The Nash-Moser Iteration Technique with Application to Characteristic Free-Boundary Problems

Ben Stevens

Abstract These notes are an overview of the Nash-Moser iteration technique for solving PDEs (or other non-linear problems) via linearisation, where the linearised equations admit estimates with a loss of regularity with respect to the source term, coefficients and/or boundary/initial data. We first introduce the abstract setting along with a version of the iteration scheme due to Hörmander (Arch Ration Mech Anal 62(1):1–52, 1976). We then introduce some modifications which allow the scheme to be applied to some characteristic free-boundary problems for hyperbolic conservation laws. We focus on the case of supersonic vortex sheets in 2D as considered by Coulombel and Secchi in Ann Sci Éc Norm Supér (4) 41(1):85–139, 2008.

2010 Mathematics Subject Classification 76N10 (35L65 35L67 35Q35 35R35)

1 Introduction

1.1 Summary

These notes are an overview of the Nash-Moser iteration technique for solving PDEs (or other non-linear problems) via linearisation, where the linearised equations admit estimates with a loss of regularity with respect to the source term, coefficients

B. Stevens (✉)
Mathematical Institute, University of Oxford, Oxford, OX1 3LB, UK
e-mail: ben.stevens@maths.ox.ac.uk

G.-Q.G. Chen et al. (eds.), *Hyperbolic Conservation Laws and Related Analysis with Applications*, Springer Proceedings in Mathematics & Statistics 49,
DOI 10.1007/978-3-642-39007-4_13, © Springer-Verlag Berlin Heidelberg 2014

and/or boundary/initial data. In these situations, Picard iteration (or the contraction mapping principle) fails, but a modified form of Newton-Raphson iteration, involving the application of smoothing operators to overcome the loss of regularity, may succeed in finding a solution for given data close to some special data for which a solution is known to exist. This technique is known as Nash-Moser iteration, or in some contexts as the Nash-Moser inverse function theorem. It was originally used by Nash in [21] for solving the isometric embedding problem. Moser in [20] and Schwartz in [23] simplified the method at the expense of a loss of regularity and showed how it could be applied in a more general setting. Hörmander, in his paper [15] on the boundary problems of physical geodesy, improved on Moser's scheme by reducing the loss of regularity, using a scheme more similar to Nash's original. More recently, Alinhac in [2] used a modified version of Hörmander's scheme to prove the short-time existence of rarefaction waves for a class of conservation laws and Coulombel and Secchi in [8] introduced an additional modification to prove the short-time existence of vortex sheets for the two dimensional isentropic Euler equations provided the Mach number is sufficiently large. A scheme similar to the one used by Coulombel and Secchi is also developed by Chen and Wang in [5] and [6] to prove the short-time existence of current-vortex sheets for three-dimensional MHD under certain stability assumptions.

We aim to provide an abstract setting for the technique, whilst keeping in mind that we want to apply it to PDE problems. Hopefully in an abstract setting it will be easier to see the key hypotheses needed on the equations to be solved than in specific situations, which may involve other complications. We first introduce the scheme used by Hörmander in [15], and detailed by Alinhac and Gérard in [3], which is closer than Moser's scheme to Nash's original technique except that Hörmander uses a discrete approximation scheme rather than one based on a continuous parameter t. Whilst Hörmander works in Hölder spaces, we work in more general Banach spaces, at the price of losing a small degree of regularity. We have in mind that the linearised equations are most likely to be estimated in Sobolev spaces (or weighted Sobolev spaces), probably with exponent two. This technique has the advantage over Moser's technique of obtaining a solution which is closer in regularity to the given data, but although Nash used his method to obtain optimal regularity, we are unlikely to obtain an optimal regularity result using this method in more complicated situations.

We then introduce a more complicated scheme which allows us to deal with difficulties in solving the linearised equations, inspired by the paper on 2D compressible vortex sheets by Coulombel and Secchi [8].

Following this, we give the construction of the smoothing operators used in Nash-Moser iteration on some Sobolev spaces which are used in practice, and some inequalities useful for obtaining the tame estimates used in the iteration scheme.

Finally, we show how the generalised scheme can be applied to the case considered by Coulombel and Secchi in [8], in a slightly simplified manner but at the expense of some loss of regularity.

1.2 Newton-Raphson Iteration, Picard Iteration, and Nash-Moser Iteration

Suppose we wish to solve the nonlinear equation $T(u) = f$ for the unknown $u \in X$, given $f \in Y$, where $T : X \to Y$. So as not to ask too much, let us look for a solution u close to u_0 of the equation $T(u) = T(u_0) + f$, where f is small. One of the most classical methods for solving such a nonlinear equation via linearisation is Newton-Raphson iteration. For $n \geq 1$, we set

$$u_{n+1} = u_n - L(u_n)(T(u_n) - T(u_0) - f)$$

where $L(u)$ is a right inverse of $DT(u)$. One can check by applying T to both sides and using Taylor's theorem that $T(u_{n+1}) = T(u_0) + f$ plus terms involving $u_{n+1} - u_n$ which one would hope to converge to zero. However, for this scheme to even make sense, we need an operator $L(u) : Y \to X$ which is a right inverse of $DT(u)$. The linearised equations $DT(u)v = g$ themselves may be difficult or impossible to solve for $v \in X$, hence we may not be able to find such an operator L.

As a possible remedy to this problem, we consider the contraction mapping theorem, or Picard iteration, which uses a slightly different kind of linearisation and may be able to solve equations where the operator L as above does not exist. For example, suppose we can write our equation in the form

$$S(u)u = 0$$

where, for fixed u, $S(u)$ is a linear operator. We seek the unknown $u \in X$, where X is a complete metric space, and we assume the initial/boundary conditions have been absorbed into the definition of X. We now define the map $F : X \to X$ by $F(u) = v$, where v is the solution to the linear equation

$$S(u)v = 0.$$

If we can prove that F is well-defined, and that F is a contraction, i.e. $d_X(F(u_1), F(u_2)) \leq \kappa d_X(u_1, u_2)$, where $\kappa < 1$, for all u_1, u_2 in X, then the contraction mapping theorem implies that F has a fixed point, w. By construction, w satisfies the original nonlinear equation we wished to solve.

Note that in order to apply this method, we require that the solution v of the linear equation be in the same space as u, on which v depends through the coefficients of the equation. This is a better situation than for Newton-Raphson iteration, which requires that the operator $L(u)$ regains the regularity lost by applying the operator T.

We can also write this method as an explicit iteration scheme (effectively re-proving the contraction mapping theorem). We pick $u_0 \in X$ and for $n \geq 0$ we define u_{n+1} as the solution of the linear equation

$$S(u_n)u_{n+1} = 0.$$

We then aim to show that, for $n \geq 1$, $d_X(u_{n+1}, u_n) \leq \kappa d_X(u_n, u_{n-1})$. This will ensure u_n is a Cauchy sequence which converges to a solution of the nonlinear equation. Using the explicit iteration scheme (known as Picard iteration) allows more scope for slight modification in specific cases. For example, Majda in [16], uses this iteration scheme, modified to include a smoothing of the initial data, to prove the short-time existence of classical solutions to multidimensional systems of conservation laws with a convex entropy.

However, it is possible that we cannot solve the linearised problem above for v in the same space as u, as required by Picard iteration. It may happen that we can solve the linear equation, but only for $v \in Z$, where $X \subset Z$. For example, perhaps, given $u \in C^k$, we can only prove that a solution v to the linearised equation exists in C^{k-1}. We refer to this as a loss of regularity in solving the linearised problem.

To overcome this, the key idea of Nash was to return to Newton-Raphson iteration, but to modify the scheme to include a smoothing operation at each step to compensate for the loss of regularity. Returning to the equation $T(u) = T(u_0) + f$, standard Newton-Raphson iteration may be written as follows. For $n \geq 0$, we set

$$u_{n+1} = u_n + \dot{u}_n.$$

The difference $\dot{u}_n$ is given by

$$\dot{u}_n = L(u_n)g_n$$

for

$$g_n = f + T(u_0) - T(u_n)$$

where $L(u)$ is a right-inverse of $DT(u)$.

Now let us suppose we have a family of smoothing operators S_n that regain the regularity lost by T and L, and such that $S_n \to \text{id}$ as $n \to \infty$. Then there are two obvious ways we can modify the scheme.

The simplest is to set $u_{n+1} = u_n + S_n\dot{u}_n$, i.e. we smooth $\dot{u}_n$ after applying the operators T and L to u_n. Since $S_n \to \text{id}$ as $n \to \infty$, this scheme looks like Newton-Raphson iteration for large n, so we might expect it to converge under certain conditions. This method is used by Moser in [20] and Schwartz in [23]. Whilst this is a very simple modification, it has the drawback that a solution u obtained by this method has a much lower degree of regularity than the given data f.

The other obvious modification is to smooth u_n before we apply the operators T and L. Thus we set

$$\dot{u}_n = L(S_n u_n)g_n.$$

We also adjust our choice of g_n (which should be smoothed) given this modification. This method is used by Hörmander in [15] and a continuous-parameter version was used by Nash in his original paper [21]. We motivate how to choose g_n in Sect. 3.1, which is based on the motivation given by Alinhac and Gérard in [3]. Again, the fact that $S_n \to \mathrm{id}$ as $n \to \infty$ means the scheme looks like Newton-Raphson iteration for large n. The advantage of this method is that the solution u obtained can be quite close in regularity to the given data f, but generally the regularity obtained will not be optimal. In modifying Hörmander's method to deal with more general Banach spaces instead of just Hölder spaces, we lose an arbitrarily small degree of regularity if we can use fractional index spaces, or one degree of regularity if we are using integer index spaces. Other modifications to the scheme used in practice further reduce the degree of regularity obtained. Nevertheless, we may consider this an improvement over Moser's technique, which we can informally attribute to the fact that we have carefully constructed g_n to compensate for the introduction of the smoothing operators, whereas Moser's method involves no such modification.

1.3 Nash-Moser Iteration as an Inverse Function Theorem

It is instructive to consider a slightly different viewpoint, that is to consider Nash-Moser iteration as an inverse function theorem for a certain class of Fréchet spaces, which are a natural generalisation of Banach spaces.

Indeed, the standard version of the Inverse Function Theorem, which can be proved (under slightly stronger hypotheses than usual to make things simpler) by an application of the contraction mapping theorem with parameter, carries over analogously to an operator $T : X \to Y$ between Banach spaces. By this we mean that if the Fréchet derivative $DT(u)$ of T is invertible at a point $u \in X$, then T itself is invertible in a neighbourhood of u. Hence, if we wish to solve the equation $T(u) = T(u_0) + f$ for u near u_0, where f is small, we can simply apply the inverse function theorem.

However, it is possible in applications that we can only find an 'unbounded' inverse for $DT(u)$. For example, if we work with differential operators in the spaces C^k of k-times differentiable functions, then we might have $T : C^k \to C^{k-1}$, but we might only be able to find a right inverse $L(u)$ of $DT(u)$ on some subset of C^{k-1}, for example on C^k, so that $L(u) : C^k \to C^k$, or, even worse, on C^{k+1} so that $L(u) : C^{k+1} \to C^k$. This is solved if we work in the space $X = C^\infty$, since then $L(u)$ maps X to itself. However, this is no longer a Banach space, but a Fréchet space. Thus we are led to ask whether there is an inverse function theorem for Fréchet spaces. The answer is that if we assume the existence of a certain family of smoothing operators on our Fréchet space (which by no means exist in general, but do for most spaces of differentiable functions commonly used), then there is a sort of inverse function theorem. This requires that $DT(u)$ be invertible on a neighbourhood of u, not just at u itself.

This point of view is elegantly considered by Hamilton in [14], who refers to this special class of Fréchet spaces as 'tame' Fréchet spaces and the necessary estimates involved on the operator T as 'tame' estimates. The proof of this result uses Nash-Moser iteration, and Hamilton's proof in particular is quite close to

Nash's original method. The similarity with the usual inverse function theorem is why Nash-Moser iteration is sometimes referred to as the Nash-Moser inverse function theorem or the Nash-Moser implicit function theorem. See also the chapter 'Generalized Implicit Function Theorems' written by E. Zehnder in Nirenberg [22] for an introduction to Nash-Moser type theorems as generalisations of the standard inverse/implicit function theorem. Another implicit function theorem in the setting of Fréchet spaces is given by Ekeland in [10], whose approach does not rely on Newton-Raphson iteration but on Lebesgue's dominated convergence theorem and Ekeland's variational principle.

Whilst this viewpoint is conceptually simple, for actual applications to PDEs, working in Fréchet spaces is not necessary and complicates matters, and it is easier to consider a family of Banach spaces in which one has estimates for the linearised equations, for example $(C^k)_{k\in\mathbb{N}}$ or $(H^s)_{s\in\mathbb{R}_{\geq 0}}$.

1.4 Tame Estimates

The key estimates involved in Nash-Moser iteration are known as tame estimates. These are estimates of the following form. (Here we use the spaces C^k for definiteness.)

Let $T : C^\infty \to C^\infty$.

Then T satisfies a tame estimate if

$$||T(u)||_{C^k} \leq C_k(1 + ||u||_{C^{k+k_1}})$$

for some fixed integer k_1 and all u in some fixed bounded set $U \subset C^{k_0}$, for some k_0, where the constant $C_k > 0$ is independent of u.

The key point about this estimate is that it is affine in the norm of u on the right hand side with the variable index k.

Similarly, the second derivative of T, D^2T, is said to satisfy a tame estimate if

$$\begin{aligned}&\left|\left|D^2T(u)(v_1, v_2)\right|\right|_{C^k}\\ &\leq C_k(||v_1||_{C^{k_1+k}}\,||v_2||_{C^{k_2}} + ||v_1||_{C^{k_1}}\,||v_2||_{C^{k_2+k}} + ||v_1||_{C^{k_1}}\,||v_2||_{C^{k_2}}\,(1 + ||u||_{C^{k+k_3}}))\end{aligned}$$

for some fixed integers k_1, k_2, k_3 and all u in some fixed bounded set $U \subset C^{k_0}$, for some k_0, where the constant $C_k > 0$ is independent of u, v_1 and v_2.

Note that this estimate is also affine in the norms on the right hand side with the variable index k, and in addition it is quadratic (with no affine terms) in (v_1, v_2), which will be a key point in the iteration. The smoothing operators will control the large k norms in terms of lower ones at the price of poorer estimates and we require DT to be a good approximation for T to compensate.

Note that the framework of tame estimates fits differential operators well because of product estimates of the form

$$||fg||_{H^s} \leq C_s(||f||_{H^r}\,||g||_{H^s} + ||f||_{H^s}\,||g||_{H^r})$$

for $r > \frac{d}{2}$, where d is the dimension.

Similarly, we have estimates for compositions $G(x) = F(u(x))$ (sometimes called Moser-type inequalities) of the form

$$||\partial^\alpha G||_{L^2} \leq C_s\,||u||_{H^{|\alpha|}}$$

for u in an H^r-bounded set.

These estimates can be derived from the Sobolev embedding theorem for large index s, and details of these estimates for certain classes of Sobolev Spaces are given in Sect. 5.2.

2 The Abstract Setting

In order to describe Nash-Moser iteration in an abstract setting we will need to introduce some notation, as well as the idea of a derivative in this setting. We will simply use the notion of a directional derivative, since all we need is a linear approximation to an operator which satisfies Taylor's theorem.

2.1 Families of Banach Spaces and Differentiation

Definition 1. Let I be an interval in $\mathbb{R}$ or $\mathbb{Z}$ of the form $[0, a)$, $[0, a]$, or $[0, \infty)$, where $a > 0$.

We will say $\{X^s\}_{s\in I}$ is a *decreasing family of Banach spaces* if, for each $s \in I$, X^s is a Banach space with norm $||\cdot||_{X^s}$, and, for $s_1, s_2 \in I$ with $s_1 \leq s_2$, we have

$$X^{s_2} \subset X^{s_1} \text{ with } ||\cdot||_{X^{s_2}} \geq ||\cdot||_{X^{s_1}} \text{ on } X^{s_2}.$$

We will write

$$X^\infty = \cap_{s\in I} X^s$$

and

$$X^{\infty-m} = \cap_{s\in I, s\geq m} X^{s-m}$$

for $m \in I$.

Remark 1. Note that it is convenient to use the notation X^∞ for the intersection of all the Banach Spaces X^s with $s \in I$, even if I is a finite interval. In the case that $I = [0, \infty)$, $X^{\infty-m}$ as defined above is the same as X^∞, but if I is a finite interval then they are not the same.

Definition 2. Let $\{X^s\}_{s\in I}$ be a decreasing family of Banach spaces. Let $\alpha : U \to X^\infty$ where $U \subset \mathbb{R}$ is open, and let $t \in U$. We say α is *differentiable at t* if there exists a $w \in X^\infty$ such that

$$\left|\left|\frac{\alpha(t+h)-\alpha(t)}{h} - w\right|\right|_{X^s} \to 0 \text{ as } h \to 0\ (h \neq 0)$$

for all $s \in I$.

If such a w exists, we say w is the derivative of α at t, and write $\alpha'(t) = w$ or $\frac{\mathrm{d}\alpha}{\mathrm{d}t}(t) = w$.

We say α is *differentiable* if it is differentiable at t for all $t \in U$.

Definition 3. Let $\{X^s\}_{s\in I}$ and $\{Y^s\}_{s\in I}$ be two decreasing families of Banach spaces. Let $T : U \to Y^{\infty-m}$ for some $m \in I$, where $U \subset X^\infty$ is $||\cdot||_{X^r}$-open for some $r \in I$, and let $u \in U$. We say T is *differentiable at u* if, for each $v \in X^\infty$, the map $\alpha_v : (-\epsilon, \epsilon) \to Y^{\infty-m}$ defined on a small neighbourhood of 0 in $\mathbb{R}$ by

$$\alpha_v(t) = T(u + tv)$$

is differentiable at 0 in the sense of Definition 2, and

$$\alpha_v'(0) = DT(u)v$$

where $DT(u) : X^\infty \to Y^{\infty-m}$ is a linear map. We call $DT(u)$ the derivative of T at u.

We say T is *differentiable* if it is differentiable at u for all $u \in U$ and call DT the derivative of T.

For an integer $k \geq 2$, we say T is *k-times differentiable* with k-th derivative D^kT if the following inductive definition holds.

T is $k-1$ times differentiable with $(k-1)$-th derivative at u given by $D^{k-1}T(u) : (X^\infty)^{k-1} \to Y^{\infty-m}$ for each $u \in U$.

For each ordered set $(v_1, \ldots, v_{k-1}) \in (X^\infty)^{k-1}$, the map $S : U \to Y^{\infty-m}$ defined by

$$S(u) = D^{k-1}(u)(v_1, \ldots, v_{k-1})$$

is differentiable in the above sense.

Define the k-th derivative of T at $u \in U$ as $D^kT(u) : (X^\infty)^k \to Y^{\infty-m}$ where

$$D^kT(u)(v_1, \ldots, v_k) = DS(u)v_k.$$

Remark 2. We will not need all the properties of standard derivatives. We merely require a linear approximation to within quadratic error of a nonlinear operator. Hence we give the above fairly weak definition of differentiability and don't worry about questions such as whether the partial derivatives commute.

Proposition 1. *Let $\{X^s\}_{s\in I}$ and $\{Y^s\}_{s\in I}$ be two decreasing families of Banach spaces. Let $T : U \to Y^{\infty-m}$ for some $m \in I$, where $U \subset X^\infty$ is $||\cdot||_{X^r}$-open* *for some $r \in I$. Then Taylor's theorem holds for T. More precisely, suppose T is k-times differentiable (in the sense of Definition 3) for some $k \geq 1$, let $u \in U$, $v \in X^\infty$, and suppose the line segment $[u, u+v]$ is contained in U. Then*

$$T(u+v) = T(u) + DT(u)v + \ldots + \frac{1}{(k-1)!} D^{k-1}T(u)(v, \ldots, v) + R_{k,u}(v)$$

where

$$||R_{k,u}(v)||_{Y^s} \leq \frac{1}{k!} \sup_{t \in [0,1]} \left|\left|D^k T(u+tv)(v, \ldots, v)\right|\right|_{Y^s}$$

for all $s \in I$ such that $s + m \in I$.

Proof. Fix $s \in I$ such that $s + m \in I$. Let $\phi \in (Y^s)^*$ be a continuous linear functional on Y^s.

Define $g : J \to \mathbb{R}$ by

$$g(t) = \phi \circ T(u+tv)$$

where J is an open interval in $\mathbb{R}$ containing $[0, 1]$.

Since ϕ is a continuous linear functional on Y^s, from the definition of differentiability we have that g is k-times differentiable with

$$g^{(k)}(t) = \phi \circ D^k T(u+tv)(v, \ldots, v).$$

Applying the one-dimensional Taylor's theorem to obtain an expansion for $g(1)$ about $g(0)$, we have

$$g(1) = g(0) + g'(0) + \ldots + \frac{1}{(k-1)!} g^{k-1}(0) + \frac{1}{k!} g^k(h)h^k$$

for some $h \in [0, 1]$ (which may depend on ϕ). Hence

$$\phi \circ T(u+v) =$$

$$\phi(T(u) + DT(u)v + \ldots + \frac{1}{(k-1)!} D^{k-1}T(u)(v, \ldots, v) + \frac{1}{k!} h^k D^k T(u+hv)(v, \ldots, v))$$

Rearranging, we have

$$\left|\phi(T(u+v) - (T(u) + DT(u)v + \ldots + \frac{1}{(k-1)!} D^{k-1}T(u)(v, \ldots, v)))\right|$$

$$\leq ||\phi||_{(Y^s)^*} \left|\left|\frac{1}{k!} h^k D^k T(u+hv)(v, \ldots, v)\right|\right|_{Y^s}$$

$$\leq ||\phi||_{(Y^s)^*} \frac{1}{k!} \sup_{t \in [0,1]} \left|\left|D^k T(u+tv)(v, \ldots, v)\right|\right|_{Y^s}.$$

Now use the Hahn-Banach theorem to pick $\phi \in (Y^s)^*$ with $||\phi||_{(Y^s)^*} = 1$ such that

$$\phi(T(u+v) - (T(u) + DT(u)v + \ldots + \frac{1}{(k-1)!}D^{k-1}T(u)(v, \ldots, v)))$$
$$= \left|\left|T(u+v) - (T(u) + DT(u)v + \ldots + \frac{1}{(k-1)!}D^{k-1}T(u)(v, \ldots, v))\right|\right|_{Y^s}.$$

We then obtain

$$\left|\left|T(u+v) - (T(u) + DT(u)v + \ldots + \frac{1}{(k-1)!}D^{k-1}T(u)(v, \ldots, v))\right|\right|_{Y^s}$$
$$\leq \frac{1}{k!} \sup_{t\in[0,1]} \left|\left|D^k T(u+tv)(v, \ldots, v)\right|\right|_{Y^s}.$$

This completes the proof.

Remark 3. Note that we can apply the above proposition when $\{X^s\}_{s\in I}$ is just $\{\mathbb{R}\}_{s\in I}$ to obtain Taylor's theorem for paths in Y^∞.

2.2 *Definition of the Smoothing Operators*

Definition 4. We will say a decreasing family of Banach spaces $\{X^s\}_{s\in I}$ *satisfies the smoothing hypothesis* if there exists a family of linear operators $\{S_\theta\}_{\theta\in\mathbb{R}_{\geq 1}}$ such that

$$S_\theta : X^0 \to X^\infty$$

and, for $u \in X^s$, we have

$$||S_\theta u||_{X^r} \leq C_{r,s}\theta^{(r-s)_+} \, ||u||_{X^s} \text{ for all } r, s \in I \tag{1}$$

$$||u - S_\theta u||_{X^r} \leq C_{r,s}\theta^{-(s-r)} \, ||u||_{X^s} \text{ for all } r, s \in I \text{ with } r \leq s \tag{2}$$

$$\left|\left|\frac{\mathrm{d}}{\mathrm{d}\theta}S_\theta u\right|\right|_{X^r} \leq C_{r,s}\theta^{r-s-1} \, ||u||_{X^s} \text{ for all } r, s \in I \tag{3}$$

where the constant $C_{r,s} > 0$ remains bounded if r and s remain bounded.

Here $(a)_+$ denotes $\max\{a, 0\}$ for $a \in \mathbb{R}$ or $a \in \mathbb{Z}$.

Note $\frac{\mathrm{d}}{\mathrm{d}\theta}S_\theta u$ is the derivative of the map $\theta \mapsto S_\theta u$ in the sense of Definition 2, which we require to exist for each $u \in X^0$.

3 Hörmander's Version of Nash-Moser Iteration

3.1 Motivation for the Iteration Scheme

Here we provide some motivation for the iteration scheme used by Hörmander in [15] by comparing it to Newton-Raphson iteration. This is unnecessary for the proof of the theorem, but the iteration scheme seems a little unmotivated without it. This motivation is partly based on the motivation given in Alinhac and Gérard [3].

3.1.1 Newton-Raphson Iteration

In order to solve the equation

$$T(u) = T(u_0) + f$$

the Newton-Raphson method uses the following iteration scheme.

$$u_{n+1} = u_n - L(u_n)(T(u_n) - (T(u_0) + f))$$

for L a right inverse of DT.

One way of justifying this is as follows.

We set

$$u_{n+1} = u_n + \dot{u}_n$$

where the increment $\dot{u}_n$ is to be determined. We then have

$$T(u_{n+1}) = T(u_n) + DT(u_n)\dot{u}_n + e_n$$

which defines the error e_n incurred by using the derivative of T to obtain a linear approximation to T. By Taylor's theorem, we expect this to be small when $\dot{u}_n$ is small.

Let us choose $\dot{u}_n$ such that

$$DT(u_n)\dot{u}_n = g_n$$

i.e.

$$\dot{u}_n = L(u_n)g_n$$

where g_n is to be determined so that u_n converges to a solution u of $T(u) = T(u_0) + f$.

From the equation

$$T(u_{n+1}) = T(u_n) + g_n + e_n$$

we obtain

$$T(u_{n+1}) = T(u_0) + \sum_{m=0}^{n} g_m + \sum_{m=0}^{n} e_m$$
$$= T(u_0) + \sum_{m=0}^{n} g_m + E_n + e_n$$

where

$$E_n = \sum_{m=0}^{n-1} e_m.$$

Thus if we define g_n by

$$\sum_{m=0}^{n} g_m + E_n = f$$

we obtain

$$T(u_{n+1}) = T(u_0) + f + e_n$$

which we hope converges to $T(u_0) + f$ as $n \to \infty$ since $e_n \to 0$.

The formula for g_n implies $g_0 = f$ and

$$g_{n+1} = -e_n$$
$$= T(u_n) + g_n - T(u_{n+1}).$$

Hence

$$g_{n+1} = T(u_0) + f - T(u_{n+1}).$$

Thus we obtain the iteration scheme

$$u_{n+1} = u_n - L(u_n)(T(u_n) - (T(u_0) + f))$$

3.1.2 Nash-Moser Iteration

We still wish to use an iteration scheme of the form

$$u_{n+1} = u_n + \dot{u}_n$$

but we are now concerned with the case when the application of the operator $L(u_n)$ to g_n causes a loss of regularity with respect to u_n and g_n. By this we mean that if u_n and g_n lie in X^s, then $L(u_n)g_n$ will lie in a larger space $X^{s'}$ for $s' < s$ so that for any fixed s the norm $||u_n||_{X^s}$ will blow up as $n \to \infty$. This loss of regularity is stated precisely in (5).

To overcome this, we apply smoothing operators S_n which allow us to control $||S_n u_n||_{X^s}$ for large s in terms of $||u_n||_{X^s}$ for small s. By choosing S_n to vary with n so that $S_n \to \text{id}$ in some sense as $n \to \infty$, we hope to be able to overcome the error introduced by these smoothing operators. In this particular version of Nash-Moser iteration, we follow Hörmander in [15] and Alinhac and Gérard in [2] by choosing to apply smoothing operators before the application of the operator L. Hence we define

$$v_n = S_n u_n$$

and set

$$T(u_{n+1}) = T(u_n) + DT(v_n)\dot{u}_n + e_n$$

which defines the error e_n incurred by using the derivative of T, evaluated at v_n, to obtain a linear approximation to T. By Taylor's theorem, and the fact that $S_n \to \text{id}$, we expect this to be small when $\dot{u}_n$ is small and n is large.

Following the same process as before, we define

$$\dot{u}_n = L(v_n)g_n$$

where g_n is to be determined so that u_n converges to a solution u of $T(u) = T(u_0) + f$, and g_n should be smoothed.

From the equation

$$T(u_{n+1}) = T(u_n) + g_n + e_n$$

we obtain

$$\begin{aligned} T(u_{n+1}) &= T(u_0) + \sum_{m=0}^{n} g_m + \sum_{m=0}^{n} e_m \\ &= T(u_0) + \sum_{m=0}^{n} g_m + E_n + e_n \end{aligned}$$

where

$$E_n = \sum_{m=0}^{n-1} e_m.$$

Before we defined g_n by

$$\sum_{m=0}^{n} g_m + E_n = f$$

but since we would like g_n to be smoothed, we define g_n by

$$\sum_{m=0}^{n} g_m = S_n(f - E_n).$$

From this, we obtain

$$T(u_{n+1}) = T(u_0) + S_n f + E_n - S_n E_n + e_n$$

which we hope converges to $T(u_0) + f$ as $n \to \infty$ since $e_n \to 0$ and $S_n \to \text{id}$.

The formula for g_n implies $g_0 = S_0 f$ and

$$\begin{aligned} g_{n+1} &= S_{n+1}(f - E_{n+1}) - S_n(f - E_n) \\ &= (S_{n+1} - S_n)(f - E_n) - S_{n+1}e_n. \end{aligned}$$

Note that we may split the error e_n up into two parts,

$$e_n = e_n' + e_n''$$

where

$$e_n' = (DT(u_n) - DT(v_n))\dot{u}_n$$

is the error caused by replacing u_n by v_n and

$$e_n'' = T(u_{n+1}) - T(u_n) - DT(u_n)\dot{u}_n$$

is the standard quadratic error in the Newton-Raphson scheme.

3.2 Statement and Proof of the Theorem

Theorem 1. *Let $\{X^s\}_{s\in I}$ and $\{Y^s\}_{s\in I}$ be two decreasing families of Banach spaces, each satisfying the smoothing hypothesis. Let $u_0 \in X^\infty$ and let $T : U^{m_0} \to Y^0$ be continuous, where $U^{m_0} \subset X^{m_0}$ is a bounded open neighbourhood of u_0 in X^{m_0}, for some $m_0 \in I$. Suppose also $T : U \to Y^{\infty - m_1}$ for some fixed $m_1 \in I$, where $U := U^{m_0} \cap X^\infty$, and T satisfies the following conditions.*

1. T is twice differentiable in the sense of Definition 3 and

$$\begin{aligned} &\left|\left|D^2T(u)(v_1,v_2)\right|\right|_{Y^s} \\ &\leq C_s^1(||v_1||_{X^{s+m_1}}\,||v_2||_{X^{m_2}} + ||v_1||_{X^{m_2}}\,||v_2||_{X^{s+m_1}} + ||v_1||_{X^{m_2}}\,||v_2||_{X^{m_2}}\,(1+||u||_{X^{s+m_3}})) \end{aligned} \tag{4}$$

for all $u \in U$, $v_1, v_2 \in X^\infty$ and $s \in I$ such that $s+m_1, s+m_3 \in I$, for some fixed numbers $m_1, m_2, m_3 \in I$, where the constant $C_s^1 > 0$ is bounded for s bounded.

2. For each $u \in U$, there exists a linear map $L(u) : Y^\infty \to X^{\infty-\max\{l_1,m_4\}}$ such that

$$DT(u)L(u) = \mathrm{id}$$

and

$$||L(u)g||_{X^s} \leq C_s^2(||g||_{Y^{s+l_1}} + ||g||_{Y^{l_1}}\,||u||_{X^{s+m_4}}) \tag{5}$$

for all $u \in U$, $g \in Y^\infty$ and $s \in I$ such that $s+l_1, s+m_4 \in I$, for some fixed numbers $l_1, m_4 \in I$, where the constant $C_s^2 > 0$ is bounded for s bounded.

Let $r_0 \in I$ with $r_0 > \max\{m_0, m_4, l_1+m_1+m_2, 2m_2, \frac{l_1+m_3}{2}+m_2\}$ and let $r_0+1 < s_1 \in I$ such that $s_1 + \max\{l_1, m_4\} \in I$ be sufficiently large depending on the constants m_i.

Then there exists a constant $0 < \epsilon \leq 1$ such that if $f \in Y^{r_0+l_1}$ with

$$||f||_{Y^{r_0+l_1}} \leq \epsilon$$

we can find $u \in U^{m_0}$ which solves the equation

$$T(u) = T(u_0) + f.$$

Moreover, let $J = \{r \in I : f \in Y^{r+l_1}, r \geq r_0\}$. Then for each $r \in J$ and $s \in I$ with $s < r$, assuming that $s_1 + r - r_0 + \max\{l_1, m_4\} \in I$, we have $u \in X^s$, and there exists a constant $K_{r,s}$ independent of f such that

$$||u-u_0||_{X^s} \leq K_{r,s}\,||f||_{Y^{r+l_1}}\,.$$

Proof.

Step 1 – Setup of the iteration scheme

Let $f \in Y^{r_0+l_1}$ be such that $||f||_{Y^{r_0+l_1}} \leq \epsilon$, where $0 < \epsilon \leq 1$ will be chosen later.

Denote the smoothing operators on $(X^s)_{s\in I}$ by $\{S_\theta^X\}_{\theta\geq 1}$ and the smoothing operators on $(Y^s)_{s\in I}$ by $\{S_\theta^Y\}_{\theta\geq 1}$.

We use an iteration scheme to construct a sequence $(u_n)_{n\geq 0}$ in X^∞ which we aim to show converges to a solution $u \in U^{m_0}$ of $T(u) = T(u_0) + f$.

For $n \geq 0$, define

$$\theta_n = \theta_0 + n$$

where $\theta_0 > 1$ will be chosen later depending only on r_0, the constants m_i, l_1, and the constants in the smoothing hypothesis and in the inequalities satisfied by D^2T and L.

Note that

$$\theta_{n+1} \leq \theta_n + 1 \leq 2\theta_n.$$

We have dropped the parameter κ from the definition of θ_n in Hörmander's version since he introduced it to make e''_n as small as e'_n, but this will turn out to be automatically true under our hypotheses.

For $n \geq 0$, define

$$\begin{aligned} v_n &= S^X_{\theta_n} u_n \\ \dot{u}_n &= L(v_n) g_n \\ u_{n+1} &= u_n + \dot{u}_n \end{aligned}$$

where g_n is defined below.

Note that the overdot $\dot{}$ is simply notation indicating a sort of difference and does not denote differentiation.

For $n \geq 0$, define

$$\begin{aligned} g_0 &= S^Y_{\theta_0} f \\ g_{n+1} &= (S^Y_{\theta_{n+1}} - S^Y_{\theta_n})(f - E_n) - S^Y_{\theta_{n+1}} e_n \end{aligned}$$

where

$$E_n = \sum_{m=0}^{n-1} e_m$$

(so $E_0 = 0$), and the error e_n is defined below, for $n \geq 0$.

$$\begin{aligned} e'_n &= (DT(u_n) - DT(v_n))\dot{u}_n \\ e''_n &= T(u_n + \dot{u}_n) - T(u_n) - DT(u_n)\dot{u}_n \\ e_n &= e'_n + e''_n. \end{aligned}$$

Note that since g_0 is defined in terms of f only, and we are given u_0, from which v_0 is obtained immediately, the iteration scheme can be determined for $n \geq 0$ in the order $\dot{u}_n$, u_{n+1}, v_{n+1}, e'_n, e''_n, e_n, E_n, g_{n+1}.

Note that e_n is defined so that it measures how well $T(u_{n+1}) - T(u_n)$ is approximated by $DT(v_n)\dot{u}_n$, by which we mean

$$\begin{aligned} T(u_{n+1}) - T(u_n) &= DT(v_n)\dot{u}_n + e_n \\ &= g_n + e_n. \end{aligned}$$

Also note that the formula for g_{n+1} can be rearranged to give

$$g_{n+1} = (S^Y_{\theta_{n+1}} f - S^Y_{\theta_n} f) - (S^Y_{\theta_{n+1}} E_{n+1} - S^Y_{\theta_n} E_n).$$

We thus obtain

$$\begin{aligned} T(u_{n+1}) - T(u_0) &= \sum_{m=0}^{n} (T(u_{m+1}) - T(u_m)) \\ &= \sum_{m=0}^{n} g_m + \sum_{m=0}^{n} e_m \\ &= S^Y_{\theta_n} f - S^Y_{\theta_n} E_n + E_{n+1} \\ &= S^Y_{\theta_n} f + (E_n - S^Y_{\theta_n} E_n) + e_n \end{aligned}$$

which we hope converges to f as $n \to \infty$, since, roughly speaking, $S^Y_{\theta_n} \to \mathrm{id}$ and $e_n \to 0$.

Step 2 – Obtaining estimates for the iterates via induction

We will show the following inductive hypothesis holds.

$$||\dot{u}_n||_{X^s} \leq K\, ||f||_{Y^{r_0+l_1}}\, \theta_n^{s-r_0-1} \qquad \text{for all } s \in [0, s_1] \qquad [H_n]$$

where the constant $K > 0$ will be chosen later, with K independent of n, f and ϵ, but depending on θ_0. We will choose ϵ sufficiently small such that $K\, ||f||_{Y^{r_0+l_1}} \leq K\epsilon \leq 1$.

In what follows, $C_s > 0$ represents a constant, which is independent of n, f and ϵ, and is bounded for s bounded. It will also be independent of θ_0, which will allow us to choose θ_0 so that θ_n is large compared to C_s for s in a certain range. We will write $C > 0$ for a constant which is also independent of s.

Assume now that $[H_m]$ is true for all $0 \leq m \leq n$ and let us show that $[H_{n+1}]$ follows. (We will leave the proof of $[H_0]$ until later.)

Pick a real number $0 < \eta < 1$ such that $r_0 > \max\{m_0, m_4, l_1 + m_1 + m_2, 2m_2, \frac{l_1+m_3}{2} + m_2\} + 2\eta$.

For $s \in I$, define

$$P(s) = \begin{cases} (s - r_0)_+ & \text{for } |s - r_0| \geq \eta, \\ \eta & \text{for } |s - r_0| < \eta. \end{cases}$$

We claim that the following estimates for $0 \leq m \leq n+1$ follow directly from $[H_m]$ for $0 \leq m \leq n$.

$$||u_m - u_0||_{X^s} \leq C_s K \, ||f||_{Y^{r_0+l_1}} \, \theta_m^{P(s)} \qquad \text{for } s \in [0, s_1], \tag{6}$$

$$\left|\left|S_{\theta_m}^X (u_m - u_0)\right|\right|_{X^s} \leq C_s K \, ||f||_{Y^{r_0+l_1}} \, \theta_m^{P(s)} \qquad \text{for } s \in I, \tag{7}$$

$$\left|\left|(u_m - u_0) - S_{\theta_m}^X (u_m - u_0)\right|\right|_{X^s} \leq C_s K \, ||f||_{Y^{r_0+l_1}} \, \theta_m^{(s-r_0)} \qquad \text{for } s \in [0, s_1], \tag{8}$$

$$||u_m - v_m||_{X^s} \leq C_s \theta_m^{s-r_0} \qquad \text{for } s \in [0, s_1], \tag{9}$$

$$||v_m||_{X^s} \leq C_s \theta_m^{P(s)} \qquad \text{for } s \in I, \tag{10}$$

$$||u_m||_{X^s} \leq C_s \theta_m^{P(s)} \qquad \text{for } s \in [0, s_1]. \tag{11}$$

Indeed, for $0 \leq m \leq n$, we have

$$\begin{aligned} ||u_{m+1} - u_0||_{X^s} &= \left|\left| \sum_{l=0}^{m} \dot{u}_l \right|\right|_{X^s} \\ &\leq \sum_{l=0}^{m} ||\dot{u}_l||_{X^s} \\ &\leq K \, ||f||_{Y^{r_0+l_1}} \sum_{l=0}^{m} \theta_l^{s-r_0-1} \\ &= K \, ||f||_{Y^{r_0+l_1}} \sum_{l=0}^{m} (\theta_0 + l)^{s-r_0-1} \\ &\leq K \, ||f||_{Y^{r_0+l_1}} \sum_{l=0}^{m} (\theta_0 + l)^{Q(s)-1} \end{aligned}$$

where

$$Q(s) = \begin{cases} s - r_0 & \text{for } |s - r_0| \geq \eta, \\ \eta & \text{for } |s - r_0| < \eta. \end{cases}$$

Set $h(x) = (\theta_0 + x)^{Q(s)-1}$ for $x \in [0, \infty)$. Then

$$\begin{aligned} \sum_{l=0}^{m} (\theta_0 + l)^{Q(s)-1} &\leq \int_0^{m+1} h(x) \mathrm{d}x \\ &= \begin{cases} \frac{1}{s-r_0}((\theta_0 + m + 1)^{s-r_0} - \theta_0^{s-r_0}) & \text{for } |s - r_0| \geq \eta \\ \frac{1}{\eta}((\theta_0 + m + 1)^{\eta} - \theta_0^{\eta}) & \text{for } |s - r_0| < \eta \end{cases} \end{aligned}$$

$$= \begin{cases} \frac{1}{s-r_0}(\theta_{m+1}^{s-r_0} - \theta_0^{s-r_0}) & \text{for } |s-r_0| \geq \eta \\ \frac{1}{\eta}(\theta_{m+1}^{\eta} - \theta_0^{\eta}) & \text{for } |s-r_0| < \eta \end{cases}$$

$$\leq \begin{cases} \frac{1}{s-r_0}\theta_{m+1}^{s-r_0} & \text{for } s-r_0 \geq \eta \\ \frac{1}{r_0-s}\theta_0^{-(r_0-s)} & \text{for } s-r_0 \leq -\eta \\ \frac{1}{\eta}\theta_{m+1}^{\eta} & \text{for } |s-r_0| < \eta \end{cases}$$

This implies (6), noting that the constant C_s remains bounded for s bounded. (We introduced η to avoid a constant involving $\frac{1}{s-r_0}$ which blows up as $s \to r_0$.)

For $s \geq r_0 + \eta$, use (1) from the smoothing hypothesis and (6) to obtain

$$\begin{aligned} \left|\left|S_{\theta_m}^X(u_m - u_0)\right|\right|_{X^s} &\leq C_s\theta_m^{s-r_0-\eta} \left|\left|u_m - u_0\right|\right|_{X^{r_0+\eta}} \\ &\leq C_s K \left|\left|f\right|\right|_{Y^{r_0+l_1}} \theta_m^{s-r_0-\eta}\theta_m^{\eta} \\ &\leq C_s K \left|\left|f\right|\right|_{Y^{r_0+l_1}} \theta_m^{s-r_0}. \end{aligned}$$

For $s < r_0 + \eta$, using (1) from the smoothing hypothesis and (6), we have

$$\left|\left|S_{\theta_m}^X(u_m - u_0)\right|\right|_{X^s} \leq C_s K \left|\left|f\right|\right|_{Y^{r_0+l_1}} \theta_m^{P(s)}.$$

This proves (7).

For $s \leq r_0 + \eta$, use (2) from the smoothing hypothesis and (6) to obtain

$$\begin{aligned} \left|\left|(u_m - u_0) - S_{\theta_m}^X(u_m - u_0)\right|\right|_{X^s} &\leq C_s\theta_m^{s-r_0-\eta} \left|\left|u_m - u_0\right|\right|_{X^{r_0+\eta}} \\ &\leq C_s K \left|\left|f\right|\right|_{Y^{r_0+l_1}} \theta_m^{s-r_0-\eta}\theta_m^{\eta} \\ &\leq C_s K \left|\left|f\right|\right|_{Y^{r_0+l_1}} \theta_m^{s-r_0}. \end{aligned}$$

For $r_0 + \eta < s \leq s_1$, using (6) and (7), we have

$$\left|\left|(u_m - u_0) - S_{\theta_m}^X(u_m - u_0)\right|\right|_{X^s} \leq C_s K \left|\left|f\right|\right|_{Y^{r_0+l_1}} \theta_m^{s-r_0}$$

as required. This proves (8).

Now

$$\begin{aligned} \left|\left|u_m - v_m\right|\right|_{X^s} &= \left|\left|u_m - S_{\theta_m}^X u_m\right|\right|_{X^s} \\ &= \left|\left|(u_m - u_0) - S_{\theta_m}^X(u_m - u_0) + u_0 - S_{\theta_m}^X u_0\right|\right|_{X^s} \\ &\leq \left|\left|(u_m - u_0) - S_{\theta_m}^X(u_m - u_0)\right|\right|_{X^s} + \left|\left|u_0 - S_{\theta_m}^X u_0\right|\right|_{X^s} \\ &\leq C_s K \left|\left|f\right|\right|_{Y^{r_0+l_1}} \theta_m^{s-r_0} + C_s\theta_m^{s-r_0} \left|\left|u_0\right|\right|_{X^{\max\{r_0,s\}}} \end{aligned}$$

by applying (8) to the first term and (1) or (2) from the smoothing hypothesis to the second term. This proves (9). (Note $K\left|\left|f\right|\right|_{Y^{r_0+l_1}} \leq K\epsilon \leq 1$.)

Similarly,

$$\begin{aligned}||v_m||_{X^s} &= \left|\left|S^X_{\theta_m} u_m\right|\right|_{X^s}\\ &= \left|\left|S^X_{\theta_m}(u_m - u_0) + S^X_{\theta_m} u_0\right|\right|_{X^s}\\ &\le \left|\left|S^X_{\theta_m}(u_m - u_0)\right|\right|_{X^s} + \left|\left|S^X_{\theta_m} u_0\right|\right|_{X^s}\\ &\le \left|\left|S^X_{\theta_m}(u_m - u_0)\right|\right|_{X^s} + C_s\,||u_0||_{X^s}\end{aligned}$$

by (1) from the smoothing hypothesis. Now use (7) to obtain (10).

We have

$$||u_m||_{X^s} \le ||u_m - u_0||_{X^s} + ||u_0||_{X^s}\,.$$

Now apply (6) to obtain (11).

This completes the proof of the claim.

Note that, using (6) and (9), we have

$$\begin{aligned}||v_m - u_0||_{X^{m_0}} &\le ||v_m - u_m||_{X^{m_0}} + ||u_m - u_0||_{X^{m_0}}\\ &\le C\theta_m^{m_0 - r_0} + CK\epsilon\theta_m^{P(m_0)}\\ &\le C\theta_m^{m_0 - r_0} + CK\epsilon.\end{aligned}$$

Thus by taking ϵ sufficiently small depending on K and C, and θ_0 sufficiently large depending on C, we have $v_n,\ v_{n+1} \in U$. Also note that (6) in the case $s = m_0$ implies $u_n \in U$ for ϵ sufficiently small, and $[H_n]$ implies that $u_n + \dot{u}_n \in U$ for ϵ sufficiently small. This guarantees that e_n and $\dot{u}_{n+1}$ are well-defined. Note that the same argument also shows that the line segments $[u_n, u_n + \dot{u}_n]$ and $[u_n, v_n]$ are in U for ϵ sufficiently small.

Estimate of e'_n. We claim that for all $s \in [0, s_1 - \max\{m_1, m_3\}]$,

$$\left|\left|e'_n\right|\right|_{Y^s} \le C_s K\,||f||_{Y^{r_0 + l_1}}\,\theta_n^{M(s)-1+\eta}$$

where

$$M(s) = \max\{s + m_1 + m_2 - 2r_0, (s + m_3 - r_0)_+ + 2m_2 - 2r_0\}.$$

Indeed, we have

$$\begin{aligned}e'_n &= (DT(u_n) - DT(v_n))\dot{u}_n\\ &= (DT((u_n - v_n) + v_n) - DT(v_n))\dot{u}_n.\end{aligned}$$

Note that, since T is twice differentiable in the sense of Definition 3, the map

$$u \mapsto DT(u)\dot{u}_n$$

is differentiable in the sense of Definition 3 with derivative acting on v given by

$$D^2T(u)(\dot{u}_n, v).$$

Hence, applying Taylor's theorem, (4), $[H_n]$ and the estimates (9) and (10), we have, for $s \in [0, s_1 - \max\{m_1, m_3\}]$,

$$\begin{aligned}
||e'_n||_{Y^s} &= ||(DT((u_n - v_n) + v_n) - DT(v_n))\dot{u}_n||_{Y^s} \\
&\leq \sup_{t\in[0,1]} ||D^2T(t(u_n - v_n) + v_n)(\dot{u}_n, u_n - v_n)||_{Y^s} \\
&\leq C_s(||\dot{u}_n||_{X^{s+m_1}} ||u_n - v_n||_{X^{m_2}} + ||\dot{u}_n||_{X^{m_2}} ||u_n - v_n||_{X^{s+m_1}} \\
&+ ||\dot{u}_n||_{X^{m_2}} ||u_n - v_n||_{X^{m_2}} (1 + \sup_{t\in[0,1]} ||v_n + t(u_n - v_n)||_{X^{s+m_3}})) \\
&\leq C_s(K ||f||_{Y^{r_0+l_1}} \theta_n^{s+m_1-r_0-1}\theta_n^{m_2-r_0} + K ||f||_{Y^{r_0+l_1}} \theta_n^{m_2-r_0-1}\theta_n^{s+m_1-r_0} \\
&+ K ||f||_{Y^{r_0+l_1}} \theta_n^{m_2-r_0-1}\theta_n^{m_2-r_0}(1 + \theta_n^{P(s+m_3)} + \theta_n^{s+m_3-r_0})) \\
&\leq C_s K ||f||_{Y^{r_0+l_1}} \theta_n^{M(s)-1+\eta}.
\end{aligned}$$

Estimate of e''_n. We claim that for all $s \in [0, s_1 - \max\{m_1, m_3\}]$,

$$||e''_n||_{Y^s} \leq C_s K ||f||_{Y^{r_0+l_1}} \theta_n^{M(s)-1+\eta}.$$

Indeed, we have

$$e''_n = T(u_n + \dot{u}_n) - T(u_n) - DT(u_n)\dot{u}_n.$$

Hence, applying Taylor's theorem, (4), $[H_n]$ and the estimate (11), we have, for $s \in [0, s_1 - \max\{m_1, m_3\}]$,

$$\begin{aligned}
||e''_n||_{Y^s} &\leq \sup_{t\in[0,1]} ||D^2T(u_n + t\dot{u}_n)(\dot{u}_n, \dot{u}_n)||_{Y^s} \\
&\leq C_s(||\dot{u}_n||_{X^{s+m_1}} ||\dot{u}_n||_{X^{m_2}} + ||\dot{u}_n||^2_{X^{m_2}} (1 + \sup_{t\in[0,1]} ||u_n + t\dot{u}_n||_{X^{s+m_3}})) \\
&\leq C_s(K ||f||_{Y^{r_0+l_1}} \theta_n^{s+m_1-r_0-1} K ||f||_{Y^{r_0+l_1}} \theta_n^{m_2-r_0-1} \\
&+ K^2 ||f||^2_{Y^{r_0+l_1}} \theta_n^{2m_2-2r_0-2}(1 + \theta_n^{P(s+m_3)} + K ||f||_{Y^{r_0+l_1}} \theta_n^{s+m_3-r_0-1})) \\
&\leq \theta_n^{-1} C_s K ||f||_{Y^{r_0+l_1}} \theta_n^{M(s)-1+\eta} \\
&\leq C_s K ||f||_{Y^{r_0+l_1}} \theta_n^{M(s)-1+\eta}
\end{aligned}$$

where we have used $K ||f||_{Y^{r_0+l_1}} \leq K\epsilon \leq 1$.

Estimate of e_n. Adding the estimates for e_n' and e_n'', we obtain

$$||e_n||_{Y^s} \leq C_s K \, ||f||_{Y^{r_0+l_1}} \, \theta_n^{M(s)-1+\eta}$$

for all $s \in [0, s_1 - \max\{m_1, m_3\}]$.

Estimate of g_{n+1}. We claim that for all $s \in I$,

$$||g_{n+1}||_{Y^s} \leq C_s(K \, ||f||_{Y^{r_0+l_1}} \, \theta_n^{M(s)-1+\eta} + ||f||_{Y^{r_0+l_1}} \, \theta_n^{s-r_0-l_1-1}).$$

Indeed, we have

$$g_{n+1} = (S^Y_{\theta_{n+1}} - S^Y_{\theta_n})(f - E_n) - S^Y_{\theta_{n+1}} e_n.$$

Note that for any $w \in Y^{s'}$,

$$\left|\left|(S^Y_{\theta_{n+1}} - S^Y_{\theta_n})w\right|\right|_{Y^s} \leq C_{s',s} \theta_n^{s-s'-1} \, ||w||_{Y^{s'}}$$

by the smoothing hypothesis (3) and Taylor's theorem.

Setting $s' = r_0 + l_1$, we have

$$\left|\left|(S^Y_{\theta_{n+1}} - S^Y_{\theta_n})f\right|\right|_{Y^s} \leq C_s \theta_n^{s-r_0-l_1-1} \, ||f||_{Y^{r_0+l_1}} \, .$$

We also have

$$\left|\left|(S^Y_{\theta_{n+1}} - S^Y_{\theta_n})E_n\right|\right|_{Y^s} \leq C_{s',s} \theta_n^{s-s'-1} \, ||E_n||_{Y^{s'}} \, .$$

Now, for $s' \in [0, s_1 - \max\{m_1, m_3\}]$, we have, from the estimate for e_n,

$$\begin{aligned} ||E_n||_{Y^{s'}} &= \left|\left|\sum_{m=0}^{n-1} e_m\right|\right|_{Y^{s'}} \\ &\leq C_{s'} K \, ||f||_{Y^{r_0+l_1}} \sum_{m=0}^{n-1} \theta_m^{M(s')-1+\eta} \\ &\leq C_{s'} K \, ||f||_{Y^{r_0+l_1}} \, \theta_n^{M(s')+\eta} \end{aligned}$$

if $M(s') \geq 0$, by the integral comparison used before. Note that $M(s')$ has slope 1 for large enough s' depending on r_0 and the constants m_i, so to achieve $M(s') \geq 0$ it suffices to take s' large in relation to r_0 and the constants m_i. To do this we require s_1 sufficiently large in relation to r_0 and the constants m_i.

Hence

$$\left|\left|(S^Y_{\theta_{n+1}} - S^Y_{\theta_n})E_n\right|\right|_{Y^s} \leq C_{s'}C_{s',s}K\,||f||_{Y^{r_0+l_1}}\,\theta_n^{M(s')+s-s'-1+\eta}$$
$$\leq C_s K\,||f||_{Y^{r_0+l_1}}\,\theta_n^{M(s)-1+\eta}$$

by choosing s' sufficiently large compared to r_0 and the constants m_i so that $M(s)$ has slope 1 for $s \geq s'$. (Hence $M(s') - s' \leq M(s) - s$ for all s since $M(s) - s$ is decreasing for $s \leq s'$ and constant for $s \geq s'$.) Again, to do this we require s_1 sufficiently large in relation to r_0 and the constants m_i. This fixes s_1.

Similarly, for s' sufficiently large, we have

$$\left|\left|S^Y_{\theta_{n+1}} e_n\right|\right|_{Y^s} \leq C_{s',s}\theta_n^{s-s'}\,||e_n||_{Y^{s'}}$$
$$\leq C_{s',s}C_{s'}K\,||f||_{Y^{r_0+l_1}}\,\theta_n^{M(s')+s-s'-1+\eta}$$
$$\leq C_s K\,||f||_{Y^{r_0+l_1}}\,\theta_n^{M(s)-1+\eta}.$$

Hence the estimate for g_{n+1} holds.

Estimate of $\dot{u}_{n+1}$. We have

$$\dot{u}_{n+1} = L(v_{n+1})g_{n+1}.$$

Hence, for all $s \in I$ such that $s + l_1, s + m_4 \in I$, using (5), the estimate (10) and the estimate for g_{n+1}, we have

$$\begin{aligned}
||\dot{u}_{n+1}||_{X^s} &\leq C_s(||g_{n+1}||_{Y^{s+l_1}} + ||g_{n+1}||_{Y^{l_1}}\,(1 + ||v_{n+1}||_{X^{s+m_4}}))\\
&\leq C_s(K\,||f||_{Y^{r_0+l_1}}\,\theta_{n+1}^{M(s+l_1)-1+\eta} + ||f||_{Y^{r_0+l_1}}\,\theta_{n+1}^{s-r_0-1}\\
&+ (K\,||f||_{Y^{r_0+l_1}}\,\theta_{n+1}^{M(l_1)-1+\eta} + ||f||_{Y^{r_0+l_1}}\,\theta_{n+1}^{-r_0-1})(1 + \theta_{n+1}^{P(s+m_4)}))\\
&\leq C_s(K\,||f||_{Y^{r_0+l_1}}\,\theta_{n+1}^{M(l_1)+s-1+\eta} + ||f||_{Y^{r_0+l_1}}\,\theta_{n+1}^{s-r_0-1}) \qquad (12)
\end{aligned}$$

since $\theta_{n+1}^{P(s+m_4)} \leq \theta_{n+1}^s$ because $r_0 > m_4 + 2\eta$, and $M(l_1 + s) \leq M(l_1) + s$ because M has slope at most 1.

We want to obtain

$$||\dot{u}_{n+1}||_{X^s} \leq K\,||f||_{Y^{r_0+l_1}}\,\theta_{n+1}^{s-r_0-1}$$

for $s \in [0, s_1]$.

To make the first term sufficiently small, we require

$$-\gamma := M(l_1) + r_0 + \eta < 0.$$

Then we can choose θ_0 large enough so that

$$C_s\theta_{n+1}^{M(l_1)+s-1+\eta} = C_s\theta_{n+1}^{s-r_0-1}\theta_{n+1}^{-\gamma} \leq C_s\theta_{n+1}^{s-r_0-1}\theta_0^{-\gamma} \leq \frac{1}{2}\theta_{n+1}^{s-r_0-1}$$

for all $s \in [0, s_1]$.

We note that $M(l_1)+r_0+\eta < 0$ if and only if $r_0-\eta > l_1+m_1+m_2$, $r_0-\eta > 2m_2$ and $r_0 - \eta > m_2 + \frac{l_1+m_3}{2}$, which indeed hold by the choice of η.

To make the second term sufficiently small, we take $K \geq 2C_s$ for all $s \in [0, s_1]$.

This gives $[H_{n+1}]$.

Proof of $[H_0]$ We have

$$g_0 = S_{\theta_0}^Y f$$

and

$$v_0 = S_{\theta_0}^X u_0.$$

Hence

$$\begin{aligned}
||\dot{u}_0||_{X^s} &= \left|\left|L(S_{\theta_0}^X u_0)S_{\theta_0}^Y f\right|\right|_{X^s} \\
&\leq C_s(\left|\left|S_{\theta_0}^Y f\right|\right|_{Y^{s+l_1}} + \left|\left|S_{\theta_0}^Y f\right|\right|_{Y^{l_1}} (1 + \left|\left|S_{\theta_0}^X u_0\right|\right|_{X^{s+m_4}})) \\
&\leq C_s \left|\left|S_{\theta_0}^Y f\right|\right|_{Y^{s+l_1}} \\
&\leq C_s \,||f||_{Y^{r_0+l_1}}\, \theta_0^{(s-r_0)+} \qquad \text{by (1) and (2) from the smoothing hypothesis} \\
&\leq K\, ||f||_{Y^{r_0+l_1}}\, \theta_0^{s-r_0-1}
\end{aligned}$$

for all $s \in [0, s_1]$, assuming that K is sufficiently large compared to θ_0 and C_s for $s \in [0, s_1]$.

This is $[H_0]$.

Step 3 – Better estimates if $f \in Y^{r+l_1}$ for $r > r_0$

Let $r \in J$, so that $f \in Y^{r+l_1}$, where $r \geq r_0$.

We will show that, for all $n \geq 0$ and for all $s \in I$ such that $s + \max\{m_1, m_3\} + \max\{l_1, m_4\} \in I$, we have

$$||\dot{u}_n||_{X^s} \leq C_{r,s}\, ||f||_{Y^{r+l_1}}\, \theta_n^{s-r-1} \tag{13}$$

where the constant $C_{r,s} > 0$ is independent of n and f.

Firstly, note that we have proved $[H_n]$ for $n \geq 0$, and hence all the estimates from step 2 which were conditional on the inductive hypothesis are now valid, and we may use them as we wish.

We are going to prove the above statement by an induction argument, but not an induction on n. We are going to use the estimates from step 2 for each n separately to obtain the above inequality, and the constant will be independent of n because the constants from step 2 are independent of n.

We claim by induction on $k \geq 0$ that for all $s \in I$ such that $s + \max\{l_1, m_4\} \in I$, we have

$$||\dot{u}_n||_{X^s} \leq C_{k,r,s} \, ||f||_{Y^{r+l_1}} \, \theta_n^{s-r_0-\gamma_k-1} \qquad [G_k]$$

where the constant $C_{k,r,s} > 0$ is independent of n and f, and

$$\gamma_k = \min\{k\gamma, r - r_0\}.$$

Indeed, the estimate (12) for $\dot{u}_{n+1}$ in step 2 implies that

$$||\dot{u}_n||_{X^s} \leq C_s \, ||f||_{Y^{r_0+l_1}} \, \theta_n^{s-r_0-1} \tag{14}$$

for all $s \in I$ such that $s + \max\{l_1, m_4\} \in I$ (not just $s \in [0, s_1]$ which would follow directly from $[H_n]$).

Using this, we can obtain the following new versions of the estimates (9)–(11) for all $s \in I$ such that $s + \max\{l_1, m_4\} \in I$ (not just $s \in [0, s_1]$) via exactly the same calculations

$$||u_m - v_m||_{X^s} \leq C_s \theta_m^{s-r_0}, \tag{15}$$

$$||v_m||_{X^s} \leq C_s \theta_m^{P(s)}, \tag{16}$$

$$||u_m||_{X^s} \leq C_s \theta_m^{P(s)}. \tag{17}$$

Using the fact that $||f||_{Y^{r_0+l_1}} \leq ||f||_{Y^{r+l_1}}$, (14) immediately implies $[G_0]$.

Now we assume $[G_k]$ holds and aim to show $[G_{k+1}]$ holds.

Now we want to obtain new estimates for e_n' and e_n''. Note that in the estimates for both of these there was at least one factor involving $\dot{u}_n$ in each term. If we estimate this one factor using the new estimate given by $[G_k]$ and the other quantities using (14) and the slightly modified estimates (15)–(17), we obtain

$$||e_n||_{Y^s} \leq C_{k,r,s} \, ||f||_{Y^{r+l_1}} \, \theta_n^{M(s)-1+\eta-\gamma_k}$$

for all $s \in I$ such that $s + \max\{m_1, m_3\} + \max\{l_1, m_4\} \in I$. The constant $C_{k,r,s}$ is independent of f since we have only used the new estimate given by $[G_k]$ in one factor, and the other estimates we have used involve $||f||_{Y^{r_0+l_1}}$, which is bounded by $\epsilon \leq 1$.

This implies that for $s' \in I$ such that $s' + \max\{m_1, m_3\} + \max\{l_1, m_4\} \in I$, we have

$$\begin{aligned} ||E_n||_{Y^{s'}} &= \left\| \sum_{m=0}^{n-1} e_m \right\|_{Y^{s'}} \\ &\leq C_{k,r,s'} \, ||f||_{Y^{r+l_1}} \sum_{m=0}^{n-1} \theta_m^{M(s')-1+\eta-\gamma_k} \\ &\leq C_{k,r,s'} \, ||f||_{Y^{r+l_1}} \, \theta_n^{M(s')+\eta-\gamma_k} \end{aligned} \tag{18}$$

as long as $M(s') \geq \gamma_k$. It is possible to pick such an s' if $s_1+r-r_0+\max\{l_1, m_4\} \in I$ given the fact that $M(s_1 - \max\{m_1, m_3\}) \geq 0$ and $M(s)$ has slope 1 for $s \geq s_1 - \max\{m_1, m_3\}$.

Hence

$$\left|\left|(S^Y_{\theta_{n+1}} - S^Y_{\theta_n})E_n\right|\right|_{Y^s} \leq C_{s',k}C_{k,r,s}\theta_n^{M(s')+s-s'-1+\eta-\gamma_k}$$
$$\leq C_{k,r,s}\,||f||_{Y^{r+l_1}}\,\theta_n^{M(s)-1+\eta-\gamma_k}$$

as long as $M(s') \geq \gamma_k$ and s' is sufficiently large compared to r_0 and the constants m_i so that $M(s)$ has slope 1 for $s \geq s'$.

We also have the estimate

$$\left|\left|S^Y_{\theta_{n+1}}e_n\right|\right|_{Y^s} \leq C_{k,r,s}\,||f||_{Y^{r+l_1}}\,\theta_n^{M(s)-1+\eta-\gamma_k}.$$

In addition we can use the new estimate

$$\left|\left|(S^Y_{\theta_{n+1}} - S^Y_{\theta_n})f\right|\right|_{Y^s} \leq C_{r,s}\theta_n^{s-r-l_1-1}\,||f||_{Y^{r+l_1}}\,.$$

This gives us the following new estimate for g_{n+1}, for all $s \in I$,

$$||g_{n+1}||_{Y^s} \leq C_{k,r,s}\,||f||_{Y^{r+l_1}}\,(\theta_n^{M(s)-1+\eta-\gamma_k} + \theta_n^{s-r-l_1-1}).$$

From this we obtain, for all $s \in I$ such that $s + \max\{l_1, m_4\} \in I$,

$$||\dot{u}_n||_{X^s} \leq C_{r,s}\,||f||_{Y^{r+l_1}}\,(\theta_n^{M(l_1)+s-1+\eta-\gamma_k} + \theta_n^{s-r-1})$$
$$\leq C_{r,s}\,||f||_{Y^{r+l_1}}\,(\theta_n^{s-r_0-1-\gamma_k-\gamma} + \theta_n^{s-r-1})$$
$$\leq C_{r,s}\,||f||_{Y^{r+l_1}}\,\theta_n^{s-r_0-\gamma_{k+1}-1}$$

where we have used the fact that $M(l_1) + r_0 + \eta = -\gamma$.

This is $[G_{k+1}]$.

For large enough k, we have $k\gamma \geq r - r_0$, so $\gamma_k = r - r_0$ and this gives (13).

Step 4 – Convergence to a solution

Let $r \in J$, so that $f \in Y^{r+l_1}$, where $r \geq r_0$.

Using (13), we have

$$\sum_{m=0}^{n} ||u_{m+1} - u_m||_{X^s} = \sum_{m=0}^{n} ||\dot{u}_m||_{X^s}$$
$$\leq C_{r,s}\,||f||_{Y^{r+l_1}}\,\theta_{n+1}^{(s-r)_+}$$

for $r \neq s$.

Thus

$$\sum_{m=0}^{n} ||u_{m+1} - u_m||_{X^s}$$

converges as $n \to \infty$ for $s < r$. Hence, by completeness, $u_n \to u$ in X^s as $n \to \infty$, for all $s < r$, for some $u \in \cap_{0 \le s < r} X^s$.

Note the above calculation also implies that

$$||u_n - u_0||_{X^s} \le C_{r,s} \, ||f||_{Y^{r+l_1}}$$

for $s < r$, so we have

$$||u - u_0||_{X^s} \le C_{r,s} \, ||f||_{Y^{r+l_1}} \, .$$

Next we claim that

$$T(u_{n+1}) - T(u_0) \to f$$

in X^s as $n \to \infty$, for all $s < r$.

Indeed,

$$T(u_{n+1}) - T(u_0) = S^Y_{\theta_n} f \, + (E_n - S^Y_{\theta_n} E_n) + e_n$$

so

$$T(u_{n+1}) - T(u_0) - f \, = (S^Y_{\theta_n} f - f) + (E_n - S^Y_{\theta_n} E_n) + e_n.$$

By (2) from the smoothing hypothesis, we have

$$\left|\left|S^Y_{\theta_n} f - f\right|\right|_{Y^{s+l_1}} \le C_{r,s} \theta_n^{s-r} \, ||f||_{Y^{r+l_1}} \to 0 \text{ as } n \to \infty.$$

Also,

$$\begin{aligned}
\left|\left|E_n - S^Y_{\theta_n} E_n\right|\right|_{Y^{s+l_1}} &\le C_{s,s'} \theta_n^{s-s'} \, ||E_n||_{Y^{s'+l_1}} \qquad \text{for } s' \ge s \\
&\le C_{s,s'} \theta_n^{s-s'} C_{r,s} \theta_n^{M(s'+l_1)+\eta-(r-r_0)} \, ||f||_{Y^{r+l_1}} \\
&\quad \text{using (18), for } s' \text{ large enough such that } M(s'+l_1) \ge r - r_0 \\
&\le C_{r,s} \theta_n^{M(s'+l_1)+s-s'+\eta-(r-r_0)} \, ||f||_{Y^{r+l_1}} \\
&\le C_{r,s} \theta_n^{s-r} \, ||f||_{Y^{r+l_1}} \to 0 \text{ as } n \to \infty
\end{aligned}$$

since $M(s' + l_1) + \eta + r_0 \le M(l_1) + \eta + r_0 + s' < s'$.

Finally,

$$||e_n||_{Y^{s+l_1}} \leq C_{r,s}\theta_n^{M(s+l_1)+\eta-(r-r_0)-1}\,||f||_{Y^{r+l_1}} \to 0 \text{ as } n \to \infty$$

since $M(s+l_1)+\eta+r_0 \leq M(l_1)+\eta+r_0+s < s$.

This proves the claim.

Now since $T : U \to Y^0$ is continuous as a map from X^{m_0} to Y^0, and $r_0 > m_0$, so $u_n \to u$ in X^{m_0}, we have that $T(u_n) \to T(u)$ in Y^0, hence $T(u) = T(u_0) + f$ as required.

This completes the proof.

Remark 4. We make a remark here on the rate of convergence of the above scheme as compared to the Newton-Raphson scheme. Since we have in mind applying the result in existence proofs in PDE problems, we have made no effort to optimise the rate of convergence in the above scheme in any way. One of the key features of the Newton-Raphson scheme is that the rate of convergence is quadratic, i.e. the error at step $n+1$ is proportional to the square of the error at step n. However, we can see in the above scheme that the error $||T(u_{n+1}) - T(u_0) - f||_{X^s}$ is proportional to θ_n^{s-r} where $r > s$ is such that $f \in Y^{r+l_1}$, and θ_n increases like n. Thus according to the crude bounds we have in the above proof, the ratio of the errors at steps n and $n+1$ may tend to 1 as $n \to \infty$, although it may be possible to better by being more careful.

4 Modified Version of Nash-Moser Iteration

4.1 Changes from Hörmander's Iteration Scheme

Here, we introduce two modifications to Hörmander's scheme which will allow it to be applied as in Coulombel and Secchi [8]. The basic principle is that the error $T(u_n)-T(u_0)-f$ in the above scheme tends to zero, so we may introduce additional approximations into the scheme that can be controlled in terms of this error. One disadvantage is that we lose regularity with respect to f since we need this error to be controlled to high order.

Firstly, we note that it may be inconvenient to solve the linearised system

$$DT(u)v = g.$$

It may in fact be more convenient to solve the system

$$A(u)v = g$$

where the operator $A(u)$ is approximately equal to $DT(u)$, such that $A(u) - DT(u)$ can be controlled in terms of the error $T(u) - T(u_0) - f$. This modification was made by Alinhac in [2] when he introduced the 'good unknown'.

Secondly, it may only be possible to solve the linearised system

$$A(u)v = g$$

under certain constraints on u which are not preserved by the iteration scheme, which was the problem encountered by Coulombel and Secchi in [8]. Abstractly, we suppose that the linear system can only be solved for $u \in V$, whereas the iteration scheme only preserves $u \in U$. In fact under the iteration scheme we are trying to solve the problem

$$A(v_n)\dot{u}_n = g_n$$

where

$$v_n = S_n u_n.$$

Therefore we denote by R an operator that maps U to V and set $w_n = R(v_n)$ and solve the system

$$A(w_n)\dot{u}_n = g_n.$$

This will require that $R(u) - u$ is controlled in terms of the error $T(u) - T(u_0) - f$ and also that R and the smoothing operators satisfy some commutation estimates.

4.2 Statement and Proof of the Theorem

Theorem 2. *Let $\{X^s\}_{s\in I}$ and $\{Y^s\}_{s\in I}$ be two decreasing families of Banach spaces, each satisfying the smoothing hypothesis. Let $u_0 \in X^\infty$ and let $T : U^{m_0} \to Y^0$ be continuous, where $U^{m_0} \subset X^{m_0}$ is a bounded open neighbourhood of u_0 in X^{m_0}, for some $m_0 \in I$. Suppose also $T : U \to Y^{\infty - m_1}$ for some fixed $m_1 \in I$, where $U := U^{m_0} \cap X^\infty$. Let $f \in Y^{s_1 - \max\{m_1, m_3\}}$ with $||f||_{Y^{s_1-\max\{m_1,m_3\}}} \leq C^0$, where $s_1, m_3 \in I$ are defined below and C^0 is a constant. Assume the following conditions are satisfied, where the constants are independent of f (at least for $||f||_{Y^{s_1-\max\{m_1,m_3\}}} \leq C^0$).*

1. T is twice differentiable in the sense of Definition 3 and

$$\begin{aligned} &\left|\left|D^2T(u)(v_1, v_2)\right|\right|_{Y^s} \\ &\leq C_s^1(||v_1||_{X^{s+m_1}}\,||v_2||_{X^{m_2}} + ||v_1||_{X^{m_2}}\,||v_2||_{X^{s+m_1}} + ||v_1||_{X^{m_2}}\,||v_2||_{X^{m_2}}\,(1 + ||u||_{X^{s+m_3}})) \end{aligned} \quad (19)$$

for all $u \in U$, $v_1, v_2 \in X^\infty$ and $s \in I$ such that $s + m_1, s + m_3 \in I$, for some fixed numbers $m_1, m_2, m_3 \in I$, where we assume $\max\{m_1, m_3\} > 0$*, and the constant $C_s^1 > 0$ is bounded for s bounded. Also,*

$$||DT(u)v||_{Y^s} \leq C_s^2(||v||_{X^{s+m_1}} + ||v||_{X^{m_2}}\,(1 + ||u||_{X^{s+m_3}})) \quad (20)$$

for all $u \in U$, $v \in X^\infty$ *and* $s \in I$ *such that* $s + m_1, s + m_3 \in I$, *where the constant* $C_s^2 > 0$ *is bounded for* s *bounded.*

2. *For each* $u \in U$, *there exists an operator* $A(u) : X^\infty \to Y^{\infty - m_1}$ *such that*

$$
\begin{aligned}
&||(A(u) - DT(u))v||_{Y^s} \\
&\leq C_s^3(||v||_{X^{s+m_5}} \, ||T(u) - T(u_0) - f||_{Y^{l_3}} + ||v||_{X^{m_6}} \, ||T(u) - T(u_0) - f||_{Y^{s+l_4}} \\
&+ ||v||_{X^{m_6}} \, ||T(u) - T(u_0) - f||_{Y^{l_3}} \, (1 + ||u||_{X^{s+m_9}})
\end{aligned} \tag{21}
$$

for all $v \in X^\infty$ *and* $s \in I$ *such that* $s + m_5, s + m_9 \in I$, $s + l_4 + \max\{m_1, m_3\} \leq s_1$, *for some fixed numbers* $m_5, m_6, m_9, l_3, l_4 \in I$, *where the constant* $C_s^3 > 0$ *is bounded for* s *bounded.*

Also, for each $v \in X^\infty$ *that map defined on* U *by* $A_v : u \mapsto A(u)v$ *is differentiable with*

$$
\begin{aligned}
&||DA_v(u)h||_{Y^s} \\
&\leq C_s^4(||h||_{X^{s+m_1}} \, ||v||_{X^{m_2}} + ||h||_{X^{m_2}} \, ||v||_{X^{s+m_1}} + ||h||_{X^{m_2}} \, ||v||_{X^{m_2}} \, (1 + ||u||_{X^{s+m_3}}))
\end{aligned} \tag{22}
$$

for all $h \in X^\infty$ *and* $s \in I$ *such that* $s + m_1, s + m_3 \in I$, *where the constant* $C_s^4 > 0$ *is bounded for* s *bounded.*

3. *For each* $u \in V$, *where* $u_0 \in V \subset X^{\infty - m_7}$, *there exists a linear map* $B(u) : Y^\infty \to X^{\infty - \max\{l_1, m_4 + m_7\}}$ *such that*

$$A(u)B(u) = \text{id}$$

and

$$||B(u)g||_{X^s} \leq C_s^5(||g||_{Y^{s+l_1}} + ||g||_{Y^{l_1}} \, ||u||_{X^{s+m_4}}) \tag{23}$$

for all $u \in V$, $g \in Y^\infty$ *and* $s \in I$ *such that* $s + l_1, s + m_4 + m_7 \in I$, *for some fixed numbers* $l_1, m_4, m_7 \in I$, *where the constant* $C_s^5 > 0$ *is bounded for* s *bounded.*

4. *There exists an operator* $R : U \to V$ *such that*

$$||R(u) - u||_{X^0} \leq C \, ||T(u) - T(u_0) - f||_{Y^{l_2}} \tag{24}$$

for some fixed number $l_2 \in I$, *where we assume* $l_2 \leq l_1$ *(else increase* l_1*), and some constant* $C > 0$. *In addition*

$$||R(u)||_{X^s} \leq C_s^6(1 + ||u||_{X^{m_8}})(1 + ||u||_{X^{s+m_7}}) \tag{25}$$

for all $u \in U$ *and* $s \in I$ *such that* $s + m_7 \in I$, *for some fixed number* $m_8 \in I$, *where* $\{S_\theta^X\}_{\theta \geq 1}$ *are the smoothing operators on* $(X^s)_{s \in I}$, *and the constant* $C_s^6 > 0$ *is bounded for* s *bounded.*

We also assume the commutator estimate

$$\begin{aligned}&\left|\left|R(S_\theta^X u) - S_\theta^X R(u)\right|\right|_{X^s}\\&\leq C_{r',r,s}(\theta^{s-r}(1 + ||u||_{X^{m_8}})(1 + ||u||_{X^{r+m_7}}) + \theta^{-r'}(1 + ||u||_{X^{s+m_7}})(1 + ||u||_{X^{r'+m_8}}))\end{aligned} \tag{26}$$

for all $u \in U$ and $r', r, s \in I$ such that $r + m_7, s + m_7, r' + m_8 \in I$, where $\{S_\theta^X\}_{\theta \geq 1}$ are the smoothing operators on $(X^s)_{s \in I}$, and the constant $C_{r',r,s} > 0$ is bounded for r', r, s bounded.

Let $r_0 \in I$ with $r_0 > \max\{m_0 + \max\{m_7, m_8\}, m_4, m_9, l_1 + m_1 + m_2 + \max\{m_7, m_8\}, 2m_2 + 2\max\{m_7, m_8\}, \frac{l_1+m_3}{2} + m_2 + \max\{m_7, m_8\}, l_1 + \max\{m_5, m_6\} + (l_3 - l_1)_+, l_1 + m_6 + \max\{m_1, m_3\} + l_4\}$ and let $s_1 \in I$ with $r_0 + 1 < s_1$, $r_0 + \max\{m_1, m_3\} + l_1 \leq s_1$ and $s_1 + \max\{l_1, m_4 + m_7\} \in I$ be sufficiently large depending on the constants m_i, l_i.

Then there exists a constant $0 < \epsilon \leq 1$ such that if

$$||f||_{Y^{r_0+l_1}} \leq \epsilon$$

we can find $u \in U^{m_0}$ which solves the equation

$$T(u) = T(u_0) + f.$$

Moreover, suppose that $f \in Y^{s_2 - \max\{m_1, m_3\}}$ where $s_2 \in I$ with $s_2 \geq s_1$ and $s_2 + \max\{l_1, m_4 + m_7\} \in I$, and suppose $||f||_{Y^{s_2 - \max\{m_1,m_3\}}} \leq C_{s_2}$. Assume also that the estimate (21) holds for all $s \in [0, s_2 - l_4 - \max\{m_1, m_3\}]$. Then for each $r \in [r_0, s_2 - \max\{m_1, m_3\} - l_1]$ and $s \in I$ with $s < r$, assuming that $s_1 + r - r_0 + \max\{l_1, m_4 + m_7\} \in I$, we have $u \in X^s$, and there exists a constant $K_{r,s}$, possibly increasing with C_{s_2}, but otherwise independent of f, such that

$$||u - u_0||_{X^s} \leq K_{r,s}\, ||f||_{Y^{r+l_1}}$$

Proof.

Step 1 – Setup of the iteration scheme

Assume that $||f||_{Y^{r_0+l_1}} \leq \epsilon$, where $0 < \epsilon \leq 1$ will be chosen later.

Denote the smoothing operators on $(X^s)_{s \in I}$ by $\{S_\theta^X\}_{\theta \geq 1}$ and the smoothing operators on $(Y^s)_{s \in I}$ by $\{S_\theta^Y\}_{\theta \geq 1}$.

We use an iteration scheme to construct a sequence $(u_n)_{n \geq 0}$ in X^∞ which we aim to show converges to a solution $u \in U^{m_0}$ of $T(u) = T(u_0) + f$.

For $n \geq 0$, define

$$\theta_n = \theta_0 + n$$

where $\theta_0 > 1$ will be chosen later depending only on r_0, the constants m_i,l_i and the constants in the smoothing hypothesis and in the inequalities satisfied by DT, D^2T, A, B and R.

Note that

$$\theta_{n+1} \le \theta_n + 1 \le 2\theta_n.$$

For $n \ge 0$, define

$$\begin{aligned} v_n &= S^X_{\theta_n} u_n \\ w_n &= R(v_n) \\ \dot{u}_n &= B(w_n) g_n \\ u_{n+1} &= u_n + \dot{u}_n \end{aligned}$$

where g_n is defined below.

Note that the overdot $\dot{}$ is simply notation indicating a sort of difference and does not denote differentiation.

For $n \ge 0$, define

$$\begin{aligned} g_0 &= S^Y_{\theta_0} f \\ g_{n+1} &= (S^Y_{\theta_{n+1}} - S^Y_{\theta_n})(f - E_n) - S^Y_{\theta_{n+1}} e_n \end{aligned}$$

where

$$E_n = \sum_{m=0}^{n-1} e_m$$

(so $E_0 = 0$), and the error e_n is defined below, for $n \ge 0$,

$$\begin{aligned} e'_n &= (A(u_n) - A(w_n))\dot{u}_n, \\ e''_n &= T(u_n + \dot{u}_n) - T(u_n) - A(u_n)\dot{u}_n, \\ e_n &= e'_n + e''_n. \end{aligned}$$

Note that since g_0 is defined in terms of f only, and we are given u_0, from which v_0 is obtained immediately, the iteration scheme can be determined for $n \ge 0$ in the order $\dot{u}_n, u_{n+1}, v_{n+1}, w_{n+1}, e'_n, e''_n, e_n, E_n, g_{n+1}$.

Note that e_n is defined so that it measures how well $T(u_{n+1}) - T(u_n)$ is approximated by $A(w_n)\dot{u}_n$, by which we mean

$$\begin{aligned} T(u_{n+1}) - T(u_n) &= A(w_n)\dot{u}_n + e_n \\ &= g_n + e_n. \end{aligned}$$

Also note that the formula for g_{n+1} can be rearranged to give

$$g_{n+1} = (S^Y_{\theta_{n+1}} f - S^Y_{\theta_n} f) - (S^Y_{\theta_{n+1}} E_{n+1} - S^Y_{\theta_n} E_n).$$

We thus obtain

$$\begin{aligned} T(u_{n+1}) - T(u_0) &= \sum_{m=0}^{n} (T(u_{m+1}) - T(u_m)) \\ &= \sum_{m=0}^{n} g_m + \sum_{m=0}^{n} e_m \\ &= S^Y_{\theta_n} f - S^Y_{\theta_n} E_n + E_{n+1} \\ &= S^Y_{\theta_n} f + (E_n - S^Y_{\theta_n} E_n) + e_n \end{aligned}$$

which we hope converges to f as $n \to \infty$, since, roughly speaking, $S^Y_{\theta_n} \to \mathrm{id}$ and $e_n \to 0$.

Step 2 – Obtaining estimates for the iterates via induction

We will show the following inductive hypothesis, $[H_n]$, holds.

$$||\dot{u}_n||_{X^s} \leq K_1 \, ||f||_{Y^{r_0+l_1}} \, \theta_n^{s-r_0-1} \qquad \text{for all } s \in [0, s_1]$$

$$||T(u_n) - T(u_0) - f||_{Y^{s+l_1}} \leq K_2 \, ||f||_{Y^{r_0+l_1}} \, \theta_n^{s-r_0} \qquad \text{for } s \in [0, r_0]$$

where the constants $K_1, K_2 > 0$ will be chosen later, with K_1, K_2 independent of n, f and ϵ, but depending on θ_0, and with K_2 depending on K_1. We will choose ϵ sufficiently small such that $K_1 \, ||f||_{Y^{r_0+l_1}} \leq K_1\epsilon \leq 1$ and $K_2 \, ||f||_{Y^{r_0+l_1}} \leq K_2\epsilon \leq 1$.

In what follows, $C_s > 0$ represents a constant, which is independent of n, f and ϵ, and is bounded for s bounded. It will also be independent of θ_0, which will allow us to choose θ_0 so that θ_n is large compared to C_s for s in a certain range. We will write $C > 0$ for a constant which is also independent of s.

Assume now that $[H_m]$ is true for all $0 \leq m \leq n$ and let us show that $[H_{n+1}]$ follows. (We will leave the proof of $[H_0]$ until later.)

Pick a real number $0 < \eta < 1$ such that $r_0 > \max\{m_0 + \max\{m_7, m_8\}, m_4, m_9, l_1 + m_1 + m_2 + \max\{m_7, m_8\}, 2m_2 + 2\max\{m_7, m_8\}, \frac{l_1+m_3}{2} + m_2 + \max\{m_7, m_8\}, l_1 + \max\{m_5, m_6\} + (l_3 - l_1)_+, l_1 + m_6 + \max\{m_1, m_3\} + l_4\} + 2\eta$ and $\eta < \max\{m_1, m_3\}$.

For $s \in I$, define

$$P(s) = \begin{cases} (s - r_0)_+ & \text{for } |s - r_0| \geq \eta, \\ \eta & \text{for } |s - r_0| < \eta. \end{cases}$$

We claim that the following estimates for $0 \leq m \leq n+1$ follow directly from $[H_m]$ for $0 \leq m \leq n$.

$$||u_m - u_0||_{X^s} \leq C_s K_1 ||f||_{Y^{r_0+l_1}} \theta_m^{P(s)} \qquad \text{for } s \in [0, s_1], \tag{27}$$

$$\left|\left|S^X_{\theta_m}(u_m - u_0)\right|\right|_{X^s} \leq C_s K_1 ||f||_{Y^{r_0+l_1}} \theta_m^{P(s)} \qquad \text{for } s \in I, \tag{28}$$

$$\left|\left|(u_m - u_0) - S^X_{\theta_m}(u_m - u_0)\right|\right|_{X^s} \leq C_s K_1 ||f||_{Y^{r_0+l_1}} \theta_m^{(s-r_0)} \qquad \text{for } s \in [0, s_1], \tag{29}$$

$$||u_m - v_m||_{X^s} \leq C_s \theta_m^{s-r_0} \qquad \text{for } s \in [0, s_1], \tag{30}$$

$$||v_m||_{X^s} \leq C_s \theta_m^{P(s)} \qquad \text{for } s \in I, \tag{31}$$

$$||u_m||_{X^s} \leq C_s \theta_m^{P(s)} \qquad \text{for } s \in [0, s_1]. \tag{32}$$

Indeed, the proofs of (27)–(32) are exactly the same as the proofs of (6)–(11).

We also claim that, for $0 \leq m \leq n$, we have

$$||v_m - w_m||_{X^s} \leq C_s \theta_m^{s+\max\{m_7,m_8\}-r_0} \qquad \text{for } s \in I \text{ such that } s + m_7 \in I, \tag{33}$$

$$||u_m - w_m||_{X^s} \leq C_s \theta_m^{s+\max\{m_7,m_8\}-r_0} \qquad \text{for } s \in [0, s_1], \tag{34}$$

$$||w_m||_{X^s} \leq C_s \theta_m^{\max\{P(s), s+\max\{m_7,m_8\}-r_0\}} \qquad \text{for } s \in I \text{ such that } s + m_7 \in I. \tag{35}$$

Indeed, first we assume $s \leq r_0 + \eta$. We have

$$\begin{aligned}
&||v_m - w_m||_{X^s} \\
&= \left|\left|S^X_{\theta_m} u_m - R(S^X_{\theta_m} u_m)\right|\right|_{X^s} \\
&\leq \left|\left|S^X_{\theta_m}(u_m - R(u_m))\right|\right|_{X^s} + \left|\left|S^X_{\theta_m} R(u_m) - R(S^X_{\theta_m} u_m)\right|\right|_{X^s}.
\end{aligned}$$

Now, using the smoothing hypothesis, estimate (24) and $[H_n]$, we obtain

$$\begin{aligned}
\left|\left|S^X_{\theta_m}(u_m - R(u_m))\right|\right|_{X^s} &\leq C_s \theta_m^s ||u_m - R(u_m)||_{X^0} \\
&\leq C_s \theta_m^s ||T(u_m) - T(u_0) - f||_{Y^{l_2}} \\
&\leq C_s \theta_m^{s-r_0} K_2 ||f||_{Y^{r_0+l_1}} \\
&\leq C_s \theta_m^{s-r_0}
\end{aligned}$$

(where we have used $K_2 ||f||_{Y^{r_0+l_1}} \leq K_2 \epsilon \leq 1$).

For the second term, using the estimates (26) and (32), we obtain, choosing $r, r' \geq r_0 + \eta$,

$$
\begin{aligned}
&\left|\left|S^X_{\theta_m} R(u_m) - R(S^X_{\theta_m} u_m)\right|\right|_{X^s} \\
&\leq C_s(\theta_m^{s-r}(1 + ||u_m||_{X^{m_8}})(1 + ||u_m||_{X^{r+m_7}}) \\
&\qquad + \theta_m^{-r'}(1 + ||u_m||_{X^{s+m_7}})(1 + ||u_m||_{X^{r'+m_8}})) \\
&\leq C_s(\theta_m^{s-r}(1 + \theta_m^{r+m_7-r_0}) + \theta_m^{-r'}\theta_m^{P(s+m_7)}(1 + \theta_m^{r'+m_8-r_0})) \\
&\leq C_s\theta_m^{s+\max\{m_7,m_8\}-r_0}.
\end{aligned}
$$

If $s \geq r_0 + \eta$, then we can directly estimate, using (32) and (25),

$$
\begin{aligned}
&||v_m - w_m||_{X^s} \\
&= \left|\left|S^X_{\theta_m} u_m - R(S^X_{\theta_m} u_m)\right|\right|_{X^s} \\
&\leq \left|\left|S^X_{\theta_m} u_m\right|\right|_{X^s} + \left|\left|R(S^X_{\theta_m} u_m)\right|\right|_{X^s} \\
&\leq C_s\, ||u_m||_{X^s} + C_s(1 + ||u_m||_{X^{m_8}})(1 + ||u_m||_{X^{s+m_7}}) \\
&\leq C_s\theta_m^{s-r_0} + C_s\theta_m^{s+m_7-r_0} \\
&\leq C_s\theta_m^{s+m_7-r_0}.
\end{aligned}
$$

This proves (33). Now, using (33) and (30), for $s \in [0, s_1]$, we have

$$
\begin{aligned}
||u_m - w_m||_{X^s} &\leq ||w_m - v_m||_{X^s} + ||v_m - u_m||_{X^s} \\
&\leq C_s\theta_m^{s+\max\{m_7,m_8\}-r_0}.
\end{aligned}
$$

This proves (34).

Using (33) and (31), for $s \in I$, we have

$$
\begin{aligned}
||w_m||_{X^s} &\leq ||w_m - v_m||_{X^s} + ||v_m||_{X^s} \\
&\leq C_s\theta_m^{\max\{P(s), s+\max\{m_7,m_8\}-r_0\}}.
\end{aligned}
$$

This proves (35).

This completes the proof of the claim.

Note that, using (27) and (30), we have

$$
\begin{aligned}
||v_m - u_0||_{X^{m_0}} &\leq ||v_m - u_m||_{X^{m_0}} + ||u_m - u_0||_{X^{m_0}} \\
&\leq C\theta_m^{m_0-r_0} + CK_1\epsilon\theta_m^{P(m_0)} \\
&\leq C\theta_m^{m_0-r_0} + CK_1\epsilon.
\end{aligned}
$$

Thus by taking ϵ sufficiently small depending on K_1 and C, and θ_0 sufficiently large depending on C, we have $v_n, v_{n+1} \in U$. Similarly we can ensure $w_n \in U$ using (34). Also note that (6) in the case $s = m_0$ implies $u_n \in U$ for ϵ sufficiently small, and $[H_n]$ implies that $u_n + \dot{u}_n \in U$ for ϵ sufficiently small. Note that the same argument also shows that the line segments $[u_n, u_n + \dot{u}_n]$ and $[u_n, w_n]$ are in U for ϵ sufficiently small.

We claim that the following estimate holds.

$$||T(u_n) - T(u_0) - f||_{Y^s} \le C_s \theta_n^{s+\max\{m_1,m_3\}-r_0} \quad \text{for } s \in [0, s_1 - \max\{m_1, m_3\}]. \tag{36}$$

Indeed, for $s \in [r_0, s_1 - \max\{m_1, m_3\}]$, using Taylor's theorem, (20) and (27), we have

$$\begin{aligned}
&||T(u_n) - T(u_0) - f||_{Y^s} \\
&\le ||T(u_n) - T(u_0)||_{Y^s} + ||f||_{Y^s} \\
&\le \left|\left| \sup_{t\in[0,1]} DT(u_0 + t(u_n - u_0))(u_n - u_0) \right|\right|_{Y^s} + C^0 \\
&\le C_s(||u_n - u_0||_{X^{s+m_1}} + ||u_n - u_0||_{X^{m_2}} (1 + ||u_n - u_0||_{X^{s+m_3}})) + C^0 \\
&\le C_s \theta_n^{s+\max\{m_1,m_3\}-r_0}
\end{aligned}$$

(assuming that $\max\{m_1, m_3\} \ge \eta$). We combine this with $[H_n]$ for $s \in [0, r_0]$ to get

$$||T(u_n) - T(u_0) - f||_{Y^s} \le C_s \theta_n^{s+\max\{m_1,m_3\}-r_0}$$

for all $s \in [0, s_1 - \max\{m_1, m_3\}]$.

Estimate of e'_n. We claim that for all $s \in [0, s_1 - \max\{m_1, m_3\}]$,

$$\left|\left|e'_n\right|\right|_{Y^s} \le C_s K_1 \, ||f||_{Y^{r_0+l_1}} \, \theta_n^{M'(s)-1+\eta}$$

where

$$\begin{aligned}
M'(s) = \max\{&s + m_1 + m_2 + \max\{m_7, m_8\} - 2r_0, \\
&(s + m_3 - r_0)_+ + 2\max\{m_7, m_8\} + 2m_2 - 2r_0\}.
\end{aligned}$$

Applying the estimate (22) together with Taylor's theorem, $[H_n]$ and the estimates (34) and (35), we have, for $s \in [0, s_1 - \max\{m_1, m_3\}]$,

$$\begin{aligned}
&\left|\left|e'_n\right|\right|_{Y^s} \\
&= ||(A((u_n - w_n) + w_n) - A(w_n))\dot{u}_n||_{Y^s} \\
&\le C_s(||\dot{u}_n||_{X^{s+m_1}} \, ||u_n - w_n||_{X^{m_2}} + ||\dot{u}_n||_{X^{m_2}} \, ||u_n - w_n||_{X^{s+m_1}}
\end{aligned}$$

$$
\begin{aligned}
&+ ||\dot{u}_n||_{X^{m_2}} \, ||u_n - w_n||_{X^{m_2}} \, (1 + ||w_n||_{X^{s+m_3}} + ||u_n - w_n||_{X^{s+m_3}})) \\
&\le C_s K_1 \, ||f||_{Y^{r_0+l_1}} \, (\theta_n^{s+m_1-r_0-1} \theta_n^{m_2-r_0+\max\{m_7,m_8\}} + \theta_n^{m_2-r_0-1} \theta_n^{s+\max\{m_7,m_8\}+m_1-r_0} \\
&+ \theta_n^{m_2-r_0-1} \theta_n^{m_2+\max\{m_7,m_8\}-r_0} (1 + \theta_n^{P(s+m_3)+\max\{m_7,m_8\}} + \theta_n^{s+\max\{m_7,m_8\}+m_3-r_0})) \\
&\le C_s K_1 \, ||f||_{Y^{r_0+l_1}} \, \theta_n^{M'(s)-1+\eta}.
\end{aligned}
$$

Estimate of e_n''. We claim that for all $s \in [0, s_1 - \max\{m_1 + l_4, m_3 + l_4, m_5, m_9\}]$,

$$
\left|\left|e_n''\right|\right|_{Y^s} \le C_s K_1 \, ||f||_{Y^{r_0+l_1}} \, \theta_n^{M(s)-1+\eta}
$$

where

$$
\begin{aligned}
M(s) = \max\{&s + m_1 + m_2 + \max\{m_7, m_8\} - 2r_0, \\
&(s + m_3 - r_0)_+ + 2\max\{m_7, m_8\} + 2m_2 - 2r_0, \\
&s + \max\{m_5, m_6\} + (l_3 - l_1)_+ - 2r_0, s + m_6 + \max\{m_1, m_3\} + l_4 - 2r_0\}.
\end{aligned}
$$

Indeed, we have

$$
\begin{aligned}
e_n'' &= T(u_n + \dot{u}_n) - T(u_n) - A(u_n)\dot{u}_n \\
&= T(u_n + \dot{u}_n) - T(u_n) - DT(u_n)\dot{u}_n + (A(u_n) - DT(u_n))\dot{u}_n.
\end{aligned}
$$

Applying Taylor's theorem, (19), $[H_n]$ and the estimate (32), we have, for $s \in [0, s_1 - \max\{m_1, m_3\}]$,

$$
\begin{aligned}
&||T(u_n + \dot{u}_n) - T(u_n) - DT(u_n)\dot{u}_n||_{Y^s} \\
&\le \sup_{t\in[0,1]} \left|\left|D^2T(u_n + t\dot{u}_n)(\dot{u}_n, \dot{u}_n)\right|\right|_{Y^s} \\
&\le C_s(||\dot{u}_n||_{X^{s+m_1}} \, ||\dot{u}_n||_{X^{m_2}} + ||\dot{u}_n||^2_{X^{m_2}} \, (1 + \sup_{t\in[0,1]} ||u_n + t\dot{u}_n||_{X^{s+m_3}})) \\
&\le C_s(K_1 \, ||f||_{Y^{r_0+l_1}} \, \theta_n^{s+m_1-r_0-1} K_1 \, ||f||_{Y^{r_0+l_1}} \, \theta_n^{m_2-r_0-1} \\
&+ K_1^2 \, ||f||^2_{Y^{r_0+l_1}} \, \theta_n^{2m_2-2r_0-2} (1 + \theta_n^{P(s+m_3)} + K_1 \, ||f||_{Y^{r_0+l_1}} \, \theta_n^{s+m_3-r_0-1})) \\
&\le \theta_n^{-1} C_s K_1 \, ||f||_{Y^{r_0+l_1}} \, \theta_n^{M'(s)-1+\eta} \\
&\le C_s K_1 \, ||f||_{Y^{r_0+l_1}} \, \theta_n^{M'(s)-1+\eta}
\end{aligned}
$$

where we have used $K_1 \, ||f||_{Y^{r_0+l_1}} \le K_1\epsilon \le 1$.

For $s \in [0, s_1 - \max\{m_5, m_9, m_1 + l_4, m_3 + l_4\}]$, we have, using (21), $[H_n]$, and (36),

$$
\begin{aligned}
&||(A(u_n) - DT(u_n))\dot{u}_n||_{Y^s} \\
&\leq C_s(||\dot{u}_n||_{X^{s+m_5}} \, ||T(u_n) - T(u_0) - f||_{Y^{l_3}} + ||\dot{u}_n||_{X^{m_6}} \, ||T(u_n) - T(u_0) - f||_{Y^{s+l_4}} \\
&+ ||\dot{u}_n||_{X^{m_6}} \, ||T(u_n) - T(u_0) - f||_{Y^{l_3}} \, (1 + ||u||_{X^{s+m_9}})) \\
&\leq C_s K_1 \, ||f||_{Y^{r_0+l_1}} \, (\theta_n^{s+m_5-r_0-1} \theta_n^{(l_3-l_1)_+-r_0} + \theta_n^{m_6-r_0-1} \theta_n^{s+\max\{m_1,m_3\}+l_4-r_0} \\
&+ \theta_n^{m_6-r_0-1} \theta_n^{(l_3-l_1)_+-r_0} \theta_n^{(s+m_9-r_0)_++\eta}) \\
&\leq C_s K_1 \, ||f||_{Y^{r_0+l_1}} \, \theta_n^{M''(s)-1+\eta}
\end{aligned}
$$

where

$$
\begin{aligned}
M''(s) = \max\{&s + m_5 + (l_3 - l_1)_+ - 2r_0, s + m_6 + \max\{m_1, m_3\} + l_4 - 2r_0, \\
&s + m_6 + (l_3 - l_1)_+ - 2r_0\}
\end{aligned}
$$

where we have used $r_0 \geq m_9$.

Adding the two above estimates yields the estimate for e_n''.

Estimate of e_n. Adding the estimates for e_n' and e_n'', we obtain

$$
||e_n||_{Y^s} \leq C_s K_1 \, ||f||_{Y^{r_0+l_1}} \, \theta_n^{M(s)-1+\eta}
$$

for all $s \in [0, s_1 - \max\{m_1 + l_4, m_3 + l_4, m_5, m_9\}]$.

Estimate of g_{n+1}. We claim that for all $s \in I$,

$$
||g_{n+1}||_{Y^s} \leq C_s(K_1 \, ||f||_{Y^{r_0+l_1}} \, \theta_n^{M(s)-1+\eta} + ||f||_{Y^{r_0+l_1}} \, \theta_n^{s-r_0-l_1-1}).
$$

Indeed, we have

$$
g_{n+1} = (S^Y_{\theta_{n+1}} - S^Y_{\theta_n})(f - E_n) - S^Y_{\theta_{n+1}} e_n.
$$

Note that for any $z \in Y^{s'}$,

$$
\left|\left|(S^Y_{\theta_{n+1}} - S^Y_{\theta_n})z\right|\right|_{Y^s} \leq C_{s',s} \theta_n^{s-s'-1} \, ||z||_{Y^{s'}}
$$

by the smoothing hypothesis (3) and Taylor's theorem.

Setting $s' = r_0 + l_1$, we have

$$
\left|\left|(S^Y_{\theta_{n+1}} - S^Y_{\theta_n})f\right|\right|_{Y^s} \leq C_s \theta_n^{s-r_0-l_1-1} \, ||f||_{Y^{r_0+l_1}} \, .
$$

We also have

$$\left|\left|(S^Y_{\theta_{n+1}} - S^Y_{\theta_n})E_n\right|\right|_{Y^s} \le C_{s',s}\theta_n^{s-s'-1}\, ||E_n||_{Y^{s'}}\,.$$

Now, for $s' \in [0, s_1 - \max\{m_1 + l_4, m_3 + l_4, m_5, m_9\}]$, we have, from the estimate for e_n,

$$\begin{aligned} ||E_n||_{Y^{s'}} &= \left|\left|\sum_{m=0}^{n-1} e_m\right|\right|_{Y^{s'}} \\ &\le C_{s'}K_1\, ||f||_{Y^{r_0+l_1}} \sum_{m=0}^{n-1} \theta_m^{M(s')-1+\eta} \\ &\le C_{s'}K_1\, ||f||_{Y^{r_0+l_1}}\; \theta_n^{M(s')+\eta} \end{aligned} \tag{37}$$

if $M(s') \ge 0$, by the integral comparison used before. Note that $M(s')$ has slope 1 for large enough s' depending on r_0 and the constants m_i, l_i, so to achieve $M(s') \ge 0$ it suffices to take s' large in relation to r_0 and the constants m_i, l_i. To do this we require s_1 sufficiently large in relation to r_0 and the constants m_i, l_i.

Hence

$$\begin{aligned} \left|\left|(S^Y_{\theta_{n+1}} - S^Y_{\theta_n})E_n\right|\right|_{Y^s} &\le C_{s'}C_{s',s}K_1\, ||f||_{Y^{r_0+l_1}}\; \theta_n^{M(s')+s-s'-1+\eta} \\ &\le C_s K_1\, ||f||_{Y^{r_0+l_1}}\; \theta_n^{M(s)-1+\eta} \end{aligned}$$

by choosing s' sufficiently large compared to r_0 and the constants m_i so that $M(s)$ has slope 1 for $s \ge s'$. (Hence $M(s') - s' \le M(s) - s$ for all s since $M(s) - s$ is decreasing for $s \le s'$ and constant for $s \ge s'$.) Again, to do this we require s_1 sufficiently large in relation to r_0 and the constants m_i, l_i. This fixes s_1.

Similarly, for s' sufficiently large, we have

$$\begin{aligned} \left|\left|S^Y_{\theta_{n+1}} e_n\right|\right|_{Y^s} &\le C_{s',s}\theta_n^{s-s'}\, ||e_n||_{Y^{s'}} \\ &\le C_{s',s}C_{s'}K_1\, ||f||_{Y^{r_0+l_1}}\; \theta_n^{M(s')+s-s'-1+\eta} \\ &\le C_s K_1\, ||f||_{Y^{r_0+l_1}}\; \theta_n^{M(s)-1+\eta}. \end{aligned}$$

Hence the estimate for g_{n+1} holds.

Estimate of $T(u_{n+1}) - T(u_0) - f$ We have

$$T(u_{n+1}) - T(u_0) - f = (S^Y_{\theta_n} f - f) + (E_n - S^Y_{\theta_n} E_n) + e_n.$$

Let $s \in [0, r_0]$.

By (2) from the smoothing hypothesis, we have

$$\left|\left|S^Y_{\theta_n} f - f\right|\right|_{Y^{s+l_1}} \leq C_s \theta_n^{s-r_0} \left|\left|f\right|\right|_{Y^{r_0+l_1}} .$$

Also,

$$\begin{aligned}
\left|\left|E_n - S^Y_{\theta_n} E_n\right|\right|_{Y^{s+l_1}} &\leq C_{s,s'} \theta_n^{s-s'} \left|\left|E_n\right|\right|_{Y^{s'+l_1}} \qquad \text{for } s' \geq s \\
&\leq C_{s,s'} \theta_n^{s-s'} C_s \theta_n^{M(s'+l_1)+\eta} K_1 \left|\left|f\right|\right|_{Y^{r_0+l_1}} \\
&\quad \text{using (37), for } s' \text{ large enough such that } M(s'+l_1) \geq 0 \\
&\leq C_s \theta_n^{M(s'+l_1)+s-s'+\eta} K_1 \left|\left|f\right|\right|_{Y^{r_0+l_1}} \\
&\leq C_s \theta_n^{s-r_0} K_1 \left|\left|f\right|\right|_{Y^{r_0+l_1}}
\end{aligned}$$

since $M(s'+l_1) + \eta \leq M(l_1) + \eta + s' < s' - r_0$.

Finally,

$$\begin{aligned}
\left|\left|e_n\right|\right|_{Y^{s+l_1}} &\leq C_s \theta_n^{M(s+l_1)+\eta-1} K_1 \left|\left|f\right|\right|_{Y^{r_0+l_1}} \\
&\leq C_s \theta_n^{s-r_0-1} K_1 \left|\left|f\right|\right|_{Y^{r_0+l_1}}
\end{aligned}$$

since $M(s+l_1) + \eta \leq M(l_1) + \eta + s < s - r_0$.

Hence we have

$$\left|\left|T(u_{n+1}) - T(u_0) - f\right|\right|_{Y^{s+l_1}} \leq C_s \theta_n^{s-r_0} K_1 \left|\left|f\right|\right|_{Y^{r_0+l_1}}$$

for $s \in [0, r_0]$. Thus, by choosing K_2 sufficiently large depending on K_1 and C_s for $s \in [0, r_0]$, we have

$$\left|\left|T(u_{n+1}) - T(u_0) - f\right|\right|_{Y^{s+l_1}} \leq K_2 \left|\left|f\right|\right|_{Y^{r_0+l_1}} \theta_{n+1}^{s-r_0} \tag{38}$$

for $s \in [0, r_0]$.

The estimates

$$\left|\left|v_{n+1} - w_{n+1}\right|\right|_{X^s} \leq C_s \theta_{n+1}^{s-r_0+\max\{m_7,m_8\}} \qquad \text{for } s \in I \text{ such that } s + m_7 \in I \tag{39}$$

$$\left|\left|u_{n+1} - w_{n+1}\right|\right|_{X^s} \leq C_s \theta_{n+1}^{s-r_0+\max\{m_7,m_8\}} \qquad \text{for } s \in [0, s_1] \tag{40}$$

$$\left|\left|w_{n+1}\right|\right|_{X^s} \leq C_s \theta_{n+1}^{\max\{P(s), s+\max\{m_7,m_8\}-r_0\}} \qquad \text{for } s \in I \text{ such that } s + m_7 \in I \tag{41}$$

now hold, and are proved exactly as for the estimates (33)–(35) using the estimate (38) to go from n to $n+1$.

Estimate of $\dot{u}_{n+1}$. We have

$$\dot{u}_{n+1} = B(w_{n+1})g_{n+1}.$$

Hence, for all $s \in I$ such that $s + l_1, s + m_4 + m_7 \in I$, using (23), the estimate (35) and the estimate for g_{n+1}, we have

$$\begin{aligned}
&||\dot{u}_{n+1}||_{X^s}\\
&\leq C_s(||g_{n+1}||_{Y^{s+l_1}} + ||g_{n+1}||_{Y^{l_1}}\,(1 + ||w_{n+1}||_{X^{s+m_4}}))\\
&\leq C_s(K_1\,||f||_{Y^{r_0+l_1}}\,\theta_{n+1}^{M(s+l_1)-1+\eta} + ||f||_{Y^{r_0+l_1}}\,\theta_{n+1}^{s-r_0-1} +\\
&(K_1\,||f||_{Y^{r_0+l_1}}\,\theta_{n+1}^{M(l_1)-1+\eta} + ||f||_{Y^{r_0+l_1}}\,\theta_{n+1}^{-r_0-1})(1 + \theta_{n+1}^{\max\{P(s+m_4),s+\max\{m_7,m_8\}-r_0\}}))\\
&\leq C_s(K_1\,||f||_{Y^{r_0+l_1}}\,\theta_{n+1}^{M(l_1)+s-1+\eta} + ||f||_{Y^{r_0+l_1}}\,\theta_{n+1}^{s-r_0-1}) \qquad (42)
\end{aligned}$$

since $P(s + m_4) \leq s$ because $r_0 > m_4 + 2\eta$ and $s + \max\{m_7, m_8\} - r_0 \leq s$ because $r_0 > \max\{m_7, m_8\}$, and $M(l_1 + s) \leq M(l_1) + s$ because M has slope at most 1.

We want to obtain

$$||\dot{u}_{n+1}||_{X^s} \leq K_1\,||f||_{Y^{r_0+l_1}}\,\theta_{n+1}^{s-r_0-1}$$

for $s \in [0, s_1]$.

To make the first term sufficiently small, we require

$$-\gamma := M(l_1) + r_0 + \eta < 0.$$

Then we can choose θ_0 large enough so that

$$C_s\theta_{n+1}^{M(l_1)+s-1+\eta} = C_s\theta_{n+1}^{s-r_0-1}\theta_{n+1}^{-\gamma} \leq C_s\theta_{n+1}^{s-r_0-1}\theta_0^{-\gamma} \leq \frac{1}{2}\theta_{n+1}^{s-r_0-1}$$

for all $s \in [0, s_1]$.

We note that $M(l_1) + r_0 + \eta < 0$ if and only if $r_0 - \eta > l_1 + m_1 + m_2 + \max\{m_7, m_8\}$, $r_0 - \eta > 2m_2 + 2\max\{m_7, m_8\}$, $r_0 - \eta > m_2 + \max\{m_7, m_8\} + \frac{l_1+m_3}{2}$, $r_0 - \eta > l_1 + \max\{m_5, m_6\} + (l_3 - l_1)_+$ and $r_0 - \eta > l_1 + m_6 + \max\{m_1, m_3\} + l_4$, which indeed hold by the choice of r_0 and η.

To make the second term sufficiently small, we take $K_1 \geq 2C_s$ for all $s \in [0, s_1]$.

This gives $[H_{n+1}]$.

Proof of $[H_0]$ We have

$$g_0 = S_{\theta_0}^{Y} f$$

and

$$v_0 = S^X_{\theta_0} u_0$$

and

$$w_0 = R(v_0).$$

Hence

$$\begin{aligned}
||\dot{u}_0||_{X^s} &= \left|\left|B(R(S^X_{\theta_0} u_0)) S^Y_{\theta_0} f\right|\right|_{X^s} \\
&\leq C_s(\left|\left|S^Y_{\theta_0} f\right|\right|_{Y^{s+l_1}} + \left|\left|S^Y_{\theta_0} f\right|\right|_{Y^{l_1}} (1 + \left|\left|R(S^X_{\theta_0} u_0)\right|\right|_{X^{s+m_4}})) \\
&\leq C_s \left|\left|S^Y_{\theta_0} f\right|\right|_{Y^{s+l_1}} \\
&\leq C_s \, ||f||_{Y^{r_0+l_1}} \, \theta_0^{(s-r_0)+} \qquad \text{by (1) and (2) from the smoothing hypothesis} \\
&\leq K_1 \, ||f||_{Y^{r_0+l_1}} \, \theta_0^{s-r_0-1}
\end{aligned}$$

for all $s \in [0, s_1]$, assuming that K_1 is sufficiently large compared to θ_0 and C_s for $s \in [0, s_1]$.

Now for $s \in [0, r_0]$,

$$\begin{aligned}
||T(u_0) - T(u_0) - f||_{Y^{s+l_1}} &= ||f||_{Y^{s+l_1}} \\
&\leq K_2 \, ||f||_{Y^{r_0+l_1}} \, \theta_0^{s-r_0}
\end{aligned}$$

for all $s \in [0, r_0]$, assuming that K_2 is sufficiently large compared to θ_0.

This proves $[H_0]$.

Step 3 – Better estimates if $f \in Y^{s_2 - \max\{m_1, m_3\}}$ for $s_2 \geq s_1$

Assume $f \in Y^{s_2 - \max\{m_1, m_3\}}$ where $s_2 \in I$ with $s_2 \geq s_1$ and $s_2 + \max\{l_1, m_4 + m_7\} \in I$, and suppose $||f||_{Y^{s_2 - \max\{m_1, m_3\}}} \leq C_{s_2}$. Let $r \in I$ with $r \geq r_0$ be such that $s_1 + r - r_0 + \max\{l_1, m_4 + m_7\} \in I$. We will show that, for all $n \geq 0$ and for all $s \in [0, s_2]$, we have

$$||\dot{u}_n||_{X^s} \leq C_{r,s} \, ||f||_{Y^{r+l_1}} \, \theta_n^{s-r-1} \tag{43}$$

where the constant $C_{r,s} > 0$ is independent of n and f, except that it may increase with $||f||_{Y^{s_2 - \max\{m_1, m_3\}}}$.

Firstly, note that we have proved $[H_n]$ for $n \geq 0$, and hence all the estimates from step 2 which were conditional on the inductive hypothesis are now valid, and we may use them as we wish.

We are going to prove the above statement by an induction argument, but not an induction on n. We are going to use the estimates from step 2 for each n separately to obtain the above inequality, and the constant will be independent of n because the constants from step 2 are independent of n.

We claim by induction on $k \geq 0$ that for all $s \in [0, s_2]$, we have

$$||\dot{u}_n||_{X^s} \leq C_{k,r,s}\, ||f||_{Y^{r+l_1}}\, \theta_n^{s-r_0-\gamma_k-1} \qquad [G_k]$$

where the constant $C_{k,r,s} > 0$ is independent of n and f, and

$$\gamma_k = \min\{k\gamma, r - r_0\}.$$

Indeed, the estimate (42) for $\dot{u}_{n+1}$ in step 2 implies that

$$||\dot{u}_n||_{X^s} \leq C_s\, ||f||_{Y^{r_0+l_1}}\, \theta_n^{s-r_0-1} \tag{44}$$

for all $s \in I$ such that $s + \max\{l_1, m_4 + m_7\} \in I$ (not just $s \in [0, s_1]$ which would follow directly from $[H_n]$).

Using this, we can obtain the following new versions of the estimates (27), (30)–(32) for all $s \in [0, s_2]$ (not just $s \in [0, s_1]$) via exactly the same calculations

$$||u_m - u_0||_{X^s} \leq C_s \theta_m^{s-r_0}, \tag{45}$$

$$||u_m - v_m||_{X^s} \leq C_s \theta_m^{s-r_0}, \tag{46}$$

$$||v_m||_{X^s} \leq C_s \theta_m^{P(s)}, \tag{47}$$

$$||u_m||_{X^s} \leq C_s \theta_m^{P(s)}. \tag{48}$$

We then obtain, for all $s \in [0, s_2]$, the estimates

$$||w_m - v_m||_{X^s} \leq C_s \theta_m^{s+\max\{m_7,m_8\}-r_0}, \tag{49}$$

$$||w_m - u_m||_{X^s} \leq C_s \theta_m^{s+\max\{m_7,m_8\}-r_0}, \tag{50}$$

$$||w_m||_{X^s} \leq C_s \theta_m^{\max\{P(s), s+\max\{m_7,m_8\}-r_0\}}. \tag{51}$$

Using the fact that $||f||_{Y^{r_0+l_1}} \leq ||f||_{Y^{r+l_1}}$, (44) immediately implies $[G_0]$.

Now we assume $[G_k]$ holds and aim to show $[G_{k+1}]$ holds.

Now we want to obtain new estimates for e'_n and e''_n.

First we estimate e'_n. Note that in the estimate for e'_n there was at least one factor involving $\dot{u}_n$ in each term. If we estimate this one factor using the new estimate given by $[G_k]$ and the other quantities using (44) and the slightly modified estimates (45)–(48), we obtain

$$\left|\left|e'_n\right|\right|_{Y^s} \leq C_{k,r,s}\, ||f||_{Y^{r+l_1}}\, \theta_n^{M'(s)-1+\eta-\gamma_k}$$

for all $s \in [0, s_2 - \max\{m_1, m_3\}]$. The constant $C_{k,r,s}$ is independent of f since we have only used the new estimate given by $[G_k]$ in one factor, and the other estimates we have used involve $||f||_{Y^{r_0+l_1}}$, which is bounded by $\epsilon \leq 1$.

Now we estimate e''_n. The first part of the estimate can be modified in exactly the same way as above, to obtain

$$||T(u_n + \dot{u}_n) - T(u_n) - DT(u_n)\dot{u}_n||_{Y^s} \leq C_{k,r,s}\, ||f||_{Y^{r+l_1}}\, \theta_n^{M'(s)-1+\eta-\gamma_k}$$

for all $s \in [0, s_2 - \max\{m_1, m_3\}]$.

We proceed similarly for the second part of the estimate of e''_n to obtain

$$||(A(u_n) - DT(u_n))\dot{u}_n||_{Y^s} \leq C_{k,r,s}\, ||f||_{Y^{r+l_1}}\, \theta_n^{M''(s)-1+\eta-\gamma_k}$$

for all $s \in [0, s_2 - \max\{m_5, m_9, m_1 + l_4, m_3 + l_4\}]$.

Here, the constant depends on $||f||_{Y^{s_2-\max\{m_1,m_3\}}}$, and we need to assume that the estimate (21) holds for all $s \in [0, s_2 - l_4 - \max\{m_1, m_3\}]$.

Thus we obtain the estimate

$$||e_n||_{Y^s} \leq C_{k,r,s}\, ||f||_{Y^{r+l_1}}\, \theta_n^{M(s)-1+\eta-\gamma_k}$$

for all $s \in [0, s_2 - \max\{m_5, m_9, m_1 + l_4, m_3 + l_4\}]$.

This implies that for $s' \in [0, s_2 - \max\{m_5, m_1 + l_4, m_3 + l_4\}]$, we have

$$\begin{aligned} ||E_n||_{Y^{s'}} &= \left|\left|\sum_{m=0}^{n-1} e_m\right|\right|_{Y^{s'}} \\ &\leq C_{k,r,s'}\, ||f||_{Y^{r+l_1}} \sum_{m=0}^{n-1} \theta_m^{M(s')-1+\eta-\gamma_k} \\ &\leq C_{k,r,s'}\, ||f||_{Y^{r+l_1}}\, \theta_n^{M(s')+\eta-\gamma_k} \end{aligned} \tag{52}$$

as long as $M(s') \geq \gamma_k$. It is possible to pick such an s' if $s_1 + r - r_0 + \max\{l_1, m_4\} \in I$ given the fact that $M(s_1 - \max\{m_5, m_9, m_1 + l_4, m_3 + l_4\}) \geq 0$ and $M(s)$ has slope 1 for $s \geq s_1 - \max\{m_5, m_9, m_1 + l_4, m_3 + l_4\}$.

Hence

$$\begin{aligned} \left|\left|(S^Y_{\theta_{n+1}} - S^Y_{\theta_n})E_n\right|\right|_{Y^s} &\leq C_{s',k} C_{k,r,s} \theta_n^{M(s')+s-s'-1+\eta-\gamma_k} \\ &\leq C_{k,r,s}\, ||f||_{Y^{r+l_1}}\, \theta_n^{M(s)-1+\eta-\gamma_k} \end{aligned}$$

as long as $M(s') \geq \gamma_k$ and s' is sufficiently large compared to r_0 and the constants m_i, l_i so that $M(s)$ has slope 1 for $s \geq s'$.

We also have the estimate

$$\left|\left|S^Y_{\theta_{n+1}} e_n\right|\right|_{Y^s} \leq C_{k,r,s}\, ||f||_{Y^{r+l_1}}\, \theta_n^{M(s)-1+\eta-\gamma_k}.$$

In addition we can use the new estimate

$$\left|\left|(S^Y_{\theta_{n+1}} - S^Y_{\theta_n}) f\right|\right|_{Y^s} \leq C_{r,s}\theta_n^{s-r-l_1-1} \, ||f||_{Y^{r+l_1}} \, .$$

This gives us the following new estimate for g_{n+1}, for all $s \in I$,

$$||g_{n+1}||_{Y^s} \leq C_{k,r,s} \, ||f||_{Y^{r+l_1}} \, (\theta_n^{M(s)-1+\eta-\gamma_k} + \theta_n^{s-r-l_1-1}).$$

From this we obtain,for all $s \in [0, s_2]$,

$$\begin{aligned} ||\dot{u}_n||_{X^s} &\leq C_{r,s} \, ||f||_{Y^{r+l_1}} \, (\theta_n^{M(l_1)+s-1+\eta-\gamma_k} + \theta_n^{s-r-1}) \\ &\leq C_{r,s} \, ||f||_{Y^{r+l_1}} \, (\theta_n^{s-r_0-1-\gamma_k-\gamma} + \theta_n^{s-r-1}) \\ &\leq C_{r,s} \, ||f||_{Y^{r+l_1}} \, \theta_n^{s-r_0-\gamma_{k+1}-1} \end{aligned}$$

where we have used the fact that $M(l_1) + r_0 + \eta = -\gamma$.

This is $[G_{k+1}]$.

For large enough k, we have $k\gamma \geq r - r_0$, so $\gamma_k = r - r_0$ and this gives (43).

Step 4 – Convergence to a solution

Assume as above that $f \in Y^{s_2 - \max\{m_1, m_3\}}$ where $s_2 \in I$ with $s_2 \geq s_1$ and $s_2 + \max\{l_1, m_4, m_7\} \in I$, and suppose $||f||_{Y^{s_2-\max\{m_1,m_3\}}} \leq C_{s_2}$. Let $r \geq r_0$.

Using (13), we have

$$\begin{aligned} \sum_{m=0}^{n} ||u_{m+1} - u_m||_{X^s} &= \sum_{m=0}^{n} ||\dot{u}_m||_{X^s} \\ &\leq C_{r,s} \, ||f||_{Y^{r+l_1}} \, \theta_{n+1}^{(s-r)+} \end{aligned}$$

for $r \neq s$, with $r, s \in [0, s_2]$.

Thus

$$\sum_{m=0}^{n} ||u_{m+1} - u_m||_{X^s}$$

converges as $n \to \infty$ for $s < r$. Hence, by completeness, $u_n \to u$ in X^s as $n \to \infty$, for all $s < r$, for some $u \in \cap_{0 \leq s < r} X^s$.

Note the above calculation also implies that

$$||u_n - u_0||_{X^s} \leq C_{r,s} \, ||f||_{Y^{r+l_1}}$$

for $s < r$, so we have

$$||u - u_0||_{X^s} \leq C_{r,s} \, ||f||_{Y^{r+l_1}} \, .$$

Next we claim that

$$T(u_{n+1}) - T(u_0) \to f$$

in X^s as $n \to \infty$, for all $s < r$.

Indeed,

$$T(u_{n+1}) - T(u_0) = S^Y_{\theta_n} f + (E_n - S^Y_{\theta_n} E_n) + e_n$$

so

$$T(u_{n+1}) - T(u_0) - f = (S^Y_{\theta_n} f - f) + (E_n - S^Y_{\theta_n} E_n) + e_n.$$

By (2) from the smoothing hypothesis, we have

$$\left|\left|S^Y_{\theta_n} f - f\right|\right|_{Y^{s+l_1}} \le C_{r,s}\theta_n^{s-r}\,||f||_{Y^{r+l_1}} \to 0 \text{ as } n \to \infty.$$

Also,

$$\begin{aligned}
\left|\left|E_n - S^Y_{\theta_n} E_n\right|\right|_{Y^{s+l_1}} &\le C_{s,s'}\theta_n^{s-s'}\,||E_n||_{Y^{s'+l_1}} \qquad \text{for } s' \ge s \\
&\le C_{s,s'}\theta_n^{s-s'} C_{r,s}\theta_n^{M(s'+l_1)+\eta-(r-r_0)}\,||f||_{Y^{r+l_1}} \\
&\quad \text{using (18), for } s' \text{ large enough such that } M(s'+l_1) \ge r - r_0 \\
&\le C_{r,s}\theta_n^{M(s'+l_1)+s-s'+\eta-(r-r_0)}\,||f||_{Y^{r+l_1}} \\
&\le C_{r,s}\theta_n^{s-r}\,||f||_{Y^{r+l_1}} \to 0 \text{ as } n \to \infty
\end{aligned}$$

since $M(s' + l_1) + \eta + r_0 \le M(l_1) + \eta + r_0 + s' < s'$.

Finally,

$$||e_n||_{Y^{s+l_1}} \le C_{r,s}\theta_n^{M(s+l_1)+\eta-(r-r_0)-1}\,||f||_{Y^{r+l_1}} \to 0 \text{ as } n \to \infty$$

since $M(s + l_1) + \eta + r_0 \le M(l_1) + \eta + r_0 + s < s$.

This proves the claim.

Now since $T : U \to Y^0$ is continuous as a map from X^{m_0} to Y^0, and $r_0 > m_0$, so $u_n \to u$ in X^{m_0}, we have that $T(u_n) \to T(u)$ in Y^0, hence $T(u) = T(u_0) + f$ as required.

This completes the proof.

5 Applying the Theorem in Sobolev Spaces

This section assumes familiarity with the standard Sobolev spaces $W^{k,p}(\Omega)$ of functions on the domain Ω with weak derivatives up to order k in $L^p(\Omega)$, and Sobolev embedding theorems – see for example the chapter of Evans [11] entitled

'Sobolev Spaces', or see Adams and Fournier [1] for a more complete reference. We do however give the definition of fractional Sobolev spaces below, since these are slightly less standard. See, for example, Adams and Fournier [1] for much more detail.

5.1 The Smoothing Operators in H^s

Definition 5. For $d \in \mathbb{N}$ and $0 \leq s \in \mathbb{R}$ we define the *Sobolev space of order s*, $H^s(\mathbb{R}^d)$, by

$$H^s(\mathbb{R}^d) = \{u \in L^2(\mathbb{R}^d) : (1 + |\xi|^2)^{\frac{s}{2}} \hat{u}(\xi) \in L^2(\mathbb{R}^d)\}$$

where $\hat{u}$ denotes the Fourier transform of u, which we also denote by $\mathscr{F}[u]$. We endow H^s with norm $||\cdot||_{H^s}$ given by

$$||u||_{H^s} = \left|\left|(1 + |\xi|^2)^{\frac{s}{2}} \hat{u}(\xi)\right|\right|_{L^2}.$$

Then $H^s(\mathbb{R}^d)$ is a Banach space for each s and $(H^s(\mathbb{R}^d), ||\cdot||_{H^s})_{s\geq 0}$ is a decreasing family of Banach spaces, in the sense of Definition 1.

Notation. For $\phi \in C^\infty(\mathbb{R}^d)$ (with values in $\mathbb{R}$), write $\phi_\epsilon = \epsilon^{-d}\phi(\frac{x}{\epsilon})$.

Notation. We write $\mathscr{S}(\mathbb{R}^d)$ for the Schwartz space of smooth functions which decay faster than the reciprocal of any polynomial, and use the well-known fact that the Fourier transform is an automorphism of $\mathscr{S}(\mathbb{R}^d)$.

Proposition 2. *The decreasing family of Banach spaces* $(H^s(\mathbb{R}^d), ||\cdot||_{H^s})_{s\geq 0}$ *satisfies the smoothing hypothesis 4. Moreover, the smoothing operators can be taken as* $S_\theta u = \rho_{\frac{1}{\theta}} * u$ *for* $\theta \geq 1$, *where* $\rho \in \mathscr{S}(\mathbb{R}^d)$ *is a specially constructed mollifier.*

Proof. Let $\hat{\rho} \in C_c^\infty(\mathbb{R}^d)$ with $0 \leq \hat{\rho} \leq 1$ be an even function such that $\hat{\rho} = 1$ on $B_{\frac{1}{2}}(0)$ and $\hat{\rho} = 0$ outside $B_1(0)$, where $B_r(x)$ denotes the closed ball of radius r about x.

Define ρ to be the inverse Fourier transform of $\hat{\rho}$, which is real since $\hat{\rho}$ is even, and $\rho \in \mathscr{S}(\mathbb{R}^d)$, since $\hat{\rho} \in \mathscr{S}(\mathbb{R}^d)$.

For $u \in H^0(\mathbb{R}^d) = L^2(\mathbb{R}^d)$, we define

$$S_\theta u = \rho_{\frac{1}{\theta}} * u.$$

Let $0 \leq r, s \in \mathbb{R}$ and $u \in H^s(\mathbb{R}^d)$.

Note that, by properties of the Fourier transform,

$$\widehat{S_\theta u}(\xi) = \widehat{\rho_{\frac{1}{\theta}}}(\xi)\hat{u}(\xi) = \hat{\rho}(\frac{\xi}{\theta})\hat{u}(\xi).$$

Hence

$$\begin{aligned}\int_{\mathbb{R}^d}(1+|\xi|^2)^r\left|\widehat{S_\theta u}(\xi)\right|^2\mathrm{d}\xi &= \int_{\mathbb{R}^d}(1+|\xi|^2)^r\hat{\rho}(\frac{\xi}{\theta})^2\hat{u}(\xi)^2\mathrm{d}\xi\\ &= \int_{\mathbb{R}^d}(1+|\xi|^2)^{r-s}\hat{\rho}(\frac{\xi}{\theta})^2(1+|\xi|^2)^s\hat{u}(\xi)^2\mathrm{d}\xi\\ &\le ||u||^2_{H^s}\sup_{\xi\in\mathbb{R}^d}(1+|\xi|^2)^{r-s}\hat{\rho}(\frac{\xi}{\theta})^2\\ &\le ||u||^2_{H^s}(1+\theta^2)^{(r-s)+}\\ &\le C_{r,s}\,||u||^2_{H^s}\,\theta^{2(r-s)+}\end{aligned}$$

since $0\le\hat{\rho}\le 1$ and $\hat{\rho}(\frac{\xi}{\theta})=0$ for $\xi\ge\theta$.

This proves (1), and also that $S_\theta : H^0(\mathbb{R}^d)\to\cap_{s\ge 0}H^s(\mathbb{R}^d)$.

Now

$$\begin{aligned}\int_{\mathbb{R}^d}(1+|\xi|^2)^r\left|\widehat{u-S_\theta u}(\xi)\right|^2\mathrm{d}\xi &= \int_{\mathbb{R}^d}(1+|\xi|^2)^r(1-\hat{\rho}(\frac{\xi}{\theta}))^2\hat{u}(\xi)^2\mathrm{d}\xi\\ &= \int_{\mathbb{R}^d}(1+|\xi|^2)^{r-s}(1-\hat{\rho}(\frac{\xi}{\theta}))^2(1+|\xi|^2)^s\hat{u}(\xi)^2\mathrm{d}\xi\\ &\le ||u||^2_{H^s}\sup_{\xi\in\mathbb{R}^d}(1+|\xi|^2)^{r-s}(1-\hat{\rho}(\frac{\xi}{\theta}))^2\\ &\le ||u||^2_{H^s}(1+(\frac{\theta}{2})^2)^{(r-s)}\\ &\le C_{r,s}\,||u||^2_{H^s}\,\theta^{2(r-s)}\end{aligned}$$

assuming $r\le s$, since $0\le\hat{\rho}\le 1$ and $1-\hat{\rho}(\frac{\xi}{\theta})=0$ for $\xi\le\frac{\theta}{2}$.

This proves (2).

Finally, for small $h\in\mathbb{R}$, we have

$$\begin{aligned}\mathscr{F}\left[\frac{S_{\theta+h}u-S_\theta u}{h}\right](\xi) &= \frac{\hat{\rho}(\frac{\xi}{\theta+h})-\hat{\rho}(\frac{\xi}{\theta})}{h}\hat{u}(\xi)\\ &= \left(-\frac{1}{\theta^2}\sum_{i=1}^d\xi_i\,\partial_i\,\hat{\rho}(\frac{\xi}{\theta})+R(h,\theta,\xi)\right)\hat{u}(\xi)\end{aligned}$$

by Taylor's theorem, where

$$|R(h,\theta,\xi)| \leq h \sup_{\theta \leq \phi \leq \theta + h} \frac{\mathrm{d}^2}{\mathrm{d}\phi^2} \hat{\rho}(\frac{\xi}{\phi}).$$

This implies

$$\int_{\mathbb{R}^d} (1 + |\xi|^2)^r \, |R(h,\theta,\xi)|^2 \, |\hat{u}(\xi)|^2 \to 0 \text{ as } h \to 0$$

so that $S_\theta u$ is differentiable with respect to θ with derivative the inverse Fourier transform of

$$-\frac{1}{\theta^2} \sum_{i=1}^{d} \xi_i \partial_i \hat{\rho}(\frac{\xi}{\theta}) \hat{u}(\xi).$$

We also see that

$$\begin{aligned}
\left\| \frac{\mathrm{d}}{\mathrm{d}\theta} S_\theta u \right\|_{H^r}^2 &= \int_{\mathbb{R}^d} (1 + |\xi|^2)^r (\frac{1}{\theta^2} \sum_{i=1}^{d} \xi_i \partial_i \hat{\rho}(\frac{\xi}{\theta}))^2 \, |\hat{u}(\xi)|^2 \\
&\leq ||u||_{H^s}^2 \sup_{\xi \in \mathbb{R}^d} (1 + |\xi|^2)^{r-s} (\frac{1}{\theta^2} \sum_{i=1}^{d} \xi_i \partial_i \hat{\rho}(\frac{\xi}{\theta}))^2 \\
&\leq C_{r,s} \, ||u||_{H^s}^2 \, \theta^{2(r-s-1)}
\end{aligned}$$

since $\partial_i \hat{\rho}(\frac{\xi}{\theta})$ is zero for $\xi \leq \frac{\theta}{2}$ and $\xi \geq \theta$.

This proves (3).

5.2 *Tame Estimates in Sobolev Spaces*

The results in this section are fairly standard, and are based on standard Sobolev embeddings. Results of this type can be found in classical references on Sobolev spaces, for example Adams and Fournier [1]. However, we try and formulate them in a form which is most useful for obtaining tame estimates in the applications we have in mind.

The following lemma is very useful for proving chain and product rules in Sobolev spaces.

Lemma 1. *Let $p \in [1,\infty]$, $\Omega \subset \mathbb{R}^d$, for $d \geq 1$, be a domain where the standard Sobolev embedding holds and let $m > \frac{d}{p}$ be an integer. Let $0 \leq m_i \leq m$ be integers for $1 \leq i \leq n$ with $\sum_{i=1}^{n} m_i \geq (n-1)m$ and let $u_i \in W^{m_i,p}(\Omega)$. Then $\prod_{i=1}^{n} u_i \in L^p(\Omega)$ and*

$$\left\| \prod_{i=1}^{n} u_i \right\|_{L^p(\Omega)} \leq C \prod_{i=1}^{n} ||u_i||_{W^{m_i,p}(\Omega)} .$$

Proof. For $p = \infty$ the result is obvious and in fact only requires $m \geq 0$, so we will assume $p < \infty$.

We will use the following Sobolev embeddings. Let $k \geq 1$ be an integer and $u \in W^{k,p}(\Omega)$. Then for $q \geq p$,

$$||u||_{L^q(\Omega)} \leq C \, ||u||_{W^{k,p}(\Omega)}$$

provided

$$\frac{1}{q} > \frac{1}{p} - \frac{k}{d}$$

and $kp \leq d$. (Note it is the case $kp = d$ that requires the inequality to be strict.) If $kp > d$ then

$$||u||_{L^\infty(\Omega)} \leq C \, ||u||_{W^{k,p}(\Omega)} .$$

Suppose $m_i p > d$ for some i. By renumbering if necessary, we may assume $m_n p > d$. Then

$$\begin{aligned} \left\| \prod_{i=1}^{n} u_i \right\|_{L^p(\Omega)} &\leq \left\| \prod_{i=1}^{n-1} u_i \right\|_{L^p(\Omega)} ||u_n||_{L^\infty(\Omega)} \\ &\leq C \left\| \prod_{i=1}^{n-1} u_i \right\|_{L^p(\Omega)} ||u_n||_{W^{m_n,p}(\Omega)} . \end{aligned}$$

Also note that since $m_n \leq m$, we have $\sum_{i=1}^{n-1} m_i \geq (n-2)m$. Hence we are reduced to proving the result with n replaced by $n-1$. Thus we may assume $m_i p \leq d$ for all i.

Suppose $m_i = 0$ for some i. By renumbering if necessary, we may assume $m_n = 0$. Then $\sum_{i=1}^{n-1} m_i \geq (n-1)m$ and $0 \leq m_i \leq m$ implies $m_i = m > \frac{d}{p}$ for all $i < n$, hence

$$\begin{aligned} \left\| \prod_{i=1}^{n} u_i \right\|_{L^p(\Omega)} &\leq \prod_{i=1}^{n-1} ||u_i||_{L^\infty(\Omega)} \, ||u_n||_{L^p(\Omega)} \\ &\leq C \prod_{i=1}^{n-1} ||u_i||_{W^{m_i,p}(\Omega)} \, ||u_n||_{W^{m_n,p}(\Omega)} . \end{aligned}$$

Thus we may assume $m_i > 0$ for all i.

Now, using Hölder's inequality,

$$\left\|\prod_{i=1}^{n} u_i\right\|_{L^p(\Omega)} \leq \prod_{i=1}^{n} ||u_i||_{L^{\frac{p}{\lambda_i}}(\Omega)}$$

where $\sum_{i=1}^{n} \lambda_i = 1$ and $0 \leq \lambda_i \leq 1$ for all i. Hence, using Sobolev embedding, we have

$$\left\|\prod_{i=1}^{n} u_i\right\|_{L^p(\Omega)} \leq C \prod_{i=1}^{n} ||u_i||_{W^{m_i,p}(\Omega)}$$

provided

$$\frac{\lambda_i}{p} > \frac{1}{p} - \frac{m_i}{d}$$

for all i. But, summing the above inequalities, it is possible, assuming $0 < m_i \leq \frac{d}{p}$, to choose such $0 \leq \lambda_i \leq 1$ with $\sum_{i=1}^{n} \lambda_i = 1$ if and only if

$$\frac{n}{p} - \frac{\sum_{i=1}^{n} m_i}{d} < \frac{1}{p} \iff \sum_{i=1}^{n} m_i > (n-1)\frac{d}{p}.$$

But this does indeed hold since $\sum_{i=1}^{n} m_i \geq (n-1)m$ and $m > \frac{d}{p}$.

Corollary 1 (Leibniz's Rule or The Product Rule). *Let $p \in [1, \infty]$, $\Omega \subset \mathbb{R}^d$, for $d \geq 1$, be a domain where the standard Sobolev embedding holds and let $m > \frac{d}{p}$ be an integer. Let $0 \leq m_i \leq m$ be integers for $1 \leq i \leq n$ and $0 \leq k \leq m$ be an integer, with $\sum_{i=1}^{n} m_i \geq (n-1)m + k$. Let $u_i \in W^{m_i,p}(\Omega)$. Then $\prod_{i=1}^{n} u_i \in W^{k,p}(\Omega)$ with weak derivatives given by the classical Leibniz rule and*

$$\left\|\prod_{i=1}^{n} u_i\right\|_{W^{k,p}(\Omega)} \leq C \prod_{i=1}^{n} ||u_i||_{W^{m_i,p}(\Omega)}.$$

Proof. Let γ^i be multi-indices with $\sum_{i=1}^{n} \gamma_i = \alpha$, where $|\alpha| \leq k$. Note that $\sum_{i=1}^{n}(m_i - |\gamma_i|) \geq (n-1)m$, hence we may apply the above result to obtain

$$\left\|\prod_{i=1}^{n} \partial^{\gamma^i} u_i\right\|_{L^p(\Omega)} \leq C \prod_{i=1}^{n} ||u_i||_{W^{m_i,p}(\Omega)}.$$

Assuming u_i are smooth, we immediately obtain the result, since $\partial^\alpha \prod_{i=1}^n u_i$ is a sum of terms of the form $\prod_{i=1}^n \partial^{\gamma^i} u_i$ by the classical chain rule. For non-smooth u_i we use approximation by smooth functions together with this inequality.

Corollary 2 (The Chain Rule). *Let $p \in [1, \infty]$, $\Omega \subset \mathbb{R}^d$, for $d \geq 1$, be a domain where the standard Sobolev embedding holds and let $m > \frac{d}{p}$ be an integer. Let $F \in C_b^m(\mathbb{R}^{d'})$ and $u : \Omega \to \mathbb{R}^{d'}$ with $u \in W^{m,p}(\Omega)$. Let α be a multi-index with $1 \leq |\alpha| \leq m$. Let $0 < \beta \leq \alpha, 0 < \gamma^j \leq \alpha$ $(1 \leq j \leq |\beta|)$, be multi-indices with $\sum_{j=1}^{|\beta|} \left|\gamma^j\right| = |\alpha|$. Then*

$$\left|\left|(\partial^\beta F)(u) \prod_{j=1}^{|\beta|} \partial^{\gamma^j} u_{i_j}\right|\right|_{L^p(\Omega)} \leq C \, ||u||_{W^{m,p}(\Omega)}^{|\beta|}$$

where u_{i_j} denotes a component of u depending on j. Moreover, the function $F(u) \in L^\infty(\Omega)$ has a weak α-derivative in $L^p(\Omega)$ given as in the classical chain rule by sums of terms of the above form which satisfies the inequality

$$||\partial^\alpha (F(u))||_{L^p(\Omega)} \leq C \, ||u||_{W^{m,p}(\Omega)} \, (1 + ||u||_{W^{m,p}(\Omega)})^{m-1}.$$

In addition, if $F(0) = 0$, then $F(u) \in W^{m,p}(\Omega)$ with

$$||F(u)||_{W^{m,p}(\Omega)} \leq C \, ||u||_{W^{m,p}(\Omega)} \, (1 + ||u||_{W^{m,p}(\Omega)})^{m-1}.$$

Proof. Note that $\sum_{j=1}^{\beta} (m - \left|\gamma^j\right|) = |\beta| \, m - |\alpha| \geq (|\beta| - 1)m$, hence we may use the above result and the fact that $\partial^\beta F$ is bounded to obtain

$$\begin{aligned}\left|\left|(\partial^\beta F)(u) \prod_{j=1}^{|\beta|} \partial^{\gamma^j} u_{i_j}\right|\right|_{L^p(\Omega)} &\leq C \prod_{j=1}^{|\beta|} \left|\left|\partial^{\gamma^j} u_{i_j}\right|\right|_{W^{m-\left|\gamma^j\right|,p}(\Omega)} \\ &\leq C \, ||u||_{W^{m,p}(\Omega)}^{|\beta|} .\end{aligned}$$

Assuming u_i are smooth, we immediately obtain the required inequalities, since $\partial^\alpha (F(u))$ is a sum of terms of the form $(\partial^\beta F)(u) \prod_{j=1}^{|\beta|} \partial^{\gamma^j} u_{i_j}$ by the classical product and chain rules. For non-smooth u_i we use approximation by smooth functions together with this inequality.

Finally, if $F(0) = 0$, then we have

$$\begin{aligned}|F(u)| &= \left|\int_0^1 DF(tu)u \, \mathrm{d}t\right| \\ &\leq C \, |u|\end{aligned}$$

since DF is bounded. Thus $F(u) \in L^p(\Omega)$ with $||F(u)||_{L^p(\Omega)} \leq C\,||u||_{L^p(\Omega)}$. Together with the previous part, this implies the final statement of the result.

Corollary 3. *Let $D, D' \subset \mathbb{R}^{d'}$ be open with $D' \subset\subset D$. Let $F \in C^m(D)$ and $u : \Omega \to D'$ with $u \in W^{m,p}(\Omega)$. Then the above chain rule holds with these new F and u.*

Proof. Since u takes values in D', we may modify F outside D' by multiplying by a smooth cut-off function which is identically 1 on D' and 0 outside D'' for some $D'' \subset\subset D$, so we may assume $F \in C_b^m(\mathbb{R}^{d'})$, and then we can apply the above result.

Proposition 3 (The Derivative of a Differential Operator on Sobolev Spaces). *Let $p \in [1,\infty]$, $\Omega \subset \mathbb{R}^d$, for $d \geq 1$, be a domain where the standard Sobolev embedding holds and let $m \geq 0$ be an integer. Let I be a subinterval of $\mathbb{N}_0$ containing 0 and $l+m$, where $l > \frac{d}{p}$ is an integer, and set $X^s = W^{s,p}(\Omega, \mathbb{R}^{d'})$ for $s \in I$ and $Y^s = W^{s,p}(\Omega, \mathbb{R}^{d''})$. Let $U \subset X^\infty$ be $||\cdot||_{X^r}$-open for some $r \in I$ with $r \geq l+m$ and assume $0 \in U$. Define*

$$T : U \to Y^{\infty - m}$$

by

$$T(u)(x) = F(\{\partial^\alpha u(x) : 0 \leq |\alpha| \leq m\})$$

where $F : \mathbb{R}^{d'} \times \cdots \times \mathbb{R}^{d'} \to \mathbb{R}^{d''}$ is smooth and bounded with bounded derivatives on the range of $\{\partial^\alpha u : 0 \leq |\alpha| \leq m\}$ for $u \in U$ (so we may assume F is smooth and bounded with bounded derivatives), and $F(0) = 0$. The above rather complicated notation is merely a convenient way of expressing that $F(\{\partial^\alpha u : 0 \leq |\alpha| \leq m\})$ is a smooth function of u and its partial derivatives up to order m, which can be evaluated at x to give a function of x.

Write v_α^i for the argument of F which is evaluated at $\partial^\alpha u^i(x)$ in the above formula.

Then T is twice differentiable with derivatives given by

$$(DT(u)h)(x) = \sum_{0 \leq i \leq d'} \sum_{0 \leq \beta \leq m} \partial^\beta h^i(x) \frac{\partial F}{\partial v_\beta^i}(\{\partial^\alpha u(x)\})$$

$$D^2T(u)(h,h')(x) = \sum_{0 \leq i,j \leq d'} \sum_{0 \leq \beta,\gamma \leq m} \partial^\beta h^i(x) \partial^\gamma h'^j(x) \frac{\partial^2 F}{\partial v_\gamma^j \partial v_\beta^i}(\{\partial^\alpha u(x)\})$$

and the following inequalities hold.

$$||DT(u)h||_{Y^s} \leq C_s(||h||_{X^{s+m}} + ||h||_{X^l}(1 + ||u||_{X^{s+m}}))$$

$$\left|\left|D^2T(u)(h,h')\right|\right|_{Y^s}$$

$$\leq C_s(||h||_{X^{s+m}} \left|\left|h'\right|\right|_{X^l} + ||h||_{X^l} \left|\left|h'\right|\right|_{X^{s+m}} + ||h||_{X^l} \left|\left|h'\right|\right|_{X^l}(1 + ||u||_{X^{s+m}}))$$

for all $u \in U$, $h, h' \in X^\infty$ and $s \in I$ such that $s + m \in I$, where the constant $C_s > 0$ is bounded for s bounded.

Proof. First we assume all functions are smooth, or else we can use approximation by smooth functions. Note that by the chain rule, $||F(\{\partial^\alpha u(x) : 0 \le |\alpha| \le m\})||_{Y^s} \le C_s\, ||u||_{X^{s+m}}$, since $r > \frac{d}{p}$ (and the constant depends on U), hence T is well-defined. Using Taylor's Theorem, for $u \in U$, $t \in (-1, 1)$ and $||h||_{X^r}$ small enough such that the line segment $[u - h, u + h]$ lies in U, we have

$$\begin{aligned}
&\frac{1}{t}(T(u + th) - T(u))(x) \\
&= \frac{1}{t}(F(\{\partial^\alpha u(x) + \partial^\alpha h(x)\}) - F(\{\partial^\alpha u(x)\})) \\
&= \sum_{0 \le i \le d'} \sum_{0 \le \beta \le m} \frac{\partial F}{\partial v^i_\beta}(\{\partial^\alpha u(x)\}) \partial^\beta h^i(x) \\
&\quad + t \sum_{0 \le i, j \le d'} \sum_{0 \le \beta, \gamma \le m} \partial^\beta h^i(x) \partial^\gamma h^j(x) \int_0^1 (1 - \tau) \frac{\partial^2 F}{\partial v^j_\gamma \partial v^i_\beta}(\{\partial^\alpha u(x) + \tau \partial^\alpha h(x)\}) \mathrm{d}\tau.
\end{aligned}$$

Applying the chain rule to $\frac{1}{t}$ times the last term, which may be thought of as a function of (u, h), we see that $\frac{1}{t}$ times the last term is in Y^s for $s \in I$ such that $s + m \in I$ hence the last term converges to zero in Y^s as $t \to 0$. Similarly

$$\begin{aligned}
&\frac{1}{t}(DT(u + th')h - DT(u)h)(x) \\
&= \frac{1}{t}\Big(\sum_{0 \le i \le d'} \sum_{0 \le \beta \le m} \partial^\beta h^i(x) \frac{\partial F}{\partial v^i_\beta}(\{\partial^\alpha u(x) + t \partial^\alpha h'(x)\}) \\
&\qquad - \sum_{0 \le i \le d'} \sum_{0 \le \beta \le m} \partial^\beta h^i(x) \frac{\partial F}{\partial v^i_\beta}(\{\partial^\alpha u(x)\})\Big) \\
&= \sum_{0 \le i, j \le d'} \sum_{0 \le \beta, \gamma \le m} \partial^\beta h^i(x) \partial^\gamma h'^j(x) \frac{\partial^2 F}{\partial v^j_\gamma \partial v^i_\beta}(\{\partial^\alpha u(x)\}) \\
&\quad + t \sum_{0 \le i, j, k \le d'} \sum_{0 \le \beta, \gamma, \delta \le m} \partial^\beta h^i(x) \partial^\gamma h'^j(x) \partial^\delta h'^k(x) \\
&\qquad \times \int_0^1 (1 - \tau) \frac{\partial^3 F}{\partial v^k_\delta \partial v^j_\gamma \partial v^i_\beta}(\{\partial^\alpha u(x) + \tau \partial^\alpha h'(x)\}) \mathrm{d}\tau.
\end{aligned}$$

Applying the same argument we see the last term converges to zero as $t \to 0$.

Now, using the chain rule we have

$$\left| \frac{\partial F}{\partial v^i_\beta}(\{\partial^\alpha u(x)\}) \right|_{W^{s,p}(\Omega)} \leq C_s \, ||u||_{W^{s+m,p}(\Omega)}$$

for integer $s \geq 1$ and $u \in U$, where $|\cdot|_{W^{s,p}(\Omega)}$ denotes the Sobolev semi-norm of order s (the sum of the L^p norms of the weak derivatives of order s).

Define

$$H(x) = \frac{\partial F}{\partial v^i_\beta}(\{\partial^\alpha u(x)\}).$$

For integer $s \geq 0$, using the product rule and the above, we have

$$\begin{aligned}
&\left| \frac{\partial F}{\partial v^i_\beta}(\{\partial^\alpha u(x)\})\partial^\beta h^i(x) \right|_{W^{s,p}(\Omega)} \\
&\leq C_s\Big(\sum_{1\leq\delta\leq s} \left|\left|\partial^\delta H \partial^{s-\delta}\partial^\beta h^i(x)\right|\right|_{L^p(\Omega)} + \left|\left|H\partial^s\partial^\beta h^i(x)\right|\right|_{L^p(\Omega)}\Big) \\
&\leq C_s(||DH||_{W^{l-1,p}(\Omega)} \, ||h||_{W^{s+m,p}(\Omega)} + ||DH||_{W^{s-1,p}(\Omega)} \, ||h||_{W^{l,p}(\Omega)} \\
&\qquad + ||H||_{L^\infty(\Omega)} \, ||h||_{W^{s+m,p}(\Omega)}) \\
&\leq C_s(||h||_{W^{s+m,p}(\Omega)} + ||h||_{W^{l,p}(\Omega)} \, (1 + ||u||_{W^{s+m,p}(\Omega)}))
\end{aligned}$$

for any $h \in X^\infty$ and $u \in U$, where we have used $r \geq l + m$.

In a similar manner, we obtain the inequality for the second derivative of T.

6 Application to Compressible Vortex Sheets in 2D

Here we show how the paper [8] of Coulombel and Secchi fits into the above framework. In fact the above framework is specifically devised to fit this case and the original ideas are contained in the paper by Coulombel and Secchi and earlier papers. For the sake of brevity, to follow this section it is necessary to refer to their paper. Note though that a significant portion of the work of the full result of Coulombel and Secchi is in solving the linearised equations with an appropriate energy estimate, which can be found in [7]. We believe that the abstract framework below should also fit the scheme used by Trakhinin in [25], since his scheme is very similar to the one used by Coulombel and Secchi.

We make some simplifications to the scheme of Coulombel and Secchi – firstly we take the boundary condition for the continuity of density (which is a linear condition) as part of the definition of the function spaces. Secondly, we treat the

Eikonal equations in a slightly simpler way which is less optimal with respect to regularity. It appears that although we need more regularity on the approximate solution, we only require it to be small in a lower-order Sobolev space.

The aim of their paper is to show short-time structural stability of plane vortex sheets for the 2D isentropic Euler equations of gas dynamics. This means the following. We start with two constant states $\bar{U}^+ = (\bar{\rho}, \bar{v}^+, 0)$, $\bar{U}^- = (\bar{\rho}, \bar{v}^-, 0)$ with pressure given by $\bar{p} = p(\bar{\rho})$ and sound speed given by $\bar{c} = \sqrt{p'(\bar{\rho})}$. When patched together either side of $\{x_2 = 0\}$ these form a weak solution of the 2D isentropic Euler equations equal to $\bar{U}^+$ in $\{x_2 > 0\}$ and equal to $\bar{U}^-$ in $\{x_2 < 0\}$, since the Rankine-Hugoniot jump conditions are satisfied across $\{x_2 = 0\}$. Since the normal velocity is continuous whereas the tangential velocity jumps this is called a vortex-sheet solution and it is characteristic in the sense that the boundary matrix for the system evaluated at this state is singular. We then impose smooth initial data close to this state (satisfying the Rankine-Hugoniot conditions with continuous normal velocity) which includes perturbing the discontinuity slightly so it is the graph of a function. The aim is to show the short-time existence of a solution with the same structure – that is, smooth either side of a surface of discontinuity across which the Rankine-Hugoniot conditions are satisfied with continuous normal velocity. This requires a stability assumption on the background state, $\left|\bar{u}^+ - \bar{u}^-\right| > 2\sqrt{2}\bar{c}$, and also a smallness assumption on the initial data. After some reductions the problem is reduced to finding a local inverse of a nonlinear operator, so that Nash-Moser iteration may be applied. The preliminary work includes changing coordinates to fix the free surface, which involves adding the Eikonal equations to the system to be solved, and introducing an approximate solution so that the initial data can be taken as zero. The main work is then to obtain a tame estimate for the linearised equations, after which a modified version of Nash-Moser iteration as above can be applied.

Notation. We will use the notation of [8], and to avoid conflict of notation with the above we will write u, v, w, f, g used in the above in bold face as $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{f}, \mathbf{g}$. We will also write $\mathscr{U}$ and $\mathscr{V}$ instead of U and V used above.

6.1 The Function Spaces

For $T > 0$, define

$$\Omega_T = \{(t, x_1, x_2) \in \mathbb{R}^3 : t < T, x_2 > 0\}$$
$$\omega_T = \{(t, x_1) \in \mathbb{R}^2 : t < T\}.$$

For integer $s \geq 0$ and real $\gamma \geq 1$ define the weighted Sobolev space

$$H^s_\gamma(\Omega_T) = \{\exp(\gamma t)v : v \in H^m(\Omega_T)\}$$

where $H^s(\Omega_T)$ is the usual Sobolev space of order s. We define $H^s_\gamma(\omega_T)$ similarly. The norm on $u \in H^s_\gamma(\Omega_T)$ is given by

$$||u||_{H^s_\gamma(\Omega_T)} = ||\exp(-\gamma t)u||_{H^s(\Omega_T)}\,.$$

Next, we define, for integer $s \geq 0$,

$$\mathscr{F}^s_\gamma(\Omega_T) = \{u \in H^s_\gamma(\Omega_T) : u = 0 \text{ for } t < 0\}$$

and we define $\mathscr{F}^s_\gamma(\omega_T)$ similarly. Now, adapting the notation of [8] to our framework, we define

$$X^s = \{\mathbf{u} \in (\mathscr{F}^{s+3}_\gamma(\Omega_T))^3 \times (\mathscr{F}^{s+3}_\gamma(\Omega_T))^3 \times \mathscr{F}^{s+3}_\gamma(\Omega_T) \times \mathscr{F}^{s+3}_\gamma(\Omega_T)$$
$$: \Psi^+|_{x_2=0} = \Psi^-|_{x_2=0}, \rho^+|_{x_2=0} = \rho^-|_{x_2=0}\}$$

where we write

$$\mathbf{u} = (V^+, V^-, \Psi^+, \Psi^-)$$

and

$$V = (\rho, v, u)$$

and define

$$\psi := \Psi^+|_{x_2=0} = \Psi^-|_{x_2=0}.$$

Note that we omit the superscripts $^+$ and $^-$ in formulae which apply to both. We have chosen X^0 to consist of products of Sobolev spaces of order 3 because of the embedding $H^s(\mathbb{R}^d) \subset W^{1,\infty}(\mathbb{R}^d)$ for $s > \frac{d}{2} + 1$, and in this case the dimension d is 3 (two space and one time).

We define the norm $||\cdot||_{X^s}$ on X^s as the usual product norm (the sum of the norms of the components). Then $\{X^s\}_{s\in I}$ is a decreasing family of Banach spaces, where $I = [0, s_3]$ is an interval in $\mathbb{N}_0$, for integer $s_3 > 0$ which we will fix later sufficiently large.

Similarly, we define

$$Y^s = \{\mathbf{g} \in (\mathscr{F}^{s+3}_\gamma(\Omega_T))^3 \times (\mathscr{F}^{s+3}_\gamma(\Omega_T))^3 \times \mathscr{F}^{s+3}_\gamma(\Omega_T) \times \mathscr{F}^{s+3}_\gamma(\Omega_T)\}$$

where we write

$$\mathbf{g} = (f^+, f^-, h^+, h^-)$$

and

$$f = (f_1, f_2, f_3).$$

6.2 The Smoothing Operators

Note that in order to define the smoothing operators on $\{X^s\}_{s\in I}$ (which can then be used on $\{Y^s\}_{s\in I}$ as well), we must make some modifications from those on $H^s(\mathbb{R}^d)$. Firstly, we must replace $\mathbb{R}^d$ by a domain with a Lipschitz boundary with finite covering, which is easily done via an extension operator. Next, we must ensure that the property $\mathbf{u} = 0$ for $t < 0$ is preserved under the action of the smoothing operators, which was done by Alinhac in [2], and finally we must ensure that the two properties $\Psi^+|_{x_2=0} = \Psi^-|_{x_2=0} = \psi$ and $\rho^+|_{x_2=0} = \rho^-|_{x_2=0}$ are preserved. See [8] for the details of this construction using a lifting operator.

6.3 The Background Solution and the Approximate Solution

Although we will not introduce the original problem considered in [8] (since we wish to show the use of Nash-Moser iteration only), we need to introduce the background or stationary solution and approximate solution for reference.

The background solution is given in the form

$$(\overline{U}^{\pm} = (\overline{\rho}^{\pm} = \overline{\rho}, \pm\overline{v}, \overline{u}^{\pm} = 0), \overline{\Phi}^{\pm} = \pm x_2)$$

where $\overline{\rho}, \overline{v}$ are constants with $\overline{\rho} > 0$.

We assume the existence of an 'approximate solution' $(U^{a+}, U^{a-}, \Phi^{a+}, \Phi^{a-})$ with $U^a - \overline{U}, \Phi^a - \overline{\Phi} \in H^{s_4+3}(\Omega_T)$ having compact support, which has the following properties. Here, s_4 is a sufficiently large integer with $s_4 \geq s_3 + 2$. In fact $s_4 = s_3 + 2$ will do.

$$\begin{aligned}
\partial_t^j \mathbb{L}(U^a, \Phi^a)|_{t=0} &= 0 \text{ for } 0 \leq j \leq s_3 + 3 \\
\partial_t \Phi^a + v^a \partial_{x_1} \Phi^a - u^a &= 0 \\
\Phi^{a+}|_{x_2=0} = \Phi^{a-}|_{x_2=0} &=: \phi^a \\
\rho^{a+} - \rho^{a-} &= 0 \\
\partial_{x_2} \Phi^{a+} &\geq \frac{3}{4} \\
\partial_{x_2} \Phi^{a-} &\leq -\frac{3}{4} \\
\rho^{a\pm} &\geq \delta_0 \\
\left|\left|U^a - \overline{U}\right|\right|_{H^7(\Omega_T)} + \left|\left|\Phi^a - \overline{\Phi}\right|\right|_{H^7(\Omega_T)} &\leq \delta_1
\end{aligned}$$

for some $\delta_0 > 0$, where we are allowed to choose constant $\delta_1 > 0$ as small as we like (which restricts the size of the initial data in the original problem). The first order differential operator $\mathbb{L}$ is defined in the next section.

6.4 The Nonlinear Operator and Equations

6.4.1 The Operator T and the Set $\mathscr{U}$

We set $m_0 = 4$ and define

$$\mathscr{U}^4 = \{\mathbf{u} \in X^4 : ||\mathbf{u}||_{X^4} \leq \delta_2\}$$

where $\delta_2 > 0$ is chosen sufficiently small. In particular, we need

$$||\Psi^{\pm}||_{W^{1,\infty}(\Omega_T)} \leq \frac{1}{2}$$
$$||\rho^{\pm}||_{L^{\infty}(\Omega_T)} \leq \frac{\delta_0}{2}$$

which is possible via Sobolev embedding. This ensures that $\partial_{x_2}(\Phi^{a\pm} + \Psi^{\pm})$ and $\rho^{a\pm} + \rho^{\pm}$ are bounded away from zero.

We define the operator $T : \mathscr{U}^4 \to Y^0$ by

$$T(\mathbf{u}) = \begin{pmatrix} \mathscr{L}(V^+, \Psi^+) \\ \mathscr{L}(V^-, \Psi^-) \\ \mathscr{E}(V^+, \Psi^+) \\ \mathscr{E}(V^-, \Psi^-) \end{pmatrix}.$$

Here,

$$\mathscr{L}(V, \Psi) = \mathbb{L}(U^a + V, \Phi^a + \Psi) - \mathbb{L}(U^a, \Phi^a)$$

and

$$\mathbb{L}(U, \Phi) = \partial_t U + A_1(U)\partial_{x_1} U + \frac{1}{\partial_{x_2}\Phi}(A_2(U) - \partial_t \Phi - \partial_{x_1}\Phi A_1(U))\partial_{x_2} U.$$

The matrices $A_1(U)$ and $A_2(U)$ are smooth functions of U for $U_1 > 0$, where U_1 is the first component of U (the 'ρ' component). See [8] for the exact expressions of these matrices. Also,

$$\mathscr{E}(V, \Psi) = \partial_t \Psi + (v^a + v)\partial_{x_1}\Psi - u + v\partial_{x_1}\Phi^a.$$

We note also that $T : \mathscr{U} \to Y^{\infty-1}$, where $\mathscr{U} = \mathscr{U}^4 \cap X^{\infty}$.

6.4.2 The Equations

Define

$$f^a = \begin{cases} -\mathbb{L}(U^a, \Phi^a) \text{ for } t \geq 0, \\ 0 \text{ for } t < 0. \end{cases}$$

Then by the properties of the approximate solution, we have $f^a \in \mathscr{F}_\gamma^{s_3+3}(\Omega_T)$ and together with the definition of $\mathbb{L}$ we obtain

$$||f^a||_{Y^{s_3}} \leq C\delta_1 =: \epsilon.$$

Set

$$\mathbf{f} = \begin{pmatrix} f^{a+} \\ f^{a-} \\ 0 \\ 0 \end{pmatrix}.$$

For ϵ sufficiently small, we wish to solve the equation

$$T(\mathbf{u}) = \mathbf{f}$$

which is equivalent to

$$T(\mathbf{u}) = T(\mathbf{u_0}) + \mathbf{f}$$

if we set

$$\mathbf{u_0} = 0$$

since $T(0) = 0$.

6.5 *The Linearised Operator, Modified Linearised Operator, Modified State and Linearised Equations*

6.5.1 The Operator *DT*

Notation. To make the notation easier, let us use $\tilde{\mathbf{u}}$ instead of $\mathbf{v}$ to represent a vector to which we apply $DT(\mathbf{u})$, with the obvious notation

$$\tilde{\mathbf{u}} = (\tilde{V}^+, \tilde{V}^-, \tilde{\Psi}^+, \tilde{\Psi}^-)$$

and $\tilde{\Psi}^{\pm}|_{x_2=0} = \tilde{\psi}$.

Then we have

$$DT(\mathbf{u})\tilde{\mathbf{u}} = \begin{pmatrix} \mathscr{L}'(V^+, \Psi^+)(\tilde{V}^+, \tilde{\Psi}^+) \\ \mathscr{L}'(V^-, \Psi^-)(\tilde{V}^-, \tilde{\Psi}^-) \\ \mathscr{E}'(V^+, \Psi^+)(\tilde{V}^+, \tilde{\Psi}^+) \\ \mathscr{E}'(V^-, \Psi^-)(\tilde{V}^-, \tilde{\Psi}^-) \end{pmatrix}$$

where $\mathscr{L}'$ is the derivative of $\mathscr{L}$ and $\mathscr{E}'$ is the derivative of $\mathscr{E}$. Calculating these, we obtain

$$\mathscr{L}'(V, \Psi)(\tilde{V}, \tilde{\Psi}) = \mathbb{L}'(U^a + V, \Phi^a + \Psi)(\tilde{V}, \tilde{\Psi})$$

where $\mathbb{L}'$ is the derivative of $\mathbb{L}$ and is given by

$$\begin{aligned} \mathbb{L}'(U, \Phi)(\tilde{V}, \tilde{\Psi}) = \partial_t \tilde{V} + A_1(U)\partial_{x_1}\tilde{V} + \frac{1}{\partial_{x_2}\Phi}(A_2(U) - \partial_t\Phi - \partial_{x_1}\Phi A_1(U))\partial_{x_2}\tilde{V} \\ +(DA_1(U)\tilde{V})\partial_{x_1}U - \frac{\partial_{x_2}\tilde{\Psi}}{(\partial_{x_2}\Phi)^2}(A_2(U) - \partial_t\Phi - \partial_{x_1}\Phi A_1(U))\partial_{x_2}U \\ +\frac{1}{\partial_{x_2}\Phi}(DA_2(U)\tilde{V} - \partial_t\tilde{\Psi} - \partial_{x_1}\tilde{\Psi}A_1(U) - \partial_{x_1}\Phi DA_1(U)\tilde{V})\partial_{x_2}U. \end{aligned}$$

Also,

$$\mathscr{E}'(V, \Psi)(\tilde{V}, \tilde{\Psi}) = \partial_t\tilde{\Psi} + (v^a + v)\partial_{x_1}\tilde{\Psi} - \tilde{u} + \tilde{v}\partial_{x_1}\Phi^a + \tilde{v}\partial_{x_1}\Psi.$$

6.5.2 The Operator A

We define

$$A(\mathbf{u})\tilde{\mathbf{u}} = \begin{pmatrix} \mathbb{L}_e'(U^{a+} + V^+, \Phi^{a+} + \Psi^+)\check{V}^+ \\ \mathbb{L}_e'(U^{a-} + V^-, \Phi^{a-} + \Psi^-)\check{V}^- \\ \mathscr{E}'(V^+, \Psi^+)(\tilde{V}^+, \tilde{\Psi}^+) \\ \mathscr{E}'(V^-, \Psi^-)(\tilde{V}^-, \tilde{\Psi}^-) \end{pmatrix}$$

where, as in [8], we have introduced the 'good unknown', which we denote by $\check{V}$ instead of $\dot{V}$ to avoid conflict of notation, as

$$\check{V} = \tilde{V} - \frac{\tilde{\Psi}}{\partial_{x_2}(\Phi^a + \Psi)}\partial_{x_2}(U^a + V).$$

The operator $\mathbb{L}'_e$ is defined as

$$\mathbb{L}'_e(U,\Phi)\check{V} = \partial_t \check{V} + A_1(U)\partial_{x_1}\check{V} + \frac{1}{\partial_{x_2}\Phi}(A_2(U) - \partial_t\Phi - \partial_{x_1}\Phi A_1(U))\partial_{x_2}\check{V}$$
$$+(DA_1(U)\check{V})\partial_{x_1}U + \frac{1}{\partial_{x_2}\Phi}(DA_2(U)\check{V} - \partial_{x_1}\Phi DA_1(U)\check{V})\partial_{x_2}U.$$

Note that, with $(U,\Phi) = (U^a + V, \Phi^a + \Psi)$, we have

$$\begin{aligned}\mathbb{L}'(U,\Phi)(\tilde{V},\tilde{\Psi}) - \mathbb{L}'_e(U,\Phi)\check{V} &= \frac{\tilde{\Psi}}{\partial_{x_2}\Phi}\partial_{x_2}(\mathbb{L}(U,\Phi))\\ &= \frac{\tilde{\Psi}}{\partial_{x_2}\Phi}\partial_{x_2}(\mathscr{L}(V,\Psi) - f^a).\end{aligned}$$

6.5.3 The Set $\mathscr{V}$ and the Operator R

We set $m_7 = 1$ and define

$$\mathscr{V} = \{\mathbf{u} \in X^{\infty-1} : \mathscr{E}(V^+,\Psi^+) = 0, \mathscr{E}(V^-,\Psi^-) = 0, ||\mathbf{u}||_{X^3} \leq \delta_5\}$$

where $0 < \delta_5$ is to be chosen sufficiently small.

We define the operator $R : \mathscr{U} \to \mathscr{V}$ by

$$R(\mathbf{u}) = \begin{pmatrix} \rho^+ \\ v^+ \\ \partial_t\Psi^+ + (v^{a+} + v^+)\partial_{x_1}\Psi^+ + v^+\partial_{x_1}\Phi^{a+} \\ \rho^- \\ v^- \\ \partial_t\Psi^- + (v^{a-} + v^-)\partial_{x_1}\Psi^- + v^-\partial_{x_1}\Phi^{a-} \\ \Psi^+ \\ \Psi^- \\ \psi \end{pmatrix}.$$

One can check that indeed $R(\mathbf{u}) \in \mathscr{V}$. In particular, one can see that $||R(\mathbf{u})||_{X^3}$ can be controlled in terms of $||\mathbf{u}||_{X^4} \leq \delta_2$ for $\mathbf{u} \in \mathscr{U}$.

We then calculate

$$R(\mathbf{u}) - \mathbf{u} = \begin{pmatrix} 0 \\ 0 \\ \mathscr{E}(V^+, \Psi^+) \\ 0 \\ 0 \\ \mathscr{E}(V^-, \Psi^-) \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

6.5.4 The Linearised Equations

Given $\mathbf{u} \in \mathscr{V}$ and $\mathbf{g} \in Y^\infty$, we wish to solve the equation

$$A(\mathbf{u})\tilde{\mathbf{u}} = \mathbf{g}$$

for $\tilde{\mathbf{u}} \in X^{\infty - \max\{l_1, m_4 + m_7\}}$. Let us write

$$\mathbf{g} = \begin{pmatrix} f^+ \\ f^- \\ h^+ \\ h^- \end{pmatrix}$$

where $h^+|_{x_2=0} = h^-_{x_2=0} = g$. Then we want to solve the system

$$\begin{pmatrix} \mathbb{L}'_e(U^{a+} + V^+, \Phi^{a+} + \Psi^+)\check{V}^+ \\ \mathbb{L}'_e(U^{a-} + V^-, \Phi^{a-} + \Psi^-)\check{V}^- \\ \mathscr{E}'(V^+, \Psi^+)(\check{V}^+ + \frac{\tilde{\Psi}^+}{\partial_{x_2}(\Phi^{a+} + \Psi^+)}\partial_{x_2}(U^{a+} + V^+), \tilde{\Psi}^+) \\ \mathscr{E}'(V^-, \Psi^-)(\check{V}^- + \frac{\tilde{\Psi}^-}{\partial_{x_2}(\Phi^{a-} + \Psi^-)}\partial_{x_2}(U^{a-} + V^-), \tilde{\Psi}^-) \end{pmatrix} = \begin{pmatrix} f^+ \\ f^- \\ h^+ \\ h^- \end{pmatrix}$$

where in the last two equations we have written $\tilde{V}^+$ in terms of the 'good unknown' $\check{V}$ and $\tilde{\Psi}$. The introduction of the 'good unknown' allows us to split the solution of this system into two steps. First we solve the system

$$\mathbb{L}'_e(U^{a\pm} + V^\pm, \Phi^{a\pm} + \Psi^\pm)\check{V}^\pm = f^\pm \tag{53}$$

with boundary conditions

$$\begin{aligned} &\check{\rho}^+|_{x_2=0} + \frac{\tilde{\psi}}{\partial_{x_2}(\Phi^{a+} + \Psi^+)|_{x_2=0}}\partial_{x_2}(\rho^{a+} + \rho^+)|_{x_2=0} \\ &-\check{\rho}^-|_{x_2=0} - \frac{\tilde{\psi}}{\partial_{x_2}(\Phi^{a-} + \Psi^-)|_{x_2=0}}\partial_{x_2}(\rho^{a-} + \rho^-)|_{x_2=0} = 0 \end{aligned} \tag{54}$$

$$\partial_t \tilde{\psi} + (v^{a\pm} + v^{\pm})|_{x_2=0}\partial_{x_1}\tilde{\psi}$$
$$-(\check{u}^{\pm}|_{x_2=0} + \frac{\tilde{\psi}}{\partial_{x_2}(\Phi^{a\pm} + \Psi^{\pm})|_{x_2=0}}\partial_{x_2}(u^{a\pm} + u^{\pm})|_{x_2=0})$$
$$+(\check{v}^{\pm}|_{x_2=0} + \frac{\tilde{\psi}}{\partial_{x_2}(\Phi^{a\pm} + \Psi^{\pm})|_{x_2=0}}\partial_{x_2}(v^{a\pm} + v^{\pm})|_{x_2=0})\partial_{x_1}(\psi^a + \psi)$$
$$= h^{\pm}|_{x_2=0} \qquad (55)$$

for the unknowns $(\check{V}^{\pm}, \tilde{\psi})$. Note the first boundary condition is $\tilde{\rho}^+|_{x_2=0} - \tilde{\rho}^-|_{x_2=0} = 0$ written in terms of the 'good unknown', and the second boundary condition is a rewriting of
$\mathscr{E}'(V^{\pm}, \Psi^{\pm})(\tilde{V}^{\pm}, \tilde{\Psi}^{\pm})\big|_{x_2=0} = h^{\pm}|_{x_2=0}$ in terms of the 'good unknown', where we replace $\tilde{\Psi}^{\pm}|_{x_2=0}$ with $\tilde{\psi}$.

Secondly, having solved the above system for $(\check{V}^{\pm}, \tilde{\psi})$, we solve the two separate equations

$$\mathscr{E}'(V^{\pm}, \Psi^{\pm})(\check{V}^{\pm} + \frac{\tilde{\Psi}^{\pm}}{\partial_{x_2}(\Phi^{a\pm} + \Psi^{\pm})}\partial_{x_2}(U^{a\pm} + V^{\pm}), \tilde{\Psi}^{\pm}) = h^{\pm} \qquad (56)$$

for $\tilde{\Psi}^{\pm}$. By restricting to $\{x_2 = 0\}$, we see that $\tilde{\Psi}^{\pm}|_{x_2=0}$ satisfy the same equations as $\tilde{\psi}$ given in the boundary conditions above, hence by uniqueness of solutions we have $\tilde{\Psi}^{\pm}|_{x_2=0} = \tilde{\psi}$.

Finally, we can rearrange to obtain $\tilde{V}$ from $\check{V}$ and $\tilde{\Psi}$.

6.6 Solution of the Linearised Equations

Assume $\mathbf{u} \in \mathscr{V}$ and $\mathbf{g} \in Y^{\infty}$. We wish to solve the equation

$$A(\mathbf{u})\tilde{\mathbf{u}} = \mathbf{g}$$

for $\tilde{\mathbf{u}}$, using the steps described above.

The key to the whole iteration scheme is the solution of the linearised problem (53)–(55).

We have the following result, stated in [8]. Assume that the stationary solution satisfies the supersonic condition

$$\overline{v} > \sqrt{2}c(\overline{\rho}).$$

Assume that U, Φ are such that $U - \overline{U}, \Phi - \overline{\Phi} \in H_\gamma^{s+3}(\Omega_T)$ for integer $s \in [3, s_3]$ with

$$\left|\left|(U - \overline{U}, \nabla(\Phi - \overline{\Phi}))\right|\right|_{H_\gamma^5(\Omega_T)} + \left|\left|(U - \overline{U}, \partial_{x_2} U, \nabla(\Phi - \overline{\Phi}))|_{x_2=0}\right|\right|_{H_\gamma^4(\omega_T)} \le \delta_4 \tag{57}$$

for some $\delta_4 > 0$, where $\Phi^+|_{x_2=0} = \Phi_-|_{x_2=0} = \phi$.

Assume also that (U, Φ) satisfy the eikonal equation

$$\partial_t \Phi + v\partial_{x_1}\Phi - u = 0.$$

Assume in addition that the coefficients $(U - \overline{U}, \Phi - \overline{\Phi})$ have fixed compact support – a technical condition which can be achieved by truncating the coefficients without affecting the solution due to the finite speed of propagation of the Euler equations.

Then if δ_4 is sufficiently small, given

$$(f^\pm, g^\pm) \in \mathscr{F}_\gamma^{s+1}(\Omega_T) \times \mathscr{F}_\gamma^{s+1}(\omega_T)$$

we have a unique solution

$$(\check{V}^\pm, \tilde{\psi}) \in \mathscr{F}_\gamma^{s}(\Omega_T) \times \mathscr{F}_\gamma^{s+1}(\omega_T)$$

of (53)–(55), replacing $h^\pm|_{x_2=0}$ with $g^\pm$, provided $\gamma \ge 1$ is sufficiently large depending on s_3. Moreover, the following estimate holds, for some constant $C_s > 0$,

$$\begin{aligned}
&\left|\left|\check{V}\right|\right|_{H_\gamma^s(\Omega_T)} + ||\tilde{\psi}||_{H_\gamma^{s+1}(\omega_T)} \\
&\le C_s(||f||_{H_\gamma^{s+1}(\Omega_T)} + ||g||_{H_\gamma^{s+1}(\omega_T)} \\
&\qquad + (||f||_{H_\gamma^4(\Omega_T)} + ||g||_{H_\gamma^4(\omega_T)}) \left|\left|(U - \overline{U}, \Phi - \overline{\Phi})\right|\right|_{H_\gamma^{s+3}(\Omega_T)}).
\end{aligned}$$

Here, we set $U = U^a + V$, $\Phi = \Phi^a + \Psi$, where $(U, \Psi) \in \mathscr{V}$. Note that the smallness condition (57) holds provided δ_5 and δ_1 are sufficiently small. Also note that the Eikonal equation holds since the approximate solution satisfies the Eikonal equation and by the definition of $\mathscr{V}$. We are given f and h and set $g^\pm = h^\pm|_{x_2=0}$. Unfortunately this method, which is slightly simpler than the one described in [8], results in a further loss of regularity due to taking the trace of h. So in fact given

$$(f^\pm, h^\pm) \in \mathscr{F}_\gamma^{s+1}(\Omega_T) \times \mathscr{F}_\gamma^{s+2}(\omega_T)$$

we have a unique solution

$$(\check{V}^{\pm}, \tilde{\psi}) \in \mathscr{F}^s_\gamma(\Omega_T) \times \mathscr{F}^{s+1}_\gamma(\omega_T)$$

satisfying the estimate

$$\left|\left|\check{V}\right|\right|_{H^s_\gamma(\Omega_T)} + ||\tilde{\psi}||_{H^{s+1}_\gamma(\Omega_T)} \leq C_s(||f||_{H^{s+1}_\gamma(\Omega_T)} + ||h||_{H^{s+2}_\gamma(\omega_T)}$$
$$+(||f||_{H^4_\gamma(\Omega_T)} + ||h||_{H^5_\gamma(\Omega_T)})(\left|\left|(U^a - \overline{U}, \Phi^a - \overline{\Phi})\right|\right|_{H^{s+3}_\gamma(\Omega_T)} + ||(V, \Psi)||_{H^{s+3}_\gamma(\Omega_T)})).$$

Having solved this system, it remains to solve the Eqs. (56) for $\tilde{\Psi}^{\pm}$. Each of these equations is a first order scalar linear equation, so has a unique solution (for smooth enough coefficients and source term). More precisely, assuming that

$$||(U^a + V, \Phi^a + \Psi)||_{H^3_\gamma(\Omega_T)}$$

is small enough (which is guaranteed by taking δ_4 small enough), we have a unique solution

$$\tilde{\Psi} \in \mathscr{F}^s_\gamma(\Omega_T)$$

of (56). Moreover, the following estimate holds, for some constant $C_s > 0$ (which may depend on the bound on $||(U^a + V, \Phi^a + \Psi)||_{H^3_\gamma(\Omega_T)}$),

$$||\tilde{\Psi}||_{H^s_\gamma(\Omega_T)} \leq$$
$$C_s(||h||_{H^s_\gamma(\Omega_T)} + \left|\left|\check{V}\right|\right|_{H^s_\gamma(\Omega_T)} + \left|\left|\check{V}\right|\right|_{H^3_\gamma(\Omega_T)}(\left|\left|\Phi^a - \overline{\Phi}\right|\right|_{H^{s+1}_\gamma(\Omega_T)} + ||\Psi||_{H^{s+1}_\gamma(\Omega_T)})$$
$$+ ||\tilde{\Psi}||_{H^3_\gamma(\Omega_T)}(\left|\left|(U^a - \overline{U}, \Phi^a - \overline{\Phi})\right|\right|_{H^{s+1}_\gamma(\Omega_T)} + ||(V, \Psi)||_{H^{s+1}_\gamma(\Omega_T)})).$$

Taking $s = 3$ and assuming δ_4 is sufficiently small, we obtain

$$||\tilde{\Psi}||_{H^3_\gamma(\Omega_T)} \leq C(||h||_{H^3_\gamma(\Omega_T)} + \left|\left|\check{V}\right|\right|_{H^3_\gamma(\Omega_T)}).$$

Thus

$$||\tilde{\Psi}||_{H^s_\gamma(\Omega_T)} \leq C_s(||h||_{H^s_\gamma(\Omega_T)} + \left|\left|\check{V}\right|\right|_{H^s_\gamma(\Omega_T)}$$
$$+(||h||_{H^3_\gamma(\Omega_T)} + \left|\left|\check{V}\right|\right|_{H^3_\gamma(\Omega_T)})(\left|\left|(U^a - \overline{U}, \Phi^a - \overline{\Phi})\right|\right|_{H^{s+1}_\gamma(\Omega_T)} + ||(V, \Psi)||_{H^{s+1}_\gamma(\Omega_T)})).$$

From the previous estimate for $\check{V}$ with $s = 3$, we obtain

$$\left|\left|\check{V}\right|\right|_{H^3_\gamma(\Omega_T)} \leq C_s(||f||_{H^4_\gamma(\Omega_T)} + ||h||_{H^5_\gamma(\omega_T)})$$

(where the constant will depend on $\left|\left|(U^a - \overline{U}, \Phi^a - \overline{\Phi})\right|\right|_{H^6_\gamma(\Omega_T)} + ||(V, \Psi)||_{H^6_\gamma(\Omega_T)} \leq \delta_1$). Thus we obtain

$$||\tilde{\Psi}||_{H^s_\gamma(\Omega_T)} \leq$$
$$C_s(||f||_{H^{s+1}_\gamma(\Omega_T)} + ||h||_{H^{s+2}_\gamma(\Omega_T)} + (||f||_{H^4_\gamma(\Omega_T)} + ||h||_{H^5_\gamma(\Omega_T)})(1 + ||(V, \Psi)||_{H^{s+3}_\gamma(\Omega_T)}))$$

(where the constant will depend on $\left|\left|(U^a - \overline{U}, \Phi^a - \overline{\Phi})\right|\right|_{H^{s+3}_\gamma(\Omega_T)}$). Combining and writing $\tilde{V}$ in terms of $\check{V}$ and $\tilde{\Psi}$, we obtain

$$||\tilde{V}||_{H^s_\gamma(\Omega_T)} + ||\tilde{\Psi}||_{H^s_\gamma(\Omega_T)} + ||\tilde{\psi}||_{H^{s+1}_\gamma(\omega_T)} \leq$$
$$C_s(||f||_{H^{s+1}_\gamma(\Omega_T)} + ||h||_{H^{s+2}_\gamma(\Omega_T)} + (||f||_{H^4_\gamma(\Omega_T)} + ||h||_{H^5_\gamma(\Omega_T)})(1 + ||(V, \Psi)||_{H^{s+3}_\gamma(\Omega_T)})).$$

Hence, for $\mathbf{u} \in \mathscr{V}$, we have $B(\mathbf{u}) : Y^\infty \to X^{\infty-4}$ and

$$||B(\mathbf{u})\mathbf{g}||_{X^s} \leq C_s(||\mathbf{g}||_{Y^{s+2}} + ||\mathbf{g}||_{Y^2}\,(1 + ||\mathbf{u}||_{X^{s+3}}))$$

for all s such that $s + 4 \in I$. Thus we have $l_1 = 2$ and $m_4 = 3$.

6.7 Estimates of the Operators

6.7.1 Estimate of *R*

Clearly from the definition of R and T, we have

$$||R(\mathbf{u}) - \mathbf{u}||_{X^0} \leq ||T(\mathbf{u}) - T(\mathbf{u_0}) - \mathbf{f}||_{Y^0}\,.$$

Thus $l_2 = 0$.

Also, using Sobolev embedding and that R is a first order differential operator, we have the tame estimate

$$||R(\mathbf{u})||_{X^s} \leq C_s(1 + ||u||_{X^0})(1 + ||u||_{X^{s+1}})$$

for $s \in [0, s_3 - 1]$. Thus $m_8 = 0$, and as we have already stated, $m_7 = 1$.

Now we estimate the commutator

$$\left|\left|R(S_\theta^X \mathbf{u}) - S_\theta^X R(\mathbf{u})\right|\right|_{X^s}$$

for $\mathbf{u} \in \mathscr{U}$.

We have

$$\begin{aligned}
&\mathscr{E}(S_\theta V, S_\theta \Psi) - S_\theta \mathscr{E}(V, \Psi)\\
&= \partial_t(S_\theta \Psi) + (v^a + S_\theta v)\partial_{x_1}(S_\theta \Psi) + (S_\theta v)\partial_{x_1}\Phi^a\\
&- S_\theta \partial_t \Psi - S_\theta(v^a \partial_{x_1}\Psi) - S_\theta(v\partial_{x_1}\Psi) - S_\theta(v\partial_{x_1}\Phi)\\
&= \partial_t(S_\theta \Psi - \Psi) + (\partial_t \Psi - S_\theta \partial_t \Psi)\\
&+ (v^a \partial_{x_1}\Psi - S_\theta(v^a \partial_{x_1}\Psi)) + v^a \partial_{x_1}(S_\theta \Psi - \Psi)\\
&+ (v\partial_{x_1}\Psi - S_\theta(v\partial_{x_1}\Psi)) + v\partial_{x_1}(S_\theta \Psi - \Psi) + (S_\theta v - v)\partial_{x_1} S_\theta \Psi\\
&+ (v\partial_{x_1}\Phi^a - S_\theta(v\partial_{x_1}\Phi^a)) + (S_\theta v - v)\partial_{x_1}\Phi^a.
\end{aligned}$$

Hence, using the property (2) of the smoothing operators and product estimates for Sobolev norms, we have, for $r-3, s-3 \in [0, s_3 - 1]$ with $r \geq s$ and $r' \in [3, s_3]$,

$$\begin{aligned}
&||\mathscr{E}(S_\theta V, S_\theta \Psi) - S_\theta \mathscr{E}(V, \Psi)||_{H^s_\gamma(\Omega_T)}\\
&\leq C_{r,s}(\theta^{s-r}\,||\Psi||_{H^{r+1}_\gamma(\Omega_T)}\\
&+ \theta^{s-r}((||v^a - \overline{v}||_{H^2_\gamma(\Omega_T)} + 1)\,||\Psi||_{H^{r+1}_\gamma(\Omega_T)} + ||v^a - \overline{v}||_{H^r_\gamma(\Omega_T)}\,||\Psi||_{H^3_\gamma(\Omega_T)})\\
&+ (||v^a - \overline{v}||_{H^s_\gamma(\Omega_T)} + 1)\theta^{3-r'}\,||\Psi||_{H^{r'}_\gamma(\Omega_T)}\\
&+ \theta^{s-r}(||v||_{H^2_\gamma(\Omega_T)}\,||\Psi||_{H^{r+1}_\gamma(\Omega_T)} + ||v||_{H^r_\gamma(\Omega_T)}\,||\Psi||_{H^3_\gamma(\Omega_T)})\\
&+ ||v||_{H^s_\gamma(\Omega_T)}\,\theta^{3-r'}\,||\Psi||_{H^{r'}_\gamma(\Omega_T)}\\
&+ \theta^{s-r}\,||v||_{H^r_\gamma(\Omega_T)}\,||\Psi||_{H^3_\gamma(\Omega_T)} + \theta^{2-r'}\,||v||_{H^{r'}_\gamma(\Omega_T)}\,||\Psi||_{H^{s+1}_\gamma(\Omega_T)}\\
&+ \theta^{s-r}(||v||_{H^2_\gamma(\Omega_T)}\,\left|\left|\Phi^a - \overline{\Phi}\right|\right|_{H^{r+1}_\gamma(\Omega_T)} + ||v||_{H^r_\gamma(\Omega_T)}\,\left|\left|\Phi^a - \overline{\Phi}\right|\right|_{H^3_\gamma(\Omega_T)})\\
&+ \theta^{s-r}\,||v||_{H^r_\gamma(\Omega_T)}\,\left|\left|\Phi^a - \overline{\Phi}\right|\right|_{H^3_\gamma(\Omega_T)} + \theta^{2-r'}\,||v||_{H^{r'}_\gamma(\Omega_T)}\,\left|\left|\Phi^a - \overline{\Phi}\right|\right|_{H^{s+1}_\gamma(\Omega_T)})\\
&\leq C_{r,s}(\theta^{s-r}(1 + ||v||_{H^2_\gamma(\Omega_T)} + ||\Psi||_{H^3_\gamma(\Omega_T)})(1 + ||v||_{H^r_\gamma(\Omega_T)} + ||\Psi||_{H^{r+1}_\gamma(\Omega_T)})\\
&+ \theta^{3-r'}(1 + ||v||_{H^s_\gamma(\Omega_T)} + ||\Psi||_{H^{s+1}_\gamma(\Omega_T)})(||v||_{H^{r'}_\gamma(\Omega_T)} + ||\Psi||_{H^{r'}_\gamma(\Omega_T)})).
\end{aligned}$$

Hence, for $r', r, s \in I$ with $r \geq s$,

$$\left|\left|R(S_\theta^X \mathbf{u}) - S_\theta^X R(\mathbf{u})\right|\right|_{X^s}$$
$$\leq C_{r,s}(\theta^{s-r}(1 + ||\mathbf{u}||_{X^0})(1 + ||\mathbf{u}||_{X^{r+1}}) + \theta^{-r'}(1 + ||\mathbf{u}||_{X^{s+1}})(1 + ||\mathbf{u}||_{X^{r'}})).$$

6.7.2 Estimate of the Derivatives of T

Since T is a first order differential operator, that is, $T(\mathbf{u})$ can be written as a smooth bounded function of $\mathbf{u}$ and its first order derivatives for $\mathbf{u} \in \mathscr{U}^1$, we immediately see that $T : \mathscr{U}^1 \to Y^0$ is continuous and it satisfies (20) and (19) with $m_1 \geq 1, m_2 \geq 0, m_3 \geq 1$. Note that we have used the Sobolev embedding $H^3_\gamma(\Omega_T) \subset W^{1,\infty}(\Omega_T)$. We will in fact need to estimate the derivative of A before we fix m_1, m_2, m_3.

6.7.3 Estimate of $A - DT$

We estimate

$$\left|\left|\frac{\tilde{\Psi}}{\partial_{x_2}(\Phi^a + \Psi)}\partial_{x_2}(\mathscr{L}(V,\Psi) - f^a)\right|\right|_{H^s_\gamma(\Omega_T)} \leq$$
$$C_s\left|\left|\tilde{\Psi}\right|\right|_{H^s_\gamma(\Omega_T)}||\mathscr{L}(V,\Psi) - f^a||_{H^3_\gamma(\Omega_T)} + C_s\left|\left|\tilde{\Psi}\right|\right|_{H^2_\gamma(\Omega_T)}||\mathscr{L}(V,\Psi) - f^a||_{H^{s+1}_\gamma(\Omega_T)}$$
$$+ C_s\left|\left|\tilde{\Psi}\right|\right|_{H^2_\gamma(\Omega_T)}||\mathscr{L}(V,\Psi) - f^a||_{H^3_\gamma(\Omega_T)}(1 + \left|\left|\Phi^a + \Psi - \overline{\Phi}\right|\right|_{H^{s+1}_\gamma(\Omega_T)}).$$

Hence

$$||(A(\mathbf{u}) - DT(\mathbf{u}))\tilde{\mathbf{u}}||_{Y^s}$$
$$\leq C_s(||\tilde{\mathbf{u}}||_{X^s}\,||T(\mathbf{u}) - T(\mathbf{u_0}) - \mathbf{f}||_{Y^0} + ||\tilde{\mathbf{u}}||_{X^0}\,||T(\mathbf{u}) - T(\mathbf{u_0}) - \mathbf{f}||_{Y^{s+1}}$$
$$+ ||\tilde{\mathbf{u}}||_{X^0}\,||T(\mathbf{u}) - T(\mathbf{u_0}) - \mathbf{f}||_{Y^0}(1 + ||\mathbf{u}||_{X^{s+1}}))$$

where the constant C_s depends on $\left|\left|\Phi^a - \overline{\Phi}\right|\right|_{H^{s+1}_\gamma(\Omega_T)}$. Thus $m_5 = 0$, $m_6 = 0$, $m_9 = 1, l_3 = 0, l_4 = 1$.

6.7.4 Estimate of the Derivative of A

Note that

$$\mathbb{L}'_e(U,\Phi)\check{V} = \mathbb{L}'(U,\Phi)(\tilde{V},\tilde{\Psi}) - \frac{\tilde{\Psi}}{\partial_{x_2}\Phi}\partial_{x_2}(\mathbb{L}(U,\Phi))$$

where $(U, \Phi) = (U^a + V, \Phi^a + \Psi)$. The first term is a component of DT. For fixed $(\tilde{U}, \tilde{\Psi})$, the second term is $\tilde{\Psi}$ multiplied by a differential operator of order 2. Hence

$$\begin{aligned}
&||DA(\mathbf{u})_{\tilde{\mathbf{u}}}\mathbf{h}||_{Y^s} \\
&\leq C_s(||\mathbf{h}||_{X^{s+1}}\,||\tilde{\mathbf{u}}||_{X^0} + ||\mathbf{h}||_{X^0}\,||\tilde{\mathbf{u}}||_{X^{s+1}} + ||\mathbf{h}||_{X^0}\,||\tilde{\mathbf{u}}||_{X^0}\,(1 + ||\mathbf{u}||_{X^{s+1}}) \\
&+ ||\tilde{\mathbf{u}}||_{X^0}\,(||\mathbf{h}||_{X^{s+2}} + ||\mathbf{h}||_{X^1}\,(1 + ||\mathbf{u}||_{X^{s+2}})) + ||\tilde{\mathbf{u}}||_{X^s}\,||\mathbf{h}||_{X^1}) \\
&\leq C_s(||\mathbf{h}||_{X^{s+2}}\,||\tilde{\mathbf{u}}||_{X^1} + ||\mathbf{h}||_{X^1}\,||\tilde{\mathbf{u}}||_{X^{s+2}} + ||\mathbf{h}||_{X^1}\,||\tilde{\mathbf{u}}||_{X^1}\,(1 + ||\mathbf{u}||_{X^{s+2}})).
\end{aligned}$$

Thus we fix $m_1 = 2, m_2 = 1, m_3 = 2$.

6.8 Conclusion

We have seen that the hypotheses of the theorem are satisfied with $m_0 = 4, m_1 = 2, m_2 = 1, m_3 = 2, m_4 = 3, m_5 = 0, m_6 = 0, m_7 = 1, m_8 = 0, m_9 = 1, l_1 = 2, l_2 = 0, l_3 = 0, l_4 = 1$. Hence we may take $r_0 = 6$. Note that in the proof we required $s_1 > r_0 + 1, s_1 \geq r_0 + \max\{m_1, m_3\} + l_1$ and $M(s_1 - \max\{m_1 + l_4, m_3 + l_4, m_5, m_9\}) \geq 0$ (with slope 1 which is satisfied for $s_1 > r_0 + 1$ automatically). One can check that $M(s) = s - 8$ hence we require $s_1 - 3 \geq 8$, so $s_1 \geq 11$. Now we require $s_1 + \max\{l_1, m_4 + m_7\} \in I$, hence $s_3 \geq 11 + 4 = 15$, and thus $s_4 \geq 17$ will do.

Thus we conclude that if we are given the approximate solution $(U^{a+}, U^{a-}, \Phi^{a+}, \Phi^{a-})$ with $U^a - \overline{U}, \Phi^a - \overline{\Phi} \in H^{20}(\Omega_T)$ which satisfies the conditions described above, with

$$\left|\left|U^a - \overline{U}\right|\right|_{H^7(\Omega_T)} + \left|\left|\Phi^a - \overline{\Phi}\right|\right|_{H^7(\Omega_T)}$$

sufficiently small, then we have a unique solution $(V^+, V^-, \Psi^+, \Psi^-) \in \mathscr{F}^7_\gamma(\Omega_T)$ to the following equations (for both $^+$ and $^-$ components),

$$\mathbb{L}(U^a + V, \Phi^a + \Psi) = 0$$
$$\partial_t(\Phi^a + \Psi) + (v^a + v)\partial_{x_1}(\Phi^a + \Psi) - (u^a + u) = 0.$$

In fact, since $f^a \in Y^{s_2-2}$, where $s_2 = 12 \leq s_3 - 3$, we may use the last part of the theorem to conclude that $(V^+, V^-, \Psi^+, \Psi^-) \in \mathscr{F}^{11}_\gamma(\Omega_T)$.

7 Further Applications and Open Problems

There are several other situations involving characteristic discontinuities for the Euler equations or the equations of ideal magnetohydrodynamics where it may be possible to obtain a tame estimate for the linearised equations, and thus apply the above Nash-Moser iteration scheme. In these contexts a characteristic discontinuity is a surface of discontinuity in the fluid across which the Rankine-Hugoniot jump conditions are satisfied with zero mass transfer. The first step is usually to perform a normal modes analysis by linearising about a background state (constant either side of a plane across which the Rankine-Hugoniot jump conditions are satisfied) and to determine criteria which rule out exponentially growing solutions. The aim is then to show short-time existence of solutions with the same structure as the background state (that is, smooth either side of a surface of discontinuity across which the Rankine-Hugoniot jump conditions are satisfied) where the initial data is a small perturbation of the background state, under the assumption that the background state satisfies the stability criteria. We call this structural stability.

One obvious open problem is to extend the above result by Coulombel and Secchi in [8] on the 2D isentropic Euler equations to the 2D full Euler equations. Miles showed in [17] that the stability criterion on the background solution $\bar{U}^{\pm}$ (using notation as above) in this case is

$$|[\bar{u}]| > ((\bar{c}^{+})^{\frac{2}{3}} + (\bar{c}^{-})^{\frac{2}{3}})^{\frac{3}{2}}$$

(where $[u] = u^{+} - u^{-}$) under the simplifying assumption

$$\bar{\rho}^{+}(\bar{c}^{+})^2 = \bar{\rho}^{-}(\bar{c}^{-})^2.$$

The main difficulty is to solve, and to deduce a tame estimate for, the linearised equations, assuming this stability criterion, after which we would expect the application of Nash-Moser iteration to be similar. In fact Morando and Trebeschi have obtained an L^2 estimate with derivative loss for the linearised equations under this stability criterion – see [18]. We note that vortex sheets in 3D Euler are always unstable according to normal modes analysis – see Miles and Fejer [13].

A modification of the Nash-Moser scheme similar to the one above has been used successfully by Chen and Wang in [5] and [6] for current-vortex sheets in ideal compressible magnetohydrodynamics under the assumption that the jump in the non-parallel component of the magnetic field dominates the jump in tangential velocity. This stability criterion was first found by Trakhinin by forming a new symmetric form of the equations – see [24] – although it is almost certainly stricter than necessary. One of the key observations made by Chen and Wang is that, using this new symmetric form of the equations, the linearised problem for current-vortex sheets is endowed with a well-structured decoupled formulation into a standard initial-boundary value problem for a symmetric hyperbolic system and a separate scalar PDE for the front. Chen and Wang then modify the iteration scheme to reconstruct the extensions of the front, $\Psi^{\pm}$, with $\Psi^{+} = \Psi^{-}$ on the boundary, which is why their scheme does not exactly fit into the above framework, but would require a small modification. In fact Trakhinin in [25] obtained the same result on current-vortex sheets, but instead of modifying significantly the iteration scheme of

Coulombel and Secchi, he solved the original linearised equations having used his new symmetric form only to help with the treatment of the linearised equations, which results in his approach being longer, although it should fit into the above framework. The normal modes analysis to determine the expected weakest possible stability criteria for current-vortex sheets in compressible magnetohydrodynamics leads to high order algebraic equations which seem impossible to solve analytically, and is detailed by Fejer in [12], where some special cases are considered.

The stability criterion for current-vortex sheets in incompressible magnetohydrodynamics is easier to determine – see e.g. Axford [4]. In 2D, the condition is

$$2(\left|\bar{H}_+\right|^2 + \left|\bar{H}_-\right|^2) > |[\bar{u}]|^2 .$$

In 3D, there are two conditions

$$\begin{aligned} 2(\left|\bar{H}_+\right|^2 + \left|\bar{H}_-\right|^2) &> |[\bar{u}]|^2 \\ 2\left|\bar{H}_+ \times \bar{H}_-\right|^2 &> \left|\bar{H}_+ \times [\bar{u}]\right|^2 + \left|\bar{H}_- \times [\bar{u}]\right|^2 \end{aligned}$$

although in fact the first follows from the second under the additional assumption $\bar{H}_+ \times \bar{H}_- \neq 0$.

Given these stability criteria, one would hope to be able to obtain a tame estimate for the linearised equations and then use Nash-Moser iteration as above to prove nonlinear structural stability of incompressible current-vortex sheets. In [19], Morando, Trakhinin and Trebeschi obtain an energy estimate for the linearised 3D equations under the above stability criteria. Also, using a different approach, Coulombel et al. [9] have derived a priori high order energy estimates directly for the nonlinear equations in 3D, using the incompressible version of Trakhinin's stability criterion – see Coulombel et al. [9]. However, the full problem of nonlinear structural stability of incompressible current-vortex sheets is still open.

The case of current-vortex sheets in 2D isentropic magnetohydrodynamics, where the magnetic fields are parallel on either side of the discontinuity, has been considered by Wang and Yu in [26]. They obtain a low order energy estimate for the linearised equations with loss of derivatives, under some quite restrictive assumptions to simplify the algebra and make the treatment similar to that of 2D isentropic Euler.

Another open problem is the case of current-entropy waves for the full magnetohydrodynamics equations, where the normal component of the magnetic field is no longer zero on the surface of discontinuity, but the velocity and magnetic field are continuous, with only the pressure, entropy and density experiencing a jump. There are strong indications that such waves ought to be stable under certain conditions, but the normal modes analysis again results in high-order algebraic equations which are difficult to study analytically.

Acknowledgements My research is supported by a UK EPSRC grant to the Department of Mathematics at Oxford University. I would like to thank my supervisor, Gui-Qiang G. Chen, for helpful discussions on this problem.

References

1. R.A. Adams, J.J.F. Fournier, *Sobolev Spaces*. Volume 140 of Pure and Applied Mathematics (Amsterdam), 2nd edn. (Elsevier/Academic, Amsterdam, 2003)
2. S. Alinhac, Existence d'ondes de raréfaction pour des systèmes quasi-linéaires hyperboliques multidimensionnels. Commun. Partial Differ. Equ. **14**(2), 173–230 (1989)
3. S. Alinhac, P. Gérard, *Pseudo-differential Operators and the Nash-Moser Theorem* (American Mathematical Society, Providence, 2007)
4. W.I. Axford, The stability of plane current-vortex sheets. Q. J. Mech. Appl. Math. **13**(3), 314–324 (1960)
5. G.-Q. Chen, Y.-G. Wang, Existence and stability of compressible current-vortex sheets in three-dimensional magnetohydrodynamics. Arch. Ration. Mech. Anal. **187**(3), 369–408 (2008)
6. G.-Q. Chen, Y.-G. Wang, Characteristic discontinuities and free boundary problems for hyperbolic conservation laws, in *Nonlinear Partial Differential Equations – The Abel Symposium 2010*, Oslo, ed. by H. Holden, K. Karlsen. Volume 7 of Abel Symposia (Springer, 2012)
7. J.-F. Coulombel, P. Secchi, The stability of compressible vortex sheets in two space dimensions. Indiana Univ. Math. J. **53**(4), 941–1012 (2004)
8. J.-F. Coulombel, P. Secchi, Nonlinear compressible vortex sheets in two space dimensions. Ann. Sci. Éc. Norm. Supér. (4) **41**(1), 85–139 (2008)
9. J.-F. Coulombel, A. Morando, P. Secchi, P. Trebeschi, A priori estimates for 3D incompressible current-vortex sheets. Commun. Math. Phys. **311**(1), 247–275 (2012)
10. I. Ekeland, An inverse function theorem in Fréchet spaces. Ann. Inst. H. Poincaré Anal. Non Linéaire **28**(1), 91–105 (2011)
11. L.C. Evans, *Partial Differential Equations*. Volume 19 of Graduate Studies in Mathematics, 2nd edn. (American Mathematical Society, Providence, 2010)
12. J.A. Fejer, Hydromagnetic stability at a fluid velocity discontinuity between compressible fluids. Phys. Fluids **7**, 499–503 (1964)
13. J.A. Fejer, J.W. Miles, On the stability of a plane vortex sheet with respect to three-dimensional disturbances. J. Fluid Mech. **15**, 335–336 (1963)
14. R.S. Hamilton, The inverse function theorem of Nash and Moser. Bull. Am. Math. Soc. (N.S.) **7**(1), 65–222 (1982)
15. L. Hörmander, The boundary problems of physical geodesy. Arch. Ration. Mech. Anal. **62**(1), 1–52 (1976)
16. A. Majda, *Compressible Fluid Flow and Systems of Conservation Laws in Several Space Variables* (Springer, New York, 1984)
17. J.W. Miles, On the disturbed motion of a plane vortex sheet. J. Fluid Mech. **4**, 538–552 (1958)
18. A. Morando, P. Trebeschi, Two-dimensional vortex sheets for the nonisentropic Euler equations: linear stability. J. Hyperbolic Differ. Equ. **5**(3), 487–518 (2008)
19. A. Morando, Y. Trakhinin, P. Trebeschi, Stability of incompressible current-vortex sheets. J. Math. Anal. Appl. **347**(2), 502–520 (2008)
20. J. Moser, A new technique for the construction of solutions of nonlinear differential equations. Proc. Natl. Acad. Sci. USA **47**, 1824–1831 (1961)
21. J. Nash, The imbedding problem for Riemannian manifolds. Ann. Math. (2) **63**, 20–63 (1956)
22. L. Nirenberg, *Topics in Nonlinear Functional Analysis*. Volume 6 of Courant Lecture Notes in Mathematics (New York University Courant Institute of Mathematical Sciences, New York, 2001). Chapter 6 by E. Zehnder, Notes by R. A. Artino, Revised reprint of the 1974 original.

23. J. Schwartz, *Nonlinear Functional Analysis* (Gordon and Breach, New York, 1969)
24. Y. Trakhinin, Existence of compressible current-vortex sheets: variable coefficients linear analysis. Arch. Ration. Mech. Anal. **177**(3), 331–366 (2005)
25. Y. Trakhinin, The existence of current-vortex sheets in ideal compressible magnetohydrodynamics. Arch. Ration. Mech. Anal. **191**(2), 245–310 (2009)
26. Y.-G. Wang, F. Yu, Stabilization effect of magnetic fields on two-dimensional compressible current-vortex sheets. Arch. Ration. Mech. Anal. (2013). doi:10.1007/s00205-012-0601-9