}. \tag{7.42}$$

Then (7.37), (7.38) become

$$\|\Psi(u)g\|_{Z^s} \leq (A\lambda + B\lambda^{1-p}\|u\|_{Z^{s_0}}^p)\|g\|_{Y^s} + C\lambda^{1-p}\|u\|_{Z^{s_0}}^{p-1}\|u\|_{Z^s}\|g\|_{Y^{s_0}} \tag{7.43}$$

for all u in the rescaled ball

$$\|u\|_{Z^{s_0}} \leq R\lambda. \tag{7.44}$$

From (7.43), (7.44) we get the bound

$$\|\Psi(u)g\|_{Z^s} \leq (A\lambda + B\lambda R^p)\|g\|_{Y^s} + CR^{p-1}\|u\|_{Z^s}\|g\|_{Y^{s_0}} \tag{7.45}$$

for u in the ball (7.44). Then Theorem A.1 solves the nonlinear problem for all data g in the ball

$$\|g\|_{Y^{s_0}} \leq \delta(\lambda), \tag{7.46}$$

where now the radius is

$$\delta(\lambda) = \min\Big\{ \frac{R\lambda}{L(\lambda)}, \frac{1}{L(\lambda)} \Big\}, \quad L(\lambda) = \lambda(A + BR^p) + CR^{p-1}. \tag{7.47}$$

For $\lambda \geq 1/R$, one has

$$\delta(\lambda) = \frac{1}{L(\lambda)} = \frac{1}{\lambda(A + BR^p) + CR^{p-1}}, \tag{7.48}$$

which is a decreasing function of λ, so that $\delta(\lambda) \leq \delta(1/R)$ for all $\lambda \geq 1/R$. For $0 < \lambda \leq 1/R$, one has

$$\delta(\lambda) = \frac{R\lambda}{L(\lambda)} = \frac{R\lambda}{\lambda(A + BR^p) + CR^{p-1}} = \frac{R}{A + BR^p + CR^{p-1}\lambda^{-1}}, \tag{7.49}$$

which is an increasing function of λ, so that $\delta(\lambda) \leq \delta(1/R)$ for all $\lambda \in (0, 1/R]$. In other words, the largest radius $\delta(\lambda)$ we can get by the rescaling (7.42) is attained at $\lambda = 1/R$. Note that $\lambda = 1/R$ is the value of λ corresponding to the unit ball $\|u\|_{Z^{s_0}} \leq 1$ in the rescaled norm (7.44). For $\lambda = 1/R$ we get the radius

$$\delta_R := \delta(1/R) = \frac{1}{AR^{-1} + (B + C)R^{p-1}}. \tag{7.50}$$

Second, we check whether taking u in a smaller ball can give a better balance among the constants, and therefore a larger radius for the data. From (7.37), (7.38) we deduce that, for every $r \in (0, R]$,

$$\|\Psi(u)g\|_{X^s} \leq (A + Br^p)\|g\|_{Y^s} + Cr^{p-1}\|u\|_{X^s}\|g\|_{Y^{s_0}} \tag{7.51}$$

for all u in the ball

$$\|u\|_{X^{s_0}} \leq r. \tag{7.52}$$

Apply the best rescaling of the form (7.42), which is

$$\frac{1}{r}\|u\|_{X^{s_0}} =: \|u\|_{Z^s}. \tag{7.53}$$

Then, by the discussion above, we obtain the radius

$$\delta_r = \delta(1/r) = \frac{1}{Ar^{-1} + (B + C)r^{p-1}}. \tag{7.54}$$

To maximize the radius δ_r in (7.54), we minimize its denominator $\varphi(r) := Ar^{-1} + (B + C)r^{p-1}$ over $r \in (0, R]$. For $p = 1$, φ is decreasing in $(0, \infty)$, and then the largest δ_r is attained at the largest r, namely $r = R$. For $p > 1$, φ is decreasing in $(0, r_0)$ and increasing in (r_0, ∞), where

$$r_0 := \Big(\frac{A}{(p-1)(B + C)} \Big)^{\frac{1}{p}}. \tag{7.55}$$

Hence $\min\{\varphi(r) : r \in (0, R]\}$ is attained at $r = r_0$ if $r_0 \le R$, and at $r = R$ if $R \le r_0$, namely at $r = \min\{r_0, R\}$ in both cases. Therefore the best radius is

$$\max_{r\in(0,R]} \delta_r = \begin{cases} \delta_R & \text{for } p = 1, \\ \delta_R & \text{for } p > 1 \text{ and } R \le r_0, \\ \delta_{r_0} & \text{for } p > 1 \text{ and } r_0 \le R. \end{cases} \tag{7.56}$$

In fact, to apply the result of this discussion to a specific operator, the only point one has to check is whether $r_0 \le R$ or vice versa.

In this way we get the best radius ignoring the contribution coming from $\Phi''(u)$, which is a condition of the form $\delta \le M^{-1}L^{-2}$ (see Theorem A.1). Then one has to check that introducing this additional constraint to the radius δ does not change its optimal size. The heuristic discussion of Section 2 shows that, in many situations, this is the case.

Remark 7.4. We see how the discussion of Remark 7.3 applies to our specific problem.

By (7.19) (ignoring the harmless constant ρ_1) and (7.24) (ignoring g_1, which will be zero in the datum of the original nonlinear problem) one has

$$A \sim \varepsilon^{\sigma}, \quad B \sim C \sim \varepsilon^{\sigma+\sigma_{\mathrm{a}}-1-d/2}, \quad R \sim \varepsilon^{1/p}.$$

This gives $r_0 \sim \varepsilon^{(1+d/2-\sigma_{\mathrm{a}})/p} \lesssim R$, and therefore the best choice is to restrict u to the smaller ball $\|u\|_{X^{s_0+2}} \lesssim r_0$ and then to rescale as in (7.26), corresponding to $\lambda = 1/r_0$.

In the previous case, by (5.59) and (5.62) one has

$$A \sim 1, \quad B + C \sim \varepsilon^{-pq}, \quad R \sim \varepsilon^{q},$$

with $q = 1/p + d/2$. This gives $r_0 \sim \varepsilon^q \sim R$, and therefore the best rescaling for the linearized operator is (6.9), corresponding to $\lambda = 1/R$.

8. Proof of Theorem 3.5

Let $p > 1$, and define

$$E_{a,1} := H^{s_0+a}(\mathbb{R}^d), \tag{8.1}$$

$$E_{a,2} := \big\{\tilde u \in C([0,T], H^{s_0+a}(\mathbb{R}^d)) \cap C^1([0,T], H^{s_0+a-2}(\mathbb{R}^d)) : \tilde u(0,x) = 0\big\}, \tag{8.2}$$

$$E_a := E_{a,1} \times E_{a,2}, \tag{8.3}$$

$$F_{a,1} := C([0,T], H^{s_0+a}(\mathbb{R}^d)), \tag{8.4}$$

$$F_{a,2} := H^{s_0+a+1}(\mathbb{R}^d), \tag{8.5}$$

$$F_a := F_{a,1} \times F_{a,2}. \tag{8.6}$$

We consider norms (7.26) on E_a, namely

$$\|(\mathrm{a}, \tilde u)\|_{E_a} := \varepsilon^{(\sigma_{\mathrm{a}}-1-d/2)/p}\big(\varepsilon^{\sigma}\|\mathrm{a}\|_{H^{s_0+a}} + \varepsilon^{-d/2}\|\tilde u\|_{C^1_\varepsilon H^{s_0+a}_\varepsilon}\big), \tag{8.7}$$

and, on F_a, we define

$$\|g\|_{F_a} = \|(g_1, g_2)\|_{F_a} := \varepsilon^{-\sigma-d/2}\|g_1\|_{C^0 H_\varepsilon^{s_0+a}} + \|g_2\|_{H^{s_0+a+1}} \tag{8.8}$$

(note that $\|\mathrm{a}\|_{H^{s_0+a}}$ and $\|g_2\|_{H^{s_0+a+1}}$ in (8.7) and (8.8) are the standard Sobolev norms, without ε). For $(\mathrm{a}, \tilde{u}) \in E_a$ and $g = (g_1, g_2) \in F_a$, we define

$$S_j(\mathrm{a}, \tilde{u}) := (S_j^1 \mathrm{a}, S_j^\varepsilon \tilde{u}), \quad S_j g := (S_j^\varepsilon g_1, S_j^1 g_2), \tag{8.9}$$

where S_j^ε, S_j^1 are the crude Fourier truncations $\varepsilon|\xi| \le 2^j$, $|\xi| \le 2^j$ respectively, namely

$$S_j^\varepsilon f(x) := (2\pi)^{-d/2} \int_{\varepsilon|\xi|\le 2^j} \hat{f}(\xi) e^{i\xi\cdot x}\, d\xi, \quad S_j^1 f(x) := (2\pi)^{-d/2} \int_{|\xi|\le 2^j} \hat{f}(\xi) e^{i\xi\cdot x}\, d\xi.$$

Thus S_j in (8.9) satisfy all (A.2)–(A.8) with constants independent of ε.

We consider the operator $\widetilde{\Phi}$ defined in (7.10). The ball (7.28) becomes

$$\|(\mathrm{a}, \tilde{u})\|_{E_2} \le \rho_2. \tag{8.10}$$

For all $(\mathrm{a}, \tilde{u})$ in the ball (8.10), by (7.29) the linearized problem $\widetilde{\Phi}'(\mathrm{a}, \tilde{u})(\mathrm{b}, \tilde{h}) = g$ has the solution $(\mathrm{b}, \tilde{h}) =: \widetilde{\Psi}(\mathrm{a}, \tilde{u})g$, which satisfies, for all $a \ge 0$,

$$\|\widetilde{\Psi}(\mathrm{a}, \tilde{u})g\|_{E_a} \lesssim_s \varepsilon^{\sigma+(\sigma_\mathrm{a}-1-d/2)/p}(\|g\|_{F_a} + \|(\mathrm{a}, \tilde{u})\|_{E_{a+s_0+3}}\|g\|_{F_0}), \tag{8.11}$$

where we assume that $s_0 \ge 1$ and $s_0 > d/2$. The second derivative of $\widetilde{\Phi}$ is

$$\widetilde{\Phi}''(\mathrm{a}, \tilde{u})[(\mathrm{b}_1, \tilde{h}_1), (\mathrm{b}_2, \tilde{h}_2)] = \begin{pmatrix} P''(u)[h_1, h_2] \\ 0 \end{pmatrix},$$

where $u = \varepsilon^\sigma \mathcal{S}\mathcal{T}_{\varepsilon,c}\mathrm{a} + \tilde{u}$ and $h_i = \varepsilon^\sigma \mathcal{S}\mathcal{T}_{\varepsilon,c}\mathrm{b}_i + \tilde{h}_i$, $i = 1, 2$. By (7.35) and (8.8), for $(\mathrm{a}, \tilde{u})$ in the ball (8.10), one has, for $a \ge 0$,

$$\begin{aligned} &\|\widetilde{\Phi}''(\mathrm{a}, \tilde{u})[(\mathrm{b}_1, \tilde{h}_1), (\mathrm{b}_2, \tilde{h}_2)]\|_{F_a} \\ &\quad = \varepsilon^{-\sigma-d/2}\|P''(u)[h_1, h_2]\|_{C^0 H_\varepsilon^{s_0+a}} \\ &\quad \lesssim_s \varepsilon^{-\sigma+(1+d/2-\sigma_\mathrm{a})/p}\big\{\|(\mathrm{b}_1, \tilde{h}_1)\|_{E_{a+1}}\|(\mathrm{b}_2, \tilde{h}_2)\|_{E_0} \\ &\qquad + \|(\mathrm{b}_1, \tilde{h}_1)\|_{E_0}\|(\mathrm{b}_2, \tilde{h}_2)\|_{E_{a+1}} \\ &\qquad + \|(\mathrm{a}, \tilde{u})\|_{E_{a+1}}\|(\mathrm{b}_1, \tilde{h}_1)\|_{E_0}\|(\mathrm{b}_2, \tilde{h}_2)\|_{E_0}\big\}. \end{aligned} \tag{8.12}$$

Hence $\widetilde{\Phi}$ satisfies the assumptions of Theorem A.1 with

$$\begin{gathered} a_0 = 0, \quad \mu = a_1 = 2, \quad \beta = \alpha = s_0 + 3 > 4, \quad a_2 > 2\beta - 2, \\ U = \big\{(\mathrm{a}, \tilde{u}) \in E_2 : \|(\mathrm{a}, \tilde{u})\|_{E_2} \le \rho_2\big\}, \quad \delta_1 = \rho_2, \quad M_3(a) = L_3(a) = 0, \\ M_1(a) = M_2(a) = C_a \varepsilon^{-\sigma+(1+d/2-\sigma_\mathrm{a})/p}, \\ L_1(a) = L_2(a) = C_a \varepsilon^{\sigma-(1+d/2-\sigma_\mathrm{a})/p}. \end{gathered} \tag{8.13}$$

For any function $\mathrm{a}_0 = \mathrm{a}_0(x) \in H^{s_0+\beta+1}(\mathbb{R}^d)$, the pair $g = (0, \mathrm{a}_0) \in F_\beta$ trivially satisfies the first inequality in (A.12) with $A = 1$ (in fact, the inequality is an identity), because a_0 does not depend on the time variable. Hence, by Theorem A.1, for every $g = (0, \mathrm{a}_0)$ in the ball

$$\|\mathrm{a}_0\|_{H^{s_0+\beta+1}} = \|g\|_{F_\beta} \leq \delta, \tag{8.14}$$

with

$$\delta = C\varepsilon^{-\sigma+(1+d/2-\sigma_{\mathrm{a}})/p} \tag{8.15}$$

given by (A.14), there exists $(\mathrm{a}, \tilde{u}) \in E_\alpha$ such that $\widetilde{\Phi}(\mathrm{a}, \tilde{u}) = \widetilde{\Phi}(0, 0) + g = (0, \mathrm{a}_0)$. By (7.10), this means that $\mathrm{a} = \mathrm{a}_0$ and the sum $u = \varepsilon^\sigma \mathcal{S}\mathcal{T}_{\varepsilon,c}\mathrm{a}_0 + \tilde{u}$ solves the nonlinear Cauchy problem (5.7) on the time interval $[0, T]$ with initial datum $u(0) = u_0 = \varepsilon^\sigma \mathcal{T}_{\varepsilon,c}\mathrm{a}_0$. By (A.13),

$$\|(\mathrm{a}, \tilde{u})\|_{E_\alpha} \leq C\varepsilon^{\sigma-(1+d/2-\sigma_{\mathrm{a}})/p}\|g\|_{F_\beta},$$

namely

$$\varepsilon^\sigma\|\mathrm{a}_0\|_{H^{s_0+\beta}} + \varepsilon^{-d/2}\|\tilde{u}\|_{C^1_\varepsilon H^{s_0+\beta}_\varepsilon} \leq C\varepsilon^\sigma\|\mathrm{a}_0\|_{H^{s_0+\beta+1}},$$

whence

$$\|\tilde{u}\|_{C^1_\varepsilon H^{s_0+\beta}_\varepsilon} \leq C\varepsilon^{\sigma+d/2}\|\mathrm{a}_0\|_{H^{s_0+\beta+1}}.$$

All $\|\mathrm{a}_0\|_{H^{s_0+\beta+1}} \leq 1$ belong to the ball (8.14) if $1 \leq \delta$, and this holds for ε sufficiently small if

$$\sigma > \frac{1 + d/2 - \sigma_{\mathrm{a}}}{p}.$$

The higher regularity part of Theorem 3.5 is also deduced from Theorem A.1.

Finally, given $s_1 > \max\{6, d + 4\}$, we define $s_0 := (s_1 - 4)/2$, so that $s_0 > \max\{1, d/2\}$, and the proof of Theorem 3.5 is complete.

Remark 8.1 (Confirmation of the heuristics discussion of Section 2 in Theorem 3.5). The radius δ given by the Nash–Moser Theorem A.1 is the minimum among $1/L$, δ_1/L, $1/(L^2M)$; here (see (8.13)) these three quantities are all of order $\varepsilon^{-\sigma+(1+d/2-\sigma_{\mathrm{a}})/p}$. In particular, the "quadratic condition" $\delta \leq 1/(L^2M)$, coming from the use of the second derivative $\Phi''(u)$ in the Nash–Moser iteration, does not modify δ. This is a confirmation of the heuristic discussion of Section 2.

For completeness, now we perform the same analysis in the case $p = 1$. We consider the same function spaces (8.1)–(8.6) as above, but now we use norms (7.30) on E_a, namely (see also (7.17))

$$\|(\mathrm{a}, \tilde{u})\|_{\mathcal{E}_a} := \varepsilon^{\sigma-1}\|\mathrm{a}\|_{H^{s_0+a}} + \varepsilon^{-1-d/2}\|\tilde{u}\|_{C^1_\varepsilon H^{s_0+a}_\varepsilon}, \tag{8.16}$$

and, on F_a, we define

$$\begin{aligned}\|g\|_{\mathcal{F}_a} &= \|(g_1, g_2)\|_{\mathcal{F}_a} \\ &:= \varepsilon^{-\sigma-\sigma_{\mathrm{a}}}\|g_1\|_{C^0 H^{s_0+a}_\varepsilon} + \|g_2\|_{H^{s_0+a+1}}.\end{aligned} \tag{8.17}$$

By (7.32), (7.33), and (7.36), for $(\mathrm{a}, \tilde{u})$ in the ball

$$\|(\mathrm{a}, \tilde{u})\|_{\mathcal{E}_2} \le \rho_2, \tag{8.18}$$

for $a \ge 0$ one has

$$\|\widetilde{\Psi}(\mathrm{a}, \tilde{u}) g\|_{\mathcal{E}_a} \lesssim_s \varepsilon^{\sigma+\sigma_{\mathrm{a}}-1-d/2}(\|g\|_{\mathcal{F}_a} + \|(\mathrm{a}, \tilde{u})\|_{\mathcal{E}_{a+s_0+3}} \|g\|_{\mathcal{F}_0}) \tag{8.19}$$

and

$$\begin{aligned} &\|\widetilde{\Phi}''(\mathrm{a}, \tilde{u})[(\mathrm{b}_1, \tilde{h}_1), (\mathrm{b}_2, \tilde{h}_2)]\|_{\mathcal{F}_a} \\ &\quad = \varepsilon^{-\sigma-\sigma_{\mathrm{a}}} \|P''(u)[h_1, h_2]\|_{C^0 H_\varepsilon^{s_0+a}} \\ &\quad \lesssim_s \varepsilon^{1-\sigma} \{\|(\mathrm{b}_1, \tilde{h}_1)\|_{\mathcal{E}_{a+1}} \|(\mathrm{b}_2, \tilde{h}_2)\|_{\mathcal{E}_0} \\ &\qquad\qquad + \|(\mathrm{b}_1, \tilde{h}_1)\|_{\mathcal{E}_0} \|(\mathrm{b}_2, \tilde{h}_2)\|_{\mathcal{E}_{a+1}} \\ &\qquad\qquad + \|(\mathrm{a}, \tilde{u})\|_{\mathcal{E}_{a+1}} \|(\mathrm{b}_1, \tilde{h}_1)\|_{\mathcal{E}_0} \|(\mathrm{b}_2, \tilde{h}_2)\|_{\mathcal{E}_0}\}. \end{aligned} \tag{8.20}$$

Hence $\widetilde{\Phi}$ satisfies the assumptions of Theorem A.1 with

$$\begin{gathered} a_0 = 0, \quad \mu = a_1 = 2, \quad \beta = \alpha = s_0 + 3 > 4, \quad a_2 > 2\beta - 2, \\ U = \{(\mathrm{a}, \tilde{u}) \in E_2 : \|(\mathrm{a}, \tilde{u})\|_{\mathcal{E}_2} \le \rho_2\}, \\ \delta_1 = \rho_2, \quad M_3(a) = L_3(a) = 0, \\ M_1(a) = M_2(a) = C_a \varepsilon^{1-\sigma}, \quad L_1(a) = L_2(a) = C_a \varepsilon^{\sigma+\sigma_{\mathrm{a}}-1-d/2}. \end{gathered}$$

Hence, by Theorem A.1, for every $g = (0, \mathrm{a}_0)$ in the ball

$$\|\mathrm{a}_0\|_{H^{s_0+\beta+1}} = \|g\|_{\mathcal{F}_\beta} \le \delta \tag{8.21}$$

with

$$\delta = C \varepsilon^{-\sigma+1+d-2\sigma_{\mathrm{a}}} \tag{8.22}$$

given by (A.14), there exists $(\mathrm{a}, \tilde{u}) \in E_\alpha$ such that $\widetilde{\Phi}(\mathrm{a}, \tilde{u}) = (0, \mathrm{a}_0)$. By (A.13), the solution $(\mathrm{a}, \tilde{u})$ satisfies

$$\|(\mathrm{a}, \tilde{u})\|_{\mathcal{E}_\alpha} \le C \varepsilon^{\sigma+\sigma_{\mathrm{a}}-1-d/2} \|g\|_{\mathcal{F}_\beta},$$

namely

$$\varepsilon^\sigma \|\mathrm{a}_0\|_{H^{s_0+\beta}} + \varepsilon^{-d/2} \|\tilde{u}\|_{C_\varepsilon^1 H_\varepsilon^{s_0+\beta}} \le C \varepsilon^{\sigma+\sigma_{\mathrm{a}}-d/2} \|\mathrm{a}_0\|_{H^{s_0+\beta+1}},$$

whence

$$\|\tilde{u}\|_{C_\varepsilon^1 H_\varepsilon^{s_0+\beta}} \le C \varepsilon^{\sigma+\sigma_{\mathrm{a}}} \|\mathrm{a}_0\|_{H^{s_0+\beta+1}}.$$

All $\|\mathrm{a}_0\|_{H^{s_0+\beta+1}} \le 1$ belong to the ball (8.21) if $1 \le \delta$, and this holds for ε sufficiently small if

$$\sigma > 1 + d - 2\sigma_{\mathrm{a}}.$$

A. Nash–Moser–Hörmander implicit function theorem

In this section we state the Nash–Moser–Hörmander theorem of [1].

Let $(E_a)_{a\geq 0}$ be a decreasing family of Banach spaces with continuous injections $E_b \hookrightarrow E_a$,

$$\|u\|_{E_a} \leq \|u\|_{E_b} \quad \text{for } a \leq b. \tag{A.1}$$

Set $E_\infty = \bigcap_{a\geq 0} E_a$ with the weakest topology making the injections $E_\infty \hookrightarrow E_a$ continuous. Assume that there exist linear smoothing operators $S_j \colon E_0 \to E_\infty$ for $j = 0, 1, \ldots,$ satisfying the following inequalities, with constants C bounded when a and b are bounded, and independent of j,

$$\|S_j u\|_{E_a} \leq C\|u\|_{E_a} \quad \text{for all } a, \tag{A.2}$$

$$\|S_j u\|_{E_b} \leq C 2^{j(b-a)}\|S_j u\|_{E_a} \quad \text{if } a < b, \tag{A.3}$$

$$\|u - S_j u\|_{E_b} \leq C 2^{-j(a-b)}\|u - S_j u\|_{E_a} \quad \text{if } a > b, \tag{A.4}$$

$$\|(S_{j+1} - S_j)u\|_{E_b} \leq C 2^{j(b-a)}\|(S_{j+1} - S_j)u\|_{E_a} \quad \text{for all } a, b. \tag{A.5}$$

Set

$$R_0 u := S_1 u, \quad R_j u := (S_{j+1} - S_j)u, \quad j \geq 1. \tag{A.6}$$

Thus

$$\|R_j u\|_{E_b} \leq C 2^{j(b-a)}\|R_j u\|_{E_a} \quad \text{for all } a, b. \tag{A.7}$$

Bound (A.7) for $j \geq 1$ is (A.5), while, for $j = 0$, it follows from (A.1) and (A.3). We also assume that

$$\|u\|_{E_a}^2 \leq C \sum_{j=0}^{\infty} \|R_j u\|_{E_a}^2 \quad \text{for all } a \geq 0, \tag{A.8}$$

with C bounded for a bounded (the "orthogonality property" for the smoothing operators).

Suppose that we have another family F_a of decreasing Banach spaces with smoothing operators having the same properties as above. We use the same notation for the smoothing operators.

Theorem A.1 ([1]). (Existence) *Let a_1, a_2, α, β, a_0, μ be real numbers with*

$$0 \leq a_0 \leq \mu \leq a_1, \quad a_1 + \frac{\beta}{2} < \alpha < a_1 + \beta, \quad 2\alpha < a_1 + a_2. \tag{A.9}$$

Let U be a convex neighborhood of 0 in E_μ. Let Φ be a map from U to F_0 such that $\Phi \colon U \cap E_{a+\mu} \to F_a$ is of class C^2 for all $a \in [0, a_2 - \mu]$, with

$$\begin{aligned}\|\Phi''(u)[v, w]\|_{F_a} &\leq M_1(a)(\|v\|_{E_{a+\mu}}\|w\|_{E_{a_0}} + \|v\|_{E_{a_0}}\|w\|_{E_{a+\mu}}) \\ &\quad + \big\{M_2(a)\|u\|_{E_{a+\mu}} + M_3(a)\big\}\|v\|_{E_{a_0}}\|w\|_{E_{a_0}}\end{aligned} \tag{A.10}$$

for all $u \in U \cap E_{a+\mu}$, $v, w \in E_{a+\mu}$, *where* $M_i\colon [0, a_2 - \mu] \to \mathbb{R}$, $i = 1, 2, 3$, *are positive, increasing functions. Assume that* $\Phi'(v)$, *for* $v \in E_\infty \cap U$ *belonging to some ball* $\|v\|_{E_{a_1}} \le \delta_1$, *has a right inverse* $\Psi(v)$ *mapping* F_∞ *to* E_{a_2}, *and that*

$$\begin{aligned} \|\Psi(v)g\|_{E_a} &\le L_1(a)\|g\|_{F_{a+\beta-\alpha}} \\ &\quad + \big\{L_2(a)\|v\|_{E_{a+\beta}} + L_3(a)\big\}\|g\|_{F_0} \quad \textit{for all } a \in [a_1, a_2], \end{aligned} \tag{A.11}$$

where $L_i\colon [a_1, a_2] \to \mathbb{R}$, $i = 1, 2, 3$ *are positive, increasing functions.*

Then for all $A > 0$ *there exists* $\delta > 0$ *such that, for every* $g \in F_\beta$ *satisfying*

$$\sum_{j=0}^{\infty} \|R_j g\|_{F_\beta}^2 \le A^2 \|g\|_{F_\beta}^2, \quad \|g\|_{F_\beta} \le \delta, \tag{A.12}$$

there exists $u \in E_\alpha$ *solving* $\Phi(u) = \Phi(0) + g$. *The solution* u *satisfies*

$$\|u\|_{E_\alpha} \le C L_{123}(a_2)(1 + A)\|g\|_{F_\beta}, \tag{A.13}$$

where $L_{123} = L_1 + L_2 + L_3$ *and* C *is a constant depending on* a_1, a_2, α, β. *The constant* δ *is*

$$\begin{aligned} &\delta = 1/B, \\ &B = C' L_{123}(a_2) \max\big\{1/\delta_1, 1 + A, (1 + A) L_{123}(a_2) M_{123}(a_2 - \mu)\big\}, \end{aligned} \tag{A.14}$$

where $M_{123} = M_1 + M_2 + M_3$ *and* C' *is a constant depending on* a_1, a_2, α, β.

(Higher regularity) *Moreover, let* $c > 0$ *and assume that* (A.10) *holds for all* $a \in [0, a_2 + c - \mu]$, $\Psi(v)$ *maps* F_∞ *to* E_{a_2+c}, *and* (A.11) *holds for all* $a \in [a_1, a_2 + c]$. *If* g *satisfies* (A.12) *and, in addition,* $g \in F_{\beta+c}$ *with*

$$\sum_{j=0}^{\infty} \|R_j g\|_{F_{\beta+c}}^2 \le A_c^2 \|g\|_{F_{\beta+c}}^2 \tag{A.15}$$

for some A_c, *then the solution* u *belongs to* $E_{\alpha+c}$, *with*

$$\|u\|_{E_{\alpha+c}} \le C_c\big\{\mathcal{G}_1(1 + A)\|g\|_{F_\beta} + \mathcal{G}_2(1 + A_c)\|g\|_{F_{\beta+c}}\big\}, \tag{A.16}$$

where

$$\mathcal{G}_1 := \widetilde{L}_3 + \widetilde{L}_{12}(\widetilde{L}_3 \widetilde{M}_{12} + L_{123}(a_2)\widetilde{M}_3)(1 + z^N), \quad \mathcal{G}_2 := \widetilde{L}_{12}(1 + z^N), \tag{A.17}$$

$$z := L_{123}(a_1) M_{123}(0) + \widetilde{L}_{12}\widetilde{M}_{12}, \tag{A.18}$$

$\widetilde{L}_{12} := \widetilde{L}_1 + \widetilde{L}_2$, $\widetilde{L}_i := L_i(a_2 + c)$, $i = 1, 2, 3$; $\widetilde{M}_{12} := \widetilde{M}_1 + \widetilde{M}_2$, $\widetilde{M}_i := M_i(a_2 + c - \mu)$, $i = 1, 2, 3$; N *is a positive integer depending on* c, a_1, α, β; *and* C_c *depends on* a_1, a_2, α, β, c.

B. Commutator and product estimates

In the next lemmas we give "asymmetric" inequalities for the Sobolev norm of commutators and products of functions on $\mathbb{R}^d$, with $W^{m,\infty}$ norms (m integer) on one function and H^s norms (s real) on the other function. Estimate (B.1) is related to the Kato–Ponce inequality (see e.g. [2, 3, 8]), but it is not clear how to deduce (B.1) directly from Kato–Ponce. Hence we give here a proof of (B.1), entirely based on well-known estimates.

Lemma B.1. *Let $s \geq 0$ be real, and let m be the smallest positive integer such that $m \geq s$. Then there exists C_s such that*

$$\|\Lambda^s(uv) - u\Lambda^s v\|_{L^2} \leq C_s(\|u\|_{W^{1,\infty}}\|v\|_{H^{s-1}} + \|u\|_{W^{m,\infty}}\|v\|_{L^2}) \tag{B.1}$$

for all $u \in W^{m,\infty}(\mathbb{R}^d)$, all $v \in H^{s-1}(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. The constant C_s is increasing in s, and it is bounded for s bounded.

The same estimate holds with Λ^s replaced by $\Lambda^{s-1}\partial_x^\alpha$, $|\alpha| = 1$, namely

$$\|\Lambda^{s-1}\partial_x^\alpha(uv) - u\Lambda^{s-1}\partial_x^\alpha v\|_{L^2} \leq C_s(\|u\|_{W^{1,\infty}}\|v\|_{H^{s-1}} + \|u\|_{W^{m,\infty}}\|v\|_{L^2}). \tag{B.2}$$

Proof. We use the standard paraproduct decomposition $uv = T_u v + (u - T_u)v$ (following Métivier [9]), and split

$$\Lambda^s(uv) - u\Lambda^s v = [\Lambda^s, T_u]v + \Lambda^s((u - T_u)v) - (u - T_u)\Lambda^s v.$$

The commutator $[\Lambda^s, T_u]$ satisfies

$$\|[T_u, \Lambda^s]v\|_{L^2} \leq C_s\|u\|_{W^{1,\infty}}\|v\|_{H^{s-1}} \tag{B.3}$$

by [9, Theorem 6.1.4]. The second term satisfies

$$\|\Lambda^s((u - T_u)v)\|_{L^2} = \|(u - T_u)v\|_{H^s} \leq \|(u - T_u)v\|_{H^m} \leq C_m\|u\|_{W^{m,\infty}}\|v\|_{L^2} \tag{B.4}$$

by [9, Theorem 5.2.8]. By duality, the third term is also bounded by the right-hand side of (B.4): for all $h \in L^2$, by Cauchy–Schwarz,

$$\begin{aligned}\langle (u - T_u)\Lambda^s v, h\rangle_{L^2} = \langle v, \Lambda^s(u - T_u)^* h\rangle_{L^2} &\leq \|v\|_{L^2}\|(u - T_u)^* h\|_{H^s} \\ &\leq \|v\|_{L^2}\|(u - T_u)^* h\|_{H^m},\end{aligned}$$

where $(u - T_u)^*$ is the adjoint of $(u - T_u)$ with respect to the L^2 scalar product. Split

$$(u - T_u)^* = (u^* - T_{u^*}) + (T_{u^*} - (T_u)^*). \tag{B.5}$$

The first component in the right-hand side of (B.5) satisfies

$$\|(u^* - T_{u^*})h\|_{H^m} \leq C_m\|u^*\|_{W^{m,\infty}}\|h\|_{L^2} = C_m\|u\|_{W^{m,\infty}}\|h\|_{L^2}$$

by [9, Theorem 5.2.8]. The second component in the right-hand side of (B.5) satisfies

$$\|(T_{u^*} - (T_u)^*)h\|_{H^m} \le C_m \|u\|_{W^{m,\infty}} \|h\|_{L^2}$$

by [9, Theorem 6.2.4]. Hence $\|(u - T_u)^* h\|_{H^m}$ is bounded by $C_m \|u\|_{W^{m,\infty}} \|h\|_{L^2}$, and

$$\langle (u - T_u)\Lambda^s v, h\rangle_{L^2} \le C_m \|u\|_{W^{m,\infty}} \|v\|_{L^2} \|h\|_{L^2}$$

for all $h \in L^2$. This implies that

$$\|(u - T_u)\Lambda^s v\|_{L^2} \le C_m \|u\|_{W^{m,\infty}} \|v\|_{L^2}. \tag{B.6}$$

The sum of (B.3), (B.4), and (B.6) gives (B.1).

Similarly, one proves that (B.3), (B.4), and (B.6) also hold with Λ^s in the left-hand side replaced by $\Lambda^{s-1}\partial_x^\alpha$, $|\alpha| = 1$. Then (B.2) follows. ∎

Lemma B.2. *Let $s \ge 0$ be real, and let m be the smallest positive integer such that $m \ge s$. Then*

$$\|uv\|_{H^s} \le 2\|u\|_{L^\infty} \|v\|_{H^s} + C_s \|u\|_{W^{m,\infty}} \|v\|_{L^2} \tag{B.7}$$

for all $u \in W^{m,\infty}(\mathbb{R}^d)$, all $v \in H^s(\mathbb{R}^d)$. The constant C_s is increasing in s, and it is bounded for s bounded.

Moreover, for all $0 < \varepsilon \le 1$,

$$\|uv\|_{H^s_\varepsilon} \le 2\|u\|_{L^\infty} \|v\|_{H^s_\varepsilon} + C_s \|u\|_{W^{m,\infty}_\varepsilon} \|v\|_{L^2} \tag{B.8}$$

with the same constant C_s as in (B.7) (in particular, C_s is independent of ε).

Proof. By the triangular inequality and (B.1),

$$\begin{aligned}\|uv\|_{H^s} = \|\Lambda^s(uv)\|_{L^2} &\le \|\Lambda^s(uv) - u\Lambda^s v\|_{L^2} + \|u\Lambda^s v\|_{L^2} \\ &\le C_s(\|u\|_{W^{1,\infty}} \|v\|_{H^{s-1}} + \|u\|_{W^{m,\infty}} \|v\|_{L^2}) \\ &\quad + \|u\|_{L^\infty} \|v\|_{H^s}.\end{aligned} \tag{B.9}$$

By standard interpolation, with $\lambda = 1/m$, for all $K \ge 1$ one has

$$\begin{aligned}\|u\|_{W^{1,\infty}} \|v\|_{H^{s-1}} &\le \|u\|_{L^\infty}^{1-\lambda} \|u\|_{W^{m,\infty}}^{\lambda} \|v\|_{H^s}^{1-\lambda} \|v\|_{H^{s-m}}^{\lambda} \\ &= \frac{1}{K}(\|u\|_{L^\infty} \|v\|_{H^s})^{1-\lambda} (\|u\|_{W^{m,\infty}} \|v\|_{H^{s-m}} K^m)^\lambda \\ &\le \frac{1}{K}(\|u\|_{L^\infty} \|v\|_{H^s} + \|u\|_{W^{m,\infty}} \|v\|_{H^{s-m}} K^m) \\ &\le \frac{1}{K}\|u\|_{L^\infty} \|v\|_{H^s} + K^{m-1} \|u\|_{W^{m,\infty}} \|v\|_{L^2}\end{aligned}$$

($\|v\|_{H^{s-m}} \le \|v\|_{L^2}$ because $s - m \le 0$). We fix K larger than or equal to the constant C_s in (B.9), and we obtain (B.7).

Inequality (B.8) is a straightforward consequence of (B.7), (4.4), (4.11), and the trivial rescaling identity for the product $R_\varepsilon(uv) = (R_\varepsilon u)(R_\varepsilon v)$. ∎

Remark B.3. Let s, m be as in Lemmas B.1, B.2. Then $m \leq [s] + 1$, where $[s]$ is the integer part of s (it is $m = [s]$ for s a positive integer, and $m = [s] + 1$ otherwise). As a consequence, (B.1), (B.7), and (B.8) hold with $[s] + 1$ in place of m.

We prove here some elementary inequalities that we have used above.

Lemma B.4. *For every real $s > 0$ there exists $C_s \geq 1$ such that*

$$(a+b)^s \leq 2a^s + C_s b^s \quad \text{for all } a, b \geq 0.$$

The constant C_s is increasing in s, with $C_s = 1$ for $0 < s \leq 1$, and $C_s \to \infty$ as $s \to \infty$.

Proof. For $b = 0$ the inequality is trivial. For $b > 0$, divide by b^s and set $\lambda = a/b$. The inequality holds with best constant $C_s = \max\{(1+\lambda)^s - 2\lambda^s : \lambda \geq 0\}$, which is $C_s = 1$ for $0 < s \leq 1$, and $C_s = 2 \cdot (2^{\frac{1}{s-1}} - 1)^{-(s-1)}$ for $s > 1$. ■

Lemma B.5. *For every $s > 0$ there exists $C_s \geq 1$ (increasing in s) such that*

$$(1 + (a+b)^2)^s \leq 4(1+a^2)^s + C_s b^{2s} \quad \text{for all } a, b \geq 0.$$

Proof. For all $\lambda > 0$ one has $2ab = 2(a\lambda^{1/2})(b\lambda^{-1/2}) \leq a^2\lambda + b^2/\lambda$, whence

$$\begin{aligned} 1 + a^2 + 2ab + b^2 &\leq 1 + a^2(1+\lambda) + b^2(1+1/\lambda) \\ &\leq (1+a^2)(1+\lambda) + b^2(1+1/\lambda). \end{aligned}$$

By Lemma B.4,

$$(1+(a+b)^2)^s \leq 2(1+\lambda)^s(1+a^2)^s + C_s(1+1/\lambda)^s b^{2s}.$$

Then we fix $\lambda = 2^{1/s} - 1$, so that $(1+\lambda)^s = 2$ and $(1+1/\lambda)^s = 2 \cdot (2^{1/s} - 1)^{-s}$. ■

In the proof of Lemma 7.1 we have used Lemma B.5 in the form

$$(1 + |\eta|^2 + 2|\eta||\xi_0| + |\xi_0|^2)^s \leq 4(1+|\eta|^2)^s + C_s|\xi_0|^{2s}, \quad \eta, \xi_0 \in \mathbb{R}^d. \tag{B.10}$$

Also, by (B.10) one directly proves the inequality

$$\|uv\|_{H^s} \leq C_{s_0}\|u\|_{H^{s_0}}\|v\|_{H^s} + C_s\|u\|_{H^s}\|v\|_{H^{s_0}}, \tag{B.11}$$

for $s \geq 0$, $s_0 > d/2$, which, by rescaling, implies inequality (4.13).

Lemma B.6. *For all $s \geq 0$ real, all functions u, v on $\mathbb{R}^d$, one has*

$$\|u\partial_x v\|_{H^{s-1}} \lesssim_s \|u\|_{L^\infty}\|v\|_{H^s} + \|u\|_{W^{[s]+1,\infty}}\|v\|_{L^2}, \tag{B.12}$$

$$\|u\varepsilon\partial_x v\|_{H^{s-1}_\varepsilon} \lesssim_s \|u\|_{L^\infty}\|v\|_{H^s_\varepsilon} + \|u\|_{W^{[s]+1,\infty}_\varepsilon}\|v\|_{L^2}, \tag{B.13}$$

where ∂_x denotes any ∂_x^α, $|\alpha| = 1$.

Proof. Write $u\partial_x v$ as $\partial_x(uv) - (\partial_x u)v$. For $s \geq 0$, by (B.7) and Remark B.3,

$$\|\partial_x(uv)\|_{H^{s-1}} \leq \|uv\|_{H^s} \lesssim_s \|u\|_{L^\infty}\|v\|_{H^s} + \|u\|_{W^{[s]+1,\infty}}\|v\|_{L^2}. \tag{B.14}$$

For $s \geq 1$, by (B.7) and Remark B.3,

$$\begin{aligned}\|(\partial_x u)v\|_{H^{s-1}} &\lesssim_s \|\partial_x u\|_{L^\infty}\|v\|_{H^{s-1}} + \|\partial_x u\|_{W^{[s-1]+1,\infty}}\|v\|_{L^2} \\ &\lesssim_s \|u\|_{W^{1,\infty}}\|v\|_{H^{s-1}} + \|u\|_{W^{[s]+1,\infty}}\|v\|_{L^2},\end{aligned} \tag{B.15}$$

while for $0 \leq s \leq 1$,

$$\|(\partial_x u)v\|_{H^{s-1}} \leq \|(\partial_x u)v\|_{L^2} \leq \|\partial_x u\|_{L^\infty}\|v\|_{L^2} \leq \|u\|_{W^{1,\infty}}\|v\|_{L^2}. \tag{B.16}$$

The sum of (B.14) and (B.16) gives (B.12) for $s \in [0, 1]$. For $s \geq 1$, the sum of (B.14) and (B.15) gives (B.12) because, by interpolation,

$$\|u\|_{W^{1,\infty}}\|v\|_{H^{s-1}} \leq \|u\|_{L^\infty}\|v\|_{H^s} + \|u\|_{W^{[s]+1,\infty}}\|v\|_{H^{s-1-[s]}}$$

and $\|v\|_{H^{s-1-[s]}} \leq \|v\|_{L^2}$. Inequality (B.13) can be proved similarly, or it can be deduced from (B.12) by rescaling. ∎

Lemma B.7. *For all $s \geq 0$ real, one has*

$$\|[\Lambda^s, u]\partial_x v\|_{L^2} \lesssim_s \|u\|_{W^{1,\infty}}\|v\|_{H^s} + \|u\|_{W^{[s]+2,\infty}}\|v\|_{L^2}, \tag{B.17}$$

$$\|[\Lambda^s_\varepsilon, u]\varepsilon\partial_x v\|_{L^2} \lesssim_s \|u\|_{W^{1,\infty}_\varepsilon}\|v\|_{H^s_\varepsilon} + \|u\|_{W^{[s]+2,\infty}_\varepsilon}\|v\|_{L^2}, \tag{B.18}$$

where ∂_x denotes any ∂_x^α, $|\alpha| = 1$.

Proof. Write

$$[\Lambda^s, u]\partial_x v = [\Lambda^s\partial_x, u]v - \Lambda^s((\partial_x u)v).$$

By (B.2), $\|[\Lambda^s\partial_x, u]v\|_{L^2}$ is bounded by the right-hand side of (B.17); by (B.7), $\|(\partial_x u)v\|_{H^s}$ is bounded by the right-hand side of (B.17). Thus (B.17) is proved. Inequality (B.18) follows from (B.17) by rescaling. ∎

Lemma B.8. *For all $s \geq 0$, $s_0 > d/2$, one has*

$$\|[\Lambda^s, u]v\|_{L^2} \lesssim_s \|u\|_{H^{s_0+1}}\|v\|_{H^{s-1}} + \|u\|_{H^s}\|v\|_{H^{s_0}}, \tag{B.19}$$

$$\|[\Lambda^s_\varepsilon, u]v\|_{L^2} \lesssim_s \varepsilon^{-d/2}(\|u\|_{H^{s_0+1}_\varepsilon}\|v\|_{H^{s-1}_\varepsilon} + \|u\|_{H^s_\varepsilon}\|v\|_{H^{s_0}_\varepsilon}). \tag{B.20}$$

The same inequalities also hold for $\Lambda^{s-1}\partial_x^\alpha$, $\Lambda^{s-1}_\varepsilon\varepsilon\partial_x^\alpha$, $|\alpha| = 1$, in place of Λ^s, Λ^s_ε respectively.

Proof. In the Fourier transform of $[\Lambda^s, u]v$ one has $\hat{u}(\xi)\hat{v}(\eta)\sigma(\xi, \eta)$, where

$$\sigma(\xi, \eta) = \langle\xi + \eta\rangle^s - \langle\eta\rangle^s = (1 + |\xi + \eta|^2)^{\frac{s}{2}} - (1 + |\eta|^2)^{\frac{s}{2}}.$$

For $|\xi| \leq \frac{1}{2}|\eta|$ one has $|\sigma(\xi, \eta)| \lesssim_s \langle\eta\rangle^{s-1}|\xi|$, leading to the term $\|u\|_{H^{s_0+1}}\|v\|_{H^{s-1}}$ in (B.19). For $|\eta| < 2|\xi|$ one has $|\sigma(\xi, \eta)| \lesssim_s \langle\xi\rangle^s$, leading to the term $\|u\|_{H^s}\|v\|_{H^{s_0}}$ in (B.19). Inequality (B.20) follows by rescaling. ∎

Lemma B.9. *For all $s \geq 0$ real, all functions u, v on $\mathbb{R}^d$, one has*

$$\begin{aligned} \|u\partial_x v\|_{H^{s-1}} &\lesssim_s \|u\|_{H^{s_0}}\|v\|_{H^s} + \|u\|_{H^s}\|v\|_{H^{s_0}} \\ &\quad + \|u\|_{H^{s_0+1}}(\|v\|_{H^{s-1}} + \|v\|_{L^2}), \end{aligned} \tag{B.21}$$

$$\begin{aligned} \|u\varepsilon\partial_x v\|_{H^{s-1}_\varepsilon} &\lesssim_s \varepsilon^{-d/2}\{\|u\|_{H^{s_0}_\varepsilon}\|v\|_{H^s_\varepsilon} + \|u\|_{H^s_\varepsilon}\|v\|_{H^{s_0}_\varepsilon} \\ &\quad + \|u\|_{H^{s_0+1}_\varepsilon}(\|v\|_{H^{s-1}_\varepsilon} + \|v\|_{L^2})\} \end{aligned} \tag{B.22}$$

where ∂_x denotes any ∂_x^α, $|\alpha| = 1$.

Proof. We adapt the proof of Lemma B.6. Write $u\partial_x v$ as $\partial_x(uv) - (\partial_x u)v$. For $s \geq 0$, by (B.11),

$$\|\partial_x(uv)\|_{H^{s-1}} \leq \|uv\|_{H^s} \lesssim_s \|u\|_{H^{s_0}}\|v\|_{H^s} + \|u\|_{H^s}\|v\|_{H^{s_0}}. \tag{B.23}$$

For $s \geq 1$, by (B.11),

$$\begin{aligned} \|(\partial_x u)v\|_{H^{s-1}} &\lesssim_s \|\partial_x u\|_{H^{s_0}}\|v\|_{H^{s-1}} + \|\partial_x u\|_{H^{s-1}}\|v\|_{H^{s_0}} \\ &\lesssim_s \|u\|_{H^{s_0+1}}\|v\|_{H^{s-1}} + \|u\|_{H^s}\|v\|_{H^{s_0}}, \end{aligned} \tag{B.24}$$

while for $0 \leq s \leq 1$,

$$\|(\partial_x u)v\|_{H^{s-1}} \leq \|(\partial_x u)v\|_{L^2} \leq \|\partial_x u\|_{L^\infty}\|v\|_{L^2} \lesssim \|u\|_{H^{s_0+1}}\|v\|_{L^2}. \tag{B.25}$$

Inequality (B.22) is deduced from (B.21) by rescaling. ∎

Lemma B.10. *For all $s \geq 0$ real, one has*

$$\|[\Lambda^s, u]\partial_x v\|_{L^2} \lesssim_s \|u\|_{H^{s_0+1}}\|v\|_{H^s} + \|u\|_{H^{s+1}}\|v\|_{H^{s_0}}, \tag{B.26}$$

$$\|[\Lambda^s_\varepsilon, u]\varepsilon\partial_x v\|_{L^2} \lesssim_s \varepsilon^{-d/2}(\|u\|_{H^{s_0+1}}\|v\|_{H^s} + \|u\|_{H^{s+1}}\|v\|_{H^{s_0}}) \tag{B.27}$$

where ∂_x denotes any ∂_x^α, $|\alpha| = 1$.

Proof. Write $[\Lambda^s, u]\partial_x v = [\Lambda^s\partial_x, u]v - \Lambda^s((\partial_x u)v)$. By Lemma B.8, $\|[\Lambda^s\partial_x, u]v\|_{L^2}$ is bounded by the right-hand side of (B.26); by (B.11), $\|(\partial_x u)v\|_{H^s}$ is also bounded by the right-hand side of (B.26). Thus (B.26) is proved. Inequality (B.27) follows by rescaling. ∎

Acknowledgments. We warmly thank Ivar Ekeland and Eric Séré for many stimulating discussions, in Naples and Paris, which have motivated this work. We also thank Luca Fanelli, Felice Iandoli, Gustavo Ponce, Michela Procesi, and Nicola Visciglia for interesting comments in Erice.

Funding. Supported by INdAM – GNAMPA Project 2019 "Hamiltonian dynamics and evolution PDEs" and PRIN 2015 "Variational methods, with applications to problems in Mathematical Physics and Geometry".

References

[1] P. Baldi and E. Haus, A Nash-Moser-Hörmander implicit function theorem with applications to control and Cauchy problems for PDEs. *J. Funct. Anal.* **273** (2017), no. 12, 3875–3900 Zbl 06792303 MR 3711883

[2] J. Bourgain and D. Li, On an endpoint Kato-Ponce inequality. *Differential Integral Equations* **27** (2014), no. 11-12, 1037–1072 Zbl 1340.42021 MR 3263081

[3] P. D'Ancona, A short proof of commutator estimates. *J. Fourier Anal. Appl.* **25** (2019), no. 3, 1134–1146 Zbl 1415.42016 MR 3953500

[4] I. Ekeland and É. Séré, A surjection theorem for maps with singular perturbation and loss of derivatives. *J. Eur. Math. Soc. (JEMS)* **23** (2021), no. 10, 3323–3349 Zbl 07367691 MR 4275475

[5] L. Hörmander, The boundary problems of physical geodesy. *Arch. Rational Mech. Anal.* **62** (1976), no. 1, 1–52 MR 602181

[6] C. E. Kenig, G. Ponce, and L. Vega, Small solutions to nonlinear Schrödinger equations. *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **10** (1993), no. 3, 255–288 Zbl 0786.35121 MR 1230709

[7] C. E. Kenig, G. Ponce, and L. Vega, The Cauchy problem for quasi-linear Schrödinger equations. *Invent. Math.* **158** (2004), no. 2, 343–388 Zbl 1177.35221 MR 2096797

[8] D. Li, On Kato-Ponce and fractional Leibniz. *Rev. Mat. Iberoam.* **35** (2019), no. 1, 23–100 MR 3914540

[9] G. Métivier, *Para-differential calculus and applications to the Cauchy problem for nonlinear systems*. CRM Series 5, Edizioni della Normale, Pisa, 2008 MR 2418072

[10] G. Métivier and J. Rauch, Dispersive stabilization. *Bull. Lond. Math. Soc.* **42** (2010), no. 2, 250–262 Zbl 1192.35165 MR 2601551

[11] B. Texier and K. Zumbrun, Nash-Moser iteration and singular perturbations. *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **28** (2011), no. 4, 499–527 Zbl 1237.47066 MR 2823882

Received 9 July 2020; revised 14 January 2022; accepted 24 January 2022.

Pietro Baldi
Dipartimento di Matematica e Applicazioni "R. Caccioppoli", Università di Napoli Federico II, Via Cintia, 80126 Napoli, Italy; pietro.baldi@unina.it

Emanuele Haus
Dipartimento di Matematica e Fisica, Università di Roma Tre, Largo San Leonardo Murialdo 1, 00146 Roma, Italy; ehaus@mat.uniroma3.it