Around the Nash-Moser Theorem

Ruoyu Wang *

Assessor: Clément Mouhot†

May 1, 2017

Declaration: I declare that this essay is work done as part of the Part III Examination. I have read and understood the Statement on Plagiarism for Part III and Graduate Courses issued by the Faculty of Mathematics, and have abided by it. This essay is the result of my own work, and except where explicitly stated otherwise, only includes material undertaken since the publication of the list of essay titles, and includes nothing which was performed in collaboration. No part of this essay has been submitted, or is concurrently being submitted, for any degree, diploma or similar qualification at any university or similar institution.

* DPMMS, CMS, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK;
† DPMMS, CMS, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK.
Contents

1 Introduction

2 Pseudodifferential Operators
   2.1 Symbols of Type 1, 0
   2.2 Asymptotics and Classical Symbols
   2.3 Operators, Kernels, Adjoints, Quantisation
   2.4 Symbolic Algebra of Operators
   2.5 Operators on Sobolev Spaces
   2.6 Ellipticity and Parametrices

3 Nonlinear Dyadic Analysis
   3.1 Littlewood-Paley Decomposition
   3.2 Characterisation of Sobolev and Hölder Spaces
   3.3 Operators on Hölder-Zygmund Spaces
   3.4 Dyadic Analysis of Products

4 Hölder Scale Nash-Moser Theorem
   4.1 Preliminaries
   4.2 Statement of Theorem
   4.3 Iteration Scheme
   4.4 Iterative Hypotheses
   4.5 Existence and Regularity

5 Isometric Embedding Theorem
   5.1 Isometric Embedding Problem
   5.2 Nash-Moser Approach
   5.3 Günther’s Approach

6 Abstract Nash-Moser Theorem
   6.1 Scale of Banach Spaces
   6.2 Banach Scale Nash-Moser Theorem

7 Epilogue
   7.1 Review
   7.2 Acknowledgement
   7.3 Bibliography
1 Introduction

In this essay our main target is to build up an implicit function theorem which could be applied to tamed mappings between $C^\infty(M)$ spaces for some compact manifold $M$. Note in this case $C^\infty(M)$ is only a Fréchet space and hence the standard implicit function theorem cannot be invoked. Moreover tameness of the mapping allows a fixed order loss in derivatives, that is, if we utilise a standard iteration scheme by assuming information on finite order derivatives, all information will be lost in finite time, which makes a standard iteration impracticable. Those two facts remark the necessity and difficulty of establishing a new implicit function theorem.

This new implicit function theorem, nowadays known as the Nash-Moser theorem, was firstly devised by Nash [19] in order to prove the smooth case of his famous isometric embedding theorem. Further refinements, improvements and new versions were attributed to, not exhaustively listed, Hörmander [10] [11] [12] [14], Zehnder [24], Mather, Sergeraert, Tougeron, Hamilton [8], Hermann, Craig, Dacorogna, Bourgain, Berti and Bolle, Ekeland and Sére [5] [6], and lately by Villani and Mouhot.

The Nash-Moser theorem is a significant result in analysis of nonlinear PDEs, for it allows iteration schemes with a loss of regularity in each step to be run to achieve an existence construction. It is, not exhaustively listed, used in the cases of isometric embedding problem, global existence theorems for hyperbolic second order equations, small divisor problems, and various other control and Cauchy problems.

In this essay, we are going through major ideas in the following order: Section 2 is a brief introduction to pseudodifferential operators; Section 3 is intended to introduce useful dyadic results via Littlewood-Paley theory; in Section 4 we detailed a very descriptive proof of the Hölder space based Nash-Moser theorem; Section 5 specialises in tackling the isometric embedding theorem, respectively by Nash’s original approach, which is an immediate consequence of the Nash-Moser theorem, and by Günther’s approach via ellipticity; in Section 6 we prove an abstract version of Nash-Moser theorem by introducing modification of Banach spaces.
2 Pseudodifferential Operators

In this chapter, we develop the basic calculus of pseudodifferential operators.

2.1 Symbols of Type 1, 0

Pseudodifferential operators are generalisations of classical differential operators. By treating classical differential operators as a multiplication operator by multiplying a polynomial symbol on the Fourier side, we aim to define classes of symbols larger than polynomials.

Definition 2.1 (Space of symbols). Given $m \in \mathbb{R}$, by $S^m(\mathbb{R}^N \times \mathbb{R}^N)$ we denote all $a(x, \xi) \in C^\infty(\mathbb{R}^N \times \mathbb{R}^N)$ such that for any multi-indices $\alpha$ and $\beta$ we have the global estimate

$$|\partial_\xi^\alpha \partial_x^\beta a(x, \xi)| \leq C_{\alpha, \beta} \langle \xi \rangle^{m - |\alpha|}, \quad \forall x, \xi \in \mathbb{R}^N. \quad (1)$$

Here $\langle \xi \rangle = (1 + |\xi|^2)^{1/2}$. And denote $S^{-\infty} = \cap_m S^m$. We call $a \in S^m$ a symbol of order $m$.

Remark. (i) We don’t usually distinguish between estimate by $\langle \xi \rangle^m$ and $(1 + |\xi|)^m$, for it is easy to verify

$$\frac{1}{\sqrt{2}} (1 + |\xi|) \leq (1 + |\xi|^2)^{1/2} \leq (1 + |\xi|).$$

They both denote growth control of a $m$-th order polynomial if $m$ is an integer.

(ii) Symbols of order $m \in \mathbb{N}$ have growth control by an order $m$ polynomial, and they resemble polynomial symbols of classical linear differential operators. The most important part of this definition is that it gives a growth control when $|\xi|$ is large, and it is exactly the case when $|\xi|$ is large is worth consideration, because symbols are always bounded bounded inside any compactum.

Example. (i) Constant coefficient linear differential operators, $\sum_{|\alpha| \leq m} a_{\alpha} D^\alpha$ has differential symbol as $\sum_{|\alpha| \leq m} a_{\alpha} \xi^\alpha$, where $D = -i\partial$. For $\beta \leq \alpha$, its $\beta$-derivatives in $\xi$ are bounded by $C_\beta \langle \xi \rangle^{m - |\beta|}$. Hence the symbols of $m$-th order constant coefficient linear differential operators are in $S^m$.

(ii) For a more general linear differential operators $\sum_{|\alpha| \leq m} a_{\alpha}(x) D^\alpha$, if its coefficients $a_{\alpha}(x) \in C^\infty$ are bounded in derivatives in $x$ of all order, then its differential symbol can have a global control in derivatives in $\xi$, hence in $S^m$.

(iii) Given $a(\xi) \in C^\infty(\mathbb{R}^N \setminus 0)$, which is positive homogeneous of order $m$. It is easy to deduce that $a^{(k)}(\xi)$, the $k$-th order derivative of $a(\xi)$ is homogeneous of order $m - k$. Now set a cutoff function $\chi(\xi) \in C^\infty_c(\mathbb{R}^N)$ and is 1 in a neighbourhood of 0; and set $\hat{\alpha}(\xi) = (1 - \chi(\xi))a(\xi)$, which is smooth on whole $\mathbb{R}^N$ and excludes singularity of $a$ at 0. Hence we have the estimate

$$|\partial_\xi^\alpha \hat{\alpha}(\xi)| \leq |p|^{m - |\alpha|} \langle \xi \rangle^{m - |\alpha|},$$

for a fixed $p$ outside support of $\chi$, when $|\xi| > 1$. Hence $\hat{\alpha} \in S^m$. 


(iv) Given \( a(\xi) \in \mathcal{S}(\mathbb{R}^N) \). Because at infinity, Schwartz class function \( a \) and all its derivatives are vanishing faster than polynomials of order \(-m\) for all \( m \in \mathbb{N} \), hence the estimate \([1]\) holds for all \( m \). So \( a \in S^{-\infty} \).

We now establish some properties of the family of \( S^m \):

**Proposition 2.2** (Algebra of symbols). (i) If \( a \in S^m \), then \( \partial_x^\varrho \partial_\xi^\theta a \in S^{m-|\alpha|} \).
(ii) \( S^s \subset S^t \) for \( s < t \).
(iii) If \( a \in S^{m_1}, b \in S^{m_2} \) then \( a + \lambda b \in S^{\max(m_1,m_2)} \) and \( ab \in S^{m_1+m_2} \) for any scalar \( \lambda \in \mathbb{R} \).

**Proof.** (i) Clear via \( S \).
(ii) Proposition 2.2 (Algebra of symbols).
(iii) Without loss of generality assume \( m_1 > m_2 \). By (ii) we have \( b \in S^{m_1} \) and hence
\[
|\partial_x^\varrho \partial_\xi^\theta a(x,\xi)| \leq C_{\alpha,\beta+\gamma} \langle \xi \rangle^{m-|\alpha|} \leq \tilde{C}_{\gamma,\theta} \langle \xi \rangle^{m-|\alpha|-|\gamma|}.
\]

For the second part we have
\[
|\partial_x^\varrho \partial_\xi^\theta (ab)| \leq \sum_{\gamma \leq \alpha} \sum_{\theta \leq \beta} \binom{\alpha}{\gamma} \binom{\beta}{\theta} (\partial_x^\varrho \partial_\xi^\theta a) (\partial_x^{(\alpha-\gamma)} \partial_\xi^{(\beta-\theta)} b)
= \sum_{\gamma \leq \alpha} \sum_{\theta \leq \beta} C_{\alpha,\beta,\gamma,\theta} \langle \xi \rangle^{m_1-|\gamma|} \langle \xi \rangle^{m_2-|\alpha|+|\gamma|}
= C_{\alpha,\beta} \langle \xi \rangle^{m_1+m_2-|\alpha|}
\]
as required. \(\square\)

We also quote a lemma for an easier calculus of symbols of order 0:

**Lemma 2.3** (Function of symbols). Given symbol \( a_1, \ldots, a_k \in S^0 \), a smooth function \( F \in C^\infty(\mathbb{C}^k) \), then \( F(a_1, \ldots, a_k) \in S^0 \).

**Proof.** Since we also know the real and imaginary parts of \( a_j \) are in \( S^0 \), we reduce the problem without loss of generality to the case of real symbols \( a_j \) and \( F \in C^\infty(\mathbb{R}^k) \). Consider the formulae of differentiation:
\[
\frac{\partial}{\partial x_j} F(a) = \sum_q \frac{\partial F}{\partial a_q} \partial_{x_j} a_q, \quad \frac{\partial}{\partial \xi_j} F(a) = \sum_q \frac{\partial F}{\partial a_q} \partial_{\xi_j} a_q.
\]
Prove by induction on \( p \) that estimate \([1]\) is valid for any \( \alpha + \beta \leq p \). Case \( p = 0 \) is clear, and for general \( |\alpha| + |\beta| = p + 1 \), pick a component \( j \) of \( \alpha \) or \( \beta \) which is positive, then applying the Leibniz’s formula to \( \frac{\partial}{\partial x_j} F(a) \) or \( \frac{\partial}{\partial \xi_j} F(a) \) reduces \( \partial_x^\varrho \partial_\xi^\theta F(a) \) to a finite sum of multi-indices-sum \( p \) estimate on \( \partial_{x_j} a_q \) or \( \partial_{\xi_j} a_q \), which is obvious to establish the required estimate. \(\square\)
Remark (Topological structures of $S^m$). $S^m$ is a Fréchet space under the semi-norms

$$
\|a(x, \xi)\|_{S^m} = \sup_{\mathbb{R}^N \times \mathbb{R}^N} \{ \langle \xi \rangle^{-m+|\alpha|} |\partial^\alpha_x \partial^\beta_\xi a(x, \xi)| \}.
$$

And we say $a_n \to a$ in this topology if $\|a_n - a\|_{S^m} \to 0$ for all multi-indices $\alpha, \beta$.

Lemma 2.4 (Approximation of symbols). Given symbol $a \in S^0(\mathbb{R}^N \times \mathbb{R}^N)$, have the zoom function in frequencies $a_\varepsilon = a(x, \varepsilon \xi)$. Then $a_\varepsilon \in S^0$ and when $\varepsilon \to 0$, $a_\varepsilon \to a_0$ in $S^m$ for any $m > 0$.

Proof. Firstly we show that for $s < t$ the injection $S^s \subset S^t$ is continuous, that is, if $a_n \to 0$ in $S^s$ then so is $a_n$ in $S^t$. Clear we have

$$
\|a(x, \xi)\|_{S^t} = \sup_{\mathbb{R}^N \times \mathbb{R}^N} \{ \langle \xi \rangle^{-t+|\alpha|} |\partial^\alpha_x \partial^\beta_\xi a(x, \xi)| \} 
$$

$$
\leq \sup_{\mathbb{R}^N \times \mathbb{R}^N} \{ \langle \xi \rangle^{-s+|\alpha|} |\partial^\alpha_x \partial^\beta_\xi a(x, \xi)| \} = \|a(x, \xi)\|_{S^s}
$$

for $\langle \xi \rangle^{-(t-s)} \leq 1$. Hence if we can show for $0 < m \leq 1$ we have $a_\varepsilon \to a_0$ in $S^m$, the convergence for all positive $m$ is evident.

Now consider $0 < m \leq 1$, and $\varepsilon < 1$ small. For $\alpha > 0$ we have immediately $\partial^\alpha_x \partial^\beta_\xi a_0 = 0$ and the estimate

$$
|\partial^\alpha_x \partial^\beta_\xi (a(x, \varepsilon \xi)) = \varepsilon^{|\alpha|} |\partial^\alpha_x \partial^\beta_\xi a(x, \varepsilon \xi)| \lesssim \varepsilon^{|\alpha|} (1 + \varepsilon |\xi|)^{-|\alpha|}.
$$

And hence we see

$$
\|a_\varepsilon - a_0\|_{S^m} \leq \sup_{\mathbb{R}^N \times \mathbb{R}^N} \{ (1 + |\xi|)^{-m+|\alpha|} \varepsilon^{|\alpha|} (1 + \varepsilon |\xi|)^{-|\alpha|} \}
$$

$$
= \varepsilon^m \sup_{\mathbb{R}^N \times \mathbb{R}^N} \{ (1 + |\xi|)^{-m+|\alpha|} \varepsilon^{|\alpha|} (1 + \varepsilon |\xi|)^{-|\alpha|+m} \} \leq \varepsilon^m
$$

for the two suprema are both no greater than 1. For $\alpha = 0$ we see $a_0 \neq 0$, but instead we have

$$
|\partial^\beta_x (a_\varepsilon - a_0)| = \varepsilon \int^1_0 \nabla \xi \partial^\alpha_x a(x, t\varepsilon \xi) \, dt 
$$

$$
\leq C \int^\varepsilon_0 \frac{d\tau}{1 + \tau} = C \log(1 + \varepsilon |\xi|) \leq C_m \varepsilon^m |\xi|^m,
$$

by $\log(1 + x) \leq C_m x^m$ for $x \geq 0$ and $m > 0$. Then

$$
\|a_\varepsilon - a_0\|_{S^m} = \sup_{\mathbb{R}^N \times \mathbb{R}^N} \{ (1 + |\xi|)^{-m} C_m \varepsilon^m |\xi|^m \} \leq C_m \varepsilon^m
$$

for $(1 + |\xi|)^{-m} |\xi|^m \leq 1$. Hence we obtain the control for all $\alpha, \beta$:

$$
\|a_\varepsilon - a_0\|_{S^m} \lesssim \varepsilon^m \to 0
$$

as $\varepsilon \to 0$, so for $0 < m \leq 1$ we have $a_\varepsilon \to a_0$ in $S^m$, and hence for all positive $m$. \qed
2.2 Asymptotics and Classical Symbols

The definition of symbol only contains its information on its highest order term, that is its highest order of growth in frequencies at infinity. For symbols of classical linear differential operators, \( \sum_{|\alpha| \leq m} a_\alpha \xi^\alpha \), we have the clear layering of symbols of order 0, 1, \ldots, \( m \), each part of which is respectively homogeneous of 0, 1, \ldots, \( m \). We can further extract the principal symbol, \( \sum_{|\alpha| = m} a_\alpha \xi^\alpha \), non-trivial zeroes of which determine ellipticity. We want to restore this layering feature on our pseudodifferential symbols in this section, via introducing asymptotics and classical pseudodifferential symbols.

**Definition 2.5 (Asymptotics).** Given a sequence of symbols \( a_j \in S^{m_j} \) with \( m_j \searrow -\infty \).

We say \( \sum a_j \) is the asymptotics of some function \( a(x, \xi) \), as of behaviour when \( |\xi| \to \infty \), written \( a \sim \sum a_j \)

if for any \( k \geq 0 \) we have

\[
a - \sum_{j=0}^{k} a_j \in S^{m_k+1}.
\]

**Remark.** For a symbol \( a(x, \xi) \in S^m(\mathbb{R}^N \times \mathbb{R}^N) \) we see it is asymptotically itself, the fact of which does not validate why we need to introduce an asymptotic expansion. Actually, symbols \( a(x, \xi) \) which only depend on location and frequency, are only one type of a more general class of symbols, \( a(x, y, \xi) \in S^m(\mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N) \), which aside from being a multiplication operator on the Fourier side, also perturbs the Fourier transform to some degree. For such kind of more general symbol, if it is properly supported, we can asymptotically associate \( \sigma_a(x, \xi) \in S^m(\mathbb{R}^N \times \mathbb{R}^N) \) with the general symbol \( a(x, y, \xi) \) by formula

\[
\sigma_a(x, \xi) \sim \sum_{\alpha} \frac{1}{\alpha} \partial_\xi^\alpha D_y^\alpha a(x, y, \xi)|_{x=y}.
\]

This process is called the left quantisation. Furthermore we can express adjoints and compositions of the pseudodifferential operators introduced by symbols via their asymptotics, which we will cover later.

We firstly see that given any asymptotics we can find an equivalent pseudodifferential symbol, unique up to class \( S^{-\infty} \).

**Proposition 2.6 (Symbol determined by asymptotics).** Given \( a_j \in S^{m_j} \) with \( m_j \searrow -\infty \), there exists \( a \in S^{m_0} \) such that \( a \sim \sum a_j \), unique up to class \( S^{-\infty} \). Furthermore the choice can be made such that \( \text{supp} \ a \subset \bigcup \text{supp} \ a_j \).

**Proof.** The uniqueness part is obvious. If \( a \) and \( \tilde{a} \) are both asymptotically equivalent to \( \sum a_j \), then for any \( k > 0 \) we have

\[
a - \tilde{a} = \left( a - \sum_{j=0}^{k} a_j \right) - \left( \tilde{a} - \sum_{j=0}^{k} a_j \right) \in S^{m_k+1}.
\]
as the sum of two \( S^{m_k+1} \) symbols. As \( m_k \searrow -\infty \) we have \( a - \tilde{a} \in S^{-\infty} \).

For the existence part, firstly set up a standard cutoff function \( \chi(\xi) \in C_0^\infty(\mathbb{R}^N) \) which is 1 near 0. Now construct

\[
a(x, \xi) = \sum_{j=0}^\infty \tilde{a}_j(x, \xi) = \sum_{j=0}^\infty (1 - \chi(\varepsilon_j \xi)) a_j(x, \xi),
\]
in which \( \varepsilon_j \) are chosen smartly such that they vanish to zero fast enough for some estimate on \( \tilde{a}_j \) to hold. In fact, we have

\[
|\partial_\xi^\alpha \partial_x^\beta \tilde{a}_j| \leq \sum_{\gamma \leq \alpha} \sum_{\theta \leq \beta} \left( \begin{array}{c} \alpha \\ \gamma \end{array} \right) \left( \begin{array}{c} \beta \\ \theta \end{array} \right) |(\partial_\xi^\alpha \partial_x^\beta (1 - \chi(\varepsilon_j \xi)))(\partial_\xi^{(\alpha-\gamma)} \partial_x^{(\beta-\theta)} a_j)|
\]

\[
\leq \sum_{\gamma \leq \alpha} \sum_{\theta \leq \beta} C_{\alpha,\beta,\gamma,\theta} \langle \xi \rangle^{\gamma-1} |\partial_\xi^\gamma \partial_x^\theta (1 - \chi(\varepsilon_j \xi))| \langle \xi \rangle^{1+m_j-|\alpha|}.
\]

Consider that \( 1 - \chi(\varepsilon \xi) \) is in fact a zoom function of \( 1 - \chi(\xi) \in S^0 \), we have \( 1 - \chi(\varepsilon \xi) \to 0 \) in \( S^1 \) as \( \varepsilon \to 0 \) by the aid of Lemma 2.4, that is we can pick \( \varepsilon_j \) small enough such that for any \( \gamma \leq \alpha, \theta \leq \beta \) we have

\[
\|1 - \chi(\varepsilon_j \xi)\|_{S^1}^\gamma = \sup \left\{ \langle \xi \rangle^{\gamma-1} |\partial_\xi^\gamma \partial_x^\theta (1 - \chi(\varepsilon_j \xi))| \right\} \leq 2^{-j} \left( \sum_{\gamma \leq \alpha} 1 \right) \left( \sum_{\theta \leq \beta} 1 \right) C_{\alpha,\beta,\gamma,\theta} \langle \xi \rangle^{1+m_j-|\alpha|}.
\]

So we have

\[
|\partial_\xi^\alpha \partial_x^\beta \tilde{a}_j| \leq 2^{-j} \langle \xi \rangle^{1+m_j-|\alpha|},
\]
the estimate which is achieved exactly by choosing \( \varepsilon_j \) as above.

Now we verify that \( a \) is asymptotically equivalent to \( \sum a_j \). Assume we are given arbitrary \( \alpha \) and \( \beta \), pick \( p \geq |\alpha| + |\beta| \) and pick \( k \) such that \( m_p + 1 \leq m_{k+1} \). We firstly establish

\[
|\partial_\xi^\alpha \partial_x^\beta \left( a - \sum_{j \leq p-1} \tilde{a}_j \right)| \leq \sum_{j \geq p} 2^{-j} \langle \xi \rangle^{1+m_j-|\alpha|} \leq \langle \xi \rangle^{m_{k+1}-|\alpha|},
\]
because \( \langle \xi \rangle^{1+m_j} \leq \langle \xi \rangle^{m_{k+1}} \) for \( 1 + m_j \leq 1 + m_p \leq m_{k+1} \). Now by the decomposition

\[
a - \sum_{j \leq k} a_j = \left( a - \sum_{j \leq p-1} \tilde{a}_j \right) + \sum_{k+1 \leq j \leq p-1} \tilde{a}_j + \sum_{j \leq k} (a_j - \tilde{a}_j)
\]
we obtain the full estimate

\[
|\partial_\xi^\alpha \partial_x^\beta \left( a - \sum_{j \leq k} a_j \right)| \leq \langle \xi \rangle^{m_{k+1}-|\alpha|} + \left( \sum_{k+1 \leq j \leq p-1} 2^{-j} \langle \xi \rangle^{1+m_j-|\alpha|} \right) + \sum_{j \leq k} |\partial_\xi^\alpha \partial_x^\beta (a_j - \tilde{a}_j)| \leq C_{\alpha,\beta} \langle \xi \rangle^{m_{k+1}-|\alpha|},
\]

8
because \( a_j - \tilde{a}_j \) have compact support and hence are in \( S^{-\infty} \). Hence \( \sum a_j \) is the asymptotic expansion of \( a \). Furthermore, we have

\[
\text{supp } a \subset \bigcup \text{supp } \tilde{a}_j \subset \bigcup \text{supp } a_j,
\]

by how we constructed \( a \).

Now we are in a position to define classical pseudodifferential symbols, symbols which retains the layering feature like polynomial symbols induced by classical linear differential operators.

**Definition 2.7** (Classical symbol). We say a pseudodifferential symbol \( a(x, \xi) \in S^m \) is classical if \( a \sim \sum_{j=0}^{\infty} a_j \) for some \( a_j \) which are respectively homogeneous of degree \( m - j \) for \( |\xi| \geq 1 \), that is \( a_j(x, \xi) = \lambda^{m-j} a_j(x, \xi) \) for \( |\xi| \geq 1 \) and \( \lambda \geq 1 \). Denote the class of classical pseudodifferential operators by \( CS^m \subset S^m \).

**Definition 2.8** (Principal symbol). For a classical symbol \( a \sim \sum_j a_j \) we call \( a_0 \) the principal symbol of \( a \).

**Example.** Symbol \( \sum_{|\alpha| \leq m} a_{\alpha} \xi^\alpha \) are classical symbol of degree \( m \), and \( \sum_{|\alpha|=m} a_{\alpha} \xi^\alpha \) is the principal symbol.

### 2.3 Operators, Kernels, Adjoints, Quantisation

In the last few sections we have defined and talked about properties of pseudodifferential symbols. Now in this section we will define pseudodifferential operators.

**Definition 2.9** (Pseudodifferential operators on \( S \)). Given \( a \in S^m(\mathbb{R}^{2N}) \) and \( u \in S(\mathbb{R}^N) \), the space of Schwartz class functions, define

\[
\text{Op}(a)u(x) = \int_{\mathbb{R}^{2N}} e^{i(x-y) \cdot \xi} a(x, \xi) u(y) \, dy \, d\xi,
\]

in which \( d\xi = (2\pi)^{-N} d\xi \). Here \( \text{Op}(a) \) acting on \( S \) is called the pseudodifferential operator of order \( m \), associated with symbol \( a \). Denote the class of pseudodifferential operators of order \( m \) by \( \Psi^m \), and let \( \Psi^{-\infty} = \cap_m \Psi^m \).

**Definition 2.10** (Classical pseudodifferential operators). We say \( \text{Op}(a) \) is classical of order \( m \), if \( a \in CS^m \).

**Remark.** (i) \( \text{Op}(a) \) is also written as \( a(x, D) \) by substituting the frequencies \( \xi \) by Fourier \( D \) in the symbol \( a(x, \xi) \). Also for concerns about simplicity of expression, in the context we conventionally refer \( A \) to the pseudodifferential operator associated with symbol \( a \). We denote the symbol of operator \( A \) by \( \sigma_A \).

(ii) We also can write the formula \( \text{Op}(a) \) as

\[
\text{Op}(a)u(x) = \mathcal{F}^{-1} [a(x, \xi) \mathcal{F} [u]] = \int_{\mathbb{R}^{2N}} e^{ix \cdot \xi} a(x, \xi) \hat{u}(\xi) \, d\xi,
\]

which makes perfect sense for \( u \in S \). This illustrates the nature of pseudodifferential operator, as a multiplication operator on the Fourier side.
Example. (i) Classical linear differential operators $\sum_{|\alpha| \leq m} a_{\alpha}(x)D^\alpha$ with all derivatives of $a_{\alpha}(x)$ bounded, are pseudodifferential operators of order $m$, associated with symbol $\sum_{|\alpha| \leq m} a_{\alpha}(x)\xi^\alpha$.

(ii) For $N \geq 2$, Laplace operator $\triangle = \sum_{k=1}^{N} D_{x_k}^2$ is pseudodifferential operator of order 2, with symbol $\sigma_\triangle = |\xi|^2$. Consider the operator $|D|$ induced by symbol $|\xi| \in S^1(\mathbb{R}^N \times \mathbb{R}^N)$, we see $|D| \in \Psi^1$ and moreover

$$|D|^2 = F^{-1} \circ m_{|\xi|} \circ F \circ F^{-1} \circ m_{|\xi|} \circ F = F^{-1} \circ m_{|\xi|^2} \circ F = \triangle,$$

in which $m_a$ denotes multiplication operator of factor $a$. Hence we see $|D|$ is a square root of $\triangle$, and this gives us a hint about how to define fractional power of operators.

(iii) Consider 2D Dirac operator acting on $u : \mathbb{R}^2 \to \mathbb{R}^2$ with $u \in S$:

$$W = \begin{pmatrix} 0 & i\partial_{x_1} + \partial_{x_2} \\ i\partial_{x_1} - \partial_{x_2} & 0 \end{pmatrix}, \quad \sigma_W = \begin{pmatrix} 0 & \xi_1 - i\xi_2 \\ \xi_1 + i\xi_2 & 0 \end{pmatrix}$$

Easy to see $\sigma_W \in S^1(\mathbb{R}^2 \times \mathbb{R}^2)$ and hence $W \in \Psi^1$. Because the symbol is location-invariant, that is the multiplication operator does only depend on $\xi$, we have

$$\sigma_{W^2} = \sigma_W^2 = \begin{pmatrix} \xi_1^2 + \xi_2^2 & 0 \\ 0 & \xi_1^2 + \xi_2^2 \end{pmatrix} = \sigma_\triangle$$

We see the classical result that $W^2 = \triangle$, where $\triangle$ is understood to be a two-component operator.

(iv) For $N \geq 2$, define $\langle D \rangle^m \in \Psi^m$ by its symbol $\langle \xi \rangle^m$, for any $m$. This is an isometry between $H^m$ and $L^2$.

We show this operator is well-defined on $S$.

**Proposition 2.11** (Pseudodifferential Operators on $S$). For $a \in S^m(\mathbb{R}^{2N})$ we have

$$\text{Op}(a) : S \to S$$

is well-defined, and the mapping $(a, u) \mapsto \text{Op}(a)u$ is continuous. Moreover $\text{Op} : a \mapsto \text{Op}(a)$ is injective.

**Proof.** For $u \in S$, we have the estimate

$$|\text{Op}(a)u(x)| \leq \sup \left\{ a(x, \xi) \langle \xi \rangle^{-m} \right\} \sup \left\{ \langle \xi \rangle^{m+N+1} |\hat{u}| \right\} \int_{\mathbb{R}^N} \langle \xi \rangle^{-N-1} d\xi \leq \left( \|a\|_{0,0}^{S_m} C \right) \left( \sum_{|\alpha| \leq m+N+1} C_\alpha \|\hat{u}\|_{\alpha,0}^S \right),$$

where $\|u\|_{\alpha,\beta}^S$ are the semi-norms on $S$:

$$\|u\|_{\alpha,\beta}^S = \sup_{\mathbb{R}^N} |x^\alpha \partial_x^\beta u(x)|.$$
Hence \( \text{Op}(a)u \) is continuous and bounded with respect to \( a \), and with respect to \( u \) by the well-known fact that Fourier transform is continuous between two copies of \( \mathcal{S} \).

On the injectivity of \( \text{Op} \), because \( \text{Op} \) is clearly linear, we only need to show that if \( \text{Op}(a) = 0 \) then \( a = 0 \). Let

\[
\int_{\mathbb{R}^N} e^{ix \cdot \xi} a(x, \xi) \hat{u}(\xi) \, d\xi = 0
\]

for all \( x \in \mathbb{R}^N \) and all \( u \in \mathcal{S} \). With \( x \) fixed, consider the term

\[
b(\xi) = a(x, \xi) \langle \xi \rangle^{-m} \langle \xi \rangle^{-N/2-1/2}.
\]

Indeed \( b \in L^2 \), as

\[
\|b\|_{L^2} = \int_{\mathbb{R}^N} (a \langle \xi \rangle^{-m})^2 \langle \xi \rangle^{-N-1} \, d\xi \leq (\sup \{a \langle \xi \rangle^{-m}\})^2 \int_{\mathbb{R}^N} \langle \xi \rangle^{-N-1} \, d\xi = C \|a\|_{S^m_{0,0}} < \infty.
\]

Set

\[
v(\xi) = e^{-ix \cdot \xi} \langle \xi \rangle^{m+N/2+1/2} \hat{u}(\xi),
\]

and note that between \( u \) and \( v \) this is a bijection between two copies of \( \mathcal{S} \). We see that \( (b(\xi), v(\xi)) = 0 \) for all \( v \in \mathcal{S} \). Hence \( b = 0 \) in \( \mathcal{S} \) and by the fact \( \mathcal{S} \) is dense in \( L^2 \), we see \( b = 0 \) in \( L^2 \) and hence \( a(x, \xi) = 0 \) almost everywhere. As \( a \) is continuous, we see \( a(x, \xi) = 0 \).

**Remark.** Here we use the trivial fact that

\[
\int_{\mathbb{R}^N} \langle \xi \rangle^k \, d\xi < +\infty
\]

for any \( k < -N \). Easy to verify in spherical coordinates.

**Proposition 2.12** (Kernel and symbol). If \( a \in S^m \subset \mathcal{S}' \), then

\[
K_{\text{Op}(a)}(x, y) = (2\pi)^{-N} [\mathcal{F}_{\xi \rightarrow y} a(x, \xi)] (x, y-x).
\]

Furthermore, this mapping as a bijection \( \mathcal{S}' \rightarrow \mathcal{S}' \) has an inverse

\[
a(x, \xi) = \mathcal{F}_{y \rightarrow \xi} \left[ K_{\text{Op}(a)}(x, x-y) \right].
\]

**Proof.** By definition of pseudodifferential operators we have

\[
\text{Op}(a)u(x) = \int_{\mathbb{R}^{2N}} e^{ix \cdot (y-x)} a(x, \xi) u(y) \, dy \, d\xi.
\]

Hence we have straightforward

\[
K_{\text{Op}(a)}(x, y) = \int_{\mathbb{R}^N} e^{i(x-y) \cdot \xi} a(x, \xi) \, d\xi = (2\pi)^{-N} [\mathcal{F}_{\xi \rightarrow y} a(x, \xi)] (x, y-x).
\]

It is easy to show that this formula actually defines a bijection between two copies of \( \mathcal{S}' \). By inverting it we have the second formula immediately. \( \square \)
Now consider the adjoint $A^*$ of pseudodifferential operator $A$ on $S$:

$$(Au, v) = (u, A^*v)$$

for any $u, v \in S$, where

$$(u, v) = \int_{\mathbb{R}^N} u(x) \overline{v(x)} \, dx$$

is the Hermitian inner product. The uniqueness of $A^*$ follows from linearity of Hermitian inner product. Hence $A^*$ is well-defined, if $A^*$ exists. Till this point it is yet unclear about the existence of $A^*$; and to answer this question, we quote a result from Alinhac and Gérard [2, Section I.8.2] instead of replicating the long-winded proof.

**Proposition 2.13.** Given $a \in S^m$, we have that $a^*$, the symbol of $\text{Op}(a)^*$, is also a symbol of order $m$, that is, $a^* \in S^m$. Furthermore we have asymptotics for $a^*$:

$$a^*(x, \xi) \sim \sum_\alpha \frac{1}{\alpha!} \partial^\alpha_x D^\alpha_\xi a(x, \xi).$$

**Remark.** (i) Hence if $A$ is a pseudodifferential operator of order $m$, its adjoint $A^*$ is also a pseudodifferential operator of order $m$, and $A$ admits a natural extension to an operator $S' \to S'$.

(ii) If $a(\xi)$ depends only on $\xi$, that is, it is location-invariant, we have the formula in the special case

$$a^*(\xi) \sim \overline{a(\xi)}.$$

Furthermore, by a direct computation we see in fact $a^*(\xi) = \overline{a(\xi)}$.

**Definition 2.14 (Pseudodifferential operators on $S'$).** Given $u \in S'(\mathbb{R}^N)$ and $a \in S^m(\mathbb{R}^N \times \mathbb{R}^N)$, define $A = \text{Op}(a)$ by

$$\langle Au, \varphi \rangle = \langle Au, \overline{\varphi} \rangle = \langle u, A^* \overline{\varphi} \rangle = \langle u, A^* \varphi \rangle$$

for any $\varphi \in S$, where $\langle \cdot, \cdot \rangle$ denotes pairing between $S$ and $S'$. Here $A^* : S \to S$ is the adjoint of $A$.

We also wish to very briefly introduce the notation of right quantisation.

**Definition 2.15 (Right quantisation).** Given symbol $a(x, \xi) \in S^m(\mathbb{R}^N \times \mathbb{R}^N)$ we define its right quantisation, $\text{Op}_r(a)$ or $a(D, y)$, to be the operator acting on $S$:

$$a(D, y)u(x) = \int_{\mathbb{R}^{2N}} e^{i(x-y) \cdot \xi} a(y, \xi)u(y) \, dy \, d\xi.$$

It is also an operator in $\Psi^m$. It similarly admits an extension to $S'$.
Remark. (i) The standard operator \( a(x, D) \) is called left quantisation, in contrast.

(ii) Each symbol \( a(x, \xi) \) has two influences to the function it acts on: perturbation \( \xi \) to the frequency, and perturbation \( x \) to the location. In the standard left quantisation, the perturbation in frequency comes firstly, and thereafter follows that in location. However in the right quantisation, the perturbation in location comes firstly.

(iii) To exemplify (ii), consider the quantisation of symbol \( x\xi \). The left quantisation gives \( xD_x \), and the right quantisation gives \( D_x x \). The order of changes is very clear: between left and right quantisation, there is a switch.

(iv) Reversely, given an operator of specific class, then one can either associate a left symbol and a right symbol, respectively which yield the operator as left and right quantisation. For example, given operator \( xD_x \), its left symbol is \( x\xi \) and right symbol is \( x\xi - 1 \).

(v) The right symbol occurs naturally in the process of taking the adjoint of a left symbol, as of in the omitted proof of Proposition 2.13.

2.4 Symbolic Algebra of Operators

We now discuss about the composition of two pseudodifferential operators. When the symbols are location-invariant, the composition is simply the multiplication of two multipliers which depends on \( \xi \) only, on the Fourier side, and its order is the sum of orders of two symbols. However the case becomes complicated when symbol of the first operator is \( x \)-dependent. We refer to a result from Alinhac and Gérard [2, Section I.8.2]:

**Theorem 2.16** (Composition). Given \( a_1 \in S^{m_1}, \ a_2 \in S^{m_2} \), we have \( \text{Op}(a_1) \text{Op}(a_2) = \text{Op}(b) \) for \( b \in S^{m_1+m_2} \) denoted by \( b = a_1 \# a_2 \). Moreover \( b \) has asymptotics

\[
b \sim \sum_{\alpha} \frac{1}{\alpha!} \partial^{\alpha}_{\xi} a_1 D_x^{\alpha} a_2.
\]

**Example.** (i) When \( a_2 \) depends only on \( \xi \), or \( a_1 \) depends only on \( x \), we see the simple relation

\[
a_1 \# a_2 \sim a_1(x).a_2(\xi),
\]

furthermore they are exactly equal, via a direct computation.

(ii) When \( a_1(x, D) \) and \( a_2(x, D) \) are linear differential operators, via Leibniz’s formula we immediately have

\[
a_1 \# a_2 = \sum_{\alpha} \frac{1}{\alpha!} \partial^{\alpha}_{\xi} a_1 D_x^{\alpha} a_2.
\]

The asymptotic formula is also exact in this case.

**Corollary 2.17.** Given \( a_1 \in S^{m_1} \) and \( a_2 \in S^{m_2} \) we have the commutator

\[
[A_1, A_2] = A_1 A_2 - A_2 A_1
\]

is a pseudodifferential operator of order \( m_1 + m_2 - 1 \). Its symbol is

\[
-i \{a_1, a_2\} \mod S^{m_1+m_2-2},
\]
where the Poisson brackets denote
\[
\{f, g\} = \sum_j \left( \frac{\partial f}{\partial \xi_j} \frac{\partial g}{\partial x_j} - \frac{\partial f}{\partial x_j} \frac{\partial g}{\partial \xi_j} \right).
\]

**Proof.** Write
\[
b_1 = a_1 \# a_2 \sim \sum_{\alpha} \frac{1}{\alpha!} \partial^\alpha_\xi a_1 D^\alpha_x a_2,
\]
\[
b_2 = a_2 \# a_1 \sim \sum_{\alpha} \frac{1}{\alpha!} \partial^\alpha_\xi a_2 D^\alpha_x a_1,
\]
then we see the symbol of the commutator is
\[
b_1 - b_2 \sim \sum_{\alpha \neq 0} \frac{1}{\alpha!} \left( \partial^\alpha_\xi a_1 D^\alpha_x a_2 - \partial^\alpha_\xi a_2 D^\alpha_x a_1 \right)
\]
\[
= \sum_j \left( \partial_\xi_j a_1 D_{x_j} a_2 - \partial_\xi_j a_2 D_{x_j} a_1 \right) \mod S^{m_1 + m_2 - 2}
\]
\[
= -i \{a_1, a_2\} \mod S^{m_1 + m_2 - 2},
\]
clearly.

### 2.5 Operators on Sobolev Spaces

Consider the fact that \( H^0 = L^2 \subset S' \), we see a pseudodifferential operator is naturally defined on \( L^2 \) by a restriction of itself on \( S \). We want to show a fact that if \( a \in S^0 \) then \( a(x, D) \) maps \( L^2 \) into \( L^2 \), before which we need to prove a lemma:

**Lemma 2.18** (Operators of negative order are continuous on \( L^2 \)). Given \( A \in \Psi^m \) for \( m < 0 \), then \( A \) is a linear bounded operator from \( L^2 \) to \( L^2 \).

**Proof.** Clear that \( A \) is linear. By \( \|Au\|_{L^2} = (A^*Au, u) \), and the fact that \( \|A\|_{L^2 \to L^2} = \|A^*\|_{L^2 \to L^2} \) as operator norm, we see \( \|A\|_{L^2 \to L^2} \leq \|A^*A\|_{L^2 \to L^2}^{1/2} \), where \( A^*A \in \Psi^{2m} \) by Theorem 2.16. Iteratively we have
\[
\|A\|_{L^2 \to L^2} \leq \|A^*A\|_{L^2 \to L^2}^{1/2} \leq \|A^* A\|_{L^2 \to L^2}^{1/4} \leq \cdots \leq \|R^* p R_p\|_{L^2 \to L^2}^{2^{-p}}, \tag{3}
\]
in which \( R^* p R_p \in \Psi^{2m} \), for any positive integer \( p \). That is, if we can prove in some \( \Psi^t \) with \( t \) negative enough, all operators are bounded from \( L^2 \) to \( L^2 \), then there will be some \( R^* p R_p \in \Psi^{2m} \) with \( 2^p m \leq t \) hence \( R^* p R_p \in \Psi^t \) is bounded from \( L^2 \) to \( L^2 \) and so is \( A \) by (3).

We now prove that operators in \( \Psi^{-N-1} \) are all bounded from \( L^2 \) to \( L^2 \). Given \( a(x, D) \in \Psi^{-N-1} \), write down its kernel and by the estimate
\[
|K(x, y)| = \left| \int_{\mathbb{R}^N} e^{i(x-y) \cdot \xi} a(x, \xi) d\xi \right| \leq \int_{\mathbb{R}^N} \langle \xi \rangle^{-N-1} d\xi,
\]
\[ |K(x, y)| \] is bounded by a constant on \( \mathbb{R}^{2N} \). Furthermore, it is easy to verify that \((x - y)^\alpha K(x, y)\) is the kernel of \((i)^{|\alpha|} \partial_\xi^\alpha a(x, D) \in S^{-N-1-|\alpha|} \subset S^{-N-1}\), which by the previous estimate is bounded by a constant for all \( \alpha \). Hence by summing over \( |\alpha| \leq N + 1 \) we see

\[
(x - y)^{N+1} |K(x, y)| \leq (1 + |x - y|)^{N+1} |K(x, y)|
\]

is bounded by a constant, for \((1 + |x - y|)^{N+1}\) can be expanded to a finite sum of terms bounded by a constant by the previous estimate. Hence

\[
\int_{\mathbb{R}^N} |K(x, y)| \, dx \leq \sup \langle x - y \rangle^{N+1} |K(x, y)| \int_{\mathbb{R}^N} |x - y|^{-N-1} \, dx \leq C
\]

for some constant \( C \), by a change of variable in the last integral. Similarly we have

\[
\int_{\mathbb{R}^N} |K(x, y)| \, dy \leq C
\]

Hence we have the estimate

\[
|Au(x)|^2 \leq \left( \int_{\mathbb{R}^N} |K(x, y)u^2(y)| \, dy \right) \left( \int_{\mathbb{R}^N} |K(x, y)| \, dy \right) \leq C \int_{\mathbb{R}^N} |K(x, y)u^2(y)| \, dy
\]

and the final estimate

\[
\|Au\|^2_{L^2} \leq C \int_{\mathbb{R}^N} \left( \int_{\mathbb{R}^N} |K(x, y)| \, dx \right) |u(y)|^2 \, dy \leq C^2 \|u\|^2_{L^2}
\]

establishes the fact that any operator in \( \Psi^{-N-1} \) are linear bounded operators from \( L^2 \) to \( L^2 \). Apply the argument in the first paragraph and the result follows immediately.

\[ \square \]

**Theorem 2.19** (Operators of order 0 maps \( L^2 \) into \( L^2 \)). Given \( a \in S^0 \), then \( \text{Op}(a) : L^2 \to L^2 \) is bounded.

**Proof.** Pick \( M = 2 \sup |a(x, \xi)|^2 \), and take

\[
c(x, \xi) = (M - |a(x, \xi)|^2)^{1/2},
\]

which is in \( S^0 \) by Lemma 2.3. From Proposition 2.13 and Theorem 2.16 we have

\[
c^* # c = c^* c + r_1 = \overline{c} c + r_2 = M - \overline{a} a + r_2 = M - a^* # a + r,
\]

where \( r \in S^{-1} \) is an asymptotic remainder. Hence we have

\[
C^* C = M - A^* A + R
\]

for remainder \( R \in \Psi^{-1} \). Now we have

\[
\|Au\|^2_{L^2} = (A^* Au, u) = M \|u\|^2_{L^2} - \|Cu\|^2_{L^2} + (Ru, u) \leq M \|u\|^2_{L^2} + (Ru, u) \leq (M + \|R\|_{L^2 \to L^2}) \|u\|^2_{L^2}
\]

where the operator norm of \( R \) is known to be finite by the previous lemma. \[ \square \]
Now for $s \geq 0$, set the $L^2$ based Sobolev spaces $H^s$ to be endowed with norm

$$\|u\|_{H^s} = \left( \int_{\mathbb{R}^N} |\langle \xi \rangle^s \hat{u}|^2 \, d\xi \right)^{1/2}.$$ 

Easy to verify $L^2 = H^0$ with the same norm, by Plancherel's theorem. We now investigate the action of pseudodifferential operators on $H^s$.

**Lemma 2.20** (Isometry between $H^s$ and $L^2$). For $s > 0$, the bijection $\langle D \rangle^s : H^s \to L^2$ is isometric.

**Proof.** Consider for given $u \in H^s$,

$$\|u\|_{H^s}^2 = \int_{\mathbb{R}^N} (\langle \xi \rangle^s \hat{u})(\langle \xi \rangle^s \hat{u}) \, d\xi = \|\langle D \rangle^s u\|_{L^2}^2.$$ 

It is trivial to verify that $\langle D \rangle^{-s} : L^2 \to H^s$ is the inverse to $\langle D \rangle^s$. \hfill \Box

**Proposition 2.21** (Operators on $H^s$). Given $A \in \Psi^m$, then we have $A : H^s \to H^{s-m}$, for any $s > 0$ and $s - m > 0$. Moreover $A$ is bounded.

**Proof.** Consider $u \in H^s$, then by the fact $\langle D \rangle^{-s} \langle D \rangle^s = \text{id}$ we see

$$\langle D \rangle^{s-m} A u = (\langle D \rangle^{s-m} A \langle D \rangle^{-s}) \langle D \rangle^s u.$$ 

Note that $\langle D \rangle^s u \in L^2$ by Lemma 2.20 and that $\langle D \rangle^{s-m} A \langle D \rangle^{-s} \in \Psi^0$ is bounded by Theorem 2.16. Apply Theorem 2.19 to see $\langle D \rangle^{s-m} A u \in L^2$. Apply Lemma 2.20 again to see $Au \in H^{s-m}$. The boundedness of $A$ follows immediately from boundedness of $\langle D \rangle^{s-m} A \langle D \rangle^{-s}$ and isometric nature of $\langle D \rangle^s$ and $\langle D \rangle^{s-m}$. \hfill \Box

**Remark.** (i) Note that here we can take $m < 0$. Then $A : H^s \to H^{s+(-m)}$ for any $s > 0$. In plain english: operators of class $\Psi^{-k}$ of negative order $-k < 0$, improve differentiability by $k$, in the weak sense. In contrast, operators of class $\Psi^k$ reduce differentiability by $k$. This coincides with our sense of classical differential operators of order $k$, on $H^m$, in the weak sense.

(ii) Pseudodifferential operators have the same mapping properties (boundedness, sharp regularity), between Hölder/Hölder-Zygmund spaces. A detailed discussion about this topic is attached in Section 3.

(iii) One can extend this result from Sobolev spaces to weighted Sobolev spaces, or more general modulation spaces. See Molahajloo and Pfander [16].

### 2.6 Ellipticity and Parametrices

In this section we want to define an analogy of elliptic differential operators, that is, ellipticity conditions for pseudodifferential operators.
Proposition 2.22 (Ellipticity and parametrices). Let $a \in S^m$. Two conditions

(i) $\exists b \in S^{-m}$ that $a(x, D)b(x, D) - \text{id} \in \Psi^{-\infty}$,

(ii) $\exists b \in S^{-m}$ that $b(x, D)a(x, D) - \text{id} \in \Psi^{-\infty}$

are equivalent, and they imply that

(iii) $|a(x, \xi)| \geq c |\xi|^m$ for some $c > 0$ and $|\xi| \geq C$.

Conversely, if (iii) is satisfied, then there is $b \in S^{-m}$ satisfying (i) and (ii), where $b$ is unique up to a $S^{-\infty}$ difference. We call such $b$ the parametrix of $a$, and we say $\text{Op}(a)$ is elliptic if $a$ satisfies (iii).

Proof. Given $b', b'' \in S^{-m}$ satisfying $AB' - \text{id} \in \Psi^{-\infty}$ and $B''A - \text{id} \in \Psi^{-\infty}$, we have

$$B'' - B' = B''(\text{id} - AB') + (B''A - \text{id})B' \in \Psi^{-\infty}$$

and hence $b'' - b' \in S^{-\infty}$, the uniqueness of parametrix up to class $\Psi^{-\infty}$ is proved. Moreover, (i) or (ii) implies $a(x, \xi)b(x, \xi) - 1 \in S^{-1}$ and

$$1/2 \leq |a(x, \xi)||b(x, \xi)| \leq C |a(x, \xi)||\xi|^{-m}$$

for large $|\xi|$ and we see $\text{Op}(a)$ is elliptic.

Conversely, if $a$ is an elliptic symbol we take

$$b = \langle \xi \rangle^{-m} F \left( a \langle \xi \rangle^{-m} \right),$$

where $F \in C^\infty(\mathbb{C})$ and $F(z) = 1/z$ for $|z| \geq c'$, with some $c'$. Then Lemma 2.3 implies $b \in S^{-m}$ and $ab = 1 + \chi(\xi)$ with $\chi(\xi) = 0$ for $|\xi| \geq C'$ with some $C'$. Theorem 2.16 tells that

$$a(x, D)b(x, D) = \text{id} - r(x, D), \quad r \in S^{-1}.$$ 

Set for $k \geq 0$, $b_k(x, D) = b(x, D)r(x, D)^k \in \Psi^{-m-k}$ and choose $b' \in S^{-m}$ such that $b' \sim \sum_{k \geq 0} b_k$ (Proposition 2.6 tells that we can do this). Then

$$AB' = A \left( B' - \sum_{j<k} B_j \right) + AB \sum_{j<k} R_j = (\text{id} - R) \sum_{j<k} R_j + \Psi^{-k} = \text{id} - R^k + \Psi^{-k} = \text{id} + \Psi^{-k}$$

for all $k$ and hence we have (i). Similarly we can construct $b''$ satisfying (ii).

Remark. Parametrices are almost-inversions of elliptic operators, with a possible error of an infinitely smoothing operator (one should hence note that it means the singular supports are not altered by this error). There are some immediate remarks: (i) with aid of pseudolocality, one can very easily verify the elliptic regularity result for elliptic pseudodifferential operators; (ii) one can define hypoellipticity, a definition using another growth controls of symbol, and show the existence of parametrices of hypoelliptic operators and hypoelliptic regularity. See Shubin [20].
3 Nonlinear Dyadic Analysis

3.1 Littlewood-Paley Decomposition

In this section we will establish the Littlewood-Paley decomposition via pseudodifferential operators. Set a cutoff function \( \psi \in C^\infty_c(\mathbb{R}^N), 0 \leq \psi \leq 1 \), with \( \psi(\xi) = 1 \) for \( |\xi| \leq 1/2 \), \( \psi(\xi) = 0 \) for \( |\xi| \geq 1 \). Set \( \varphi(\xi) = \psi(\xi/2) - \psi(\xi) \), we then see \( \varphi \) is supported in \( 1/2 \leq |\xi| \leq 2 \), and for all \( \xi \) we have

\[
1 = \psi(\xi) + \sum_{p \geq 0} \varphi_p(\xi), \quad \varphi_p(\xi) = \varphi(2^{-p}\xi), \quad \varphi_{-1}(\xi) = \psi(\xi)
\]

which is locally finite near \( \xi \), and moreover is pointwise a sum of up to two non-zero terms, by the fact that

\[
\text{supp} \, \psi \subset [-1, 1], \quad \text{supp} \, \varphi_p \subset \{ \xi : 2^{p-1} \leq |\xi| \leq 2^{p+1} \}.
\]

It is an immediate result that

\[
\varphi_p \varphi_q = 0, \quad \text{if } |p - q| \geq 2,
\]

\[
\psi \varphi_p = 0, \quad \text{if } p \geq 0.
\]

(4)

Also note that \( 0 \leq \varphi_p \leq 1 \). Now given any \( u \in S'(\mathbb{R}^N) \), set

\[
u_{-1} = \psi(D)u \]

\[
u_p = \varphi_p(D)u, \quad p \geq 0
\]

We obtain the so-called Littlewood-Paley decomposition:

\[
u = \sum_{p=-1}^{\infty} \nu_p.
\]

Denote the partial sum by

\[
S_k \nu = \sum_{p=-1}^{k-1} \nu_p.
\]

We remark here that \( \psi(\xi) \) and \( \varphi_p(\xi) \) are in \( C^\infty_c \subset S^{-\infty} \), hence the pseudodifferential operators induced by them make perfect sense.

Lemma 3.1 (Almost-orthogonality of terms). Given the Littlewood-Paley decomposition, we have

\[
1/2 \leq \psi^2(\xi) + \sum_{p \geq 0} \varphi_p^2(\xi) \leq 1,
\]

(5)

and for any \( u \in L^2 \) we have

\[
\sum_{p \geq -1} \|u_p\|^2_{L^2} \leq \|u\|^2_{L^2} \leq \sum_{p \geq -1} \|u_p\|^2_{L^2}.
\]

18
Proof. For the first inequality, consider that
\[
1 = \left( \psi(\xi) + \sum_{p \geq 0} \varphi_p(\xi) \right)^2 = \psi^2 + \sum_{p \geq 0} \varphi_p^2 + 2\psi \sum_p \varphi_p + 2 \sum_{p<q} \varphi_p \varphi_q \geq \psi^2 + \sum_p \varphi_p^2,
\]
for \( \psi \geq 0 \) and \( \varphi_p \geq 0 \). And also via
\[
2 \left( \psi^2 + \sum_{p \geq 0} \varphi_p^2 \right) - \left( \psi + \sum_{p \geq 0} \varphi_p \right)^2 = \psi^2 + \sum_{p \geq 0} ((-1)^p \varphi_p)^2 + 2 \sum_{p<q} (-1)^p \varphi_p (-1)^q \varphi_q - 2\psi \sum_{p \geq -1} (-1)^p \varphi_p
\]
\[
= \left( \psi - \sum_{p \geq -1} (-1)^p \varphi_p \right)^2 \geq 0.
\]
we get (5). For the second inequality, first consider
\[
\int_{\mathbb{R}^N} \psi^2(\xi) + \sum_{p \geq 0} \varphi_p^2(\xi) |\hat{u}|^2 \, d\xi = \sum_{p \geq -1} \int_{\mathbb{R}^N} |\hat{u}_p|^2 \, d\xi = \sum_{p \geq -1} \|u_p\|_{L^2}^2.
\]
Via the first inequality we have
\[
\sum_{p \geq -1} \|u_p\|_{L^2}^2 \leq \|u\|_{L^2}^2 \leq 2 \sum_{p \geq -1} \|u_p\|_{L^2}^2
\]
immediately. \( \square \)

**Lemma 3.2** (Sensitivity of terms to differentiation). With given Littlewood-Paley decomposition \( u = \sum_{p \geq -1} u_p \), we have the following estimates independent of \( p \) and \( u \), independent of parameter \( \alpha \) only:

(i) For any \( \alpha, p \geq -1 \) we have
\[
\|\partial^\alpha u_p\|_{H^0} \lesssim 2^{p|\alpha|} \|u\|_{H^0}
\]
\[
\|\partial^\alpha S_p u\|_{H^0} \lesssim 2^{p|\alpha|} \|u\|_{H^0}
\]
and
\[
\|\partial^\alpha u_p\|_{L^\infty} \lesssim 2^{p|\alpha|} \|u\|_{L^\infty}
\]
\[
\|\partial^\alpha S_p u\|_{L^\infty} \lesssim 2^{p|\alpha|} \|u\|_{L^\infty}.
\]

(ii) For any \( s \in \mathbb{R} \) and \( p \geq 0 \) we have constant \( C > 0 \) such that
\[
C^{-1}2^{ps} \|u_p\|_{H^0} \leq \|u_p\|_{H_s} \leq C2^{ps} \|u_p\|_{H^0}.
\]

(iii) For any \( k \in \mathbb{N} \) and \( p \geq 0 \), we have constant \( C > 0 \) such that
\[
C^{-1}2^{pk} \|u_p\|_{L^\infty} \leq \sum_{|\alpha|=k} \|\partial^\alpha u_p\|_{L^\infty} \leq C2^{pk} \|u_p\|_{L^\infty}.
\]
Proof. Consider the fact for $p \geq 0$, $\varphi_p$ is non-zero only when $2^p \leq |\xi| \leq 2^{p+1}$, we see at points on which $\varphi_p$ is non-zero, for any $s$ we have $C^{-1}2^{ps} \leq \langle \xi \rangle^{2s} \leq C2^{2ps}$ by a direct inequality and with increment in $C$ if necessary. We see by integration

$$C^{-1}2^{ps} \|u\|_{H^0}^2 \leq \int_{\mathbb{R}^N} \langle \xi \rangle^{2s} \varphi_p^2(\xi) |\hat{u}(\xi)|^2 \, d\xi = \|u_p\|_{H^s} \leq C2^{2ps} \|u\|_{H^0}^2.$$ 

Immediately we have

$$\|\partial^\alpha u_p\|_{H^0} \leq \|u_p\|_{H^{|\alpha|}} \leq C2^{p|\alpha|} \|u\|_{H^0},$$

and

$$\|\partial^\alpha S_p u\|_{H^0} \leq \|S_p u\|_{H^{|\alpha|}} = \int_{\mathbb{R}^N} \langle \xi \rangle^{2|\alpha|} \left( \sum_{k \geq -1} \varphi_p \right)^2 |\hat{\varphi}(\xi)|^2 \, d\xi \leq C2^{p|\alpha|} \|u\|_{H^0}$$

for the support of $\sum_k^p \varphi_p$ is with in ball of radius $2^{p+1}$. This proves (6), (7), (10).

We utilise the result that given arbitrary $v \in L^1(\mathbb{R}^N)$ we have $\|v * u\|_{L^\infty} \lesssim \|u\|_{L^\infty}$ for any $u \in L^\infty$, by Hölder inequality. Fix a $w \in C_c^\infty(\mathbb{R}^N)$, set $w_\mu(\xi) = w(\mu \xi)$, one have $F^{-1} w \in S \subset L^1$ and hence we obtain the estimate $\|F^{-1} w_\mu * u\|_{L^\infty} \lesssim \|u\|_{L^\infty}$. We set then we estimate

$$\|\partial^\alpha (w_\mu(D) u)|_{L^\infty} \lesssim \mu^{-|\alpha|} \|u\|_{L^\infty}.$$ 

Note that the constant of estimation do not depend on $\alpha$. Take $w(\xi) = \varphi_0(\xi)$ we get $w_{2^{-p}}(\xi) = \varphi(\xi)$ and hence

$$\|\partial^\alpha u_p\|_{L^\infty} \lesssim 2^{p|\alpha|} \|u\|_{L^\infty}$$

verifying (8). Take $w(\xi) = \psi(\xi)$ and we get $w_{2^{-p}} = S_p(\xi) = \psi(2^{-p} \xi)$, then see

$$\|\partial^\alpha S_p u\|_{L^\infty} \lesssim 2^{p|\alpha|} \|u\|_{L^\infty},$$

verifying (9). Finally we pick $w = \xi^\alpha \chi(\xi)$ for some standard cutoff function $\chi(\xi) \in C_c^\infty$ equal to 1 near support of $\varphi$ and see

$$\partial^\alpha \hat{u}_p = C \xi^\alpha \varphi(2^{-p} \xi) \hat{u}(\xi) = C2^{p|\alpha|} \left( 2^{-p} \xi \right)^\alpha \varphi(2^{-p} \xi) \hat{u}(\xi) = C2^{p|\alpha|} w(2^{-p} \xi) \hat{u}_p(\xi),$$

and hence

$$\|\partial^\alpha u_p\|_{L^\infty} = C2^{p|\alpha|} \|F^{-1} (w(2^{-p} \xi)) \ast u_p\|_{L^\infty} \leq C2^{p|\alpha|} \|u_p\|_{L^\infty}.$$ 

Sum it with $|\alpha| = k$ we get the right hand inequality of (11).

To obtain left hand of (11), choose $\chi \in C_c^\infty$ with $\chi \equiv 1$ near support of $\varphi$. Write $\varphi(\xi) = \left( \sum_{|\alpha| = k} \xi^\alpha \chi_\alpha \right) \varphi$, where $\chi_\alpha = (\xi^\alpha \chi(\xi)) \left( \sum_{|\alpha| = k} \xi^\alpha \right)^{-2} \in C_c^\infty$. Then see

$$\hat{u}_p = \sum_{|\alpha| = k} (2^{-p} \xi)^\alpha \xi^\alpha (2^{-p} \xi) \hat{u}_p(\xi) = 2^{-pk} \sum_{|\alpha| = k} \chi_\alpha (2^{-p} \xi) \hat{D}^\alpha u_p(\xi)$$

and take the inverse Fourier transform and multiply by $2^{pk}$ to see

$$2^{pk} \|u_p\|_{L^\infty} \leq \sum_{|\alpha| = k} \left( \|2^{pN} F^{-1} \chi_\alpha (2^{-p} \xi) \ast D^\alpha u_p(\xi)\|_{L^\infty} \lesssim \|D^\alpha u_p(\xi)\|_{L^\infty}$$

by Hölder inequality. \hfill \Box
3.2 Characterisation of Sobolev and Hölder Spaces

**Proposition 3.3** (Characterisation of Sobolev Spaces). (i) If \( u \in H^s(\mathbb{R}^N) \), then there is constant \( C \) for all \( p \geq -1 \), we have

\[
\|u_p\|_{H^0} \leq C \|u\|_{H^s} C_p^{2-2s}
\]

for some \( C_p \) that satisfies \( \sum C_p^2 \leq 1 \).

(ii) Conversely if for \( p \geq -1 \),

\[
\|u_p\|_{H^0} \leq CC_p^{2-2ps}
\]

with \( \sum C_p^2 \leq 1 \), then \( u \in H^s \) and \( \|u\|_{H^s} \lesssim C \).

**Proof.** For (i) consider the fact \((\langle D \rangle^s u)_p = \langle D \rangle^s u_p\) by a direct computation. We then see

\[
\|u_p\|_{H^0} \leq C \|\langle D \rangle^s u_p\|_{H^0} = C \|\langle D \rangle^s u\|_{H^s} = C C_p^{2-2ps} \|u\|_{H^s}
\]

by Lemma 2.20 and for first inequality by Lemma 3.2(ii). The estimate \( \sum C_p^2 \leq 1 \) follows from applying Lemma 3.1 on \( \langle D \rangle^s u \). For (ii) see

\[
\|u\|_{H^s}^2 = \|\langle D \rangle^s u\|_{H^0}^2 \leq 2 \sum_{p \geq -1} \|u_p\|_{H^s}^2 \leq 2 \sum_{p \geq -1} C^2 C_p^{2-2ps} \leq 2^{2s+1} C^2,
\]

by the first inequality by Lemma 3.2(ii). \( \square \)

Here we are in a position to define Hölder classes \( C^\alpha(\mathbb{R}^N) \):

**Definition 3.4** (Hölder spaces on \( \mathbb{R}^N \)). If \( 0 < \alpha \leq 1 \), we define the remainder norm for \( u \in C^0_b \), the space of bounded continuous functions on \( \mathbb{R}^N \):

\[
|u|'_{C^\alpha} = \sup_{x \neq y \in \mathbb{R}^N} |u(x) - u(y)| \|x - y\|^{\alpha};
\]

if \( k < \alpha \leq k + 1 \), define the remainder norm for \( u \in C^k_b \):

\[
|u|'_{C^\alpha} = \sum_{|\gamma| = k} |\partial^\gamma u|_{C^{\alpha-k}}.
\]

and define the Hölder norm of \( u \) of index \( \alpha \):

\[
\|u\|_{C^\alpha(K)} = \|u\|_{L^\infty} + |u|'_{C^\alpha}.
\]

Finiteness of Hölder norm of index \( \alpha \) characterises Hölder space of index \( \alpha \), denoted by \( C^\alpha \). We also take \( C^0 \) to be the space of bounded continuous functions, where \( L^\infty \) norm can be used.

**Remark.** When \( \alpha = p \in \mathbb{N} \), in general \( C^\alpha \neq C^p \), the former of which is Hölder space, the latter of which is the classical space of \( p \)-th order continuously differentiable functions. Across the context of this section we will avoid the situation when \( \alpha \in \mathbb{N} \), exactly because when at integer indices, Littlewood-Paley theory does not characterise \( C^\alpha \) in the Hölder sense, however, characterises a larger class in Definition 3.8.
Proposition 3.5 (Characterisation of Hölder Spaces of non-integer indices). (i) If \( u \in C^\alpha(\mathbb{R}^N), \alpha \notin \mathbb{N} \), then there is constant \( C \) for all \( p \geq -1 \) we have

\[
\| u_p \|_{L^\infty} \leq C \| u \|_{C^\alpha} 2^{-p\alpha}.
\]

(ii) Conversely, if for \( p \geq -1 \) we have \( \| u_p \|_{L^\infty} \leq M 2^{-p\alpha} \), \( \alpha \notin \mathbb{N} \), then \( u \in C^\alpha \) and \( \| u \|_{C^\alpha} \lesssim M \).

Proof. By Lemma 3.2 (iii) and the fact that \( (\partial^\alpha u)_p = \partial^\alpha u_p \) permit an immediate reduction to the case \( 0 < \alpha < 1 \). For (i), since \( \| u_1 \|_{L^\infty} \lesssim \| u \|_{L^\infty} \), it suffices to consider

\[
 u_p(x) = \int_{\mathbb{R}^N} 2^{pn} (F^{-1} \varphi) (2^p(x-y)) \, u(y) \, dy
\]

for \( p \geq 0 \). From the fact that \( \int_{\mathbb{R}^N} (F^{-1} \varphi)(z) \, dz = C \varphi(0) = 0 \) we also have

\[
 u_p(x) = \int_{\mathbb{R}^N} 2^{pn} (F^{-1} \varphi) (2^p(x-y)) (u(y) - u(x)) \, dy
\]

whence

\[
 |u_p(x)| \leq \| u \|_{C^\alpha} \int_{\mathbb{R}^N} 2^{pn} \left| (F^{-1} \varphi) (2^p(x-y)) \right| |x-y| \, dy \lesssim \| u \|_{C^\alpha} 2^{-p\alpha}
\]

and (i) is proved.

(ii) Conversely for some \( p \) to be determined, set

\[
 u = S_p u + R_p u, \quad R_p u = \sum_{q \geq p} u_q.
\]

Then we have

\[
 \| R_p u \|_{L^\infty} \leq \sum_{q \geq p} \| u_q \|_{L^\infty} \lesssim \sum_{q \geq p} M 2^{-q\alpha} \leq M 2^{-p\alpha}.
\]

Moreover we have

\[
 |S_p u(x) - S_p u(y)| \leq |x-y| \sum_{q=-1}^{p-1} \| \nabla u_q \|_{L^\infty}.
\]

By Lemma 3.2 (iii) we have

\[
 \| \nabla u_q \|_{L^\infty} \leq M 2^{q(1-\alpha)}
\]

and \( \| \nabla u_{-1} \|_{L^\infty} \lesssim M \). Hence if \( 0 < \alpha < 1 \) we have

\[
 |S_p u(x) - S_p u(y)| \lesssim M |x-y| 2^{p(1-\alpha)},
\]

for the series \( \sum \| \nabla u_q \|_{L^\infty} \) is then geometrically divergent. Regrouping the estimates for \( R_p u \) and \( S_p u \) we find

\[
 |u(x) - u(y)| \leq CM |x-y| 2^{p(1-\alpha)} + 2M 2^{-p\alpha}.
\]

Take \( p \) to be the largest integer such that \( 2^p \leq |x-y|^{-1} \), and we obtain

\[
 |u(x) - u(y)| \lesssim M |x-y|^{\alpha},
\]

and (ii) is proved. \( \square \)
Proposition 3.6 (Sobolev injection into Hölder spaces). Given $s > N/2$ and $s - N/2 \notin \mathbb{N}$, we have $H^s \subset C^{s-n/2}$ and the injection is continuous.

Proof. See

$$
|u_p| = \int e^{ix} |\hat{u}_p(x)| \, dx \leq C \int_{|x| \leq 2^{p+1}} |\hat{u}_p(x)| \, dx \leq C \|u_p\|_{H^0} \left(\sigma_N 2^{(p+1)N}\right)^{1/2}
$$

$$
\leq \tilde{C} \|u_p\|_{H^0} 2^{pN/2} \leq \tilde{C} \|u_p\|_{H^s} 2^{-p(s-N/2)},
$$

where $\sigma_N$ denote the volume of unit ball in $\mathbb{R}^N$, by Proposition 3.3 (i), with $C_p \leq 1$ used.

It is also written as $\|u_p\|_{C^0} \leq \tilde{C} \|u_p\|_{H^s} 2^{-p(s-N/2)}$ and we apply Proposition 3.5 (ii).

Proposition 3.7 (Convexity inequalities). (i) If $s = \lambda s_0 + (1 - \lambda)s_1$ for $0 \leq \lambda \leq 1$, $s_0 \leq s_1$ are real numbers, for $u \in C^\infty_c$ we have the inequality

$$
\|u\|_{H^s} \leq C \|u\|_{H^{s_0}} \|u\|_{H^{s_1}}^{1-\lambda}.
$$

(ii) If $\alpha = \lambda \alpha_0 + (1 - \lambda)\alpha_1$ for $0 \leq \lambda \leq 1$, $\alpha_0 \leq \alpha_1$ are positive non-integers, for $u \in C^{\alpha_1}$ we have the inequality

$$
\|u\|_{C^\alpha} \leq C \|u\|_{C^{\alpha_0}} \|u\|_{C^{\alpha_1}}^{1-\lambda}.
$$

Proof. For (i), observe

$$
\|u\|_{H^s}^2 = \int \langle \xi \rangle^{s_0} |\hat{u}|^{2\lambda} \langle \xi \rangle^{(1-\lambda)s_1} |\hat{u}|^{2(1-\lambda)} \, dx \leq (2\pi)^{-N} \|u\|_{H^{s_0}} \|u\|_{H^{s_1}}^{1-\lambda},
$$

by Hölder inequality. For (ii) we have

$$
\|u_p\|_{L^\infty} = \|u_p\|_{L^\infty} \|u_p\|_{L^\infty}^{1-\lambda} \leq C \|u\|_{C^{\alpha_0}} \left(2^{-p\alpha_0}\right)^{\lambda} \|u\|_{C^{\alpha_1}}^{1-\lambda} \left(2^{-p\alpha_1}\right)^{1-\lambda} = C \|u\|_{C^{\alpha_0}} \|u\|_{C^{\alpha_1}}^{1-\lambda} 2^{-p\alpha},
$$

by Proposition 3.5 (i) and then apply Proposition 3.5 (ii).

3.3 Operators on Hölder-Zygmund Spaces

In this section we want to establish results to aid our proofs in later sections: Theorem 3.13 will be used in Günther's approach to solve the isometric embedding problem in Section 5.3; Proposition 3.14 will be used to prove our Hölder scale Nash-Moser theorem in Section 4.1

Definition 3.8 (Hölder-Zygmund spaces). Given non-negative $\alpha \in \mathbb{R}$, the Hölder-Zygmund space of index $\alpha$, denoted by $C^\alpha$, is set to be the class of $u \in L^\infty$ finite under the norm

$$
\|u\|_{C^\alpha} = \sup_{p \geq 1} \left\{ 2^{p\alpha} \|u_p\|_{L^\infty} \right\}.
$$
Remark. (i) Using Littlewood-Paley decomposition, one should also be able to give a dyadic characterisation of Besov spaces.

(ii) When $\alpha > 0$ is non-integer, then $C^\alpha_\ast = C^\alpha$, the Hölder space, by Proposition 3.5.

By the trivial fact induced by definition of Hölder-Zygmund spaces:

$$\|u_p\|_{L^\infty} \leq \|u\|_{C^\alpha_\ast} 2^{-p\alpha}. \quad (14)$$

We have hence the immediate modifications of previous propositions:

**Proposition 3.9** (Properties of $C^\alpha_\ast$). (i) For $s > N/2$, $H^s \subset C^{s-N/2}_\ast$.

(ii) Given $\alpha = \lambda \alpha_0 + (1 - \lambda) \alpha_1$, with $0 \leq \lambda \leq 1$, then for all $u \in C^{\alpha_1}_\ast$ we have

$$\|u\|_{C^\alpha_\ast} \leq C\|u\|_{C^{\alpha_1}_\ast} \|u\|_{C^{\alpha_2}_\ast}^{1-\lambda}. \quad \text{(14)}$$

**Proof.** Exactly the same in previous propositions, except instead of facilitating characterisation of Proposition 3.5 we use the definition of Hölder-Zygmund norm directly. \qed

We now want to establish mapping properties of pseudodifferential operators between Hölder-Zygmund spaces. We firstly need three lemmata then will be able to state our main result.

**Lemma 3.10** (Estimates of kernels). Let $a \in S^m(\mathbb{R}^N \times \mathbb{R}^N)$ for $m \in \mathbb{R}$, and take $a(x, D) = \sum_{p \geq -1} a_p(x, D)$ to be the Littlewood-Paley decomposition of $a$. Let $k_p(x, y)$ denote the Schwartz kernel of $a_p(D)$, as in Proposition 2.12, that is

$$k_p(x, y) = \mathcal{F}_{\xi \rightarrow y}^{-1}(a_p(x, \xi)).$$

Then for any multi-indices $\alpha, \beta$, and any $M \in \mathbb{N}$ we have

$$\left| \partial^\beta_x \partial^\alpha_y k_p(x, y) \right| \leq C_{\alpha, \beta, M} |y|^{-M} 2^p(N+m-M+|\alpha|). \quad (15)$$

**Proof.** Firstly from the definition of $k_p$ we have for any $\alpha, \beta, \gamma$,

$$y^\gamma \partial^\beta_x D^\alpha_y k_p(x, y) = \int_{\mathbb{R}^N} e^{i y \cdot \xi} D^\gamma_\xi \left[ \xi^\alpha \partial^\beta_x a_p(x, \xi) \right] d\xi$$

by a symbolic rewriting. Observe that the integrand is supported within ball $\{|\xi| \leq 2^p+1\}$ which has volume $2^{Np+\sigma N}$. Furthermore, since the support is even contained in the annulus $\{2^{p-1} \leq |\xi| \leq 2^{p+1}\}$ (when $p \neq -1$) on which $C_1 2^p \leq \langle \xi \rangle \leq C_2 2^p$, we have

$$\left| D^\gamma_\xi \left[ \xi^\alpha \partial^\beta_x a_p(x, \xi) \right] \right| \leq C_{\alpha, \beta, \gamma} 2^{(m+|\alpha| - |\gamma|)}$$

by the property of $\xi^\alpha \partial^\beta_x a_p(x, \xi)$ as a symbol of order $m + |\alpha|$, via applying Proposition 2.2(iii) to product of symbols $\xi^\alpha$ and $\partial^\beta_x a_p(x, \xi)$. Hence for $|\gamma| = M$ we have

$$|y^\gamma D^\delta_x D^\alpha_y a_p(x, y)| \leq C_{\alpha, \beta, \gamma} 2^{(N+m+|\alpha| - M)}$$

Let $C_{\alpha, \beta, M} = \sup_{|\gamma|=M} C_{\alpha, \beta, \gamma}$ and we have the inequality \(15\). \qed

24
Lemma 3.11 (Boundedness of decomposition). (i) Given $a \in S^m_{1,0}$, let $a_p(x, \xi) = a(x, \xi)\varphi_k(\xi)$ be the Littlewood-Paley decomposition of $a$. Then we have

$$\|a_p(x, D)\|_{L^p \to L^p} \leq C 2^{pm}$$

(16)

for $1 \leq \rho \leq \infty$, where $C$ depends not on $p$.

(ii) Let $b \in S^m(R^N \times R^N)$ and let $b_q(D, x) = \varphi_q(D)b(D, x)$, where $b(D, x)$ means

$$b(D, y)u(x) = \int_{R^N} \int_{R^N} e^{i(x-y)}b(y, \xi)u(y) \, dyd\xi,$$

that is, a right-quantised operator. Then for all $l$ we have

$$\|b_q(D, x)\|_{L^p \to L^p} \leq C 2^{qm}$$

(17)

for $1 \leq \rho \leq \infty$, where $C$ depends not on $p$.

Proof. For (i), we represent $a_p$ in form of integral kernel $k_p$:

$$a_p(x, D)u = \int_{R^N} k_p(x, x-y)u(y) \, dy,$$

We know from (15) that

$$|k_p(x, y)| \leq C_1,$$

$$|k_p(x, y)| \leq C_2 |y|^{-N-1} 2^{p(m-1)}$$

by taking $\alpha = \beta = 0$ and $M$ to be 0 and $N + 1$ respectively. Then

$$\int_{R^N} |k_p(x, y)| \, dy \leq \tilde{C} \left( \int_{|y| \leq 2^{-p}} 2^{p(N+m)} \, dy + \int_{|y| > 2^{-p}} |y|^{-N-1} 2^{p(m-1)} \, dy \right)$$

$$\leq \tilde{C} \left( 2^{pm} \sigma_N + C' 2^{p(m-1)} \right) \leq C2^{pm}.$$

And see

$$\left\| \int_{R^N} k_p(x, x-y)u(y) \, dy \right\|_{L^p} \leq \|u\|_{L^p} \sup_{x \in R^N} \int_{R^N} |k_p(x, y)| \, dy \leq C2^{pm} \|u\|_{L^p}$$

by Young’s inequality.

For (ii), we represent

$$b_q(D, x)u = \int_{R^N} k_q(y, x-y)u(y) \, dy,$$

where $k_q(x, y) = \mathcal{F}^{-1}_{\xi \rightarrow y}(b_q(x, \xi))$ is the same kernel for the left quantisation $b_q(x, D)$. By a very similar estimate the result (ii) holds. \qed
Lemma 3.12. Let $a \in S^m(\mathbb{R}^N \times \mathbb{R}^N)$ for $m \in \mathbb{R}$, and let $a_p(x, D)$ be the Littlewood-Paley decomposition of $a(x, D)$. Then for all $l$ we have

$$
\|\varphi_q(D)a_p(x, D)\|_{L^p \to L^q} \leq C_l 2^{\min(p,q)l} 2^{-|p-q|l}
$$

(18)

for any $p, q \geq -1$, for any $1 \leq \rho \leq \infty$ where $C_l$ is independent of $p, q$.

Proof. Firstly we decompose 1 such that

$$
1 = \left( \frac{1}{1 + |\xi|^2} + \sum_{j=1}^{n} \frac{\xi_j}{1 + |\xi|^2} \xi_j \right)^l = \sum_{|\alpha| \leq l} c_\alpha(\xi) \xi^\alpha,
$$

for $c_\alpha \in S^{-2l+|\alpha|} \subset S^{-l}$ by virtue of coefficients of order 1 and 2 in the binomial expansion. Hence

$$
\varphi_q(D)a_p(x, D) = \varphi_q(D) \sum_{|\alpha| \leq l} c_\alpha(D) D_x^\alpha a_p(x, D) = \sum_{|\alpha| \leq l} c_\alpha(D) \varphi_q(D) D_x^\alpha a_p(x, D)
$$

by commutativity of $c_\alpha(D)$ and $\varphi_q(D)$ for their symbols are location-invariant. Furthermore by considering that $D_x^\alpha a(x, D)$ is an operator of order $m + |\alpha|$ by Theorem 2.16 and hence

$$
D_x^\alpha a_p(x, D) = D_x^\alpha a(x, D) \varphi_p(x) = (D_x^\alpha a)_p (x, D).
$$

Invoke (16) for $(c_\alpha)_q$ and $(D_x^\alpha a)_p$ to see

$$
\|\varphi_q(D)a_p(x, D)\|_{L^p \to L^q} \leq \sum_{|\alpha| \leq l} \|c_\alpha(D) \varphi_q(D)\|_{L^p \to L^q} \left\| (D_x^\alpha a)_p (x, D) \right\|_{L^p \to L^q} \leq C 2^{-q} 2^{l+m}
$$

which is what we want when $p \leq q$.

When $p > q$, see that

$$
a(x, D) = (a(x, D)^*)^* = (a^*(x, D))^* = b(D, x), \quad b(y, \xi) = \overline{a^*(y, \xi)},
$$

and hence decompose

$$
\varphi_q(D)a_p(x, D) = \varphi_q(D)b(D, x) \varphi_p(D) = \sum_{|\alpha| \leq l} b_q(D, x) D_x^\alpha c_\alpha(D) \varphi_p(D).
$$

Now invoke (17) for $b_q(D, x) D_x^\alpha$ and $(c_\alpha)_p$ we shall get

$$
\|\varphi_q(D)a_p(x, D)\|_{L^p \to L^q} \leq \sum_{|\alpha| \leq l} \|b_q(D, x) D_x^\alpha\|_{L^p \to L^q} \|c_\alpha(D) \varphi_p(D)\|_{L^p \to L^q} \leq C 2^{-p} 2^{q(l+m)},
$$

as required for $p > q$. \qed
Theorem 3.13 (Operators on Hölder-Zygmund spaces). Let $a \in S^m(\mathbb{R}^N \times \mathbb{R}^N)$, $m > 0$, $s > 0$ such that $s + m > 0$. Then $a(x, D) : C^{s+m}_*(\mathbb{R}^N) \to C^s_*(\mathbb{R}^N)$ is a bounded linear operator.

Proof. Firstly decompose

$$u = \sum_{p=-1}^{\infty} u_p,$$

where $u_p = \varphi_p(D)u$ and $u \in S'$. Since by (4) we have $\varphi_p \varphi_q = 0$ for $|p - q| \geq 2$, and

$$a(x, D)u = \sum_{p=-1}^{\infty} a_p(x, D)u = \sum_{p=-1}^{\infty} a_p(x, D) (u_{p-1} + u_p + u_{p+1}) = \sum_{p=-1}^{\infty} a_p(x, D) \tilde{u}_p$$

for $\tilde{u}_p = u_{p-1} + u_p + u_{p+1}$ and $u_{-2} = 0$. Suppose $u \in C^{s+m}_*$ then

$$2^{s+m}p \|\tilde{u}_p\|_{L^\infty} \leq 2^{s+m}p (\|u_{p-1}\|_{L^\infty} + \|u_p\|_{L^\infty} + \|u_{p+1}\|_{L^\infty}) \leq C\|u\|_{C^{s+m}_*}$$

by (14) where $C = 2^{-(s+m)} + 1 + 2^{s+m}$. Apply (18) to get

$$2^{sq} \|\varphi_q(D) a_p(x, D) \tilde{u}_p\|_{L^\infty} \leq C_1 2^{sq + \min\{p, q\}m - |p-q|l - p(s+m)} 2^{p(s+m)} \|\tilde{u}_p\|_{L^\infty},$$

here after simplication,

$$2^{sq + \min\{p, q\}m - |p-q|l - p(s+m)} 2^{p(s+m)} = \begin{cases} 2^{-|p-q|(l-s)}, & p \leq q \\ 2^{-|p-q|(m+s+l)}, & p > q. \end{cases}$$

Now we choose $l$ large such that $l \geq s + 1$ and $m + s + l \geq 1$, we obtain

$$\|a(x, D)u\|_{C^s_*} = \sup_{q \geq 1} 2^{sq} \|\varphi_q(D) a(x, D)u\|_{L^\infty} \leq C \sup_{q \geq 1} \sum_{p \geq -1} 2^{-|p-q|l} 2^{p(s+m)} \|\tilde{u}_p\|_{L^\infty} \leq C\|u\|_{C^{s+m}_*}.$$
(iv) \[ \| \frac{d}{d\theta} S_\theta u \|_{C^2} \leq C \theta^{\alpha-\beta-1} \| u \|_{C^\beta}, \] for all \( \alpha, \beta. \)

Proof. Fix a cutoff \( \chi \in C^\infty_c(\mathbb{R}) \), with \( \chi = 1 \) in \( |\xi| \leq 1 \) and 0 on \( |\xi| \geq 2 \), and 0 \( \leq \chi \leq 1 \).

Set regularisation operators \( S_\theta u = \sum_{p \geq -1} \chi(2^p/\theta) u_p. \)

Given \( u \in C^\beta_\ast \), immediately we have \( (S_\theta u)_p = 0 \) when \( 2^p \geq 2\theta \), and

\[ \left\| (S_\theta u)_p \right\|_{L^\infty} \leq \| u \|_{C^\alpha_\ast} 2^{-p\beta} \leq C \| u \|_{C^\beta_\ast} 2^{-p\alpha} \]

for \( \alpha < \beta \) (\( C \) is used to compensate the case \( p = -1 \)): in this case we have

\[ \| S_\theta u \|_{C^\beta_\ast} \leq C \| u \|_{C^\beta_\ast}, \]

proving (i); when \( \alpha > \beta \), we have

\[ \left\| (S_\theta u)_p \right\|_{L^\infty} \leq 2^{p(\alpha-\beta)} \| u \|_{C^\alpha_\ast} 2^{-p\alpha} \leq C \theta^{\alpha-\beta} \| u \|_{C^\beta_\ast} 2^{-p\alpha}, \]

for \( 2^p \leq \theta \) or \( (S_\theta u)_p = 0 \), and hence

\[ \| S_\theta u \|_{C^\alpha_\ast} \leq C \theta^{\alpha-\beta} \| u \|_{C^\beta_\ast}, \]

implying (ii). Observe \( (S_\theta - 1) u = \sum_p (\chi(2^p/\theta) - 1) u_p \), and for \( 2^p \leq 2\theta \) we have \( (S_\theta - 1) u = 0 \), and for \( 2^p \geq 2\theta \)

\[ \left\| (S_\theta u - u)_p \right\|_{L^\infty} \leq 2^{p(\alpha-\beta)} \| u \|_{C^\alpha_\ast} 2^{-p\alpha} \leq C \theta^{\alpha-\beta} \| u \|_{C^\beta_\ast} 2^{-p\alpha}, \]

and hence

\[ \| S_\theta u - u \|_{C^2} \leq C \theta^{\alpha-\beta} \| u \|_{C^\beta_\ast} \]

proving (iii). Eventually we see \( \frac{d}{d\theta} S_\theta u = \theta^{-1} \sum_p \chi_1(2^p/\theta) u_p \), where \( \chi_1(\xi) = -\xi \chi'(\xi) \). Note that the support of \( \chi_1 \) is within \( 1 \leq |\xi| \leq 2 \), hence \( \theta \leq 2^p \leq 2\theta \) or \( \frac{d}{d\theta} S_\theta u = 0 \). Then similarly

\[ \left\| \left( \frac{d}{d\theta} S_\theta u \right)_p \right\|_{L^\infty} \leq \theta^{-1} 2^{p(\alpha-\beta)} \| u \|_{C^\alpha_\ast} 2^{-p\alpha} \leq C \theta^{\alpha-\beta-1} \| u \|_{C^\beta_\ast} 2^{-p\alpha} \]

and

\[ \left\| \frac{d}{d\theta} S_\theta u \right\|_{C^2} \leq C \theta^{\alpha-\beta-1} \| u \|_{C^\beta_\ast}, \]

as in (iv).

\[ \square \]

Remark. We will use this family of regularisation operators with restriction \( \alpha, \beta \notin \mathbb{N} \).
Note in our case, \( C^\alpha_\ast, C^\beta_\ast \) are \( C^\alpha, C^\beta \) respectively. Also note that \( \cap_{\beta \geq 0} C^\beta_\ast = \cap_{\beta \geq 0} C^\beta = C^\infty. \)
3.4 Dyadic Analysis of Products

Lemma 3.15. Let \((a_q)_{q \geq -1}\) be a sequence of functions such that
\[
supp \hat{a}_q \subset \{\xi, |\xi| \leq C_0 2^q\}.
\]
Suppose that \(\|a_q\|_{L^\infty} \leq M 2^{-q\alpha}\) for some \(\alpha > 0\), then \(u = \sum_{q \geq -1} a_q \in C^\alpha\) with \(\|u\|_{C^\alpha} \leq CM\).

Proof. Observe that for some \(N\) we have
\[
u_p = \sum_{q \geq p - N} (a_q)_p
\]
and hence
\[
\|u_p\|_{L^\infty} \leq \sum_{q \geq p - N} \|(a_q)_p\|_{L^\infty} \leq C \sum_{q \geq p - N} \|a_q\|_{L^\infty} \leq C \sum_{q \geq p - N} M 2^{-q\alpha} \leq CM 2^{-p\alpha},
\]
by (8). Clearly we have what is required.

Theorem 3.16 (Estimates for products). Given \(u, v \in C^\alpha\), we have
\[
\|uv\|_{C^\alpha} \leq C \left(\|u\|_{C^0} \|v\|_{C^\alpha} + \|u\|_{C^\alpha} \|v\|_{C^0}\right).
\]

Proof. Decompose \(u\) and \(v\) in Littlewood-Paley formulation:
\[
u = \sum_p u_p, \quad v = \sum_p v_p.
\]
Then we have
\[
u v = \sum_{p, q} u_p v_q = \sum_q (S_q u) v_q + \sum_p u_p (S_{p+1} v) = \Sigma_1 + \Sigma_2.
\]
Note that supports of Fourier transforms of \(\Sigma_1\) and \(\Sigma_2\) are both in the ball \(\{\xi \leq C_0 2^p\}\) for some \(C_0\). This enables us to apply the lemma, by the help of the following estimates:
\[
\|(S_q u) v_q\|_{L^\infty} \leq \|S_q u\|_{L^\infty} \|v_q\|_{L^\infty} \leq M \|u\|_{L^\infty} \|v\|_{C^\alpha} 2^{-q\alpha},
\]
\[
\|u_p (S_{p+1} v)\|_{L^\infty} \leq \|u_p\|_{L^\infty} \|S_{p+1} v\|_{L^\infty} \leq M \|u\|_{C^\alpha} 2^{-p\alpha} \|v\|_{L^\infty},
\]
by (8), and hence we have by previous lemma
\[
\|\Sigma_1\|_{C^\alpha} \lesssim \|u\|_{L^\infty} \|v\|_{C^\alpha},
\]
\[
\|\Sigma_2\|_{C^\alpha} \lesssim \|u\|_{C^\alpha} \|v\|_{L^\infty}
\]
and
\[
\|uv\|_{C^\alpha} \leq \|\Sigma_1\|_{C^\alpha} + \|\Sigma_2\|_{C^\alpha} \leq C \left(\|u\|_{C^0} \|v\|_{C^\alpha} + \|u\|_{C^\alpha} \|v\|_{C^0}\right)
\]
as required.

Remark. There is an immediate adaption of this theorem to \(C^\alpha\) for \(\alpha \notin \mathbb{N}\), by restricting \(C^\alpha\) to \(\alpha \notin \mathbb{N}\). We will apply this dyadic result repeatedly in Section 4 and Section 5.
4 Hölder Scale Nash-Moser Theorem

4.1 Preliminaries

4.1.1 Fréchet Derivative

In the setting for the Nash-Moser theorem we set $M$ to be a compact $C^\infty$ manifold, and $\Phi$ a mapping from an open set $U \subset C^\infty(M, \mathbb{R}^p)$ to $C^\infty(M, \mathbb{R}^q)$. We would like to talk about differentiability of $\Phi$ as a mapping. Given two topological vector spaces $E_1$ and $E_2$ on $\mathbb{R}$, $U$ open in $\mathbb{R}$, we say a continuous mapping $\Phi : U \rightarrow E_2$ is of class $C^1$ if there is a continuous linear in second variable:

$$\Phi' : U \times E_1 \rightarrow E_2; \quad (u, v) \mapsto \Phi'(u)v,$$

such that

$$\forall u \in U, \forall v \in E_2, \lim_{t \to 0} t^{-1}(\Phi(x + tv) - \Phi(x)) = \Phi'(u)v.$$

Hence we define $C^k$ mappings recursively, via that $\Phi$ is $C^k$ is $\Phi'(u)$ is $C^{k-1}$ for all $u \in U$. Write

$$\Phi^{(k)}(u)(v_1, \ldots, v_k) = \frac{\partial^k}{\partial t_1 \cdots \partial t_k} \Phi (u + t_1v_1 + \cdots + t_kv_k)|_{t_1=\cdots=t_k=0}.$$

4.1.2 Tame Mappings

We also need to introduce the concept of a tame mapping. Firstly, note that a neighbourhood of $u_0$ in $C^\infty(M)$ is a neighbourhood of $u_0$ in some $C^\mu(M)$, we call the neighbourhood a $\mu$-neighbourhood of $u_0$. Furthermore we need to fix a decreasing family $E_s$ with norm $|\cdot|_s$, such that $\cap_s E_s = C^\infty$, for the very moment $E_s = C^s$.

**Definition 4.1** (Tame mapping). Let $u \rightarrow \Phi(u)$ be a continuous mapping from an open set $U$ in $C^\infty(M, \mathbb{R}^p)$ to $C^\infty(M, \mathbb{R}^q)$. Then $\Phi$ is said to be tame if, for any $u_0 \in U$, there is a neighbourhood $V$ of $u_0$, numbers $b \geq 0$ and $r \geq 0$, and constants $C_s$ such that for all $u \in V$, for all $s \geq b$ we have

$$|\Phi(u)|_s \leq C_s \left(1 + |u|_{s+r}\right),$$

where the constant $C_s$ depend on $s$, $u_0$ and $V$, numbers $b, r$ may only depend on $u_0$ and $V$.

4.1.3 Preliminary Lemma

**Lemma 4.2.** Given a family of $C^\infty$ functions $(p_\theta)$ with parameter $\theta > \theta_0$, and given that

$$\|p_\theta\|_{C^{a_j}} \leq M \theta^{b_j - 1}, \quad j = 1, 2$$

for some $b_1 < 0 < b_2$, $a_1 < a_2$, then if we interpolate in between $a_1$ and $a_2$ by

$$a = \nu a_1 + (1 - \nu) a_2$$

30
with respect to parameter $\nu$ defined by $\nu b_1 + (1 - \nu)b_2 = 0$, we have

$$p = \int_{\theta_0}^{\infty} p_\theta \, d\theta \in C^a, \quad \|p\|_{C^a} \leq C_a M$$

for $a_1, a_2, a \notin \mathbb{N}$.

**Proof.** It suffices to prove the theorem for $a_1, a_2$ between 0 and 1 (indeed, we can apply convexity inequality (13) to reduce the general case firstly to that between two adjacent integers $k$ and $k + 1$, then by considering all $k$-th order derivatives to reduce the problem back to case 0 and 1). Let

$$v_\theta = \int_{\theta_0}^{\theta} p_\tau \, d\tau, \quad w_\theta = \int_{\theta}^{\infty} p_\tau \, d\tau.$$ 

Integrate the given estimates we have

$$\|v_\theta\|_{C^{a_2}} \leq M \theta^{b_2}/b_2, \quad \|w_\theta\|_{C^{a_1}} \leq -M \theta^{b_1}/b_1.$$ 

If $|x - y| = \delta$ it follows that

$$|p(x) - p(y)| \leq M \left( -\theta^{b_1} \delta^{a_1}/b_1 + \theta^{b_2} \delta^{a_2}/b_2 \right).$$

For sufficiently small $\delta$ we can choose $\theta > \theta_0$ so that two terms in the parentheses are equal and it then follows that

$$|u(x) - u(y)| \leq 2M \delta^a |b_1|^{-\lambda} |b_2|^{\lambda-1}$$

and we see $\|u\|_{C^a} \leq 2M |b_1|^{-\lambda} |b_2|^{\lambda-1}$ as required. \qed

**Remark.** This important lemma fails when $a \in \mathbb{N}$. This is exactly why we have to exclude Hölder spaces of integer indices.

### 4.2 Statement of Theorem

#### 4.2.1 Assumptions

We make two assumptions on operator $\Phi$, defined on a $\mu$-neighbourhood of $u_0$ in $C^\infty(M, \mathbb{R}^p)$:

$(A_1)$ $\Phi$ is of class $C^2$ in the previous settings for operators, in a $\mu$-neighbourhood of $u_0$ and satisfies the tame estimate

$$\|\Phi''(u)(v_1, v_2)\|_{C^a} \leq C \left\{ \|v_1\|_{C^a} \|v_2\|_{C^a} (1 + \|u\|_{C^{a+d}}) \right. \right. \left. \right. \left. \right. \right. \left. \right. \left. \right. \left. \right. + \|v_1\|_{C^a} \|v_2\|_{C^{a+c}} + \|v_1\|_{C^{a+c}} \|v_2\|_{C^a} \right\},$$

for some $a, b, c \geq 0$ and all $\alpha \geq 0$.

$(A_2)$ For $u$ in a $\mu'$-neighbourhood of $u_0$ in $C^\infty(M, \mathbb{R}^p)$, there is a linear mapping $\psi(u) : C^\infty(M, \mathbb{R}^p) \to C^\infty(M, \mathbb{R}^p)$ such that $\Phi'(u)\psi(u) = \text{id}$, satisfying the tame estimate

$$\|\psi(u)g\|_{C^a} \leq C \left\{ \|g\|_{C^{a+d}} + \|g\|_{C^a} (1 + \|u\|_{C^{a+d}}) \right\},$$

for some numbers $\lambda, d \geq 0$ and all $\alpha \geq 0$. 

31
4.2.2 Statement

We now can state our main theorem of the essay, the Nash-Moser theorem, based on Hölder spaces.

Theorem 4.3 (Hölder scale Nash-Moser theorem). Suppose \( \Phi \) satisfies \((A_1)\) and \((A_2)\), and \( \alpha \) is such that
\[
\alpha \geq \mu, \quad \alpha \geq \mu', \quad \alpha \geq d, \quad \alpha > \lambda + a + c, \quad \alpha > a + 1/2 \sup(\lambda + b, 2a), \quad \alpha \notin \mathbb{N}.
\]
Then the followings hold:

(i) There is a \((\alpha + \lambda)\)-neighbourhood \( W \) of the origin such that for \( f \in W \), the equation
\[
\Phi(u) = \Phi(u_0) + f
\]
has a solution \( u = u(f) \in C^\alpha \). Moreover we have
\[
\|u(f) - u_0\|_{C^\alpha} \leq C\|f\|_{C^\alpha + \lambda}.
\]

(ii) If there is \( \alpha' > \alpha \), a non-integer, such that \( f \in W \cap C^{\alpha'+\lambda} \), then the solution \( u(f) \) constructed is in \( C^{\alpha'} \). In particular if \( f \in W \cap C^\infty \) then \( u(f) \in C^\infty \).

Remark. What makes this theorem novel? (i) It is applied to a mapping between two Fréchet spaces with a local right inverse, (ii) where the mapping and its right inverse are tame, that is, they are allowed to have fixed losses in derivatives. Those facts make a standard implicit function impracticable, and thereby lie the novelty of Nash-Moser theorem.

4.3 Iteration Scheme

4.3.1 Intuition From Original Newton Scheme

We aim to solve the perturbation problem
\[
\Phi(u) = \Phi(u_0) + f
\]

The normal Newton scheme implies in iteration scheme
\[
\Phi(u_{k+1}) = \Phi(u_k) + \Phi'(u_k) \Delta u_k + \varepsilon_k
\]

A delicate control over \( \Delta u_k = u_{k+1} - u_k \) is needed, so that the accumulated error is controlled by latest \( \varepsilon_k \). That is, by controlling \( \Delta u_n \) we let the error at each step \( n \)
\[
\Phi(u_{n+1}) = \Phi(u_0) + f + \varepsilon_n.
\]
By denoting each $\gamma_k = \Phi'(u_k)\Delta u_k$, by a recursive substitution of $\Phi(u_k)$ by $\Phi(u_{k-1})$ in formula (23) we see representation

$$\Phi(u_{n+1}) = \Phi(u_0) + \sum_{k=0}^{n} \gamma_k + \sum_{k=0}^{n} \varepsilon_k.$$ 

And hence the control is exactly

$$\sum_{k=0}^{n} \gamma_k + \sum_{k=0}^{n} \varepsilon_k = f + \varepsilon_n,$$ 

(25)

and we construct iteratively

$$\gamma_n = f - \sum_{k=0}^{n-1} \gamma_k - \sum_{k=0}^{n-1} \varepsilon_k$$

hence

$$\Phi(u_{n+1}) = \Phi(u_0) + f + \varepsilon_n \rightarrow \Phi(u_0) + f$$

$$\gamma_n = f + \Phi(u_0) - \Phi(u_n) = \varepsilon_{n-1} \rightarrow 0$$

as $\varepsilon_n \rightarrow 0$.

**Remark** (Failure of Newton scheme). Unfortunately in our case, the simple Newton scheme does not work, for even if we assume information $\|\gamma_0\|_{C^\beta}$ for $\beta \in [0, \beta^+]$, however, in the process of inverting $\gamma_k$ to $\Delta u_k = \psi(u_k)$, by estimate (A2) (21) we can only obtain a control over $\|\Delta u_k\|_{C^\beta}$ for $\beta \in [0, \beta^+ - \lambda]$: in plain english, in each iteration, we lose $\lambda$ order of information, and eventually lose all in finite time. Moreover there is another loss of $b$ order derivatives in the process of obtain $\varepsilon_k$, generated by $\Phi$ in (A1) (20). It is exactly the loss of all information in finite time that prevents us from running the Newton scheme.

### 4.3.2 Symbolism and Analysis of Nash-Moser Scheme

However a modified Newton scheme can be used. We split our argument into several steps for a more approachable explanation.

The first step is to decompose the perturbation $f$. From Proposition 3.14 there exists a family of regularisation operators $S_\theta$ with following properties

$$\|S_\theta u\|_{C^\alpha} \leq C\|u\|_{C^\beta}, \quad \text{for } \alpha \leq \beta,$$ 

(26)

$$\|S_\theta u\|_{C^\alpha} \leq \theta^{\alpha-\beta}C\|u\|_{C^\beta}, \quad \text{for } \alpha \geq \beta,$$ 

(27)

$$\|S_\theta u - u\|_{C^\alpha} \leq C\theta^{\alpha-\beta}\|u\|_{C^\beta}, \quad \text{for } \alpha \leq \beta,$$ 

(28)

$$\left\| \frac{d}{d\theta} S_\theta u \right\|_{C^\alpha} \leq C\theta^{\alpha-\beta-1}\|u\|_{C^\beta}, \quad \text{for all } \alpha, \beta.$$ 

(29)
Note that here we replace the Hölder-Zygmund spaces by Hölder spaces, this makes perfect sense when $\alpha, \beta$ are non-integers. Pick the sequence

$$\theta_n = \left( \theta_0^{1/\varepsilon} + n \right)^{\varepsilon},$$

(30)

where we choose parameter $\theta_0$ smartly later, we have the decomposition

$$f = S_{\theta_0} f + \sum_{k \geq 0} \Delta_k \dot{f}_k,$$

in which

$$\Delta_k = \theta_{k+1} - \theta_k,$$

$$\dot{f}_k = \frac{1}{\Delta_k} \left( S_{\theta_{k+1}} - S_{\theta_k} \right) f.$$

(31)
(32)

Note that by picking $\theta_n = C2^n$ we would obtain the usual Littlewood-Paley decomposition of $f$, but be alerted that here we certainly can decompose $f$ in a different way. A smart choice of $\Delta_k$, the step from $\theta_k$ to $\theta_{k+1}$, is essential to the success of this iteration scheme, discussed later. Here one should also note we have an estimate for each block $\dot{f}_k$:

$$\left\| \dot{f}_k \right\|_{C^s} \leq C \theta_k^{s-\alpha-1} \left\| f \right\|_{C^\alpha},$$

by property (29).

The second step is explication of three alterations to our original iteration formula and control of error in Newton scheme. (i) The first alteration, to the iteration formula, is that, in each iteration when we use $\Phi(u_k)$ and $\Phi'(u_k)$ to search for $\Delta_k u_k$, we use $\Phi'(v_k)$ in place of $\Phi'(u_k)$, where $v_k = S_{\theta_k} u_k$ is a regularisation of $u_k$, that is, the iteration formula becomes

$$\Phi(u_{n+1}) = \Phi(u_n) + \Phi'(v_n) \Delta u_n + \varepsilon_n.$$

This clever alteration moves the losses in $u_k$ to $v_k$, all order derivatives of which are controlled by fixed order derivatives of $u_k$; that is, assume information of $\beta$ order derivative of $u_k$ and we can control the $\beta + \lambda + b$ order derivative of $v_{k+1}$ by the properties of $S_{\theta_k}$, and the loss of derivatives in $v_k$ will tell us that $\beta$ order derivative of $u_k$ is given by $\beta + \lambda + b$ order derivative of $v_k$, controlled by $\beta$ order derivative of $u_k$. The loss of derivatives are recovered! (ii) The second change, to the control of error, is that, in our old scheme, we try find iteration for $u_n$ such that each $\Phi(u_n)$ comes closer to the same target, $\Phi(u_0) + f$, as in (24). In the new scheme, we set a different target for each $u_n$, that is, we start with easy targets, and after each round of iteration we increase the difficulty of next target. To be specific, we want to control

$$\Phi(u_{n+1}) = \Phi(u_0) + S_{\theta_n} f + n^{\text{th}} \text{ remainder}.$$
(iii) The last modification, to the control of error, is that for the $n^{th}$ remainder, instead of a simple estimate with $\varepsilon_n$ only, we ask the control to be

$$
\Phi(u_{n+1}) = \Phi(u_0) + S_{\theta_n} f + \varepsilon_n + (1 - S_{\theta_n}) E_n,
$$

where

$$
E_n = \sum_{k=0}^{n-1} \varepsilon_k
$$

is the accumulated error up to step $n - 1$. That is, not like the situation of (25), in which all error of previous steps are cancelled by $\sum_{k=0}^{n} \gamma_k$, we keep a very small bit of the accumulated errors.

The third step is to write down the iteration scheme explicitly. Starting with known $u_0, \ldots, u_n$, we want to determine $u_{n+1}$ via solving

$$
\Phi(u_{n+1}) = \Phi(u_n) + \Phi'(v_n) \Delta u_n + \varepsilon_n,
$$

where $\Delta u_n = u_{n+1} - u_n$, via choosing $\gamma_n = \Phi'(v_n) \Delta u_n$ smartly and then by $\Delta u_n = \psi'(v_n) \gamma_n$. We want to pick $\gamma_n$ to be such that

$$
\Phi(u_{n+1}) = \Phi(u_0) + S_{\theta_n} f + \varepsilon_n + (1 - S_{\theta_n}) E_n,
$$

which is for each $n$,

$$
\sum_{k=0}^{n} \gamma_k + S_{\theta_n} E_n = S_{\theta_n} f.
$$

More explicitly, we want

$$
\gamma_n = (S_{\theta_n} - S_{\theta_{n-1}}) (f - E_{n-1}) - S_{\theta_n} \varepsilon_{n-1},
$$

for each $n \geq 1$. By iterating we can find $u_k$ corresponding to errors $\varepsilon_k$ for all $k$, and we want to pick those $\varepsilon_k$ in a smart way.

The fourth step is a rough analysis of errors. In this modified Newton scheme, we have two sources of errors, the first of which is from the errors we set deliberately, that is, usual quadratic error in original Newton scheme; the second of which is the error occurring in the process of replacing $\Phi'(u_k)$ by $\Phi'(v_k)$. We want our errors $\varepsilon_k$ to be decomposed into such two categories of errors, that, $\varepsilon_k = \varepsilon'_k + \varepsilon''_k$:

$$
\varepsilon'_k = \{ \Phi'(u_k) - \Phi'(v_k) \} \Delta u_k,
$$

$$
\varepsilon''_k = \Phi(u_{k+1}) - \Phi(u_k) - \Phi'(u_k) \Delta u_k,
$$

where $\varepsilon'_k$ is the error by regularising $u_k$ by $v_k$ and $\varepsilon''_k$ is the quadratic error in original Newton scheme. We should take into account the nature of the modification $\varepsilon'_k$ when picking $\varepsilon_k$. 

35
The fifth step is a clarification of symbolism: we want to define a set of variables according to the style of decomposition (31) and (32). Set

\[
\dot{u}_n = \Delta^{-1}_n \Delta u_n, \quad g_n = \Delta^{-1}_n \gamma_n, \quad e_n = \Delta^{-1}_n \varepsilon_n, \\
e'_n = \Delta^{-1}_n \varepsilon'_n = \{\Phi'(u_n) - \Phi'(v_n)\} \dot{u}_n, \\
e''_n = \Delta^{-1}_n \varepsilon''_n = \Delta^{-1}_n \{\Phi(u_{n+1}) - \Phi(u_n) - \Phi'(u_n) \Delta u_n\}. 
\]

(33)

with

\[
\Delta_n = \theta_{n+1} - \theta_n, \\
\dot{f}_n = \frac{1}{\Delta_n} (S_{\theta_{n+1}} - S_{\theta_n}) f.
\]

and introduce the same decomposition of accumulated error

\[
\dot{E}_n = \frac{1}{\Delta_n} (S_{\theta_{n+1}} - S_{\theta_n}) E_n.
\]

Till this point, the setting of the iteration scheme is done.

4.3.3 Explicit Iteration Scheme

Here we rewrite our Nash-Moser scheme in a clear and simplified algorithmic way. For the initial setup we write

\[
g_0 = \Delta^{-1}_0 S_{\theta_0} f,
\]

and for the iteration we write with \(k \geq 0\):

\[
v_k = S_{\theta_k} u_k \\
\dot{u}_k = \psi(v_k) g_k \\
\dot{u}_{k+1} = u_k + \Delta_k \dot{u}_k \\
e_k = \Delta^{-1}_k \{\Phi(u_{k+1}) - \Phi(u_k) - \Phi'(v_k) \dot{u}_k\} \\
g_{k+1} = \Delta^{-1}_{k+1} \left((S_{\theta_{k+1}} S_{\theta_k}) (f - E_k) - \Delta_k S_{\theta_{k+1}} e_k\right).
\]

This scheme is just a reader-friendly version of the Nash-Moser scheme.

4.4 Iterative Hypotheses

We now want to make a family of hypotheses, to make sure that we have convergence of the sequence \( \{u_n\} \) under the hypothesis that \( \|f\|_{C^{\alpha+1}} \) is sufficiently small, where \( \alpha \) is given in statement of Theorem 4.3 and \( \lambda \) is given in assumption \((A_2)\) (21). Firstly we are given some small \( \delta > 0 \) and some \( \alpha > \alpha \), both of which will be fixed later. Set the hypothesis
(Hₙ) For 0 ≤ k ≤ n and s ∈ [0, ̄α], we have
\[ \| \dot{u}_k \|_{C^s} \leq \delta \theta_k^{s-\alpha-1}. \]

We now prove the iteration hypothesis \((H_n)\) for all \(n\), and we use an iterative argument, by assuming \((H_n)\) holds in our case and try to derive \((H_{n+1})\). Note that it is trivial that \((H_n) \implies (H_{n-1})\).

### 4.4.1 Validity of Tame Estimates

Firstly, we claim given \(\alpha \geq \mu, \alpha \geq \mu', 0 < \delta < \delta_0, \delta_0 \) sufficiently small and \(\theta_0 \) sufficiently large, \((H_n)\) implies that \(u_{n+1}\) and \(v_{n+1}\) are both in a sufficiently small ball around \(u_0\), where we can use assumptions \((A_1)\) and \((A_2)\). This is done in the following lemma.

**Lemma 4.4.** \((H_n)\) implies the followings:

1. \(\|u_{n+1} - u_0\|_{C^s} \leq C_\alpha \delta\)
2. \(\|S_\theta_{n+1}(u_{n+1} - u_0)\|_{C^s} \leq C_\alpha \delta \theta_{n+1}^{(a-\alpha)+}\) for \(a\) in a finite interval,
3. \(\| (1 - S_\theta_{n+1}) (u_{n+1} - u_0) \|_{C^s} \leq C_\alpha \delta \theta_{n+1}^{(a-\alpha)}, \) for \(0 \leq a \leq \bar{\alpha}\),
4. \(\|u_{n+1} - v_{n+1}\|_{C^s} \leq C_\alpha \theta_{n+1}^{(a-\alpha)+}\) for \(a\) in a finite interval,
5. \(\|v_{n+1}\|_{C^s} \leq C_\alpha \theta_{n+1}^{(a-\alpha)+}\) for \(a\) in a finite interval,

where \(x_+ = \sup \{x, 0\}\).

**Proof.** Set \(U = u_{n+1} - u_0 = \sum_{k=0}^n \Delta_j \dot{u}_j\). Set a continuous family of functions \(p_\theta\) with

\[ p_\theta = \begin{cases} \dot{u}_k, & \theta_k \leq \theta < \theta_{k+1}, \quad k \leq n \\ 0, & \theta_{n+1} \leq \theta. \end{cases} \]

Moreover we can establish an estimate on \(\|p_\theta\|_{C^s}\):
\[
\|p_\theta\|_{C^s} \leq (1 + 2^\varepsilon (a+1)) \delta \theta^{s-\alpha-1};
\]
indeed, we have \(\|p_\theta\|_{C^s} \leq \delta \theta_k^{s-\alpha-1} \leq \delta \theta^{s-\alpha-1}\) when \(s - \alpha - 1 \geq 0\), and
\[
\|p_\theta\|_{C^s} \leq 2^\varepsilon (a+1) \delta \theta_{k+1}^{s-\alpha-1} \leq 2^\varepsilon (a+1) \delta \theta^{s-\alpha-1}
\]
by the fact that \(\theta_k < \theta_{k+1} < 2^\varepsilon \theta_k\) as an immediate result from \((30)\), and that \(s - \alpha - 1 \geq -\alpha - 1\). By \((35)\) invoke the previous lemma with
\[
a_1 = \alpha - \rho, \quad a_2 = \alpha + (1 - \rho), \quad b_1 = -\rho, \quad b_2 = 1 - \rho.
\]
for some \( \rho \in (0, 1) \) such that \( a_1, a_2 \not\in \mathbb{N} \). Set \( \nu = 1 - \rho \) and see \( a = \alpha \), then we have

\[
U = \sum_{k=0}^{n} \Delta_j u_j = \int_{\theta_0}^{\theta_{n+1}} p_\theta \, d\theta \in C^\alpha
\]

and \( \|U\|_{C^\alpha} \leq C\delta \), hence (i) is clear. By property of the regularisation operators \( (28) \) we have for \( a \leq \alpha \):

\[
\| (1 - S_{\theta_{n+1}}) U \|_{C^\alpha} \leq C\delta^a_{n+1}.
\]

Consider the case when \( \alpha < a \leq \bar{\alpha} \):

\[
\|U\|_{C^\alpha} \leq \delta \sum_{k=0}^{n} \Delta_k \delta^a_{k-\alpha-1} = \delta \int_{\theta_0}^{\theta_{n+1}} p_{\theta}^{a-\alpha-1} \leq C'\delta^a_{n+1}.
\]

Furthermore by \( (28) \) we have

\[
\| (1 - S_{\theta_{n+1}}) U \|_{C^\alpha} \leq C''\|U\|_{C^\alpha} \leq C\delta^a_{n+1},
\]

and (iii) is complete. However by virtue of decomposition

\[
u_{n+1} - v_{n+1} = u_0 + U - S_{\theta_{n+1}} (u_0 + U) (1 - S_{\theta_{n+1}}) u_0 + (1 - S_{\theta_{n+1}}) U
\]

we have

\[
\|\nu_{n+1} - v_{n+1}\|_{C^\alpha} \leq \| (1 - S_{\theta_{n+1}}) u_0 \|_{C^\alpha} + \| (1 - S_{\theta_{n+1}}) U \|_{C^\alpha} \leq (C'\|u_0\|_{C^\alpha} + C''\|u_0\|_{C^\alpha}) \theta^a_{n+1} + C'''\delta\theta^a_{n+1} = C\theta^a_{n+1},
\]

by virtue of (iii); indeed by \( (28) \) we have

\[
\| (1 - S_{\theta_{n+1}}) u_0 \|_{C^\alpha} \leq C'\theta^a_{n+1} \|u_0\|_{C^\alpha}, \text{ when } a \leq \alpha,
\]

\[
\| (1 - S_{\theta_{n+1}}) u_0 \|_{C^\alpha} \leq C'\theta^a_{n+1} \|u_0\|_{C^\alpha} \leq C'\theta^a_{n+1} \|u_0\|_{C^\alpha}, \text{ when } \alpha < a \leq \bar{a},
\]

by virtue of \( \theta^a_{n+1} \leq \theta^a_{n+1} \). However note that we need information on \( \bar{a} \) which is arbitrary but finite, and the informative interval for \( a \) is \( (0, \bar{a}] \). This completes (iv). By \( (21) \) and \( (27) \) (ii) follows immediately, and by considering

\[
\nu_{n+1} = S_{\theta_{n+1}} (u_0 + U)
\]

and estimate

\[
\|\nu_{n+1}\|_{C^\alpha} \leq \|S_{\theta_{n+1}} u_0\|_{C^\alpha} + \|S_{\theta_{n+1}} u_0\|_{C^\alpha} \leq (C'\delta + C''\theta^a_{n+1})
\]

by \( (21) \) and \( (27) \). \( \square \)

**Remark.** Here in this lemma we see from (i) that \( \|\nu_{n+1} - u_0\|_{C^\alpha}, \|\nu_{n+1} - u_0\|_{C^{\nu'}} \) are both less than \( C\delta \) by virtue of \( \mu, \mu' \leq \alpha \), arbitrarily small if we pick \( \delta \) small enough. Also we have

\[
\|\nu_{n+1} - u_0\|_{C^\alpha} \leq \|S_{\theta_{n+1}} (u_{n+1} - u_0)\|_{C^\alpha} + \| (1 - S_{\theta_{n+1}}) u_0 \|_{C^\alpha} \leq C\delta + C'\theta^{-1}_{n+1} \|u_0\|_{C_{\theta_{n+1}}} \leq C\delta + C'\|u_0\|_{C_{\theta_{n+1}}}
\]

is arbitrarily small if we pick small \( \delta \) and large \( \theta_0 \). Hence we can make use of conditions \( (A_1) \) and \( (A_2) \) on \( u_{n+1} \) and \( \nu_{n+1} \).
4.4.2 Estimates on $e_k$

Lemma 4.5 (Estimate on $e'_k$). For $0 \leq k \leq n$, $s \in [0, \tilde{\alpha} - \sup \{b, c\}]$, we have

$$\|e'_k\|_{C^s} \leq C\delta \theta_k^{L(s)-1},$$

(37)

where

$$L(s) = \sup \{s + a + c - 2\alpha, (s + b - \alpha)_{+} + 2a - 2\alpha\}$$

and note that $a, b$ and $c$ are fixed parameters in tame hypothesis $(A_1)$ (20).

Proof. We have

$$e'_k = \{\Phi'(u_k) - \Phi'(v_k)\} \dot{u}_k = \int_0^1 \Phi''(v_k + t(u_k - v_k)) (\dot{u}_k, u_k - v_k) \, dt.$$ By $(A_1)$ we have

$$\|e'_k\|_{C^s} \leq \int_0^1 \|\Phi''(v_k + t(u_k - v_k)) (\dot{u}_k, u_k - v_k)\|_{C^s} \, dt \leq C\|\dot{u}_k\|_{C^s} \|u_k - v_k\|_{C^a}$$

$$\left(1 + \sup_t \|v_k + t(u_k - v_k)\|_{C^{s+b}}\right) + C'\|\dot{u}_k\|_{C^s} \|u_k - v_k\|_{C^{s+c}} + C''\|\dot{u}_k\|_{C^{s+c}} \|u_k - v_k\|_{C^a}$$

$$\leq \tilde{C} \delta \theta_k^{L(s)-1}$$

by virtue of

$$\|\dot{u}_k\|_{C^s} \|u_k - v_k\|_{C^{s+c}} \lesssim \delta \theta_k^{s+a+c-2\alpha-1},$$

$$\|\dot{u}_k\|_{C^{s+c}} \|u_k - v_k\|_{C^a} \lesssim \delta \theta_k^{s+a+c-2\alpha-1}$$

using $(H_n)$ and Lemma 4.4 (iv) for $(H_{n-1})$, and by

$$\|\dot{u}_k\|_{C^s} \|u_k - v_k\|_{C^s} \left(1 + \sup_t \|v_k + t(u_k - v_k)\|_{C^{s+b}}\right) \leq C\delta \theta_k^{a-\alpha} \theta_k^{a-\alpha}$$

$$\left(\|v_k\|_{C^{s+b}} + \|u_k - v_k\|_{C^{s+b}}\right) \leq C\delta \theta_k^{2a-2\alpha-1} \left(C_1\theta_k^{(s+b-\alpha)_+} + C_2\theta_k^{s+b-\alpha}\right)$$

$$\leq C_3 \delta \theta_k^{(s+b-\alpha)_+ + 2a - 2\alpha - 1},$$

for large $\theta_0$, obtained by invoking Lemma 4.4 (v) for $(H_{n-1})$. □

Remark (Negligibility of $e''_k$). We see

$$e''_k = \Delta_k^{-1} \{\Phi(u_{k+1}) - \Phi(u_k) - \Phi'(u_k)\Delta u_k\} = \Delta_k \int_0^1 (1 - t)\Phi''(u_k + t\Delta_k \dot{u}_k)(\dot{u}_k, \dot{u}_k) \, dt.$$ and we estimate $e''_k$ similarly:

$$\|e''_k\|_{C^s} \leq C\Delta_k \delta \theta_k^{L(s)-1}$$

is negligible for $\Delta_k \leq \delta_k^{1-\varepsilon-1}$, when $\varepsilon$ and $\delta$ small enough, compared to $e'_k$. Hence we can use (37) to estimate whole of $e_k$.
4.4.3 Estimates on \( g_{n+1} \)

**Proposition 4.6** (Estimates on \( g_{n+1} \)). For \( s \in [0, s_0] \), with any \( s_0 > 0 \) finite, we have

\[
\| g_{n+1} \|_{C^s} \leq C \left( \delta \theta_{n+1}^{L(s)-1} + \theta_{n+1}^{s-a-\lambda-1} \| f \|_{C^{\alpha+\lambda}} \right).
\]

*Proof.* Consider the iterative formula

\[
g_{n+1} = \frac{\Delta_n}{\Delta_{n+1}} \left\{ \dot{f}_n - \dot{E}_n - S_{\theta_{n+1}} e_n \right\}.
\]

By (37) we have

\[
\left\| \dot{f}_n \right\|_{C^s} \leq C \theta_{n}^{s-a-\lambda-1} \| f \|_{C^{\alpha+\lambda}} \tag{38}
\]

and

\[
\left\| (\dot{E}_n)_n \right\|_{C^s} \leq C \theta_{n}^{s-r-1} \| E_n \|_{C^r},
\]

where \( r = \tilde{a} - \sup \{ b, c \} \). Now pick \( \tilde{a} \) large enough such that \( L(r) > 0 \) and \( L'(s) = 1 \) on \((r, s_0) \) (that is, when \( s \geq \alpha - b \) or \( s + c - a \geq (s + b - \alpha)_+ \)). We then have

\[
\| E_n \|_{C^r} \leq \sum_{k=0}^{n-1} \Delta_k \| e_k \|_{C^r} \leq \sum_{k=0}^{n-1} \Delta_k C \delta \theta_{n}^{L(r)-1} \leq C \delta \theta_{n}^{L(r)-1} \sum_{k=0}^{n-1} \Delta_k = C \delta \theta_{n}^{L(r)} \tag{39}
\]

and hence

\[
\left\| (\dot{E}_n)_n \right\|_{C^s} \leq C \theta_{n}^{s-r-1+L(r)} = C \delta \theta_{n}^{L(s)-1}
\]

for \( L(r) \leq L(s) + r - s \) when \( s \leq r \), and \( L(r) = L(s) + r - s \) when \( s > r \) for \( L'(s) = 1 \). We also have

\[
\| S_{\theta_{n+1}} e_n \|_{C^s} \leq C' \| e_n \|_{C^s} \leq C \delta \theta_{n}^{L(s)-1}
\]

by (27) and previous estimate on \( e_n \). Summing up three estimates we have the required estimates for \( g_{n+1} \). \( \square \)

4.4.4 Estimates on \( \dot{u}_{n+1} \)

**Proposition 4.7** (Estimates on \( \dot{u}_{n+1} \)). For \( s \in [0, \tilde{a}] \), we have the estimate \( (H_{n+1}) \), that

\[
\| \dot{u}_{n+1} \|_{C^s} \leq \delta \theta_{k}^{s-a-1},
\]

if \( \delta, \| f \|_{C^{\alpha+\lambda}, \varepsilon} \) are sufficiently small, and \( \tilde{a}, \theta_0 \) sufficiently large.

*Proof.* With formula for \( \dot{u}_{n+1} \),

\[
\dot{u}_{n+1} = \psi'(v_{n+1}) g_{n+1},
\]

apply tame estimate \( (A_2) \) [21] to obtain

\[
\| \dot{u}_{n+1} \|_{C^s} \leq C \{ \| g_{n+1} \|_{C^{\alpha+\lambda}} + \| g_{n+1} \|_{C^{\lambda}} (1 + \| v_{n+1} \|_{C^{\varepsilon+d}}) \}
\]

\[
\leq C' \delta \theta_{n+1}^{L(s+\lambda)-1} + C' \theta_{n+1}^{s-a-1} \| f \|_{C^{\alpha+\lambda}} + C' \left( \delta \theta_{n+1}^{L(\lambda)-1} + \| f \|_{C^{\alpha+\lambda} \theta_{n+1}^{a-1}} \right) \theta_{n+1}^{(s+d-a)\varepsilon}. \tag{40}
\]

It is immediate to see if the followings hold:
(i) \( L(s + \lambda) < s - \alpha, \)
(ii) \( \|f\|_{C^{\alpha + \lambda}}/\delta \) is small,
(iii) \( L(\lambda) + (s + d - \alpha)_+ < s - \alpha, \)
(iv) \( (s + d - \alpha)_+ - \alpha \leq s - \alpha, \)
(v) \( \theta_0 \) is sufficiently large,

then \((H_{n+1})\) is satisfied. Note that \( \theta_0 \) is made large here to cancel out the constants with one order of polynomial decreases, to meet the form of \((H_{n+1})\). Condition \( \alpha \geq d \) validates (iv). Conditions \( \alpha > \lambda + a + c \) and \( \alpha > a + 1/2 \sup \{\lambda + b, 2a\} \) imply \( L(\lambda) < -\alpha \), hence implies (i), and together with (iv) imply (iii). All what is left is to pick \( \|f\|_{C^{\alpha + \lambda}} \) small.

Remark. Note here the parameters are chosen depending on \( a, b, c, d, \alpha \) only, not depending on \( u_n \). Till now every parameter is fixed, except \( \|f\|_{C^{\alpha + \lambda}} \).

4.4.5 Validation of Case \((H_0)\)

Proposition 4.8 (Validation of case \((H_0)\)). We have \((H_0)\) holds, that is, for any \( s \in [0, \tilde{\alpha}] \) we have
\[
\|\dot{u}_0\|_{C^s} \leq \delta \theta_0^{s-\alpha-1},
\]
when \( \|f\|_{C^{\alpha + \lambda}} \) is small enough.

Proof. By initiation \( g_0 = \Delta_0^{-1} S_{\theta_0} f \), and hence have \( \dot{u}_0 = \Delta_0^{-1} (S_{\theta_0} u_0) S_{\theta_0} f \) and see
\[
\|\dot{u}_0\|_{C^s} \leq C \Delta_0^{-1} \left\{ \|S_{\theta_0} f\|_{C^{s+\lambda}} + \|S_{\theta_0} f\|_{C^s} (1 + \|S_{\theta_0} u_0\|_{C^{s+d}}) \right\}
\leq C \Delta_0^{-1} \left\{ C_1 \theta_0^{(s-\alpha)+} \|f\|_{C^{\alpha + \lambda}} + C_2 \|f\|_{C^{\alpha + \lambda}} (1 + C_3 \theta_0^s \|u_0\|_{C^d}) \right\}
\leq C \Delta_0^{-1} \left\{ \theta_0^{(s-\alpha)+} + 1 + \theta_0^s \|u_0\|_{C^d} \right\} \|f\|_{C^{\alpha + \lambda}},
\]
by virtue of tame estimate \((A_2)\) \([21]\) and property of regularisation operator \((26)\). Note everything before \( \|f\|_{C^{\alpha + \lambda}} \) are already fixed. Easy to see \( \|\dot{u}_0\|_{C^s} \leq \delta \theta_0^{s-\alpha-1} \) when
\[
\|f\|_{C^{\alpha + \lambda}} \leq \delta \theta_0^{-\alpha-1} \Delta_0 \tilde{C}^{-1} (2 + \|u_0\|_{C^d}),
\]
as required. 

Remark. (i) Here we have proved \((H_n)\) hold for all \( n \), if \( \delta, \|f\|_{C^{\alpha + \lambda}}, \varepsilon \) are sufficiently small, and \( \tilde{\alpha}, \theta_0 \) sufficiently large.
(ii) Here we can take \( \|f\|_{C^{\alpha + \lambda}} \) arbitrarily small because \( f \) is our small perturbation term in the statement of Nash-Moser theorem.

4.5 Existence and Regularity

We claim the following results:
4.5.1 Existence of Solution in $C^\alpha$

**Proposition 4.9** (Existence of $C^\alpha$-limit). If $(H_n)$ hold for all $n$, then we have $u_n = u_0 + \sum_{k=0}^{n-1} \Delta_k \dot{u}_k$ converges pointwise to $u$, moreover, $u \in C^\alpha$.

**Proof.** From assumption that $(H_n)$ hold for all $n$, we deduce that

$$\|\dot{u}_k\|_{C^s} \leq \delta \theta_k^{s-\alpha-1}$$

for all $k$. To apply Lemma 4.2 we firstly need to convert $\dot{u}_k$ into a family of parameter continuum $\theta$ from the discrete nature of $\theta_k$. We define

$$p_\theta = \dot{u}_k, \quad \text{if } \theta_k \leq \theta < \theta_{k+1},$$

and $(p_\theta)_{\theta > \theta_0}$ is a family of $C^\infty$-functions. Clearly we still have (35):

$$\|p_\theta\|_{C^s} \leq (1 + 2\varepsilon(\alpha+1)) \delta \theta^{s-\alpha-1}.$$

Invoke Lemma 4.2 with (36) and we see

$$u = u_0 + \int_{\theta_0}^{\theta} p_\theta \, d\theta \in C^\alpha,$$

by $u_0 \in C^\infty$. Note there $u$ is written in the continuous integral form instead of discrete sum for

$$u = u_0 + \sum_{k=0}^{\infty} (\theta_{k+1} - \theta_k) \dot{u}_k = u_0 + \int_{\theta_0}^{\theta} p_\theta \, d\theta,$$

by virtue of the nature of $p_\theta$. Hence the limit solution is in $C^\alpha$. $\square$

**Proposition 4.10** (Convergence in $C^{\alpha-\gamma}$). By assuming $(H_n)$ hold for all $n$, we have $u_n \to u$ in all $C^{\alpha-\gamma}$ for $\gamma > 0$, where $u$ is the limit solution in previous proposition.

**Proof.** With $n > m$, consider that

$$\|u_n - u_m\|_{C^{\alpha-\gamma}} = \left\|\sum_{k=m}^{n-1} \Delta_k \dot{u}_k\right\|_{C^{\alpha-\gamma}} \leq \sum_{k=m}^{n-1} \Delta_k \|\dot{u}_k\|_{C^{\alpha-\gamma}}$$

$$= \int_{\theta_m}^{\theta_n} \|p_\theta\|_{C^{\alpha-\gamma}} \, d\theta \leq (1 + 2\varepsilon(\alpha+1)) \delta \int_{\theta_m}^{\theta_n} \theta^{-\gamma-1} \, d\theta = C \gamma^{-1} (\theta_n^{-\gamma} - \theta_m^{-\gamma}) \leq \frac{2C}{\gamma} \theta_m^{-\gamma} \to 0$$

as $m \to \infty$. Hence $\{u_n\}$ are Cauchy hence convergent in $C^{\alpha-\gamma}$, as of the nature of Banach spaces $C^{\alpha-\gamma}$. $\square$

**Proposition 4.11** (Existence of Solution). For every $\gamma > 0$, we have

$$\Phi(u_k) \to \Phi(u_0) + f,$$

in $C^{\alpha+\lambda-\gamma}$. Moreover, limit of $\Phi(u_k)$ is bounded in $C^{\alpha+\lambda}$. 42
Proof. See
\[ \Phi(u_{k+1}) - \Phi(u_0) - f = (S_{\theta_n} - 1) f + (1 - S_{\theta_n}) E_n + \Delta_n e_n. \]
A direct computation shows \( L(\alpha + \lambda) = -a \). Consider \( \|f\|_{C^{\alpha + \lambda}} \) is small and fixed; as \( n \to \infty \), via (39) we have
\[ \|E_n\|_{C^{\alpha + \lambda}} \leq C\delta \theta_n^{-a} \to 0 \]
and via (37) we have
\[ \|\Delta_n e_n\|_{C^{\alpha + \lambda}} \leq C\delta \theta_n^{-a-1} \Delta_n \leq \tilde{C} \delta \theta_n^{-a} \to 0. \]
Hence \( \Phi(u_k) \) are bounded in \( C^{\alpha + \lambda} \). Furthermore, see
\[ \|\Phi(u_{k+1}) - \Phi(u_0) - f\|_{C^{\alpha + \lambda} - \gamma} \leq C' \theta_n^{-\gamma} \|f\|_{C^{\alpha + \lambda}} + C'' \theta_n^{-\gamma} \|E_n\|_{C^{\alpha + \lambda}} + C''' \theta_n^{-\gamma} \|\Delta_n e_n\|_{C^{\alpha + \lambda}} \to 0 \]
as \( n \to \infty \).

Remark. Here we showed that if the iteration hypothesis \( (H_n) \) hold for all \( n \), then the existence of the solution and the convergence in lower regularity spaces are proved. However we remark here the convergence is based on Lemma 4.2, which is a unique result which is valid to Hölder spaces only. Later when we drop the restriction of Nash-Moser theorem on Hölder spaces, we can not make use of this lemma.

4.5.2 Regularity

We now want to prove the regularity part, (ii) of Theorem 4.3. We now verify the modified hypotheses for \( \alpha' < \bar{\alpha} \) when we assume additional information on data \( f \) that \( f \in W \cap C^{\alpha' + \lambda} \).

\( (H'_n) \) For \( 0 \leq k \leq n \) and \( s \in [0, \bar{\alpha}] \), we have
\[ \|\dot{u}_k\|_{C^s} \leq \delta \theta_k^{s - \alpha' - 1}. \] (41)
To see this, we firstly consider that in (40) that terms in which \( \|f\|_{C^{\alpha + \lambda}} \) does not occur has \( \theta_{n+1} \)-exponents strictly less than \( s - \alpha - 1 \) by the fact
\[ L(s + \lambda) - 1 < s - \alpha - 1, \quad L(\lambda) - 1 + (s + d - \alpha)_+ \leq L(\lambda) - 1 + s_+ < s - \alpha - 1. \]
Also, the \( \|f\|_{C^\beta} \) only originates from (38). If we replace (38) by
\[ \|\dot{f}_n\|_{C^s} \leq C\theta_n^{s - \alpha' - \lambda - 1} \|f\|_{C^{\alpha' + \lambda}}, \]
we have then recovered an estimate
\[ \|\dot{u}_{n+1}\|_{C^s} \leq C \left( \theta_{n+1}^{s - \alpha - 1} + \theta_{n+1}^{s - \alpha' - 1} \|f\|_{C^{\alpha' + \lambda}} \right). \]
for some small $\gamma > 0$. In fact, ignoring the trivial case $n = 0$, we have for all $n$ that
\[
\|\dot{u}_n\|_{C^s} \leq C\theta_n^{s-\rho-1},
\]
where $\rho = \min\{\alpha + \gamma, \alpha'\}$. Here we remark that $\gamma$ depends only on a lower bound for $\alpha$, and hence we can repeat the process above by $k$ times to get for all $n$, within informative interval $s \in [0, \tilde{\alpha}]$ we have
\[
\|\dot{u}_n\|_{C^s} \leq C\theta_n^{s-\rho_k-1},
\]
where $\rho_k = \min\{\alpha + k\gamma, \alpha'\}$. When $k$ is sufficiently large we have $\rho_k = \alpha'$ and we obtain iterative hypotheses $(H'_n)$ for all $n$.

Note that $(H'_n)$ are exactly iterative hypotheses for index $\alpha$ replaced by $\alpha'$, with constant possibly altered; however in the proof of Proposition 4.9 the scale of the constant is not important (it is important to run the iterative hypotheses $(H_n)$, and here we derive $(H'_n)$ from $(H_n)$ directly without an iterative process). We apply Proposition 4.9 to see the regularity result $u \in C^{\alpha'}$. The desired result is achieved.

Remark. (i) Till this point, our main result Theorem 4.3 is fully proved. One should be alerted that this theorem evades all integer indices of Hölder spaces.

(ii) The Nash-Moser theorem can also be achieved via Bony’s paradifferential calculus. See Bony [4] and Hörmander [14].
5 Isometric Embedding Theorem

5.1 Isometric Embedding Problem

5.1.1 Statement of Problem

In this essay, we only pay attention to the embedding for compact manifolds. The problem reads:

Can every smooth compact manifold be smoothly embedded into $\mathbb{R}^N$ for some $N$?

We ask, whether for any manifold $M$ with smooth metric $g$ of dimension $n$, we can find an embedding $u : M \to \mathbb{R}^N$ for some $N$, such that

$$g_{ij} = \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j},$$

for any local coordinates near all $x \in M$.

Remark. (i) We will present two methods for tackling the isometric embedding problem for compact manifolds: in Section 5.2 we see the isometric embedding theorem as a natural consequence of our Nash-Moser theorem via the original proof of Nash, presented in the form of Alinhac and Gérard [2, Chapter III], and in Section 5.3 we will go through Günther’s approach [7].

(ii) The case for non-compact manifolds is also provable. See Nash [19].

5.1.2 Reduction to Embedding of $\mathbb{T}^n$

We want to reduce the embedding problem of a general smooth compact $(M, g)$ of dimension $n$ back to that of a torus. This is done by smoothly embedding $M$ into $\mathbb{T}^N$ for some $N > n$. We will explain why we can do this. Firstly we invoke Whitney embedding theorem, without proof:

**Theorem 5.1** (Strong Whitney embedding theorem). Every smooth manifold $M$ of dimension $n$ can be smoothly embedded into $\mathbb{R}^{2n}$.

We firstly invoke the Whitney embedding theorem and get a smooth embedding $\iota : M \to \iota(M) \subset \mathbb{R}^{2n}$. $\iota$ is continuous, and $M$ is compact, hence $\iota(M)$ is compact in $\mathbb{R}^{2n}$, where we invoke Heine-Borel theorem to get that $\iota(M)$ is bounded. Hence we can cut a $2d$-cube in $\mathbb{R}^{2n}$ containing $\iota(M)$, and treat this cube as a torus. Hereby $M$ is smoothly embedded into $\mathbb{T}^{2n}$. Hence we reduced the case for general compact $M$ back to that for $\mathbb{T}^n$, for any $n$.

Remark. (i) We remark that in the process of embedding $M$ into $\mathbb{T}^{2n}$, the Riemannian metric $g$ is not necessarily invariant under the embedding. However, it does not matter because we can apply our theorem to new $g$ after the embedding extended to the whole $\mathbb{T}^n$. 

45
via Tietze extension theorem and the solution embedding composited with inclusion map on the right is the isometric embedding, which is associated with \((M, g)\).

(ii) The benefit of the reduction to torus is that on torus we have a global chart.

### 5.1.3 Free Embedding

**Definition 5.2** (Free embedding). For a smooth manifold \(M\) of dimension \(n\), an embedding \(u_0 : M \to \mathbb{R}^N\) is called a free embedding if for each \(x \in M\) we have that, the \(n(n + 3)/2\) vectors, \(\partial_i u_0(x)\) and \(\partial_i \partial_j u_0(x)\) in \(\mathbb{R}^N\), are linearly independent in local coordinates.

**Proposition 5.3** (Existence of free embedding of \(T^n\)). There exists a free embedding \(u_0 : T^n \to \mathbb{R}^{2n + n(2n + 1)}\).

**Proof.** First embed \(T^n = S^1 \times \cdots \times S^1\) into \((\mathbb{R}^2)^n\) in the normal way, via \(\tilde{u}\):

\[
\tilde{u}(x_1 + k_1 2\pi, \ldots, x_n + k_n 2\pi) = (\cos(x_1), \sin(x_1), \ldots, \cos(x_n), \sin(x_n)).
\]

Write \(\varphi : \mathbb{R}^{2n} \to \mathbb{R}^{2n + n(2n + 1)}\):

\[
\varphi(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_n, \bar{x}), \quad \bar{x} = (x_i x_j)_{1 \leq i \leq j \leq 2n}.
\]

By a direct computation one sees \(\partial_i \varphi\) and \(\partial_i \partial_j \varphi\) are linearly independent, and by the fact that \(\tilde{u}\) is an embedding, it is easy to see

\[
u = \varphi \circ \tilde{u} : T^n \to \mathbb{R}^{2n + n(2n + 1)}
\]

is a free embedding. \(\square\)

**Remark.** A free embedding enables one to establish an invertible linear system near \(u_0\), and then we can find a simple right inverse to the objective operator in the perturbation problem, which allows us to run Nash-Moser scheme.

### 5.2 Nash-Moser Approach

**Theorem 5.4** (Local embedding theorem). Let \(g_0\) be a metric induced by a free embedding \(u_0\) of \(\mathbb{T}^n\). For \(\rho > 2\) and \(\rho \notin \mathbb{N}\), and any metric \(g\) of class \(C^\rho\) near \(g_0\), there is a map \(u : \mathbb{T}^n \to \mathbb{R}^N\) of class \(C^\rho\) for which we have

\[
g_{ij} = \frac{\partial u}{\partial x_i} \cdot \frac{\partial u}{\partial x_j} = (u', u')_{ij}.
\]

**Proof.** We will apply Nash-Moser theorem in this proof. Construct the objective operator

\[
\phi(u) = u' \cdot u'.
\]

The problem is reduced to solving perturbation problem

\[
\phi(u) = \phi(u_0) + (g - g_0)
\]
where \( g - g_0 \) is known to be sufficiently small, and symmetric for \( g \) and \( g_0 \) being Riemannian metric. Observe that

\[
\phi'(u)(v) = \langle u', v' \rangle + \langle v', u' \rangle
\]

\[
\phi''(u)(v, w) = \langle u', v' \rangle + \langle v', u' \rangle
\]

and

\[
\|\phi''(u, v, w)\|_{C_0} \leq C \|\langle u', v' \rangle\|_{C_0} \leq C (\|w\|_{C^1} \|v\|_{C^{\alpha+1}} + \|w\|_{C^{\alpha+1}} \|v\|_{C^1})
\]

(43)

show that \( \phi \) is tame, in the sense that \( \phi \) satisfies assumption (A_1) (20). Moreover we want to consider whether there is a right inverse to objective operator \( \phi \), near \( u_0 \). We consider the problem to solve \( \phi'(u)(v) = r \), given some symmetric metric \( r \), and \( u \) near \( u_0 \). The problem reads

\[
\phi'(u)(v) = \langle u', v' \rangle + \langle v', u' \rangle = r.
\]

By choosing \( v \) such that

\[
t u'.v = 0,
\]

\[
t u''.v = -r/2,
\]

(44)

(45)

then we have \( \langle v, u' \rangle = 0 \) and \( \langle v, u'' \rangle = -r/2 \) by symmetry, and furthermore

\[
t u'.v + \langle v', u' \rangle = (\langle u', v \rangle)' - \langle v, u'' \rangle + (\langle v, u' \rangle)' - \langle v, u'' \rangle = r
\]

then the problem is solved. However, the choosing of \( v \) by (44) and (45) is in fact to solve

\[
M(u', u'').v = \langle u', u'' \rangle . v = \langle 0, -r/2 \rangle
\]

where we dropped the transpose of \( 0 \) and \(-r/2 \) for their symmetry. Recall that by definition of free embedding \( u_0 \) we have \( \det (\langle u', u'' \rangle) > 0 \). However by continuity of determinant we know \( \det (\langle u', u'' \rangle) > 0 \) for \( u \) sufficiently close to \( u_0 \), which is exactly the assumption we made on \( u \). Hence \( M(u', u'') \) is invertible hence define for \( u \) sufficiently near \( u_0 \),

\[
\psi(u)(r) = M(u', u'').^{-1}. \langle t, 0, -r/2 \rangle
\]

and we see \( \phi(u) \psi(u)(r) = r \), hence \( \psi \) is the right inverse of \( \phi \). Moreover we see

\[
\|\psi(u)(r)\|_{C_0} = \left\| \tilde{M}(u'').r \right\|_{C_0}
\]

where \( \tilde{M} \) is \(-1/2 \) times the matrix \( M^{-1} \) with first \( d \) rows and columns deleted, note that entries of \( \tilde{M} \) depend only on \( u'' \) with \( \left\| \tilde{M} \right\|_{C_0} \leq C \|u\|_{C^{\alpha+2}} \). Herby for \( u \) in some \( \mu' \)-neighbourhood of \( u_0 \), with \( \mu' > 2 \), we have

\[
\|\psi(u)(r)\|_{C_0} \leq C_1 \left( \left\| \tilde{M} \right\|_{C_0} \|r\|_{C^0} + \left\| \tilde{M} \right\|_{C_0} \|r\|_{C^0} \right)
\]

\[
\leq C_2 (\|u\|_{C^2} \|r\|_{C^0} + \|u\|_{C^{\alpha+2}} \|r\|_{C^0}) \leq C (\|r\|_{C_0} + \|r\|_{C^0} \|u\|_{C^{\alpha+2}})
\]

(46)
shows \( \psi \) is tame with loss of two derivatives in \( u \) hence satisfies (A2) \([21]\). Note in the last inequality we used the fact that \( \| u \|_{C^2} \) is bounded by a constant, exactly because \( u \) is in some \( \mu' \)-neighbourhood of \( u_0 \), with \( \mu' > 2 \).

By the previous arguments, we set the parameters

\[
a = 1, \quad b = 0, \quad c = 1, \quad d = 2, \quad \lambda = 0, \quad \mu = \mu' = 2 + 1/2 (\rho - 2)
\]

for \( g \in C^\rho \) with \( \rho > 2 \) and \( \rho \notin \mathbb{N} \). Take \( \alpha = \rho \), it is trivial to verify

\[
\alpha \geq \mu, \quad \alpha \geq \mu', \quad \alpha > d, \quad \alpha > \lambda + a + c, \quad \alpha > a + 1/2 \sup(\lambda + b, 2a).
\]

Invoke Nash-Moser Theorem 4.3 with assumptions (A1) on \( \phi \) and (A2) on \( \psi \), we get when \( \| g - g_0 \|_{C^\rho} \) is sufficiently small, we obtain

\[
\phi(u) = \phi(u_0) + g - g_0 = g
\]

has a solution \( u \in C^\rho \).

**Remark.** (i) This is the important theorem in which we invoke the Nash-Moser Theorem 4.3. Note that here we have not finished the proof for embedding of arbitrary metric tensor \( g \); we have only proved metrics sufficiently near to our freely embedded \( g_0 \) admit a solution to the embedding problem. The rest is done in next few pages.

(ii) We stick to a freely embedded metric \( g_0 \) exactly because near \( g_0 \) we have a nice inverse for \( \phi' \).

(iii) Note that we obtain (43) and (46) by the dyadic analysis result (19).

(iv) Note the injectivity of the solution is not proved. However by the upcoming proposition and lemma we can see increasing the dimension \( N \) suffices the injectivity condition.

**Lemma 5.5.** (i) Given two \( C^1 \)-embedding \( u_j : \mathbb{T}^n \to \mathbb{R}^N_j \) for \( j = 1, 2 \), define embedding

\[
u : \mathbb{T}^n \to \mathbb{R}^{N_1+N_2}, \quad \nu(x) = (t_1u_1(x), t_2u_2(x)),
\]

with scalars \( t_1, t_2 \). Denote by \( g, g_1, g_2 \) the metrics induced by embedding \( \nu, u_1, u_2 \) respectively, then we have

\[
g = t_1^2 g_1 + t_2^2 g_2.
\]

(ii) Assume every metric \( g \) can be represented by an \( u : \mathbb{T}^n \to \mathbb{R}^N \), then every metric \( g \) can be represented by an injective \( \tilde{u} : \mathbb{T}^n \to \mathbb{R}^{N+n} \).

**Proof.** (i) Let \( \langle \cdot, \cdot \rangle \) denote the normal inner product in Euclidean spaces. Then we see immediate properties

\[
\langle (a, b), (c, d) \rangle = \langle a, c \rangle \langle b, d \rangle.
\]

Now consider

\[
g_{ij} = (u^* \text{id}) \left( \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right) = \text{id} \left( u_i' \frac{\partial}{\partial x_i}, u_j' \frac{\partial}{\partial x_j} \right) = \text{id} \left( t_1u_{1i}' \frac{\partial}{\partial x_i}, t_1u_{1j}' \frac{\partial}{\partial x_j} \right) + \text{id} \left( t_2u_{2i}' \frac{\partial}{\partial x_i}, t_2u_{2j}' \frac{\partial}{\partial x_j} \right) = t_1^2 (u_{1i}^* \text{id}) \left( \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right) + t_2^2 (u_{2i}^* \text{id}) \left( \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right) = t_1^2 (g_1)_{ij} + t_2^2 (g_2)_{ij},
\]
as required.

(ii) Firstly we see there is an injective $C^\infty$ mapping $u_1 : \mathbb{T}^n \to \mathbb{R}^n$ given by the ordinary embedding. Let $g_1 = t' u_1' u_1'$, and fix $t$ such that $t^2 g_1 < g$; then $g - t^2 g_1$ can be represented by some $u_2 : \mathbb{T}^n \to \mathbb{R}^N$ by our assumption. Easy to check $\tilde{u}(x) = (t u_1, u_2)$ represents $g$ by virtue of (i).

We quote a proposition with an abbreviated proof borrowed from Tao [21].

**Proposition 5.6 (Density of metrics induced by embedding).** On a torus, the set of metrics $g$ representable as

$$g_{ij} = \frac{\partial u}{\partial x^i} \cdot \frac{\partial u}{\partial x^j},$$

for some $C^\infty$-mappings $u : \mathbb{T}^n \to \mathbb{R}^N$, is dense in the set of all metrics.

**Proof.** It suffices to prove that given any metric $g$ on $\mathbb{T}^n$, we have a smooth symmetric perturbation metric $h$ on $\mathbb{T}$ such that for all $\varepsilon$ there are $u_\varepsilon$ which are $C^\infty$,

$$(g + \varepsilon^2 h)_{ij} = \frac{\partial u_\varepsilon}{\partial x^i} \cdot \frac{\partial u_\varepsilon}{\partial x^j}$$

is representable.

By the spectral theorem for finite-dimensional matrices, we see all positive definite tensors $g'_{ij}$ can be decomposed as finite positive linear sum of symmetric rank one tensors $v_i v_j$ where $v \in \mathbb{R}^N$. If necessary add additional rank one tensors then one can make sure that any nearby metric $g'_ij$ is also a positive linear combination of $v_i v_j$. Write $v_i = \partial_i \psi$ for some linear $\psi : \mathbb{R}^N \to \mathbb{R}^N$ and we see

$$g_{ij} = \sum_{k=1}^m \left( \eta^k \right)^2 \partial_i \psi^k \partial_j \psi^k$$

by compactness of $\mathbb{T}$ and smooth partition of unity, where $\psi^k$ is $\psi$ restricted to partition and $\eta^k$ is the cutoff for partition, both mapping from $\mathbb{R}^N$ to $\mathbb{R}$. For any $\varepsilon > 0$ consider the map $u^k_\varepsilon : \mathbb{T}^n \to \mathbb{R}^2$ by

$$u^k_\varepsilon(x) = \varepsilon \eta^k(x) \cdot \left( \cos \left( \psi^k(x)/\varepsilon \right), \sin \left( \psi^k(x)/\varepsilon \right) \right).$$

Compute

$$\frac{\partial u^k_\varepsilon}{\partial x^i} \cdot \frac{\partial u^k_\varepsilon}{\partial x^j} = \left( \eta^k \right)^2 \partial_i \psi^k \partial_j \psi^k + \varepsilon^2 \eta^k_i \eta^k_j.$$  

Take $u_\varepsilon = (u^1_\varepsilon, u^2_\varepsilon, \ldots, u^m_\varepsilon)$ and we see

$$\frac{\partial u_\varepsilon}{\partial x^i} \cdot \frac{\partial u_\varepsilon}{\partial x^j} = g_{ij} + \varepsilon^2 h_{ij}$$

where

$$h = \sum_{k=1}^m \eta^k_i \eta^k_j,$$

as required. \qed
Remark. (i) This proposition is does not prove the isometric embedding theorem since \( u \) does not necessarily converge as \( \varepsilon \) converges to 0.

(ii) It means we can approximate a given metric to an arbitrary accuracy, by good metrics (metrics induced by isometric embedding). Note that we do not have any free embedding structures in these approximations. However we can use these approximations to reduce the distance between given metric and a fixed freely embedded metric, by using an approximation as a stepping stone.

**Theorem 5.7** (Isometric embedding theorem). Given any \( C^\rho \) metric \( g \) on \( \mathbb{T}^n \), \( \rho > 2 \), \( \rho \notin \mathbb{N} \), there exists an \( C^\rho \)-embedding \( u : \mathbb{T}^n \to \mathbb{R}^N \) such that

\[
g_{ij} = \frac{\partial u}{\partial x_i} \cdot \frac{\partial u}{\partial x_j}.
\]

**Proof.** We know there is a freely embedded metric \( g_0 \). Fix \( t \) such that \( t^2 g_0 < g \). Apply Proposition \ref{prop:approximation} to approximate \( g - t^2 g_0 \) we can find an smooth mapping \( u_2 : \mathbb{T}^n \to \mathbb{R}^{N_2} \) such that metric \( g_2 = g - t^2 g_0 - t^2 f \) is represented by \( u_2 \), with \( f \) small. Note here we can make \( f \) arbitrarily small by improving approximation \( u_2 \) by density. Then we have the decomposition of our given metric \( g \):

\[
g = t^2 (g_0 + f) + (g - t^2 g_0 - t^2 f) = t^2 g_1 + g_2.
\]

where \( g_1 = g_0 + f \in C^\rho \). Note that \( g_1 \) is sufficiently close to freely embedded metric \( g_0 \) by sufficiently small \( f \), hence by Theorem \ref{thm:representation} we have a \( C^\rho \) representation \( u_1 : \mathbb{T}^n \to \mathbb{R}^{N_1} \) representing \( g_1 \). Now set \( u : \mathbb{T}^n \to \mathbb{R}^N \) for \( N = N_1 + N_2 \) to be

\[
u(x) = (tu_1(x), u_2(x))
\]

and we see that \( u \) is \( C^\rho \) and \( u \) represents \( g \) because

\[
g_u = t^2 g_1 + g_2
\]

by virtue of Lemma \ref{lem:decomposition} (i). By Lemma \ref{lem:decomposition} (ii) increase the dimension of \( u \) in case injectivity is required, and we obtain the desired isometric embedding. \( \square \)

### 5.3 Günther’s Approach

In this section want to get Theorem \ref{thm:representation} without invoking Nash-Moser theorem. Instead, we are going to practice a fixed point argument facilitating the ellipticity of the system, to solve the perturbation problem \ref{eq:perturbation}, following Günther \ref{gunther}.

We start from the given free embedding \( u_0 \) with induced metric \( g_0 \). Write \( u = u_0 + v \), \( g = g_0 + h \), \( \partial_i = \frac{\partial}{\partial x_i} \), the perturbation problem \( \partial_i u. \partial_j u = g_{ij} \) reads

\[
\partial_i u_0. \partial_j v + \partial_j u_0. \partial_i v + \partial_i v. \partial_j v = h_{ij}
\]

in global coordinates of \( \mathbb{T}^n \).
Lemma 5.8. The perturbation problem \([47]\) is equivalent to finding \(v \in C^p_\ast\) for \(\rho > 2\) such that
\[
\partial_j (\partial_i u_0.v) + \partial_j F_i(v) + \partial_i (\partial_j u_0.v) + \partial_i F_j(v) - 2 \left( \partial^2_{ij} u_0 \right).v = h_{ij} - R_{ij},
\]
for
\[
R_{ij}(v) = (1 - \triangle)^{-1} r_{ij}(v) = (1 - \triangle)^{-1} \left( \partial_i v.\partial_j v + 2\partial^2_{ij}.v. \triangle v - 2 \sum_{k=1}^{n} \partial^2_{ik} v.\partial^2_{jk} v \right),
\]
\[
F_i(v) = (1 - \triangle)^{-1} f_i(v) = (1 - \triangle)^{-1} (-\triangle v.\partial_i v),
\]
in which \(\triangle\) is the intrinsic Laplace-Beltrami operator on \((\mathbb{T}^n, \text{id})\), that is, \(\triangle = \sum_k \partial^2_{kk}\) in global coordinates.

Proof. Firstly see
\[
(1 - \triangle) (\partial_i v.\partial_j v) = -\partial_i (-\triangle v.\partial_j v) - \partial_j (-\triangle v.\partial_i v)
\]
\[
+ \left( \partial_i v.\partial_j v + 2\partial^2_{ij}.v. \triangle v - 2 \sum_{k=1}^{n} \partial^2_{ik} v.\partial^2_{jk} v \right) = -\partial_i f_j(v) - \partial_j f_i(v) + r_{ij}(v)
\]
by a direct computation via chain rule and commutativity between \(\triangle\) and \(\partial\). Note that \((1 - \triangle)\) is an elliptic pseudodifferential operator of degree 2, hence by Proposition [2.22] we have its inverse \((1 - \triangle)^{-1}\)
\[
(1 - \triangle)^{-1} u = \int_{\mathbb{R}^n} e^{ix.\xi} (1 + |\xi|^2)^{-1} \hat{u}(\xi) \, d\xi
\]
as a pseudodifferential operator of degree \(-2\). Moreover it commutes with \(\partial\), because both symbols are location-invariant. Hence we have
\[
(\partial_i v.\partial_j v) = -(1 - \triangle)^{-1} (\partial_i f_j(v) - \partial_j f_i(v)) + R_{ij}(v) = -\partial_i F_j(v) - \partial_j F_i(v) + R_{ij}(v).
\]
And the rest follows immediately. 

Lemma 5.9 (Estimates of terms). For \(\|v\|_{C^p_\ast} \leq R\) and \(\|w\|_{C^p_\ast} \leq R\) we have the following estimates:
\[
\|r_{ij}(v)\|_{C^{p-2}_\ast} \leq 2\tilde{C}_\rho^p\|v\|_{C^p_\ast},
\]
\[
\|f_i(v)\|_{C^{p-2}_\ast} \leq 2\tilde{C}_\rho^f\|v\|_{C^p_\ast},
\]
\[
\|r_{ij}(v) - r_{ij}(w)\|_{C^{p-2}_\ast} \leq 2\tilde{C}_\rho^r\|v - w\|_{C^p_\ast},
\]
\[
\|f_i(v) - f_i(w)\|_{C^{p-2}_\ast} \leq 2\tilde{C}_\rho^f\|v - w\|_{C^p_\ast},
\]
\[
\|R_{ij}(v)\|_{C^p_\ast} \leq 2\tilde{C}_\rho^r\|v\|_{C^p_\ast},
\]
\[
\|F_i(v)\|_{C^p_\ast} \leq 2\tilde{C}_\rho^f\|v\|_{C^p_\ast},
\]
\[
\|R_{ij}(v) - R_{ij}(w)\|_{C^p_\ast} \leq 2\tilde{C}_\rho^r\|v - w\|_{C^p_\ast},
\]
\[
\|F_i(v) - F_i(w)\|_{C^p_\ast} \leq 2\tilde{C}_\rho^f\|v - w\|_{C^p_\ast},
\]
for constants \(\tilde{C}_\rho^p, \tilde{C}_\rho^f, \tilde{C}_\rho^r, \tilde{C}_\rho^*, C^p_\rho, C^r_\rho, C^f_\rho, C^*_\rho\) which depend on \(\rho\) only.
\textbf{Proof.} By Proposition 3.16 we have

\[
\|\partial_t v, \partial_j v\|_{C^{\sigma-2}} \lesssim \|\partial_t v\|_{C^\sigma} \|\partial_j v\|_{C^\sigma} + \|\partial_t v\|_{C^\sigma-2} \|\partial_j v\|_{C^\sigma} \\
\quad \lesssim \|v\|_{C^\sigma} \|v\|_{C^\sigma} + \|v\|_{C^\sigma} \|v\|_{C^\sigma} \lesssim 2\|v\|_{C^\sigma} \|v\|_{C^\sigma} \lesssim 2R\|v\|_{C^\sigma}.
\]

Similarly we have

\[
\|\partial^2_{ij} v, \Delta v\|_{C^{\sigma-2}} \lesssim 2\|v\|_{C^\sigma} \|v\|_{C^\sigma} \lesssim 2R\|v\|_{C^\sigma},
\]

\[
\|\partial^2_{ik} v, \partial k v\|_{C^{\sigma-2}} \lesssim 2\|v\|_{C^\sigma} \|v\|_{C^\sigma} \lesssim 2R\|v\|_{C^\sigma}
\]

and then by formula (49) we immediately have

\[
\|r_{ij}(v)\|_{C^{\sigma-2}} \lesssim 2R\tilde{C}_{p}^{\sigma} \|v\|_{C^\sigma}.
\]

Then for \(f_i(v)\) we have similarly

\[
\|f_i(v)\|_{C^\sigma} = ||-\Delta v, \partial_i v\|_{C^{\sigma-2}} \lesssim \|v\|_{C^\sigma} \|v\|_{C^\sigma} + \|v\|_{C^\sigma} \|v\|_{C^\sigma} \lesssim 2R\tilde{C}_{p}^{f} \|v\|_{C^\sigma}.
\]

Denote \(q = v - w\), and easy to see \(\|q\|_{C^\sigma} \lesssim 2R\). By Proposition 3.16 we have

\[
\|\Delta q, \partial_i w\|_{C^{\sigma-2}} \lesssim \|\Delta q\|_{C^\sigma} \|\partial_i w\|_{C^\sigma} + \|\Delta q\|_{C^\sigma} \|\partial_i w\|_{C^\sigma} \lesssim \|q\|_{C^\sigma} \|w\|_{C^\sigma} + \|q\|_{C^\sigma} \|w\|_{C^\sigma} \lesssim 2R\|q\|_{C^\sigma}.
\]

Similarly we have

\[
\|\Delta w, \partial j q\|_{C^{\sigma-2}} \lesssim \|w\|_{C^\sigma} \|q\|_{C^\sigma} \lesssim 2R\|q\|_{C^\sigma},
\]

whence

\[
\|f_i(v) - f_i(w)\|_{C^{\sigma-2}} \lesssim \|\Delta q, \partial_i w\|_{C^{\sigma-2}} + \|\Delta w, \partial j q\|_{C^{\sigma-2}} + \|\Delta q, \partial j q\|_{C^{\sigma-2}} \lesssim 2R\tilde{C}_{p}^{f} \|v - w\|_{C^\sigma},
\]

as required.

Via a similar scheme, we estimate

\[
\|\partial_i q, \partial_j w\|_{C^{\sigma-2}} \lesssim \|\partial_i \|_{C^\sigma} \|w\|_{C^\sigma} + \|\partial_i \|_{C^\sigma} \|w\|_{C^\sigma} \lesssim 2R\|q\|_{C^\sigma},
\]

\[
\|\partial_i w, \partial j q\|_{C^{\sigma-2}} \lesssim \|\partial_i \|_{C^\sigma} \|q\|_{C^\sigma} + \|\partial_i \|_{C^\sigma} \|q\|_{C^\sigma} \lesssim 2R\|q\|_{C^\sigma},
\]

\[
\|\partial_i q, \partial_j q\|_{C^{\sigma-2}} \lesssim 2\|\partial_i \|_{C^\sigma} \|q\|_{C^\sigma} \|q\|_{C^\sigma} \lesssim 4R\|q\|_{C^\sigma},
\]

and

\[
\|\partial_i v, \partial_j v - \partial_i w, \partial_j w\|_{C^{\sigma-2}} \lesssim \|\partial_i q, \partial_j w\|_{C^{\sigma-2}} + \|\partial_i w, \partial j q\|_{C^{\sigma-2}} + \|\partial_i q, \partial j q\|_{C^{\sigma-2}} \lesssim 2R\|v - w\|_{C^\sigma};
\]

\[52\]
estimate
\[ \| \partial^2_{ij} q \cdot \Delta w \|_{C^{s-2}_\rho} \lesssim \| q \|_{C^s_\rho} \| w \|_{C^s_\rho} + \| q \|_{C^s_\rho} \| w \|_{C^s_\rho} \lesssim 2R \| q \|_{C^s_\rho}, \]
\[ \| \partial^2_{ij} w \cdot \Delta q \|_{C^{s-2}_\rho} \lesssim \| w \|_{C^s_\rho} \| q \|_{C^s_\rho} + \| w \|_{C^s_\rho} \| q \|_{C^s_\rho} \lesssim 2R \| q \|_{C^s_\rho}, \]
\[ \| \partial^2_{ij} q \cdot \Delta q \|_{C^{s-2}_\rho} \lesssim 2 \| q \|_{C^s_\rho} \| q \|_{C^s_\rho} \lesssim 4R \| q \|_{C^s_\rho}, \]
and
\[ \| \partial^2_{ij} v \cdot \Delta v - \partial^2_{ij} w \cdot \Delta w \|_{C^{s-2}_\rho} \leq \| \partial^2_{ij} q \cdot \Delta w \|_{C^{s-2}_\rho} + \| \partial^2_{ij} w \cdot \Delta q \|_{C^{s-2}_\rho} + \| \partial^2_{ij} q \cdot \Delta q \|_{C^{s-2}_\rho} \lesssim 2R \| v - w \|_{C^s_\rho}; \]

estimate
\[ \| \partial^2_{ik} q \cdot \partial^2_{jk} w \|_{C^{s-2}_\rho} \lesssim \| q \|_{C^s_\rho} \| w \|_{C^s_\rho} + \| q \|_{C^s_\rho} \| w \|_{C^s_\rho} \lesssim 2R \| q \|_{C^s_\rho}, \]
\[ \| \partial^2_{ik} w \cdot \partial^2_{jk} q \|_{C^{s-2}_\rho} \lesssim \| w \|_{C^s_\rho} \| q \|_{C^s_\rho} + \| w \|_{C^s_\rho} \| q \|_{C^s_\rho} \lesssim 2R \| q \|_{C^s_\rho}, \]
\[ \| \partial^2_{ik} q \cdot \partial^2_{jk} q \|_{C^{s-2}_\rho} \lesssim 2 \| q \|_{C^s_\rho} \| q \|_{C^s_\rho} \lesssim 4R \| q \|_{C^s_\rho}, \]
and
\[ \| \partial^2_{ik} v \cdot \partial^2_{jk} v - \partial^2_{ik} w \cdot \partial^2_{jk} w \|_{C^{s-2}_\rho} \leq \| \partial^2_{ik} q \cdot \partial^2_{jk} w \|_{C^{s-2}_\rho} + \| \partial^2_{ik} w \cdot \partial^2_{jk} q \|_{C^{s-2}_\rho} + \| \partial^2_{ik} q \cdot \partial^2_{jk} q \|_{C^{s-2}_\rho} \lesssim 2R \| v - w \|_{C^s_\rho}. \]

eventually we have
\[
\| r_{ij}(v) - r_{ij}(w) \|_{C^{s-2}_\rho} \leq \| \partial_i v \partial_j v - \partial_i w \partial_j w \|_{C^{s-2}_\rho} + \| \partial^2_{ij} v \cdot \Delta v - \partial^2_{ij} w \cdot \Delta w \|_{C^{s-2}_\rho} \\
+ \| \partial^2_{ik} v \cdot \partial^2_{jk} v - \partial^2_{ik} w \cdot \partial^2_{jk} w \|_{C^{s-2}_\rho} \leq 2R \delta \| v - w \|_{C^s_\rho},
\]
as required.

From Theorem 3.13 we know the pseudodifferential operator \((1 - \Delta)^{-1}\) is a bounded linear operator from \(C^{s-2}_\rho\) to \(C^s_\rho\), hence
\[ \| (1 - \Delta)^{-1} u \|_{C^s_\rho} \leq C \| u \|_{C^{s-2}_\rho}. \]

By letting \( u \) be \( r_{ij}(v), f_i(v), r_{ij}(v) - r_{ij}(w) \) or \( f_i(v) - f_i(w) \), the rest follow immediately. \( \square \)

**Theorem 5.10** (Local embedding theorem). Let \( g_0 \) be a metric induced by a free embedding \( u_0 \) of \( \mathbb{T}^n \). For \( \rho > 2 \), and any metric \( g \) of class \( C^s_\rho \) near \( g_0 \), there is a map \( u : \mathbb{T}^n \to \mathbb{R}^N \) of class \( C^s_\rho \) for which we have
\[ g_{ij} = \partial_i u \partial_j u. \]
Proof. To solve perturbation problem (48), it suffices to solve
\[ \partial_i u_0 \cdot v = -F_i(v), \]
\[ \partial^2_{ij} u_0 \cdot v = \frac{1}{2} (R_{ij} - h_{ij}), \]
which is to solve
\[ M(u_0) \cdot v = \left( \begin{array}{c} t u_0' \\ u_0'' \end{array} \right) \cdot v = \left( \begin{array}{c} -F_i(v) \\ \frac{1}{2} (R_{ij}(v) - h_{ij}) \end{array} \right). \]
As \( u_0 \) is a free embedding, \( M(u_0) \) is invertible. Construct solution operator \( G \) with
\[ G(v) = M(u_0)^{-1} \left( \begin{array}{c} -F_i(v) \\ \frac{1}{2} (R_{ij}(v) - h_{ij}) \end{array} \right). \]
and the problem is reduced to finding a fixed point of \( G \). Consider
\[ G(v - w) = M(u_0)^{-1} \left( \begin{array}{c} -F_i(v) + F_i(w) \\ \frac{1}{2} (R_{ij}(v) - R_{ij}(w)) \end{array} \right). \]
We want \( G \) to be \( BC^\rho(R) \to BC^\rho(R) \), where \( BC^\rho(R) \) denotes ball of radius \( R \) in \( C^\rho \) around 0. By the estimate
\[ \|G(v)\|_{C^\rho} \lesssim \sum_{i,j} \left( \|F_i(v)\|_{C^\rho} + \|R_{ij}(v)\|_{C^\rho} + \|h_{ij}\|_{C^\rho} \right) \lesssim 2RC^\rho\|v\|_{C^\rho} + 2RC^\rho\|v\|_{C^\rho} + \|h\|_{C^\rho} \leq C_2R^2 + C_3\|h\|_{C^\rho}, \]
where we see \( \|G(v)\|_{C^\rho} < R \) when
\[ R < C_2^{-1} \left( 1 + \sqrt{1 - 4C_2C_3\|h\|_{C^\rho}} \right) / 4 \]
provided \( \|h\|_{C^\rho} \), the perturbation in metric, is sufficiently small. By Lemma 5.9, we have the estimate for \( v, w \in BC^\rho(R) \):
\[ \|G(v - w)\|_{C^\rho} \leq 2RC^\rho \left( C^f_\rho + C^r_\rho \right) \|v - w\|_{C^\rho}. \]
When \( R < C^{-1} (C^f_\rho + C^r_\rho)^{-1} / 4 \) we see \( G : BC^\rho(R) \to BC^\rho(R) \) is a contraction. However \( BC^\rho(R) \) is clearly Banach. Invoke Banach contraction theorem and the existence of \( v \in C^\rho \) is straightforward. The mapping constructed \( u = u_0 + v \) obviously solves the embedding problem.

Remark. (i) This is almost the same statement as of Theorem 5.4, and next steps towards the full embedding theorem are exactly the same.

(ii) Note that the difference between this theorem is that we no longer have the constraint that \( \rho \notin \mathbb{N} \). This is because we proved a flawed Nash-Moser theorem with holes within the indices of Hölder spaces (so that we can use Lemma 4.2). Later we will prove a
more general theorem which extends to H"older-Zygmund spaces and the same result can be achieved using a complete Nash-Moser theorem.

(iii) Note that in G"unther’s original paper, he stucked to two de Rham Laplacians on vector fields and symmetric tensors, however here we simplified his geometrically strict arguments down to plain analytic descriptions. Another improvement is that he invoked interior Schauder estimates to recover information lost due to the second order elliptic operator \((1 - \Delta)\), however here we use the H"older-Zygmund (Besov) mapping properties, Theorem 3.13 of parametrix \((1 - \Delta)^{-1}\) as a pseudodifferential operator of order \(-2\).
6 Abstract Nash-Moser Theorem

6.1 Scale of Banach Spaces

We want to generalise a modification process of a family of Banach spaces, just like extending Hölder spaces to Hölder-Zygmund spaces by interpolation. Let \( E^a \) be a decreasing family of Banach spaces for indices \( a \geq 0 \), with injections \( E^b \hookrightarrow E^a \) for \( b \geq a \), such that operator norm of the injection is no greater than 1. We assume for this family \( \{E^a\} \) we are given a family of regularisation operators:

**Definition 6.1 (Regularisation operators).** A family of continuous linear operators \( S_\theta \) for parameter \( \theta \geq 1 \)

\[
S_\theta : E^0 \rightarrow E^\infty = \bigcap_{a=1}^\infty E^a,
\]

are said to be the regularisation operators on family of Banach spaces \( \{E^a\} \) if for \( a, b < \infty \) we have constants not depending on \( \theta \) only for the conditions to hold:

(A) \( \|S_\theta u\|_{E^b} \leq C\|u\|_{E^a} \), for \( b \leq a \).

(B) \( \left\| \frac{d}{d\theta} S_\theta u \right\|_{E^b} \leq C\theta^{b-a-1}\|u\|_{E^a} \), for all \( a, b \).

(C) If \( u \in E^0 \) and \( S_\theta u \rightarrow v \) in \( E^0 \) as \( \theta \rightarrow \infty \), then \( v = u \).

**Proposition 6.2 (Properties of regularisation operators).** We have the followings properties for \( S_\theta \):

(i) \( \theta \mapsto S_\theta \in L(E^0, E^b) \) is \( C^1 \) with respect to \( \theta \) for any \( b < \infty \).

(ii) \( \|S_\theta u\|_{E^b} \leq C(b - a)^{-1}\theta^{b-a}\|u\|_{E^a} \), for \( a < b \).

(iii) \( \|u - S_\theta u\|_{E^b} \leq C(a - b)^{-1}\theta^{b-a}\|u\|_{E^a} \), for \( a > b \).

(iv) \( \|u\|_{E^c} \leq C((b - a)\lambda(1 - \lambda))\|u\|_{E^a}^\lambda\|u\|_{E^b}^{1-\lambda} \), for \( c = \lambda a + (1 - \lambda)b \) with \( 0 < \lambda < 1 \).

Here \( L(E^0, E^b) \) indicates the space of linear bounded operators from \( E^0 \) to \( E^b \).

**Proof.** We see (i) is immediate from (B). For (ii) see

\[
\|S_\theta u\|_{E^b} \leq \int_1^\theta \left\| \frac{d}{d\tau} S_\tau u \right\|_{E^b} d\tau \leq \int_1^\theta C\tau^{b-a-1}\|u\|_{E^a} d\tau \leq C(b - a)^{-1}\theta^{b-a}\|u\|_{E^a}
\]

by (B) and dropping the lower index term of the integral, and we can do this only when \( a < b \). Note that here we can interchange norm in and out integral for the integral is with respect to parameter \( \theta \) instead of \( u \). For (iii) see \( u = \lim_{\theta \rightarrow \infty} S_\theta u \) via (C) and

\[
\|u - S_\theta u\|_{E^b} \leq \int_\theta^\infty \left\| \frac{d}{d\tau} S_\tau u \right\|_{E^b} d\tau \leq \int_\theta^\infty C\tau^{b-a-1}\|u\|_{E^a} d\tau \leq C(a - b)^{-1}\theta^{b-a},
\]

56
by (B) and for \( b > a \), the integral converges and the upper index term of integral vanishes. Lastly for (iv), observe

\[
\|u\|_{E^c} \leq \|S_\theta u\|_{E^c} + \|u - S_\theta u\|_{E^c} \leq C \left( (c - a)^{-1} \theta^{c-a} \|u\|_{E^a} + (b - c)^{-1} \theta^{c-b} \|u\|_{E^b} \right)
\]

by (ii) with \( a < c \) and (iii) with \( b > c \). Take \( \theta = \left( \|u\|_{E^b} / \|u\|_{E^a} \right)^{1/(b-a)} \), which is no smaller than 1 by the contracting injection of \( E^b \hookrightarrow E^a \), and (iv) follows immediately.

\textbf{Definition 6.3} (Weak scales). Given \( a > 0 \) fixed, we denote by \( E^*_a \) the set of all \( u \in E^0 \) such that there is some \( M \), the following conditions hold

\begin{align*}
\text{(D)} & \quad \|u\|_{E^0} \leq M, \\
\text{(E)} & \quad \left\| \frac{d}{d\theta} S_\theta u \right\|_{E^0} \leq M \theta^{-a-1}, \\
\text{(F)} & \quad \left\| \frac{d}{d\theta} S_\theta u \right\|_{E^{a+1}} \leq M.
\end{align*}

And set the norm of \( E^*_a \):

\[
\|u\|_{E^*_a} = \inf_{\theta \geq 1} \left\{ \|u\|_{E^0}, \left\| \frac{d}{d\theta} S_\theta u \right\|_{E^0}^\theta, \left\| \frac{d}{d\theta} S_\theta u \right\|_{E^{a+1}} \right\}. \tag{50}
\]

The weak scale \( E^*_a \) is Banach. We call \( \{E^a\} \) the strong scale and \( \{E^*_a\} \) the weak scale.

\textbf{Example} (Examples of regularisation operators). (i) Take the H"older spaces \( \{C^a\} \) to be the strong scale with respect to regularisation operators defined in Proposition 3.14, and we obtain the H"older-Zygmund spaces \( \{C^*_a\} \) as the weak scale. That is, strong and weak scales coincide for non-integer indices.

(ii) Take the strong scale to be the Sobolev spaces \( H^a \), with respect to the same family of regularisation operators in Proposition 3.14, we obtain a new space \( H^*_a \). Note that \( \{H^*_a\} \) is strictly weaker than \( \{H^a\} \), that is, \( H^*_a \) is strictly larger than \( H^a \).

\textbf{Proposition 6.4} (Injections of function spaces). For any \( b < a \), we have \( E^a \subset E^*_a \subset E^b \), and \( E^*_a \subset E^*_b \).

\textit{Proof.} From (B) the we clearly have \( E^a \subset E^*_a \). Given \( u \in C^*_a \) for \( 0 < b < a + 1 \), we have from (E) and (F) that

\[
\left\| \frac{d}{d\theta} S_\theta u \right\|_{E^0} \leq \theta^{-a-1} \|u\|_{E^*_a}, \quad \left\| \frac{d}{d\theta} S_\theta u \right\|_{E^{a+1}} \leq \|u\|_{E^*_a}
\]

and from Proposition 6.2 (iv) we see

\[
\left\| \frac{d}{d\theta} S_\theta u \right\|_{E^b} \leq C \left( \left\| \frac{d}{d\theta} S_\theta u \right\|_{E^0}^\lambda \left\| \frac{d}{d\theta} S_\theta u \right\|_{E^{a+1}}^{1-\lambda} \right) \leq C \theta^{b-a+1} \|u\|_{E^*_a}
\]

57
by taking interpolating parameter \( \lambda = (a + 1 - b) / (a + 1) \). Moreover specify \( 0 < b < n \) we see \( \| \frac{d}{d\theta} S_\theta u \|_{E^b} = o(\theta^{-1}) \) and hence

\[
\| S_\theta u \|_{E^b} \leq \| S_1 u \|_{E^b} + \left\| \int_1^\infty \frac{d}{d\theta} S_\theta u \, dx \right\|_{E^b} \leq \| S_1 u \|_{E^b} + \int_1^\infty \left\| \frac{d}{d\theta} S_\theta u \right\|_{E^b} \, dx < \infty
\]

Hence \( S_\theta u \) converges in \( E^b \). However \( S_\theta u \to u \) in \( E^0 \), hence \( u \in E^b \) and hence \( E_a^* \subset E^b \).

Naturally for \( a < b \) we have

\[
E_a^* \subset E^b \subset E^b\, \ast,
\]

as required.

**Lemma 6.5.** Given \([1, \infty) \ni t \mapsto u(t) \in E^0\) such that for some \( a_1, a_2 \) with \( 0 \leq a_1 < a_2 < \infty \), we have for \( t \geq 1 \),

\[
\| u(t) \|_{E^a_j} \leq C t^\lambda_{j-1}, \quad j = 1, 2,
\]

where \( \lambda_1 < 0 < \lambda_2 \). Define \( a = (\lambda_2 a_1 - \lambda_1 a_2) / (\lambda_2 - \lambda_1) \), then \( a \in (a_1, a_2) \) and

\[
U = \int_1^\infty u(t) \, dt \in E_a^*.
\]

Moreover we have \( \| U \|_{E^a_2} \leq C' C \) and

\[
\left\| \frac{d}{d\theta} S_\theta U \right\|_{E^b} \leq C_b \theta^{b-a-1}
\]

for any \( b \geq 0 \).

**Proof.** By (B) we have

\[
\left\| \frac{d}{d\theta} S(\theta) u(t) \right\|_{E^b} \leq C_b \min_{j=1,2} \theta^{b-a_j-1} t^{\lambda_j-1}.
\]

The two estimates are equal when \( t^{\lambda_2 - \lambda_1} = \theta^{a_2 - a_1} \). When \( t \) is below this bound the better estimate is obtained with \( j = 2 \) and \( j = 1 \) when \( t \) is above the bound. Since

\[
b - a_j - 1 + (a_2 - a_1) \lambda_j / (\lambda_2 - \lambda_1) = b - 1 - a.
\]

Integrate with respect to \( t \) to see

\[
\left\| \frac{d}{d\theta} S_\theta U \right\|_{E^b} \leq C_b \theta^{b-a-1}.
\]

Take \( b = a \) to see

\[
\| U \|_{E^a_2} \leq C' C
\]

and \( U \in E_a^* \).
Remark. This lemma exactly restores the facility of Lemma 4.2 in the case of weak scale, and it is not surprising with those new spaces we can restore our main theorem because in the process of generalising the Hölder scale problem to this general Banach scale problem, the only reasoning we cannot naturally preserve is the Lemma 4.2.

Corollary 6.6 (Additional properties of regularisation operators). We have a strengthened version of Definition 6.1 (B) and Proposition 6.2 (ii) that

\[
(B') \quad \left\| \frac{d}{d\theta} S_{\theta} u \right\|_{E^b} \leq C_{b} \theta^{b-a-1} \left\| u \right\|_{E^2}, \text{ for all } a, b,
\]

\[
(ii') \quad \left\| u - S_{\theta} u \right\|_{E^b} \leq C_{a, b} \theta^{b-a} \left\| u \right\|_{E^a} \text{ for } a > b.
\]

where both constants do not depend on \( \theta \).

Proof. For \( u \in E^a \) write

\[
u = S_1 u + \int_1^\infty \frac{d}{dt} S_t u \, dt = S_1 u + U.
\]

Invoke Lemma 6.5 with

\[
a_1 = 0, \quad \lambda_1 = -a, \quad a_2 = a + 1, \quad \lambda_2 = 1,
\]

together with \( \left\| S_1 u \right\|_{E^b} \leq C \left\| u \right\|_{E^0} \leq C \left\| u \right\|_{E^2} \), given by Proposition 6.2 (ii) to obtain

\[
\left\| \frac{d}{d\theta} S_{\theta} u \right\|_{E^b} \leq C_{b} \theta^{b-a-1} \left\| u \right\|_{E^2}.
\]

Furthermore integrate \( \frac{d}{d\theta} S_{\theta} u \) from 1 to \( \theta \) with respect to \( \theta \) to obtain \( (ii') \).

We define the weak convergence:

Definition 6.7 (Weak convergence). We say a sequence \( u_n \in E^a \) is weakly convergent to \( u \in E^a \) if for every \( b < a \) we have \( u_n \to u \) in \( E^b \).

Remark. We have an equivalent definition by convergence \( u_n \to u \) in \( E^b \) for every \( b < a \) by help of the injection between weak and strong spaces.

We end this section with a version of Leray-Schauder fixed point theorem:

Lemma 6.8 (Leray-Schauder). Let \( E \) be a Banach space, \( K \) a continuous map mapping the closed unit ball \( B_r \subset E \) into \( B_{r/2} \), with \( K(B_r) \) contained in a compact set. If \( v \in B_{r/2} \) we have that equation \( u + K(u) = v \) has a solution \( u \in B_r \) with \( \| u \|_E \leq \| v \|_E + r/2 \).

Proof. We extend the definition of \( K \) to \( E \) by defining \( K(u) = K(r^{-1} u / \| u \|_E) \) whenever \( \| u \|_E > r \). If \( 0 \leq \lambda \leq 1 \) and \( u \) is a fixed point of the map \( u \mapsto \lambda (v - K(u)) \), then \( \| u \|_E \leq \| v \|_E + r/2 \), which is independent of \( \lambda \). Hence the Leray-Schauder theorem proves that there is a fixed point for every \( \lambda \in [0, 1] \), and for \( \lambda = 1 \) proves this lemma. \( \square \)


6.2 Banach Scale Nash-Moser Theorem

We need to set some assumptions. Firstly we need two families of Banach spaces, \( \{ E^a \} \), \( \{ F^a \} \) and regularisation operators \( S^b_\theta, S^b_\Phi \) on them as described in Definition 6.1, and denote their weak scales by \( \{ E^a \} \), \( \{ F^a \} \) respectively. We assume further that the injection \( F^b_\theta \hookrightarrow F^a \) is compact for all \( b > a \).

We also have assumptions on tame operators. Let \( \alpha \) and \( \beta > 0 \) be fixed, \( [a_1, a_2] \) be an interval with \( 0 \leq a_1 < \alpha < a_2 \), \( V \) be a convex \( E^\alpha \) neighbourhood of 0, and \( \Phi : V \cap E^{a_2} \rightarrow F^\beta \). Assume

\((A_1')\) \( \Phi \) is of class \( C^2 \) in the previous settings for operators and satisfies the tame estimate for some \( \gamma > 0 \),

\[
\| \Phi''(u)(v, w) \|_{F^\beta + \gamma} \leq C \sum_{j \leq \beta_{\max}} \left( 1 + \| u \|_{E^\alpha_j} \right) \| v \|_{E^\alpha_j''} \| w \|_{E^\alpha_j''}.
\]

\((A_2')\) For \( v \in C \cap E_\infty \), \( \Phi'(v) \) has a right inverse \( \psi(v) : F^\infty \rightarrow E^{a_2} \), continuous as mapping \((v, g) \rightarrow \psi(v)g\), such that

\[
\| \psi(v)g \|_{E^\alpha} \leq C \left( \| g \|_{F^\beta + a - a} + \| g \|_{F^\alpha} \| v \|_{F^\beta + a} \right)
\]

for \( a_1 \leq a \leq a_2 \).

We state the main theorem:

**Theorem 6.9** (Abstract Nash-Moser theorem). Let \( a_2 \) be now less than all \( m_j', m_j'', m_j''' \), and

\[
\max \{ m_j' - \alpha, 0 \} + \max \{ m_j'', a_1 \} + m_j''' < 2\alpha,
\]

for every \( j \), and \( \alpha - \beta < a_1 \). For every \( f \in F^\beta_\gamma \) with \( \| f \|_{F^\beta} \) sufficiently small, one can find a sequence \( u_j \in V \cap E^{a_2} \) which has a weak limit \( u \in E^\alpha \) and \( \Phi(u_j) \) converges weakly in \( F^\beta \) to \( \Phi(0) + f \).

**Proof.** Let \( g \in F^\beta_\gamma \), take \( g_j = (S_{\theta_{j+1}} - S_{\theta_j})g \) and we have

\[
g = \sum_{j=0}^\infty \Delta_j g_j, \quad \| g_j \|_{F^{\beta}} \leq C_\beta \theta_j^{\alpha - \beta - 1} \| g \|_{F^\beta_\gamma}
\]

for \( \Delta_j = \theta_{j+1} - \theta_j \) and \( \theta_j = \left( \theta_0^{1/\varepsilon} + n \right)^\varepsilon \). Assume the perturbation \( \| g \|_{F^\beta_\gamma} \) is small and we claim that we can invoke the iteration scheme with initial \( u_0 = 0 \),

\[
u_{j+1} = u_j + \Delta_j \dot{u}_j, \quad \dot{u}_j = \psi(v_j)g_j, \quad v_j = S_{\theta_j}u_j,
\]

and we claim the following estimates, there is constant \( \delta \),

\[
\| \dot{u}_j \|_{E^\alpha} \leq \delta \| g \|_{F^\beta_\gamma} \theta_j^{\alpha - \alpha - 1}, \quad a_1 \leq a \leq a_2
\]

\[
\| v_j \|_{E^\alpha} \leq C_2 \| g \|_{F^\beta_\gamma} \theta_j^{\alpha - \alpha}, \quad \alpha < a \leq a_2
\]

\[
\| u_j - v_j \|_{E^\alpha} \leq C_3 \| g \|_{F^\beta_\gamma} \theta_j^{\alpha - \alpha}, \quad a \leq a_2.
\]
To validate the estimates, we conduct an iterative argument similar to our iterative hypotheses in the case of Hölder spaces. Suppose we have constructed $u_j$ for all $j \leq k$, with (53) and (54) holding for $j \leq k$, then automatically (52) holds for $j < k$ by aid of $\dot{u}_j = \Delta_j^{-1}[(u_{j+1} - v_{j+1}) - (u_j - v_j) + (v_{j+1} - v_j)]$ (Note $\theta_{j+1}/\theta_j \leq 2^\varepsilon$). At the boundary case $j = k$, use $(A'_2)$ to derive

$$\|\psi(v_k)g_k\|_{E^a} \leq C\left(\theta_k^{a-\alpha-1}\|g\|_{L^2} + \theta_k^{-\beta-1}\|g\|_{L^2}C_2\|g\|_{L^2}^{2} \theta_k^{\beta+a-\alpha}\right)$$

Here we have $\beta + a_1 > \alpha$ by (51) and hence (52) holds for case $j = k$ as well, when $\delta < C$ and $\|g\|_{E^a}$ is sufficiently small ($\delta$ does not depend on $C_2$). Since $u_{k+1} = \sum_{j=0}^{k} \Delta_j \dot{u}_j$ we obtain from Lemma 6.5 via discretisation that

$$\|u_{k+1}\|_{E^a} \leq C\delta\|g\|_{E^a}.$$  (55)

From Corollary 6.6 (ii) we see for $a < \alpha$ we have

$$\|u_{k+1} - v_{k+1}\|_{E^a} \leq C\|C_{a,\alpha}\|\|\delta\theta_{k+1}^{a-\alpha}\|\|g\|_{E^a} \leq C\theta_{k+1}^{a-\alpha}\|g\|_{E^a} \leq C\|g\|_{E^a}$$  (56)

if $a$ is not close to $\alpha$ if $C_3/\delta$ is large enough. In the limit case $a = a_2$, by adding (52) we obtain (56) at $a = a_2$. By the convexity inequality (12) we obtain (54) for $j \leq k + 1$ with same constant in the whole interval $a \in [0, a_2]$. By the property of regularisation operators and (55), (53) for $j = k + 1$ is also clear. Hence the iteration runs on and on such that (52)--(54) hold for all $j$. We have now proved that the construction of the infinite sequence $u_j, v_j, \dot{u}_j$ is possible for (55) and (54) are in $V$ when $\|g\|_{E^a}$ is sufficiently small. It follows from (52) that $u_k$ has a weak limit in $E^a_\alpha$. What remains is to examine the limit of $\Phi(u_k)$. Write

$$\Phi(u_{j+1} - \Phi(u_j)) = \Delta_j(e_j' + e_j'' + g_j)$$

where we adopt the notations $e_j'$ and $e_j''$ from (33) and (34). We want to obtain some estimates similar to our error estimate for the case of Hölder spaces. Consider

$$e_j'' = \int_0^1 \Phi''(v_j + t(u_j - v_j))(\dot{u}_j, u_j - v_j) \, dt$$

and $(A'_2)$, (52)--(54), (51) to derive that if $e > 0$ is sufficiently small that

$$\max\{m_j' - \alpha, 0\} + \max\{m_j'', a_1\} + m_j''' + e < 2\alpha$$

we will have

$$\|e_j''\|_{F^{a+\gamma}} \leq C\theta_j^{-1-e}\|g\|^2_{L^2}.$$ 

For any $N > 0$ we can pick $\varepsilon$ small to control the growth of $\theta_j$, that is, $\Delta_j = O(\theta_j^{-N})$. For large $N$ we obtain that

$$\|e_j''\|_{F^{a+\gamma}} \leq C\theta_j^{-1-e}\|g\|^2_{L^2}.$$
It follows that $\Phi(u_k)$ converges weakly to $\Phi(0) + T(g) + g$ where

$$T(g) = \sum_{j=0}^{\infty} \Delta_j \left( e'_j + e''_j \right).$$

The sum is uniformly convergent in $F_{\beta+\gamma}$ norm when $\|g\|_{F_\beta^a}$ is sufficiently small. Hence $T(g)$ is a continuous map from a neighbourhood of 0 in $F_\beta^a$ to a compact subset of $F_\beta^a$, and

$$\|T(g)\|_{F_\beta^a} \leq C\|g\|^2_{F_\beta^a}.$$

Invoke Leray-Schauder theorem we see $g + T(g)$ takes on all values in a neighbourhood of 0 and take $g + T(g) = f$ for $f \in F_\beta^a$ with $\|f\|_{F_\beta^a}$ sufficiently small we have $\|g\|_{F_\beta^a}$ small enough to extract $u_k$ as above such that $\Phi(u_k)$ converges to $\Phi(0) + f$ weakly.

**Remark** (Consequences of abstract Nash-Moser theorem). (i) If we take $E^a = F^a = C^a_\gamma$, the Hölder-Zygmund spaces, then we obtain an extension of Theorem 4.3 to integer indices (Hölder-Zygmund spaces). To get the existence part of Theorem 4.3 then exclude the integer indices.

(ii) Take $E^a = F^a = H^a_\gamma$, where $H^a_\gamma$ is the weak scale of Sobolev spaces $H^a$ as of modification (50). We can then establish a Nash-Moser theorem for weak scales $H^a_\gamma$, stating that for a tamed map $\Phi$ losing $\gamma$ derivatives, for perturbation $f$ with $\|f\|_{H^a_\gamma}$ sufficiently small, we can obtain $u_k \in H^{a-\gamma}_\gamma$ such that $u_k$ converges weakly to $u$ in $H^{a-\gamma}_\gamma$ and $\Phi(u_k)$ converges weakly to $\Phi(u_0) + f$ in $H^a_\gamma$. Indeed $H^a_\gamma$ is a strictly weaker scale that $H^a_\gamma \neq H^a$ for any $a$, that is $H^a_\gamma$ is strictly larger than $H^a$. By the injection between weak and strong scales, we can obtain an almost-sharp Sobolev space based theorem: for $\|f\|_{H^a}$ small, we have $u_k, u \in H^{a-\gamma}_\gamma$ for which $u_k \to u$ in any $H^{a-\gamma-\varepsilon}_\gamma$ with $\varepsilon > 0$, and $\Phi(u_k)$ converges weakly to $\Phi(u_0) + f$ in $H^a_\gamma$. One can obtain a sharp version ($\varepsilon = 0$) with other techniques, for example Baldi and Haus [3] Theorem 2.1, by facilitating an orthogonality property similar to our Proposition 3.3.

(iii) In general, one can prove a Nash-Moser theorem between any tame Fréchet spaces, which are graded Fréchet spaces with a tame structure, with reference to Hamilton [8].

(iv) There is a relatively recent result by Ekeland [5], in which an implicit function theorem between two graded Fréchet spaces is devised. Note that the assumptions of this new theorem are much weakened: (i) the objective mapping $\Phi$ should be continuous and Gâteaux-differentiable (not necessarily $C^2$, $C^1$ or even Fréchet differentiable); (ii) Fréchet spaces should be graded (not necessarily tame, not necessarily with regularisation operators). Its proof does not utilise Newton iterations, but uses dominated convergence theorem and Ekeland’s variational principle. The Nash-Moser theorems are special cases of this more general theorem. See also Ekeland and Séré [6].
7 Epilogue

7.1 Review

This essay has its skeleton and its order of presentation as in Alinhac and Gérard [2]; indeed, presentations of pseudodifferential operators, dyadic analysis results, our major Hölder scale Nash-Moser theorem (Theorem 4.3), local embedding theorems (Theorem 5.4, Theorem 5.10) are largely attributed to [2]. In Section 6 the modification of spaces and Banach scale Nash-Moser theorem are very much attributed to Hörmander [12] [14]. Mapping properties of pseudodifferential operators on Hölder-Zygmund spaces are attributed to the textbook of Abels [1]. Auxiliary results around the isometric embedding are well explained in Tao’s online notes [21]. Historic statements of mathematical works around this topic presented in introduction section are borrowed from [3] and [6].

It is also noted, though originality is hard to achieve in very limited time on this historical topic, some originality and deep understanding of this topic are still reflected in essay, by the following evidences: (i) a natural adaption from Hölder spaces to Hölder-Zygmund spaces of most dyadic results and of Günther’s approach; (ii) most of the proofs are with more details for better clarity; (iii) a very detailed estimate, as an original computation, is obtained in Lemma 5.9; (iv) many remarks provide deep intuition into and around the topic and refer to historic and recent developments (see remarks on pp. 16, 17, 24, 27, 44, 45, 54, 62).

7.2 Acknowledgement

Firstly I thank Clément Mouhot for kindly setting this essay topic, recommendation of references and assessing the essay. I thank Anthony Ashton for related discussions on pseudodifferential operators. I thank Matteo Capoferri for bringing me to the conflicts between constructiveness of Nash-Moser scheme and poor numerical performance of solving isometric embedding problems. I thank Fritz Hiesmayr for brief discussion on embedding of torus into flat spaces. I thank Ilia Kamotski for discussion on the Nash-Moser theorem and dyadic results in Hölder spaces. I thank Zhuo Min Lim for various discussions around the topic. I thank Nikolai Saveliev and Dmitri Vassiliev for related discussion on the reduction of isometric embedding problem back to the case of a torus.

7.3 Bibliography


