# On the Nash-Moser iteration technique

Paolo Secchi

*Dedicated to Professor Y. Shibata in occasion of his 60-th birthday*

**Abstract.** The aim of this work is to provide a brief presentation of the Nash-Moser iteration technique for the resolution of nonlinear equations, where the linearized equations admit estimates with a loss of regularity with respect to the given data.

**Mathematics Subject Classification (2010).** Primary 35-02; Secondary 58C15, 47H14, 47N20, 35Q35, 35R35, 35L60.

**Keywords.** Nash-Moser theorem, implicit function theorems, nonlinear hyperbolic systems, compressible Euler equations, characteristic free boundary.

## 1. Introduction

The aim of this work is to provide a brief presentation of the Nash-Moser iteration method for the resolution of nonlinear equations, where the linearized equations admit estimates with a loss of regularity with respect to the given data. This method was originally introduced by Nash in [16] for solving the isometric embedding problem. Moser in [14, 15] simplified the method at the expense of a loss of regularity, and showed how to apply it in a more general setting. Since then, this iteration technique is known as the Nash-Moser method. Hörmander, in his paper [10] on the boundary problem of physical geodesy, improved on Moser's scheme by reducing the loss of regularity, using a scheme more similar to Nash's original.

Our personal interest is motivated by the study of certain characteristic free-boundary problems for systems of nonlinear conservation laws that arise in fluid dynamics. Interesting and challenging problems arise when the unknown free-boundary is weakly but not strongly stable, i.e. the Kreiss-Lopatinskiĭ condition only holds in weak form. A typical difficulty in the analysis of weakly stable problems is the loss of regularity in the a priori estimates of solutions. Short-time existence results have been obtained for various weakly stable nonlinear problems, typically by the use of a Nash-Moser scheme to compensate for the loss of derivatives in the linearized energy estimates.

Alinhac [1] used a modified version of Hörmander's scheme to prove the short-time existence of rarefaction waves for a class of conservation laws. More recently, Coulombel and Secchi [6] introduced an additional modification to prove the existence of compressible vortex sheets for the two dimensional Euler equations, provided the Mach number is sufficiently large. A scheme similar to the one used in [6] is also considered by Trakhinin in the paper [19] on current-vortex sheets (see also [5]) and by Secchi and Trakhinin in the paper [18] on the plasma-vacuum interface problem.

In this paper we aim to provide a brief presentation of the Nash-Moser iteration method, whilst keeping in mind that our main interest is the application to PDE problems. We present a simplified version of the scheme of [6]; our exposition is also much indebted with [2]. We refer to [2, 9, 11] for a general description of the method. Other related classical references are [3, 4, 13, 17], see also the recent paper [7].

Finally, it is interesting to recall that the isometric embedding problem, originally solved by Nash with this method, was solved much later with an ordinary fixed-point argument, see [8]. Nevertheless, the Nash-Moser method remains a fundamental tool of nonlinear analysis for the study of perturbation problems.

Given $\mathcal{F} : X \to Y$, with $X, Y$ Banach spaces, suppose we wish to solve the nonlinear equation

$$\mathcal{F}(u) = f. \tag{1.1}$$

We assume $\mathcal{F}(0) = 0$; here $f$ is a given "small" perturbation and we look for a solution $u$ close to 0.

Assume $\mathcal{F}$ is continuously differentiable and the differential $d\mathcal{F}(\cdot)$ is invertible in a neighborhood of $u = 0$, so that $\mathcal{F}$ is locally invertible. One of the most classical methods for solving such a nonlinear equation via linearization is Newton iteration method, where the approximating sequence is defined by

$$\begin{aligned} &u_0 = 0, \\ &u_{k+1} = u_k + (d\mathcal{F}(u_k))^{-1}(f - \mathcal{F}(u_k)), \quad k \geq 1. \end{aligned} \tag{1.2}$$

It is well-known that Newton's method has a fast convergence rate:

$$\|u_{k+1} - u_k\|_X \leq C\|u_k - u_{k-1}\|_X^2.$$

However, for this scheme to make sense, we need the inverse $(d\mathcal{F}(u))^{-1}$. In fact, the linearized equation

$$d\mathcal{F}(u)v = g \tag{1.3}$$

may be difficult or impossible to solve for $v \in X$, hence we may not be able to define $(d\mathcal{F}(u))^{-1}$ in a neighborhood of $u = 0$.

In order to introduce the typical situations in which we may recourse to the Nash-Moser method, let us change the formulation as follows.
Instead of single spaces $X, Y$, suppose we are given scales of Banach spaces $X_0 \supset X_1 \supset \cdots \supset X_m \supset \ldots$ with increasing norms $\|\cdot\|_{X_m}$, $m \geq 0$, and

spaces $Y_0 \supset Y_1 \supset \cdots \supset Y_m \supset \ldots$ with increasing norms $\|\cdot\|_{Y_m}$, $m \geq 0$. For instance $X_m = H^m$ (Sobolev spaces) or $X_s = C^s$ (Hölder spaces). Having in mind possible applications to nonlinear hyperbolic equations where the natural function spaces are the Sobolev spaces $H^m$, we will consider Banach scales $X_m, Y_m$ with discrete indices $m$ (instead of a continuous parameter $s \in \mathbb{R}$).
We again wish to solve (1.1) where now $\mathcal{F} : X_m \to Y_m$ for every $m \geq 0$, but $d\mathcal{F}(\cdot)$ is only invertible between $Y_m$ and $X_{m-s}$, with a loss of regularity of order $s$. To be more specific, let us suppose that we have a solution $v$ of (1.3) satisfying, for some given $u$, an estimate of the form

$$\|v\|_{X_m} \leq C\|g\|_{Y_{m+s}}$$

for all $m$ (in a finite interval), $s$ being a fixed number. In this case we say that the equation is solved with a "loss of $s$ derivatives"(clearly in arbitrary families of spaces $X_m, Y_m$ this expression cannot make sense).
Trying to apply again Newton's method (1.2) we would get

$$\begin{aligned} &u_k \in X_m, \\ &u_{k+1} = u_k + (d\mathcal{F}(u_k))^{-1}\underbrace{(f - \mathcal{F}(u_k))}_{\in Y_m} \in X_{m-s} \end{aligned}$$

with a finite loss of regularity at each step. Iteration is then impossible.

Furthermore, this loss $s$ may be doubled by a loss $s'$ due to the cost of solving (1.3) in terms of information about the coefficients of the equation, that is about $u$. Let us suppose, for example, that the solution $v$ satisfies an estimate of the form

$$\|v\|_{X_m} \leq C\left(\|g\|_{Y_{m+s}} + \|g\|_{Y_{m_0}}\|u\|_{X_{m+s'}}\right) \tag{1.4}$$

for all $m$, with $m_0, s$ and $s'$ fixed. In Newton's method, which uses $d\mathcal{F}(u_k)$ to calculate $u_{k+1}$, the solution to one step becomes the coefficient of the next, and the loss $s'$ is added to $s$.
We will see that the nature of this double loss of derivatives determines the applicability of the Nash-Moser technique. Roughly speaking, it is sufficient for the losses $s$ and $s'$ to be fixed, in which case (1.4) is said to be a "tame" estimate.

To overcome this difficulty, the key idea of Nash was to modify Newton's scheme (1.2) by including a smoothing operator at each step to compensate for the loss of regularity. Let us set $u_{k+1} = u_k + \delta u_k$ and write (1.2) as

$$\delta u_k = (d\mathcal{F}(u_k))^{-1} g_k, \qquad g_k = f - \mathcal{F}(u_k).$$

Now let us suppose we have a family of smoothing operators $\{S_X(\theta)\}_{\theta \geq 1}$

$$S_X(\theta) : X_0 \to X_\infty := \cap_{m \geq 0} X_m$$

satisfying $S_X(\theta) \to Id$ as $\theta \to \infty$ and other properties that will be detailed later on[1]. We modify the scheme by setting[2]

$$\delta u_k = (d\mathcal{F}(S_X(\theta_k)u_k))^{-1} S_Y(\theta_k) g_k, \tag{1.5}$$

where $\{\theta_k\}_{k\geq 1}$ is a sequence such that $\theta_k \to \infty$ as $k \to \infty$. Since $S_X(\theta_k) \to Id_X$ and $S_Y(\theta_k) \to Id_Y$ as $k \to \infty$, the scheme looks like Newton iteration for large $k$, so we might expect it to converge under certain conditions. In fact, balancing in appropriate way the fast convergence rate of Newton's scheme and loss of regularity gives the convergence of the approximating sequence.

In applications, very often one observes a rather big difference of regularity between the solution $u$ and the data $f$. That is, generally the regularity of the solution is not optimal compared with the regularity of the data that appear exaggerated. Thus, other modifications to the scheme are introduced in order to reduce this difference of regularity. In this paper, we will not insist on this feature of the method.

Depending on the particular problem under consideration, sometimes *after* having solved the equation by the Nash-Moser technique, one can try to get the optimal regularity of solution, see [12].

## 2. Statement of the Nash-Moser theorem

We first give the main assumptions on the function $\mathcal{F}$.

**Assumption 2.1.** *For all $u \in U \cap X_\infty$, where $U$ is a bounded open neighborhood of $0$ in $X_{m_0}$ for some $m_0 \geq 0$, the function $\mathcal{F} : X_m \to Y_m$ is twice differentiable and satisfies the tame estimate*

$$\begin{aligned} \|d^2\mathcal{F}(u)(v_1, v_2)\|_{Y_m} \leq C\big(&\|v_1\|_{X_{m+r}}\|v_2\|_{X_{m_0}} + \|v_1\|_{X_{m_0}}\|v_2\|_{X_{m+r}} \\ &+ \|v_1\|_{X_{m_0}}\|v_2\|_{X_{m_0}}(1 + \|u\|_{X_{m+r'}})\big) \end{aligned} \tag{2.1}$$

*for all $m \geq 0$ and for all $v_1, v_2 \in X_\infty$, for some fixed integers $r, r' \geq 0$. The constant $C$ is bounded for $m$ bounded.*

**Assumption 2.2.** *For all $u \in U \cap X_\infty$ there exists a linear mapping $\Psi(u) : Y_\infty \to X_\infty$ such that $d\mathcal{F}(u)\Psi(u) = Id$, and satisfying the tame estimate*

$$\|\Psi(u)g\|_{X_m} \leq C\left(\|g\|_{Y_{m+s}} + \|g\|_{Y_{m_0}}\|u\|_{X_{m+s'}}\right) \tag{2.2}$$

*for all $m \geq 0$ and some fixed integers $s, s' \geq 0$. The constant $C$ is bounded for $m$ bounded.*

The method requires a family of smoothing operators; for its construction in Sobolev and Hölder spaces we refer the reader to [1, 2].

[1] Similar smoothing operators $S_Y(\theta) : Y_0 \to Y_\infty := \cap_{m\geq 0} Y_m$ are introduced as well.

[2] Following Hörmander's method [11, 2], in the sequel our iteration scheme will be a little more elaborated than in (1.5).

**Definition 2.3.** The decreasing family of Banach spaces $\{X_m\}_{m\ge 0}$ satisfies *the smoothing hypothesis* if there exists a family $\{S_\theta\}_{\theta\ge 1}$ of operators $S_\theta : X_0 \to X_\infty := \cap_{m\ge 0} X_m$ such that

$$\|S_\theta u\|_{X_\beta} \le C\,\theta^{(\beta-\alpha)_+}\,\|u\|_{X_\alpha} \qquad \forall\,\alpha,\beta \ge 0\,, \tag{2.3a}$$

$$\|S_\theta u - u\|_{X_\beta} \le C\,\theta^{\beta-\alpha}\,\|u\|_{X_\alpha} \qquad 0 \le \beta \le \alpha\,, \tag{2.3b}$$

$$\|\frac{d}{d\theta} S_\theta u\|_{X_\beta} \le C\,\theta^{\beta-\alpha-1}\,\|u\|_{X_\alpha} \qquad \forall\,\alpha,\beta \ge 0\,. \tag{2.3c}$$

Here we use the classical notation $(\beta-\alpha)_+ := \max(0,\beta-\alpha)$. The constants in the inequalities are uniform with respect to $\alpha,\beta$, when $\alpha,\beta$ belong to some bounded interval.

In the decreasing family of Banach spaces $\{Y_m\}_{m\ge 0}$ we will introduce similar smoothing operators $S_Y(\theta) : Y_0 \to Y_\infty := \cap_{m\ge 0} Y_m$. To not overload the notation such smoothing operators $S_Y(\theta)$ will be again denoted by $S_\theta$.

Now we can give our statement of the Nash-Moser theorem.

**Theorem 2.4.** *Let $\{X_m\}_{m\ge 0}$ and $\{Y_m\}_{m\ge 0}$ be two decreasing family of Banach spaces, each satisfying the smoothing hypothesis, and assume that both Assumptions* 2.1 *and* 2.2 *hold. Let $m'$ be a positive integer such that $m' \ge m_0 + \max\{r,r'\} + \max\{s,s'\}$.*

*i) There exists a constant $0 < \epsilon \le 1$ such that if $f \in Y_{m'+s+1}$ with*

$$\|f\|_{Y_{m'+s+1}} \le \epsilon,$$

*the equation $\mathcal{F}(u) = f$ has a solution $u \in X_{m'}$, in the sense that there exists a sequence $\{u_n\} \subset X_\infty$ such that $u_n \to u$ in $X_{m'}$, $\mathcal{F}(u_n) \to f$ in $Y_{m'+s}$, as $n \to \infty$.*

*ii) Moreover, if there exists $m'' > m'$ such that $f \in Y_{m''+s+1}$, then the solution constructed $u \in X_{m''}$.*

As for the regularity of $u$ see Remark 3.10 at the end of Section 3.6.

## 3. Proof of Theorem 2.4

### 3.1. Description of the iterative scheme

The scheme starts from $u_0 = 0$. Assume that $u_k$ are already given for $k = 0,\dots,n$. We consider

$$u_{n+1} = u_n + \delta u_n\,, \tag{3.1}$$

where the differences $\delta u_n$ will be specified later on. Given $\theta_0 \ge 1$, let us set $\theta_n := (\theta_0^2 + n)^{1/2}$, and consider the smoothing operators $S_{\theta_n}$. We decompose

$$\mathcal{F}(u_{n+1}) - \mathcal{F}(u_n) \,=\, d\mathcal{F}(u_n)(\delta u_n) + e'_n \,=\, d\mathcal{F}(S_{\theta_n} u_n)(\delta u_n) + e'_n + e''_n\,,$$

where $e'_n$ denotes the usual "quadratic"error of Newton's scheme, and $e''_n$ the "substitution"error. Let us also set

$$e_n := e'_n + e''_n. \tag{3.2}$$

The iteration proceeds as follows. Assume that $f$ is as in the statement of the theorem. Given

$$\begin{aligned}&u_0 := 0\,, \quad f_0 := S_{\theta_0} f\,, \quad E_0 := 0\,,\\ &u_1, \dots, u_n\,, \quad f_1, \dots, f_{n-1}\,, \quad e_0, \dots, e_{n-1},\end{aligned}$$

we first compute for $n \geq 1$

$$E_n := \sum_{k=0}^{n-1} e_k\,. \tag{3.3}$$

These are the accumulated errors at the step $n$. Then we compute $f_n$ from the equations:

$$\sum_{k=0}^{n} f_k + S_{\theta_n} E_n = S_{\theta_n} f\,, \tag{3.4}$$

and we solve the linear equation

$$d\mathcal{F}(S_{\theta_n} u_n)\, \delta u_n = f_n, \tag{3.5}$$

finding $\delta u_n$. Finally, we compute $e_n$ from

$$\mathcal{F}(u_{n+1}) - \mathcal{F}(u_n) = f_n + e_n\ . \tag{3.6}$$

For $n = 0$ we only consider (3.5), (3.6) and compute $u_1, e_0$. Adding (3.6) from 0 to $N$, and combining with (3.4) gives

$$\mathcal{F}(u_{N+1}) - f = (S_{\theta_N} - I) f + (I - S_{\theta_N}) E_N + e_N\,. \tag{3.7}$$

Because $S_{\theta_N} \to I$ as $N \to +\infty$, and since we expect $e_N \to 0$, we will formally obtain the solution of the problem (1.1) from

$$\mathcal{F}(u_{N+1}) \to f.$$

### 3.2. Introduction of the iterative scheme

We recall that the sequence $\{\theta_n\}_{n\geq 0}$ is defined by $\theta_n := (\theta_0^2 + n)^{1/2}$, for some $\theta_0 \geq 1$. Let us denote $\Delta_n := \theta_{n+1} - \theta_n$. In particular, the sequence $\{\Delta_n\}$ is decreasing, and tends to zero. Moreover, one has

$$\forall n \in \mathbb{N}\,, \quad \frac{1}{3\theta_n} \leq \Delta_n = \sqrt{\theta_n^2 + 1} - \theta_n \leq \frac{1}{2\theta_n}\,.$$

Let us take an integer $\alpha \geq m_0 + 1$, a small number $0 < \delta < 1$, and an integer $\tilde\alpha > \alpha$ that will be chosen later on. Our inductive assumption reads:

$$(H_{n-1}) \qquad \begin{cases} \forall\, k = 0, \dots, n-1\,, \quad \forall\, m \in [m_0, \tilde\alpha] \cap \mathbb{N}\,, \\ \|\delta u_k\|_{X_m} \leq \delta\, \theta_k^{m-\alpha-1}\, \Delta_k\,. \end{cases}$$

The next task is to prove that for a suitable choice of the parameters $\theta_0 \geq 1$, and $\delta > 0$, and for $f$ small enough, $(H_{n-1})$ implies $(H_n)$. In the end, we shall prove that $(H_0)$ holds for $f$ sufficiently small.

From now on, we assume that $(H_{n-1})$ holds. Let us show some basic consequences:

**Lemma 3.1.** *If $\theta_0$ is big enough, chosen independently of $\alpha$, then for every $k = 0, \dots, n$, and for every integer $m \in [m_0, \tilde{\alpha}]$, we have*

$$\|u_k\|_{X_m} \leq \delta\, \theta_k^{(m-\alpha)_+}, \qquad m \neq \alpha, \tag{3.8a}$$

$$\|u_k\|_{X_\alpha} \leq \delta\, \log\theta_k. \tag{3.8b}$$

The proof follows from the triangle inequality, and from the classical comparisons between series and integrals. The choice of how large should be $\theta_0$ is independent of $\alpha$.

**Lemma 3.2.** *If $\theta_0$ is big enough, chosen independently of $\alpha$, then for every $k = 0, \dots, n$, and for every integer $m \in [m_0, \tilde{\alpha} + \max\{r', s'\}]$, we have*

$$\|S_{\theta_k} u_k\|_{X_m} \leq C\, \delta\, \theta_k^{(m-\alpha)_+}, \qquad m \neq \alpha, \tag{3.9a}$$

$$\|S_{\theta_k} u_k\|_{X_\alpha} \leq C\, \delta\, \log\theta_k. \tag{3.9b}$$

*For every $k = 0, \dots, n$, and for every integer $m \in [m_0, \tilde{\alpha}]$, we have*

$$\|(I - S_{\theta_k})u_k\|_{X_m} \leq C\, \delta\, \theta_k^{m-\alpha}. \tag{3.10}$$

The proof follows from Lemma 3.1 and the properties of the smoothing operators, respectively (2.3a) for (3.9) and (2.3b) for (3.10). We remark that the choice of how large should be $\theta_0$ is independent of $\alpha$.

### 3.3. Estimate of the errors

**3.3.1. Estimate of the quadratic errors.** We start by proving an estimate for the quadratic error $e_k'$ of the iterative scheme. Recall that this error is defined by

$$e_k' := \mathcal{F}(u_{k+1}) - \mathcal{F}(u_k) - d\mathcal{F}(u_k)\delta u_k, \tag{3.11}$$

**Lemma 3.3.** *Assume that $\alpha \geq m_0 + 1$ also satisfies $\alpha \geq m_0 + r' - r + 1$. There exist $\delta > 0$ sufficiently small, and $\theta_0 \geq 1$ sufficiently large, both chosen independently of $\alpha$, such that for all $k = 0, \dots, n-1$, and for all integer $m \in [m_0, \tilde{\alpha} - \max\{r, r'\}]$, one has*

$$\|e_k'\|_{Y_m} \leq C\, \delta^2\, \theta_k^{L_1(m)-1}\, \Delta_k, \tag{3.12}$$

*where $L_1(m) := \max\{m + m_0 + r - 2\alpha - 2; (m + r' - \alpha)_+ + 2m_0 - 2\alpha - 2\}$.*

*Proof.* The quadratic error given in (3.11) may be written as

$$e_k' = \int_0^1 (1-\tau)\, d^2\mathcal{F}(u_k + \tau\, \delta u_k)(\delta u_k, \delta u_k)\, d\tau.$$

From $(H_{n-1})$ and (3.8a), we have

$$\sup_{\tau\in[0,1]} \|u_k + \tau\, \delta u_k\|_{X_{m_0}} \leq 2\,\delta,$$

so for $\delta$ sufficiently small we can apply the tame estimate (2.1). Using $(H_{n-1})$ and (3.8) we obtain (3.12). If $m + r' = \alpha$ we use (3.8b) with $\log\theta_k \leq \theta_k$; it yields an estimate of the form (3.12) if $\alpha \geq m_0 + r' - r + 1$. □

**3.3.2. Estimate of the substitution errors.** Now we estimate the substitution error $e''_k$ of the iterative scheme, defined by

$$e''_k := d\mathcal{F}(u_k)\delta u_k - d\mathcal{F}(S_{\theta_k}u_k)\delta u_k \,. \tag{3.13}$$

**Lemma 3.4.** *Assume that $\alpha \geq m_0 + 1$ also satisfies $\alpha \geq m_0 + r' - r + 1$. There exist $\delta > 0$ sufficiently small, and $\theta_0 \geq 1$ sufficiently large, both chosen independently of $\alpha$, such that for all $k = 0, \dots, n-1$, and for all integer $m \in [m_0, \tilde{\alpha} - \max\{r, r'\}]$, one has*

$$\|e''_k\|_{Y_m} \leq C\,\delta^2\,\theta_k^{L(m)-1}\,\Delta_k \,, \tag{3.14}$$

*where $L(m) := \max\{m + m_0 + r - 2\alpha; (m + r' - \alpha)_+ + 2m_0 - 2\alpha\}$.*

*Proof.* The substitution error given in (3.13) may be written as

$$e''_k = \int_0^1 d^2\mathcal{F}(S_{\theta_k}u_k + \tau(I - S_{\theta_k})u_k)(\delta u_k, (I - S_{\theta_k})u_k)\,d\tau \,.$$

As in the calculation for the quadratic error, we first show that we can apply (2.1) for $\delta$ sufficiently small. Then, the estimate (3.14) follows from $(H_{n-1})$, (3.8), (3.9) and (3.10). □

Adding (3.12), (3.14) gives the estimate for the sum of errors defined in (3.2):

**Lemma 3.5.** *Assume that $\alpha \geq m_0 + 1$ also satisfies $\alpha \geq m_0 + r' - r + 1$. There exist $\delta > 0$ sufficiently small, and $\theta_0 \geq 1$ sufficiently large, both chosen independently of $\alpha$, such that for all $k = 0, \dots, n-1$ and all integer $m \in [m_0, \tilde{\alpha} - \max\{r, r'\}]$, one has*

$$\|e_k\|_{Y_m} \leq C\,\delta^2\,\theta_k^{L(m)-1}\,\Delta_k \,, \tag{3.15}$$

*where $L(m)$ is defined in Lemma 3.4.*

The preceding lemma immediately yields the estimate of the accumulated error $E_n$ defined in (3.3):

**Lemma 3.6.** *Assume that $\alpha \geq m_0 + 1$ also satisfies $\alpha \geq m_0 + r' - r + 1$. Let $\tilde{\alpha} = 2\alpha + \max\{r, r'\} + 1 - m_0 - r$. There exist $\delta > 0$ sufficiently small, $\theta_0 \geq 1$ sufficiently large, both chosen independently of $\alpha$, such that*

$$\|E_n\|_{Y_p} \leq C\,\delta^2\,\theta_n^{L(p)} \,, \tag{3.16}$$

*where we have set $p := \tilde{\alpha} - \max\{r, r'\}$.*

*Proof.* For the estimate in $Y_p$ of the accumulated error we choose $p$ to be as large as possible, namely $p = \tilde{\alpha} - \max\{r, r'\}$. Moreover $\tilde{\alpha}$ is taken sufficiently large so that $L(p) \geq 1$. Then it follows from (3.15) that

$$\|E_n\|_{Y_p} \leq C\,\delta^2 \sum_{k=0}^{n-1} \theta_k^{L(p)-1}\,\Delta_k \leq C\,\delta^2\,\theta_n^{L(p)},$$

which gives (3.16). We can check that $L(p) \geq 1$ if

$$\tilde{\alpha} \geq \min\{2\alpha + \max\{r, r'\} + 1 - m_0 - r, 3\alpha + \max\{r, r'\} + 1 - 2m_0 - r'\}$$
$$= 2\alpha + \max\{r, r'\} + 1 - m_0 - r \ ,$$

which explains our choice for $\tilde{\alpha}$. □

### 3.4. Estimate of $f_n$

Going on with the iteration scheme, the next lemma gives the estimates of the source term $f_n$, defined by equations (3.4):

**Lemma 3.7.** *Let $\alpha$ and $\tilde{\alpha}$ be given as in Lemma 3.6. There exist $\delta > 0$ sufficiently small and $\theta_0 \geq 1$ sufficiently large, both chosen independently of $\alpha$, such that for all integer $m \in [m_0, \tilde{\alpha} + s]$, one has*

$$\|f_n\|_{Y_m} \leq C\,\Delta_n \big\{\theta_n^{m-\alpha-s-1}\,\|f\|_{Y_{\alpha+s}} + \delta^2\,\theta_n^{L(m)-1}\big\} \,. \tag{3.17}$$

*Proof.* From (3.4) we have

$$f_n = (S_{\theta_n} - S_{\theta_{n-1}})f - (S_{\theta_n} - S_{\theta_{n-1}})E_{n-1} - S_{\theta_n} e_{n-1} \,.$$

Using (2.3c) gives

$$\|(S_{\theta_n} - S_{\theta_{n-1}})f\|_{Y_m} \leq C\,\Delta_{n-1}\theta_{n-1}^{m-\alpha-s-1}\,\|f\|_{Y_{\alpha+s}} \tag{3.18}$$

for all $m \geq 0$. Using (2.3c), (3.16) gives

$$\|(S_{\theta_n} - S_{\theta_{n-1}})E_{n-1}\|_{Y_m} \leq C\,\Delta_{n-1}\delta^2\,\theta_{n-1}^{m-p-1+L(p)} \leq C\,\Delta_{n-1}\delta^2\,\theta_{n-1}^{L(m)-1} \,, \tag{3.19}$$

because $m - p + L(p) \leq L(m)$ for all $m \geq 0$. Moreover, from (2.3a), (3.15) we get

$$\|S_{\theta_n} e_{n-1}\|_{Y_m} \leq C\,\Delta_{n-1}\delta^2\,\theta_{n-1}^{L(m)-1} \,, \tag{3.20}$$

for all $m \geq m_0$. Finally, using $\theta_{n-1} \leq \theta_n \leq \sqrt{2}\theta_{n-1}$ and $\Delta_{n-1} \leq 3\,\Delta_n$ in (3.18)–(3.20) yields (3.17). □

### 3.5. Proof of induction

We now consider problem (3.5), that gives the solution $\delta u_n$.

**Lemma 3.8.** *Let $\alpha \geq m_0 + \max\{r, r'\} + \max\{s, s'\} + 1$, and let $\tilde{\alpha}$ be given as in Lemma 3.6. If $\delta > 0$ is sufficiently small, $\theta_0 \geq 1$ is sufficiently large, both chosen independently of $\alpha$, and $\|f\|_{Y_{\alpha+s}}/\delta$ is sufficiently small, then for all $m \in [m_0, \tilde{\alpha}]$, one has*

$$\|\delta u_n\|_{X_m} \leq \delta\,\theta_n^{m-\alpha-1}\,\Delta_n \,. \tag{3.21}$$

*Proof.* Let us consider problem (3.5). By (3.9a) $S_{\theta_n} u_n$ satisfies

$$\|S_{\theta_n} u_n\|_{X_{m_0}} \leq C\delta \ .$$

So for $\delta$ sufficiently small we may apply (2.2) in order to obtain

$$\|\delta u_n\|_{X_m} \leq C\,\Big(\|f_n\|_{Y_{m+s}} + \|f_n\|_{Y_{m_0}}\|S_{\theta_n} u_n\|_{X_{m+s'}}\Big) \,. \tag{3.22}$$

Estimating the right-hand side of (3.22) by Lemma 3.7 and (3.9a) yields

$$\|\delta u_n\|_{X_m} \le C\,\Delta_n\{\theta_n^{m-\alpha-1}\,\|f\|_{Y_{\alpha+s}} + \delta^2\,\theta_n^{L(m+s)-1}\} \\ + C\,\Delta_n\,\{\theta_n^{m_0-\alpha-s-1}\,\|f\|_{Y_{\alpha+s}} + \delta^2\,\theta_n^{L(m_0)-1}\}\,\delta\,\theta_n^{(m+s'-\alpha)_+}\,. \quad (3.23)$$

One checks that, for $\alpha \ge m_0 + \max\{r,r'\} + \max\{s,s'\} + 1$, and $m \in [m_0, \tilde\alpha]$, the following inequalities hold true:

$$\begin{aligned} &L(m+s) < m-\alpha\,,\\ &m_0-\alpha-s+(m+s'-\alpha)_+ \le m-\alpha\,,\\ &L(m_0)+(m+s'-\alpha)_+ < m-\alpha\,. \end{aligned} \quad (3.24)$$

From (3.23), we thus obtain

$$\|\delta u_n\|_{X_m} \le C\left(\|f\|_{Y_{\alpha+s}} + \delta^2\right)\theta_n^{m-\alpha-1}\,\Delta_n\,, \quad (3.25)$$

and (3.21) follows, for $\delta > 0$ and $\|f\|_{Y_{\alpha+s}}/\delta$ sufficiently small. □

The crucial point of the method is seen in (3.24): the quadratic nature of the errors is reflected in the estimate (3.15) by the presence of the term "$-2\alpha$", while the tame nature of the estimates contributes linearly in $m$ (with $|L'(m)| \le 1$). It is the "$-2\alpha$" term which allows (for $\alpha$ sufficiently large) to get (3.24) and close the induction.

Lemma 3.8 shows that $(H_{n-1})$ implies $(H_n)$ provided that $\alpha \ge m_0 + \max\{r,r'\} + \max\{s,s'\} + 1$, $\tilde\alpha = 2\alpha + \max\{r,r'\} + 1 - m_0 - r$, $\delta > 0$ is small enough, $\|f\|_{Y_{\alpha+s}}/\delta$ is small enough, and $\theta_0 \ge 1$ is large enough. We fix $\alpha$, $\tilde\alpha$, $\delta > 0$, and $\theta_0 \ge 1$, and we finally prove $(H_0)$.

**Lemma 3.9.** *If $\|f\|_{Y_{\alpha+s}}/\delta$ is sufficiently small, then property $(H_0)$ holds.*

*Proof.* Let us consider problem (3.5) for $n = 0$:

$$d\mathcal{F}(0)\,\delta u_0 = S_{\theta_0} f.$$

Applying (2.2) gives

$$\|\delta u_0\|_{X_m} \le C\,\|S_{\theta_0} f\|_{Y_{m+s}} \le C\,\theta_0^{(m-\alpha)_+}\,\|f\|_{Y_{\alpha+s}}.$$

Then

$$\|\delta u_0\|_{X_m} \le \delta\,\theta_0^{m-\alpha-1}\Delta_0\,, \qquad m_0 \le m \le \tilde\alpha,$$

provided $\|f\|_{Y_{\alpha+s}}/\delta$ is taken sufficiently small. □

### 3.6. Conclusion of the proof of Theorem 2.4 i)

Given an integer $\alpha \ge m_0 + \max\{r,r'\} + \max\{s,s'\} + 1$, in agreement with the requirements of Lemma 3.8, we take $\tilde\alpha = 2\alpha + \max\{r,r'\} + 1 - m_0 - r$ as in Lemma 3.6. If $\delta > 0$ and $\|f\|_{Y_{\alpha+s}}/\delta$ are sufficiently small, $\theta_0 \ge 1$ is sufficiently large, then $(H_n)$ holds true for all $n$. Let us set $m' = \alpha - 1$. In particular, from $(H_n)$ we obtain

$$\sum_{n\ge 0} \|\delta u_n\|_{X_{m'}} < +\infty\,, \quad (3.26)$$

so the sequence $\{u_n\}$ converges in $X_{m'}$ towards some limit $u \in X_{m'}$. From (3.7) we have

$$\mathcal{F}(u_{n+1}) - f = (S_{\theta_n} - I)f + (I - S_{\theta_n})E_n + e_n\,.$$

Using (2.3b), (3.15), (3.16) we can pass to the limit in the right-hand side in $Y_{m'+s}$ and get

$$\lim_{n\to\infty} \mathcal{F}(u_{n+1}) = \mathcal{F}(u) = f.$$

Therefore $u$ is a solution of (1.1), and the proof of Theorem 2.4 i) is complete.

*Remark* 3.10. In view of (2.2) with a loss of regularity of order $s$ from $g$, given $f \in Y_{m'+s+1}$ we could wish to find a solution $u \in X_{m'+1}$ instead of $u \in X_{m'}$ as above. The regularity of $u$ follows from the condition $m' < \alpha$ for the convergence of the series (3.26). Working with spaces $X_m$ with integer index the condition yields $m' \le \alpha - 1$; in spaces with real index it would be enough $m' \le \alpha - \epsilon$, for all $\epsilon > 0$, and we would get $u \in X_{m'+1-\epsilon}$.

### 3.7. Additional regularity of the solution constructed

Let us now prove assertion ii) of the Nash-Moser theorem. Let us assume that $f \in Y_{m''+s+1}$, with $m'' > m'$. Let us set $\alpha' = m'' + 1$ and define $\tilde{\alpha}'$ accordingly, $\tilde{\alpha}' = 2\alpha' + \max\{r, r'\} + 1 - m_0 - r$. The proof is obtained by finite induction. For it we shall use the estimate $(H_n)$ which is now true for all $n$, and the estimates that can be obtained from it.

We consider again (3.23) and remark that the exponents of $\theta_n$ of the terms not involving $f$ are strictly less than $m - \alpha - 1$, as shown in (3.24). On the other hand, the terms in (3.23) involving $f$ come from (3.17), or more precisely from (3.18). Using the fact that $f$ is now more regular, we can substitute (3.18) by

$$\|(S_{\theta_n} - S_{\theta_{n-1}})f\|_{Y_m} \le C\,\Delta_n \theta_n^{m-\alpha-s-2}\,\|f\|_{Y_{\alpha+s+1}},$$

and, accordingly, instead of (3.25) we find

$$\|\delta u_n\|_{X_m} \le C\left(\|f\|_{Y_{\alpha+s+1}} + \delta^2\right)\theta_n^{m-\alpha-2}\,\Delta_n \le C\,\theta_n^{m-\alpha-2}\,\Delta_n, \qquad \forall n \ge 0. \tag{3.27}$$

Starting from these new estimates instead of $(H_n)$, we can revisit the proof of assertion i). Note that in $e'_k, e''_k$ there is at least one factor involving $\delta u_n$ in each term. Estimating this factor by (3.27) gives

$$\|e_k\|_{Y_m} \le C\,\delta\,\theta_k^{L(m)-2}\,\Delta_k\,.$$

Going on with the repetition of the proof we obtain

$$\|E_n\|_{Y_{p+1}} \le C\,\delta\,\theta_n^{L(p)}\,,$$

$$\|f_n\|_{Y_m+1} \le C\,\Delta_n\left\{\theta_n^{m-\alpha-s-1}\,\|f\|_{Y_{\alpha+s+1}} + \delta^2\,\theta_n^{L(m)-1}\right\},$$

$$\|\delta u_n\|_{X_{m+1}} \le C\left(\|f\|_{Y_{\alpha+s+1}} + \delta\right)\theta_n^{m-\alpha-1}\,\Delta_n\,.$$

This gives the gain of one order. After a finite number of iterations of the same procedure we find

$$\|\delta u_n\|_{X_m} \le C\,\theta_n^{m-\alpha'-1}\,\Delta_n \qquad \text{for all } m \in [m_0, \tilde{\alpha}'].$$

The conclusion of the proof of assertion ii) follows as for (3.26).

## 4. Simplified case

To understand better the role of parameters in the induction of the proof, let us assume for simplicity that $m_0 = 0$, $r = r' = s = s' = 1$. Then estimate (3.15) holds with $L(m) = m + 1 - 2\alpha$. The number $p = \tilde{\alpha} - 1$ in (3.16) is chosen such that $L(p) = 1$ which yields $p = 2\alpha$, $\tilde{\alpha} = 2\alpha + 1$. To close the induction we choose $\alpha$ from (3.24) that now reads

$$\begin{aligned} & m + 2 - 2\alpha < m - \alpha\,, \\ & -\alpha - 1 + (m + 1 - \alpha)_+ \leq m - \alpha\,, \\ & 1 - 2\alpha + (m + 1 - \alpha)_+ < m - \alpha\,. \end{aligned}$$

Here it is sufficient to take $\alpha > 2$, i.e. $\alpha \geq 3$, and $(H_n)$ will hold for all $m \in [0, 2\alpha + 1]$. The quadratic nature of the errors with the presence of the term "$-2\alpha$"allows (for $\alpha$ sufficiently large) to close the induction. Thus, the same nonlinearity of the equation is exploited for the convergence of the approximating sequence.

## References

[1] S. Alinhac, *Existence d'ondes de raréfaction pour des systèmes quasi-linéaires hyperboliques multidimensionnels.* Comm. Partial Differential Equations **14** (1989), 173–230.

[2] S. Alinhac, P. Gérard. *Opérateurs pseudo-différentiels et théorème de Nash-Moser.* InterEditions, 1991.

[3] V. I. Arnol′d, *Small denominators. I. Mapping the circle onto itself.* Izv. Akad. Nauk SSSR Ser. Mat. **25** (1961), 21–86.

[4] V. I. Arnol′d, *Small denominators and problems of stability of motion in classical and celestial mechanics.* Uspehi Mat. Nauk **18** (1963), 91–192.

[5] G.-Q. Chen, Y.-G. Wang, *Existence and stability of compressible current-vortex sheets in three-dimensional magnetohydrodynamics.* Arch. Ration. Mech. Anal. **187** (2008), 369–408.

[6] J. F. Coulombel, P. Secchi, *Nonlinear compressible vortex sheets in two space dimensions.* Ann. Sci. École Norm. Sup. (4) **41** (2008), 85–139.

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

[8] M. Günther, *On the perturbation problem associated to isometric embeddings of Riemannian manifolds.* Ann. Global Anal. Geom. **7** (1989), 69–77.

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

[10] L. Hörmander, *The boundary problems of physical geodesy.* Arch. Rational Mech. Anal. **62** (1976), 1–52.

[11] L. Hörmander. *Implicit function theorems.* Stanford University Lecture Notes, 1977.

[12] T. Kato. *Abstract differential equations and nonlinear mixed problems.* Lezioni Fermiane. [Fermi Lectures]. Scuola Normale Superiore, Pisa, 1985.

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

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

[15] J. Moser, *A rapidly convergent iteration method and non-linear differential equations. II.* Ann. Scuola Norm. Sup. Pisa (3) **20** (1966), 499–535.

[16] J. Nash, *The imbedding problem for Riemannian manifolds.* Ann. of Math. (2) **63** (1956), 20–63.

[17] J. Schwartz, *On Nash's implicit functional theorem.* Comm. Pure Appl. Math. **13** (1960), 509–530.

[18] P. Secchi, Y. Trakhinin, *Well-posedness of the plasma-vacuum interface problem.* Preprint (2013), http://arxiv.org/abs/1301.5238.

[19] Y. Trakhinin, *The existence of current-vortex sheets in ideal compressible magnetohydrodynamics.* Arch. Ration. Mech. Anal. **191** (2009), 245–310.

Paolo Secchi
DICATAM, Sect. of Mathematics
University of Brescia
Via Valotti, 9
25133 Brescia, Italy
e-mail: `paolo.secchi@ing.unibs.it`