Around the Nash-Moser theorem

David Gérard-Varet

February 10, 2019

Preamble

These lecture notes correspond to the contents of a graduate course taught in University Paris Diderot in 2018. The general concern of these notes is the use of perturbative methods for solving equations in infinite dimension. Typically, we want to solve an equation of the form

$$F(u) = f, \tag{1}$$

where $F : X \to Y$ is a smooth function between two Banach spaces X and Y such that $F(0) = 0$, and where $\|f\|_Y$ is small enough. In many applications that we will consider, the mapping F will be a (nonlinear) partial differential operator, and X and Y will always be spaces of functions.

In the case where the differential at 0, denoted $F'(0)$, is an isomorphism from X to Y, the local inversion theorem provides us with a unique solution u of (1) close to 0 in X, attending that f is close enough to 0 in Y. The main tool behind the local inversion theorem is the Picard fixed point theorem: namely, a key idea is to reformulate equation (1) under the form

$$u = F'(0)^{-1} (f - F(u) + F'(0)u)$$

and show that it is a contraction from a small ball in X to itself. In this way, the solution u can be obtained as the limit of the iteration

$$u_{n+1} = F'(0)^{-1} (f - F(u_n) + F'(0)u_n). \tag{2}$$

The convergence typically holds at geometric rate, as

$$\|u_{n+1} - u_n\|_X \leq \kappa \|u_n - u_{n-1}\|_X$$

for some fixed $\kappa \in (0, 1)$, which implies $\|u_n - u\|_X \leq C\kappa^n$ for some $C > 0$.

The Picard iteration (2) can be compared to another famous iterative method, namely Newton's method. In our setting, the Newton-Raphson iteration can be written as:

$$u_{n+1} = F'(u_n)^{-1} (f - F(u_n) + F'(u_n)u_n). \tag{3}$$

On one hand, the Newton's iteration is more demanding: at each step, one must solve a different linearized equation, involving the operator $F'(u_n)$ rather than $F'(0)$. On the other hand, if $\|u_0\|_X$ and $\|f\|_Y$ are small enough, the Newton's scheme converges quadratically rather than geometrically: one has

$$\|u_{n+1} - u_n\|_X \leq c\|u_n - u_{n-1}\|_X^2 \tag{4}$$

which implies $\|u_n - u\|_X \leq C\kappa^{2^n}$ for some $C > 0$ and $\kappa = c\|u_1 - u_0\|_X$.

Overall, the common point between these methods is an assumption of invertibility of the linearized operator, with an estimate of the type:

$$F'(u)v = g \quad \Rightarrow \quad \|v\|_X \leq C\|g\|_Y$$

either for $u = 0$, or for u in a small neighborhood of 0. *However, there are many interesting and important situations where such nice invertibility assumption does not hold*. What may for instance happen is that the linearized equation is solvable, but only in a functional space which is larger than X. This larger space often corresponds to a loss of regularity. Typically, one can think of $X = X_{s_0}$ and $Y = Y_{s_0}$ as elements of two families of Banach spaces $(X_s)_{s\in\mathbb{N}}$ and $(Y_s)_{s\in\mathbb{N}}$, where s denotes a regularity index: for instance, $X_s = C_b^s(\mathbb{R}^d)$ (space of s-times continuously differentiable functions with bounded derivatives), or $X_s = H^s(\mathbb{R}^d)$ (Sobolev space of index s). Then, the bad situations are those in which F sends X_s to Y_s, but the inverse linearized operator $F'(u)^{-1}$ is only defined from Y_s to $X_{s-\sigma}$, for some $\sigma > 0$:

$$F'(u)v = g \quad \Rightarrow \quad \|v\|_{X_{s-\sigma}} \le C_\sigma \|g\|_{Y_s} \tag{5}$$

Obviously, such feature plagues the use of the standard iterative methods (2) or (3): at each step of the iteration, the regularity of the approximation u_n is lowered by σ, so that no more regularity is available after a finite number of steps.

This phenomenon of *loss of derivatives* will be the central topic of these lecture notes. It appears in many different physical and mathematical contexts: celestial mechanics (in connection to the so-called KAM theory and small divisor problems), riemannian geometry (in connection to the so-called isometric embedding problem), or in various PDE problems (for instance in the recent work by Mouhot and Villani on Landau damping [13]). We will of course examine several examples in these notes, and present the mathematical tools that are used to handle this kind of problems. These tools were mostly developped in the 50's and 60's, through the works of Kolmogorov, Arnold, Nash and Moser.

The bottom line of these works is that, while the Picard iteration becomes useless, a proper adaptation of Newton's iteration can still be used to solve (1): the very fast convergence of the Newton's scheme can somehow compensate for the loss of derivatives. Obviously, to take advantage of the former over the latter, both must be put onto a common ground: one needs to make the information about the loss of derivatives more quantitative. One way to proceed is to work in the framework of analytic functions, meaning that X_s or Y_s are spaces of analytic functions (with s being more or less the radius of analyticity). In such framework, the Cauchy formula allows to quantify the loss of derivatives in terms of loss of the radius of analyticity. This information can be implemented in the estimate (5), and then in an analogue of (4). This is in short the path followed by Kolmogorov and Arnold in the treatment of small divisor problems (see chapter 2). Another way to quantify the loss of derivatives is the Fourier transform: if a function has frequencies in a ball of radius N, a derivative "costs at most a factor N". This idea is at the basis of the Nash-Moser theorem, which allows to overcome the problem of the loss of derivatives in finite regularity (see chapter 3).

The outline of the lectures is as follows. The first chapter is a reminder on iterative methods for solving nonlinear equations (Picard, Newton). A simple example illustrating the phenomenon of loss of derivatives is provided, giving the motivation for the next chapters. The second chapter is devoted to circle diffeomorphisms. The goal of this chapter is to establish a theorem of Arnold about the analytic conjugacy of analytic diffeomorphisms of the circle. This theorem exhibits a first way to circumvent the loss of derivatives, by combining the properties of analytic functions together with the Newton's scheme. The third chapter is devoted to a discussion and proof of the famous Nash-Moser theorem, which allows to solve (1) in finite regularity. For the sake of clarity, we restrict in this chapter to a simple statement inspired by Moser's paper [14] (see also [24]). Chapter 4 is about the application of the Nash-Moser theorem to the isometric embedding problem. Chapter 5 is about some modern tools of Fourier analysis (Littlewood-Paley decomposition, paraproduct), which have revealed a good substitute to the Nash-Moser linearization technique in some PDE contexts (see for instance [3] on the Landau damping problem). Eventually, we conclude these lecture notes by discussing the famous De Giorgi-Nash-Moser theorem, on the Hölder regularity of solutions of linear scalar elliptic equations with measurable coefficients. We follow here the approach of De Giorgi [7, 5]. Although this last topic may seem a bit disconnected from the rest of the lectures, it turns out that one of the core arguments borrows to the same philosophy: the gain in regularity due to the ellipticity competes with the shrinking of the domain where interior elliptic estimates can be derived, a bit like the gain in the error of convergence competes with the loss of derivatives in the previous settings. Furthermore, one key idea is to express the gain due to elliptic regularity "in a nonlinear manner", to echo the nonlinear nature of the Newton's estimate (4)

Contents

Chapter 1

Iterative methods for solving nonlinear equations

1.1 Picard iteration

Let X, Y Banach spaces, and $F : X \to Y$, $F\ C^1$ with $F(0) = 0$.

Theorem 1.1. (local inversion theorem)
If the differential at 0*, denoted* $F'(0)$*, is an isomorphism from* X *to* Y*, there exists an open neighborhood* $U \times V$ *of* $(0,0)$ *in* $X \times Y$ *such that* F *is a* C^1 *diffeomorphism from* U *to* V*.*

Proof. The idea is to use the following global result.

Lemma 1.1. *Let* A *an isomorphism from* X *to* Y*,* $\varphi : X \to Y$ *a Lipschitz map with*

$$Lip(\varphi) < \frac{1}{||A^{-1}||_{L(Y,X)}}.$$

Then $F = A + \varphi$ *is a bijection from* X *to* Y*, with Lipschitz inverse. Moreover, if* F *is* C^1 *in an open set* U *of* X *and if* $F'(u)$ *in an isomorphim for all* $u \in U$*, then* F^{-1} *is* C^1 *on* $F(U)$*.*

Proof of the lemma.[1] To show that F is a bijection, we have to show that for any $f \in Y$, there exists a unique $u \in X$ such that $F(u) = f$. We reformulate this equation as

$$u = A^{-1}f - A^{-1}\varphi(u).$$

Hence, we have to show that for all $f \in Y$, the map $G_f(u) = A^{-1}f - A^{-1}\varphi(u)$ has a unique fixed point. This is a consequence of the Banach fixed point theorem: for all $u, v \in X$

$$\begin{aligned} \|G_f(u) - G_f(v)\|_X &= \|A^{-1}\varphi(u) - A^{-1}\varphi(v)\|_X \leq \|A^{-1}\| \, \|\varphi(u) - \varphi(v)\|_Y \\ &\leq \|A^{-1}\| \ \mathrm{Lip}(\varphi) \, \|u - v\|_X \end{aligned}$$

[1] This proof is borrowed to unpublished lecture notes of Albert Fathi, taught in ENS Lyon in 1997.

so that G is a contraction. For the Lipschitz regularity of F^{-1}, we take two elements $f = F(u)$ and $f' = F(u')$ in Y and write

$$u - u' = A^{-1}(f - f') + A^{-1}(\varphi(u) - \varphi(u'))$$

so that

$$\|u - u'\|_X \le \|A^{-1}\| \, \|f - f'\|_Y + \|A^{-1}\| \, \mathrm{Lip}(\varphi) \|u - u'\|_X,$$

and eventually

$$\|F^{-1}(f) - F^{-1}(f')\|_X \; = \; \|u - u'\|_X \; \le \; \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \mathrm{Lip}(\varphi)} \|f - f'\|_Y.$$

Hence, F is bi-Lipschitz. The second part of the lemma follows then easily from Lemma A.1, Appendix A.

Back to the proof of the local inversion theorem. We write $A = F'(0)$, $\varphi = F - A$. The application φ is C^1, with $\varphi'(0) = 0$. By continuity, there exists $r > 0$ such that $\sup_{u \in \overline{B}_r} \|\varphi'(u)\| < \frac{1}{4\|A^{-1}\|}$, where B_r (resp. $\overline{B}_r$) is the open (resp. closed) ball of radius r in X. This of course implies that over $\overline{B}_r$, $\mathrm{Lip}(\varphi) < \frac{1}{4\|A^{-1}\|}$. Let $\rho_r : X \to \overline{B}_r$, defined by

$$\rho_r(x) = x, \quad x \in \overline{B}_r, \quad \rho_r(x) = r \frac{x}{\|x\|}, \quad x \notin \overline{B}_r.$$

The proof of the following lemma is given in Appendix A:

Lemma 1.2. *For any $r > 0$, ρ_r is a Lipschitz map, with $Lip(\rho_r) \le 2$.*

We now set $\varphi_r = \varphi \circ \rho_r$. One has φ_r Lipschitz in X, with $\mathrm{Lip}(\varphi_r) < \frac{1}{2\|A^{-1}\|}$. By Lemma 1.1, $F_r = A + \varphi_r$ is a Lipschitz bijection with Lipschitz inverse. As $F_r = F$ on $\overline{B}_r$, one finds that $F(B_r)$ is open in F, and $F : B_r \to F(B_r)$ is a Lipschitz bijection with Lipschitz inverse. Moreover, up to take a smaller r, we can always assume that $F'(u)$ is an isomorphism for all $u \in B_r$. By the second part of Lemma 1.1, it yields the C^1 regularity of F^{-1}. $\square$

Remark 1.1. *The previous proof of the local inversion theorem emphasizes the key role of Picard fixed point theorem. The theorem itself relies on an iteration to construct the fixed point. In our context, the iterative process takes the form:*

$$u_{n+1} = A^{-1} f - A^{-1} \varphi(u_n) \quad \Leftrightarrow \quad F'(0) \, (u_{n+1} - u_n) + F(u_n) = f.$$

As is well-known, the convergence if of geometric type, that is $\|u_n - u\|_X \le C\kappa^n$, $\kappa \in (0,1)$.

1.2 Newton's method

Besides Picard iteration, another famous iteration for solving equations is given by Newton's (or Newton-Raphson's) method:

$$F'(u_n)(u_{n+1}-u_n)+F(u_n)=f.$$

To study its convergence, we introduce the "error" $\varepsilon_n = \|F(u_n)-f\|_Y$. In particular, taking $u_0 = 0$, we have $\varepsilon_0 = \|f\|_Y$. Let V a neighborhood of 0, and $c > 0$ such that $\sup_{u\in V} \|F'(u)^{-1}\|_{L(Y,X)} \leq c$. Then, *assuming that $u_n \in V$ for all n*, we deduce:

$$\|u_{n+1}-u_n\|_X \leq c\varepsilon_n \tag{1.1}$$

Then,

$$\begin{aligned}
\varepsilon_{n+1} &= \|F(u_{n+1})-f\|_Y \\
&\leq \|F(u_{n+1})-F'(u_n)(u_{n+1}-u_n)-F(u_n)\|_Y + \|F'(u_n)(u_{n+1}-u_n)+F(u_n)-f\|_Y \\
&= \|F(u_{n+1})-F'(u_n)(u_{n+1}-u_n)-F(u_n)\|_Y
\end{aligned}$$

Under the assumption that: for all $u, v \in V$,

$$\|F(v)-F(u)-F'(u)(v-u)\|_Y \leq C\|v-u\|^2 \tag{1.2}$$

we obtain

$$\varepsilon_{n+1} \leq C\|u_{n+1}-u_n\|^2 \leq Cc^2\varepsilon_n^2 \tag{1.3}$$

where the last inequality is deduced from (1.1). Setting $\varepsilon'_n = Cc^2\varepsilon_n$, the last inequality can be written as $\varepsilon'_{n+1} \leq (\varepsilon'_n)^2$, so that $\varepsilon'_n \leq (\varepsilon'_0)^{2^n}$. Hence, the error converges very fast to zero if $\varepsilon'_0 < 1$, that is if $\varepsilon_0 < 1/(Cc^2)$.

Exercice 1.1. *Complete the proof: show rigorously that $u_n \in V$ for all n and all estimates above.*

1.3 Loss of derivatives

The previous iterations rely crucially on the fact that $F'(0)^{-1}$ (or $F'(u)^{-1}$ for u near 0) is a continuous linear operator from Y to X. However, for many functionals F, notably many (nonlinear) partial differential operators, this property is not satisfied. We present here a crude example, as more interesting and involved ones will appear in the next chapters. We consider

$$F(u) = \partial_t u + \partial_x u + u\partial_x u$$

acting on functions $u = u(t,x)$, with $t \in [0,T]$, $x \in \mathbb{R}$. More precisely, we see F as acting from

$$X = \{\ u \in C^2_b([0,T]\times\mathbb{R}), u|_{t=0} = 0\} \quad \text{to} \quad Y = C^1_b([0,T]\times\mathbb{R}).$$

We remind that $C_b^k([0,T]\times\mathbb{R})$ is the space of functions of class C^k over $[0,T]\times\mathbb{R}$ such that

$$\|u\|_{C_b^k} \;:=\; \sup_{\alpha\in\mathbb{N}^2, |\alpha|\le k}\; \sup_{[0,T]\times\mathbb{R}} |\,\partial^\alpha u| \;<\; +\infty$$

Clearly, $F'(0)$ is given by $F'(0)v = \partial_t v + \partial_x v$, and defines a continuous linear operator from X to Y. As regards its invertibility, for all $f\in Y$, it is easy to verify that

$$v(t,x) = \int_0^t f(s, x-(t-s))\, ds$$

is the unique solution in Y satisfying $v|_{t=0} = 0$, and if we denote by $F'(0)^{-1}$ the mapping $f\to v$, one has $\|F'(0)^{-1} f\|_Y \le C\|f\|_Y$. But clearly, the lack of regularization in x prevents $F'(0)^{-1}$ to send Y into X. This phenomenon is called a *loss of derivatives*.

As mentioned in the preamble, many situations where the loss of derivatives appear involve families of Banach spaces X_s and Y_s, where s denotes a regularity index. One can think of $X_s = C^s(\mathbb{T})$ (1-periodic functions of class C^s), or $X_s = H^s(\mathbb{R}^n)$ (Sobolev functions of index s). The typical pathology corresponds to functionals $F : X_s \to Y_s$, such that $F'(0) : X_s \to Y_s$, but the inverse $F'(0)^{-1}$ only exists as a linear operator from X_s to X_{s-r} for some $r > 0$. This loss of regularity in inverting the linearized equation is incompatible with the classical Picard or Newton's schemes: regularity is lost after a finite number of iterations.

The goal of the following chapters will be to illustrate methods that allow to overcome this problem. The basic idea is that with a suitable modification, the Newton's scheme can resist this phenomenon : the quadratic gain in the error can win over the loss of derivatives. Of course, to have a competition between these two mechanisms, one must "quantify the loss of derivatives", and manage to encode it into an error estimate like (1.3). We will consider two cases:

- the case where the index loss r can be taken *arbitrarily small*, with estimates of the type: for all $s > 0$,

$$\|F'(0)^{-1} f\|_{X^{s-r}} \le C_{r,s}\|f\|_{Y^s}, \qquad \lim_{r\to 0} C_{r,s} = +\infty$$

  Such case is typical of analytic settings, and will be discussed in detail in Chapter 2 in the context of circle diffeomorphisms. Roughly, the X_s will be spaces of analytic functions with radius of analyticity s, and the Cauchy formula will allow to quantify the loss of derivatives through estimates of the type:

$$\forall r > 0, \quad \|\eta'\|_{X_{s-r}} \;\le\; \frac{1}{r}\|\eta\|_{X_s}.$$

- the case where the index loss r is *fixed*. This case enters the scope of the celebrated Nash-Moser theorem, see Chapter 3. Roughly, the idea behind this theorem is to modify the Newton's scheme by introducing truncations in frequencies. For instance, if we work with $X_s = C^s_{2\pi}$ (the space of 2π-periodic functions of class C^s), and define for $N \geq 0$ the truncation operator:

$$S_N f(x) = \sum_{|k|\leq N} \hat{f}(k) e^{ikx}, \quad \hat{f}(k) = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-it}dt$$

we can for instance take advantage of the Bernstein inequality [21], which implies

$$\|(S_N f)'\|_{L^\infty} \ \leq \ N\|S_N \ f\|_{L^\infty}$$

and in turn

$$\forall N \geq 1, \quad \|(S_N f)'\|_{X_s} \ \leq \ N\|S_N \ f\|_{X_s}$$

to quantify the loss of derivatives.

Chapter 2

Conjugacy of circle diffeomorphisms

2.1 Circle maps

Let $\mathbb{S}^1 = \{z, |z| = 1\}$ the unit circle, with the metric

$$d(z, z') = \inf\{|t - t'|, t, \quad t' \in \mathbb{R}, \quad z = e^{2i\pi t},\ z' = e^{2i\pi t'}\}$$

Let $\Pi : \mathbb{R} \to \mathbb{S}^1, \quad t \to e^{2i\pi t}$. Note that it is continuous, and that

$$\Pi_s : (s - \frac{1}{2}, s + \frac{1}{2}) \to \mathbb{S}^1 \setminus \{-e^{2i\pi s}\}, \quad \Pi_s(t) = \Pi(t)$$

is a homeomorphism for all $s \in \mathbb{R}$.

Definition 2.1. (Lift of a circle map)
Let $f : \mathbb{S}^1 \to \mathbb{S}^1$. A lift of f is a map $F : \mathbb{R} \to \mathbb{R}$ such that $\Pi \circ F = f \circ \Pi$.

Theorem 2.1. *Any continuous circle map f has a continuous lift. Moreover, if F_1 and F_2 are two continuous lifts, there exists $k \in \mathbb{Z}$ such that $F_1 = F_2 + k$.*

Proof. Uniqueness. If F_1 and F_2 are two continuous lifts of the same f, the relation $\Pi \circ F_1 = \pi \circ F_2$ shows that $F_1 - F_2$ takes integer values. But as it is continuous, it is constant (and equal to a fixed integer).

Existence. It is enough to show that there exists $F : [0, 1] \to \mathbb{R}$ continuous and satisfying $\Pi \circ F = f \circ \Pi$ over $[0, 1]$. Indeed, one can check that $\tilde{F}(t) = F(t - E(t)) + E(t)(F(1) - F(0))$ is then a continuous lift of f. The existence of such an F is deduced easily from the following property:

There exists $\varepsilon > 0$ such that for all $x \in [0, 1]$, if there exists an $F : [0, x] \to \mathbb{R}$ continuous and satisfying $\Pi \circ F = f \circ \Pi$ over $[0, x]$, then there exists $F_\varepsilon : [0, \min(x + \varepsilon, 1)] \to \mathbb{R}$ continuous and satisfying $\Pi \circ F_\varepsilon = f \circ \Pi$ over $[0, \min(x + \varepsilon, 1)]$.

To prove this property, we first express that f is uniformly continuous on $\mathbb{S}^1$: there exists $\varepsilon > 0$, such that $d(z,z') \leq \varepsilon \Rightarrow d(f(z),f(z')) < \frac{1}{4}$. Then, for F as above, the function F_ε defined by

$$F_\varepsilon(t) = F(t) \quad \text{on } [0,x], \quad F_\varepsilon(t) = \Pi^{-1}_{F(x)} \circ f \circ \Pi(t) \quad \text{on } [x, \min(x+\varepsilon, 1)]$$

is suitable. □

In what follows, all circle maps will be continuous, and the word lift will refer automatically to a continuous lift.

Definition 2.2. (Degree of a circle map)
Let f be a continuous circle map, F a lift of f. The degree of f is $deg(f) = F(1) - F(0)$.

Remark 2.1. (Properties of the degree)

- *By the uniqueness part of Theorem 2.1, $deg(f)$ depends only on f, not on the lift.*
- *By the definition of the lift, $e^{2i\pi F(1)} = f(e^{2i\pi}) = f(1) = e^{2i\pi F(0)}$ so that $deg(f)$ is an integer.*
- *One can show that for any lift F, any $x \in \mathbb{R}$ and any integer $k \in \mathbb{Z}$,*

  $$F(x+k) \;=\; F(x) \;+\; k\, deg(f)$$

  Indeed, as $F(\cdot + k)$ is a lift, there exists $n_k \in \mathbb{Z}$ such that $F(\cdot + k) = F + n_k$. Clearly, $n_0 = 0$, while $n_k - n_{k-1} = F(k) - F(k-1) = F(\cdot + k - 1)|_{x=1} - F(\cdot + k - 1)|_{x=0} = deg(f)$ by definition, because $F(\cdot + k - 1)$ is also a lift. The formula follows.
- *From the previous formula, it is easy to see that $deg(f \circ g) = deg(f)deg(g)$.*

Example 2.1. *$f(z) = z^n$, $n \in \mathbb{Z}$ is a circle map, with lift $F(t) = nt$. Hence, $deg(f) = n$.*

2.2 Circle homeomorphisms

We consider in this section circle homeomorphisms (which, as the circle is compact, are exactly the continuous and bijective circle maps). We say that a homeomorphism f is orientation-preserving (resp. orientation-reversing) if the image by f of any triplet of points in anti-clockwise order is in anti-clockwise order (resp. clockwise order). One can see that this is equivalent to f preserving (resp. reversing) the orientation of one triplet only, *cf* the proof of the next proposition. In other words, there are only two classes of circle homeomorphisms: those which preserve orientation and those which reverse orientation.

Proposition 2.1. *If f is an orientation-preserving circle homeomorphism, then $deg(f) = 1$, and any lift of F is an increasing homeomorphism from $\mathbb{R}$ to $\mathbb{R}$.*

Proof. As f is one-to-one, any lift F is one-to-one over $(0,1)$, and therefore monotonic over $[0,1]$. Moreover, if there exists $t \in (0,1)$ such that $F(t) - F(0) \in \{-1,+1\}$, one has $f(e^{2i\pi t}) = f(1)$, which again contradicts the fact that f is one-to-one. It follows that either $F((0,1)) \subset (F(0), F(0)+1)$ or $F((0,1)) \subset (F(0)-1, F(0))$. As f preserves orientation, $F|_{[0,1]}$ is increasing and $F(1) = F(0)+1$, or $\deg(f) = 1$. As $F(x+1) = F(x)+1$ (see Remark 2.1), it follows that F is an increasing homeomorphism from $\mathbb{R}$ to $\mathbb{R}$. □

Proposition 2.2. *Let f be an orientation-preserving circle homeomorphism, and F a lift of f. Then, the number $\rho_0(F) = \lim_{n\to+\infty} \frac{F^n(x)}{n}$ exists and does not depend on x. Moreover, $\rho_0(F+k) = \rho_0(F) + k$ for all $k \in \mathbb{Z}$.*

Proof. From the property $F(x+k) = F(x) + k$, see Remark 2.1, it is easily deduced that $F^n(x+k) = F^n(x) + k$ (for all x, for all $n, k \in \mathbb{Z}$).

Let us first prove that $\lim_n \frac{F^n(0)}{n}$ exists. We remind the following well-known lemma:

Lemma 2.1. *Let $(a_n)_{n\in\mathbb{N}}$ be a real sequence satisfying $a_{n+m} \le a_n + a_m$ for all n, m. Then, $\lim_{n\to+\infty} \frac{a_n}{n} = \inf_{n\in\mathbb{N}} \frac{a_n}{n}$ exists in $\mathbb{R} \cup \{-\infty\}$.*

Let $x, y \in \mathbb{R}$, and $k \in \mathbb{Z}$ such that $x + k \le y \le x + k + 1$. Applying the increasing function F^n, we obtain

$$F^n(x+k) \le F^n(y) \le F^n(x+k+1) \quad \Leftrightarrow \quad F^n(x) + k \le F^n(y) \le F^n(x) + k + 1.$$

By substracting the double inequality $x + k \le y \le x + k + 1$, we end up with

$$F^n(x) + y - x - 1 \le F^n(y) \le F^n(x) + y - x + 1 \tag{2.1}$$

Taking $x = 0$, $y = F^m(0)$, and setting $a_n = -F^n(0) + 1$, we get $a_{n+m} \le a_n + a_m$. By the lemma, this shows that $\lim_{n\to+\infty} \frac{F^n(0)}{n}$ exists in $\mathbb{R} \cup \{+\infty\}$. Moreover, as F^k is increasing for all $k \in \mathbb{N}$,

$$\begin{aligned} F^n(0) = F^{n-1}(F(0)) &= F^{n-1}\big(F(0) - E(F(0))\big) + E(F(0)) \\ &\le F^{n-1}(1) + E(F(0)) \le F^{n-1}(0) + 1 + E(F(0)). \end{aligned}$$

It implies that $F^n(0) \le C(n+1)$ for some C and for all n, which implies in turn that $\lim_{n\to+\infty} \frac{F^n(0)}{n}$ exists (and is finite).

To conclude the proof of the proposition, we write $F^n(x) = F^n(x - E(x)) + E(x)$, so that

$$F^n(0) \le F^n(x - E(x)) \le F^n(1) = F^n(0) + 1.$$

We get

$$\frac{E(x)}{n} \le \frac{F^n(x)}{n} - \frac{F^n(0)}{n} \le \frac{E(x)+1}{n}.$$

Hence for all x, $\lim_{n\to+\infty} \frac{F^n(x)}{n}$ exists and is equal to $\lim_{n\to+\infty} \frac{F^n(0)}{n}$. Finally, the relation $\rho_0(F+k) = \rho_0(F) + k$ is deduced easily from the relation $F^n(x+k) = F^n(x) + k$. □

Exercice 2.1. *Show that for any orientation preserving homeomorphism f and continuous lift F of f, the quantity $\frac{F^n(x)-x}{n}$ converges uniformly in x to $\rho_0(F)$ as $n \to +\infty$.*

The previous proposition allows to set the following

Definition 2.3. (rotation number) *Let f be an orientation-preserving circle homeomorphism, F a lift of f. Its rotation number is defined as $\rho(f) = \rho_0(F)$ (mod 1).*

Example 2.2. *Let $\alpha \in \mathbb{R}$, $R_\alpha(z) = e^{2i\pi\alpha}z$. A lift of R_α is $F(t) = t + \alpha$. Hence, $\rho_0(F) = \alpha$, and $\rho(f) = \alpha$ mod 1.*

Exercice 2.2. *Let f a circle homeomorphism that reverses orientation.*

1. *Show that any lift F of f is decreasing from $[0, 1]$ to $[F(0) - 1, F(0)]$.*

2. *Show that there exists a lift F of f having a fixed point over $[0, 1]$.*

3. *Show that for all x, $\lim_{n\to+\infty} \frac{F^n(x)}{n} = 0$.*

Proposition 2.3. (A few properties of the rotation number)

- *For all $n \in \mathbb{Z}$, $\rho(f^n) = n\rho(f)$.*

- *$f \to \rho(f)$ is continuous from the set of orientation-preserving circle homeomorphisms (endowed with the sup norm) to $\mathbb{R}/\mathbb{Z}$.*

- *$\rho(f) \in \mathbb{Q}$ if and only if f has a periodic point.*

- *Let $h : \mathbb{S}^1 \to \mathbb{S}^1$ a continuous circle map with deg$(h) = 1$, and f, g two orientation-preserving homeomorphisms such that $h \circ f = g \circ h$. Then, $\rho(f) = \rho(g)$. In particular, if h is an orientation-preserving homeomorphism, $\rho(h \circ f \circ h^{-1}) = \rho(f)$.*

Proof. Let F be a lift of f. The first item follows from the series of equalities:

$$\rho_0(F^n) = \lim_{k\to+\infty} \frac{(F^n)^k(0)}{k} = n \lim_{k\to+\infty} \frac{F^{(nk)}(0)}{nk} = n\rho_0(F)$$

For the second item, we fix $\varepsilon > 0$ and n such that $\frac{1}{n} < \varepsilon$. From (2.1), we deduce that

$$F^n(0) - 2 < F^n(t) - t < F^n(0) + 2 \quad \forall t \in \mathbb{R}. \tag{2.2}$$

For homeomorphisms g close enough to f in C^0 topology, and for a proper choice of of the lift G of g, one can make $\|F^n - G^n\|_\infty$ arbitrarily small. Hence, we can ensure that

$$F^n(0) - 2 < G^n(t) - t < F^n(0) + 2 \quad \forall t \in \mathbb{R}. \tag{2.3}$$

Using that for any $k \in \mathbb{N}$:

$$F^{nk}(0) = F^{nk}(0) - 0 = \sum_{j=0}^{k-1} F^n \circ F^{(jn)}(0) - F^{(jn)}(0)$$

we can deduce from (2.2) that

$$k(F^n(0) - 2) < F^{nk}(0) < k(F^n(0) + 2)$$

and similarly, we can deduce from (2.3)

$$k(F^n(0) - 2) < G^{nk}(0) < k(F^n(0) + 2)$$

Dividing by nk and letting k go to infinity, we get

$$\frac{F^n(0) - 2}{n} \leq \rho_0(F) \leq \frac{F^n(0) + 2}{n}, \quad \frac{F^n(0) - 2}{n} \leq \rho_0(G) \leq \frac{F^n(0) + 2}{n}$$

which implies $|\rho_0(F) - \rho_0(G)| \leq \frac{4}{n} \leq 4\varepsilon$.

We now focus on the third item. If f has a periodic point $z = e^{2i\pi x}$ of period q, then $\Pi \circ F^q(x) = \Pi(x)$, so that there exists $p \in \mathbb{Z}$ such that $F^q(x) = x + p$. Inductively, we get that $F^{nq}(x) = x + np$, and $\rho_0(F) = \frac{p}{q}$.

Conversely, if $\rho_0(F) = \frac{p}{q}$, then, using the first item of the proposition, $G = F^q - p$ is a lift of f^q that satisfies $\rho_0(G) = 0$.

- if $G(0) > 0$, one obtains that $(G^n(0))_{n\in\mathbb{N}}$ is an increasing sequence, because G is increasing. Then, either it is is bounded from above by 1 and converges to a fixed point of G, as expected, or there exists $k \geq 1$ such that $G^k(0) > 1$. In the latter case, we get $G^{2k}(0) > G^k(1) > G^k(0) + 1$. Inductively, one shows that for all $n \geq 1$, $G^{nk}(0) > n$, so that $\rho_0(G) \geq \frac{1}{k}$, reaching a contradiction.

- if $G(0) < 0$, one can proceed similarly with the decreasing sequence $(G^n(0))_{n\in\mathbb{N}}$.

As regards the last item, let F, G, H lifts of f, g, h respectively. One can check easily that $H \circ F$ and $G \circ H$ are two lifts of $h \circ f = g \circ h$, so that there exists $N \in \mathbb{Z}$ such that $H \circ F = \tilde{G} \circ H$, where $\tilde{G} = G + N$. This implies $H \circ F^n = \tilde{G}^n \circ H$ for all $n \in \mathbb{Z}$. Now,

$$\rho_0(\tilde{G}) = \lim_n \frac{\tilde{G}^n(H(x))}{n} = \lim_n \frac{H(F^n(x))}{n} = \lim_n \frac{H(F^n(x) - E(F^n(x)))}{n} + \frac{E(F^n(x))}{n}$$

where E is the integer part, and the last equality comes from the property: $H(x + k) = H(x) + k$, see Remark 2.1. As H is bounded on $[0, 1]$, the first term at the right-hand side goes to zero, while the second one goes to $\rho_0(F)$. This concludes the proof. $\square$

Exercice 2.3. *Let f be an orientation preserving homeomorphism, with $\rho(f) = \frac{p}{q}$, $p \in \mathbb{Z}^*$, $q \in \mathbb{N}^*$, $p \wedge q = 1$. Show that f has a periodic point of period q, but not of period q' with $0 \leq q' < q$.*

Exercice 2.4. *Let f be a circle homeomorphism. For all $x \in \mathbb{S}^1$, we introduce the ω-*limit set of x (under f)*: $\omega(x, f) = \cap_{n\in\mathbb{N}}\overline{\{f^k(x), k \geq n\}}$.*

1. *Show that if f is orientation preserving and has a periodic point x_0 of period q, then for any $x \in \mathbb{S}^1$, $f^{nq}(x)$ converges to a fixed point of f^q as $n \to +\infty$ (which is not necessarily x_0). Show that such a point is a periodic point of f of period q.*
2. *Deduce from the previous question and the previous exercise that if $\rho(f) = \frac{p}{q}$, $p \in \mathbb{Z}^*$, $q \in \mathbb{N}^*$, $p \wedge q = 1$, then for all $x \in \mathbb{S}^1$, $\omega(x, f)$ is a periodic orbit of period q.*

We see from the last point of the previous proposition that the rotation number is invariant by conjugacy. It is then natural to wonder if the converse holds, that is if two (orientation-preserving) homeomorphisms having the same rotation number are in the same conjugacy class. In other words, one can ask if a homeomorphism with rotation number α is conjugated to the rotation R_α. In the case where $\alpha \in \mathbb{Q}$ (mod 1), this is clearly not enough: for instance, any orientation-preserving homeomorphism which fixes 1 has rotation number 0, but is not necessarily the identity. In the case where α is irrational, the problem is deeper. What is always true is the following weaker result:

Proposition 2.4. *Let f be an orientation-preserving diffeomorphism with irrational rotation number α. Then, f is semi-conjugated to R_α, which means that there exists a continuous circle map h, with $\deg(h) = 1$ and a non-decreasing lift (but not necessarily one-to-one), such that $h \circ f = R_\alpha \circ h$.*

Proof. [1] By the Krylov-Bogolyubov theorem [22, p10], there exists a Borel probability measure μ on $\mathbb{S}^1$ which is invariant by f. As f is a homeomorphism, this amounts to the identity $\mu(f(A)) = \mu(A)$ for any Borel set A. As the rotation number of f is irrational, f has no periodic point, *cf* Proposition 2.3. Hence, for any $a \in \mathbb{S}^1$, the $f^n(a)$, $n \in \mathbb{Z}$, are all distinct. This implies

$$1 = \mu(\mathbb{S}^1) \geq \sum_{n\in\mathbb{Z}} \mu(\{f^n(a)\}) = \sum_{n\in\mathbb{Z}} \mu(\{a\}), \quad \text{so that } \mu(\{a\}) = 0.$$

Hence, μ has no atom. It follows that the circle map $h(x) = \Pi \circ \mu([1, x])$, where $[1, x]$ is the arc joining 1 to x anticlockwise, is continuous. It is also orientation-preserving and satisfies $\deg(h) = 1$. Note that a lift H of h is given by $H(y) = \mu([1, \Pi(y)]) + E(y)$. Now, for all $x \in \mathbb{S}^1$, taking $A = [1, x]$ and noticing that $f([1, x]) = [f(1), f(x)]$ because f is orientation-preserving, we deduce that $\mu([f(1), f(x)]) = \mu([1, x])$, which implies $zh \circ f(x) = h(x)$, where $z = \Pi \circ \mu([(f(1), 1])$. This means that there exists β such that $h \circ f(x) = R_\beta \circ h(x)$. But, by Proposition 2.3, $\rho(f) = \rho(R_\beta)$, which means $\alpha = \beta$, and the result follows. □

[1] This proof borrows to unpublished lecture notes of Albert Fathi, taught in ENS Lyon in 1998.

Exercice 2.5. (Proof of the Krylov-Bogolyubov theorem)
Let K a compact metric space, $T : K \to K$ continuous. Let ν_0 any finite borelian measure on K.

1. *For all $k \in \mathbb{N}$, we introduce $\nu_k = T_*^k \nu_0$, defined by $\nu_k(A) = \nu_0(T^{-k}(A))$. Explain why the sequence of averaged measures $(\mu_k = \frac{1}{l}\sum_{l=0}^{k-1} \nu_l)_{k\geq 1}$ has a subsequence $(\mu_{\varphi(k)})_{k\geq 1}$ that weakly converges to some finite borelian measure μ.*
2. *Deduce from there that for all continuous function f on K, $\frac{1}{\varphi(k)}\sum_{l=0}^{\varphi(k)-1}\int_K f\circ T^{l+1} d\nu_0$ converges to $\int_K f \circ T d\mu$ as $k \to +\infty$.*
3. *Show that it also converges to $\int_K f d\mu$, so that μ is an invariant measure for T.*

To determine if f is conjugated to R_α, one can examine under which conditions the circle map h given by the previous proposition is one-to-one. Note that as α is irrational, $R^{n_1}(x) \neq R^{n_2}(x)$ for all x and all $n_1 \neq n_2$. It follows that $h \circ f^{n_1}(x) \neq h \circ f^{n_2}(x)$ for all $n_1 \neq n_2$. If we can show that f has a dense orbit, then we will find x, $n_1 \neq n_2$ such that for any $z \neq z'$ $z < f^{n_1}(x) < f^{n_2}(x) < z'$ (the symbol $<$ refering to the trigonometric order). As h preserves orientation, it will follow that $h(z) \neq h(z')$, proving conjugacy of f to a rotation. We will see that this is the case when f is regular enough.

Definition 2.4. (Circle diffeomorphism) *We say that a circle map f is C^k, for $k \in \mathbb{N}^* \cup \{\infty\} \cup \{\omega\}$[2], if it has a lift which is C^k over $\mathbb{R}$. It is a circle diffeomorphism of class C^k if it is C^k, bijective, with a C^k inverse.*

Theorem 2.2. (Denjoy's theorem) *Let f be a circle diffeomorphism of class C^2, with irrotational rotation number α. Then, f is topologically conjugated to R_α (that is conjugated by an orientation-preserving homeomorphism).*

Remark 2.2. *If h_1 and h_2 are two homeomorphisms conjugating f to R_α, then*

$$h_1 \circ R_\alpha \circ h_1^{-1} = h_2 \circ R_\alpha \circ h_2^{-1}$$

so that $h = h_1^{-1} \circ h_2$ satisfies $R_\alpha \circ h = h \circ R_\alpha$, that is : $\forall x$, $e^{2i\pi\alpha}h(x) = h(e^{2i\pi\alpha}x)$. From there:

$$\forall n \in \mathbb{Z}, \quad e^{2i\pi n\alpha}h(1) = h(e^{2i\pi n\alpha}).$$

By density of $(e^{2i\pi n\alpha})_{n\in\mathbb{Z}}$ in $\mathbb{S}^1$, it follows that $h(z) = h(1)z$ for all $z \in \mathbb{S}^1$, so that $h_1 = h_2 \circ R_\beta$ for some β. Hence, the conjugacy is "almost unique".

Proof. [3] Let h given by Proposition 2.4. Let us assume that h is not one-to-one. Then there exists a non-trivial arc I on which h is constant. From the relation $h \circ f = R_\alpha \circ h$, we deduce that $h \circ f^n = R_\alpha^n \circ h$. In particular, h is constant on $f^n(I)$. Moreover, as $R_\alpha^{n_1}(x) \neq R_\alpha^{n_2}(x)$

[2] the notation C^ω refers to analytic functions.
[3] This proof borrows to unpublished lecture notes of Albert Fathi, taught in ENS Lyon in 1998.

for any x as soon as $n_1 \neq n_2$, we see that the arcs $f^n(I)$, $n \in \mathbb{Z}$, must be disjoint. Denoting $I_n = f^n(I)$ and l_n the length of I_n, this implies in turn that $l_n \to 0$ as $n \to \pm\infty$. Now,

$$\begin{aligned} l_n &= \int_{\Pi^{-1}(I_n)\cap[0,1]} ds = \int_{F^{-n}(\Pi^{-1}(I_n)\cap[0,1])} (F^n)'(t)dt \\ &= \int_{F^{-n}(\Pi^{-1}(I_n))\cap[F^{-n}(0),F^{-n}(1)]} (F^n)'(t)dt = \int_{\Pi^{-1}(I)\cap[0,1]} (F^n)'(t)dt \end{aligned}$$

where we used the 1-periodicity of $(F^n)'$ and the identity : $F^{-n}(\Pi^{-1}(I_n)) = \Pi^{-1}(I)$. Hence, we find that for some $\theta_n \in \Pi^{-1}(I) \cap [0,1]$ (with l_I the length of I):

$$l_n = (F^n)'(\theta_n)\, l_I = \prod_{i=1}^{n-1} F'(F^i(\theta_n))\, l_I.$$

With a similar calculation, we find that for some $\theta'_n \in \Pi^{-1}(I_{-n}) \cap [0,1]$,

$$l_I = (F^n)'(\theta'_n)\, l_{-n} = \prod_{i=1}^{n-1} F'(F^i(\theta'_n))\, l_n.$$

As $F' > 0$, this leads to

$$\begin{aligned} \ln(l_n\, l_{-n}) &= \sum_{i=0}^{n-1} \big(\ln F'(F^i(\theta_n)) - \ln F'(F^i(\theta'_n))\big) \;+\; 2\ln l_I \\ &= \sum_{i=0}^{n-1} \Big(\ln F'\left(F^i(\theta_n) - E(F^i(\theta_n))\right) - \ln F'\left(F^i(\theta'_n) - E(F^i(\theta'_n))\right)\Big) \;+\; 2\ln l_I \end{aligned}$$

where the last equality comes from the 1-periodicity of F'.

We now claim that : *there exists a strictly increasing sequence of integers n_l, such that for each $l \geq 0$, the arcs $[f^i \circ \Pi(\theta_{n_l}), f^i \circ \Pi(\theta'_{n_l})]$, $i = 0, \ldots, n_l - 1$ are disjoint.*

Let us assume temporarily that this claim holds. For all i, for all n,

$$\Pi^{-1}[f^i \circ \Pi(\theta_n), f^i \circ \Pi(\theta'_n)] \cap [0,1]$$

is either a single interval, of the type

$$[F^i(\theta_n) - E(F^i(\theta_n)), F^i(\theta'_n) - E(F^i(\theta'_n))]$$

(possibly switching θ_n and θ'_n), or a union of two disjoint intervals, of the type

$$[0, F^i(\theta_n) - E(F^i(\theta_n))] \,\cup\, [F^i(\theta'_n) - E(F^i(\theta'_n)), 2\pi].$$

By the claim, when $n = n_l$, and i varies between 0 and $n_l - 1$, we obtain a family of disjoint subintervals of $[0,1]$. Moreover, $\ln F'$ is 1-periodic, and it is C^1 (therefore with bounded variation) over $[0,1]$. From there, writing

$$\ln F'(F^i(\theta_n)) - \ln F'(F^i(\theta'_n)) = \ln F'\left(F^i(\theta_n) - E(F^i(\theta_n))\right) - \ln F'\left(F^i(\theta'_n) - E(F^i(\theta'_n))\right)$$

in the case of a single interval and

$$\ln F'(F^i(\theta_n)) - \ln F'(F^i(\theta'_n)) = \ln F'\left(F^i(\theta_n) - E(F^i(\theta_n))\right) - \ln F'(0) \\ + \ln F'(1) - \ln F'\left(F^i(\theta'_n) - E(F^i(\theta'_n))\right)$$

in the case of a union of two intervals, we get that the right-hand side of the previous equality is bounded uniformly in $n = n_l$. But the left hand-side $\ln(l_{n_l}\, l_{-n_l}) \to -\infty$ as $l \to +\infty$. We reach in this way a contradiction.

It remains to prove the claim. We remind once more that $h \circ f^i = R^i_\alpha \circ h$ for all $i \in \mathbb{Z}$. As h is constant on I (say equal to z), we have $h \circ \Pi(\theta_{n_l}) = z$ and $h \circ \Pi(\theta'_{nl}) = R_\alpha^{-n_l}(z)$. It follows that

$$h\left([f^i \circ \Pi(\theta_{n_l}), f^i \circ \Pi(\theta'_{n_l})]\right) = [R^i_\alpha(t), R^{i-n_l}_\alpha(t)]$$

It is then enough to show that the intervals at the right hand-side are disjoint for some good sequence n_l and $i = 0, \dots, n_l-1$. There is no loss of generality in taking $t = 1$. We then introduce the following sequence:

$$n_0 = 1, \quad n_{l+1} = \inf\{n \geq n_l,\ d(1, R^n_\alpha(1)) < d(1, R^{n_l}_\alpha(1))\}.$$

Note that the infimum is well-defined, by the density of the orbits of R_α (α is irrational). The sequence n_l is strictly increasing and goes to infinity. Moreover, for all $p_1 < p_2$, p_1, p_2 in $\{0, ..., n_l - 1\}$, one has

$$d(R^{p_1}_\alpha(1), R^{p_2}_\alpha(1)) = d(1, R^{p_2-p_1}_\alpha(1)) > d(1, R^{n_l}_\alpha(1)).$$

Thanks to this property, one can show that the arcs $[R^i_\alpha(1), R^{i-n_l}_\alpha(1)]$ are disjoint, which concludes the proof. $\square$

Let us point out that the smoothness assumption on f in Denjoy's theorem is crucial: Denjoy himself exhibited a counterexample with C^1 regularity. On the other hand, the conjugacy h of the theorem is only shown to be continuous. A natural question is then: *if f is smooth, is the conjugacy h itself smooth ?* We shall in the next sections discuss the special case of a perturbative and analytic framework.

2.3 Conjugacy of analytic diffeomorphisms

2.3.1 Setting of the problem

Let f be an analytic diffeomorphism of the circle, orientation-preserving. We want to study the regularity properties of a map h that conjugates f to the rotation R_α, where $\alpha = \rho(f)$ is in $\mathbb{R} \setminus \mathbb{Q}$. We will restrict to the case where f is a small perturbation of R_α. Roughly, we will assume that f has a lift F of the form $F(x) = x + \alpha + \eta(x)$, where η will be small in an analytic norm (to be specified later on). Let us note that the relation $F(x+1) = F(x) + 1$ implies the relation $\eta(x+1) = \eta(x)$, so that η is 1-periodic.

Remark 2.3. *We know from the definition of the rotation number that* $\rho_0(F) = \alpha + k$ *for some* $k \in \mathbb{Z}$*. Moreover, for* η *small enough in analytic norm, one can ensure that* $|\rho_0(F) - \rho_0(x \to x + \alpha)| \leq \frac{1}{2}$*, see the proof of the second point in Proposition 2.3. Hence,* $\rho_0(F) = \alpha$*, which puts some constraint on* η*. Namely, one has* $F^n(x) = x + n\alpha + \sum_{i=0}^{n-1} \eta(F^i(x))$*, so that*

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} \eta(F^i(x)) = 0.$$

This obviously implies that η *vanishes at some point, a property that will be used later.*

As f is a small perturbation of R_α, it is natural to look for the conjugacy h as a small perturbation of the identity. More precisely, we will look for a lift H under the form $H(x) = x + U(x)$, where the 1-periodic function U will also be small in some appropriate analytic norm.

Formally, the conjugacy relation writes: for all $x \in \mathbb{R}$,

$$\begin{aligned} F \circ H(x) = H(x + \alpha) &\Leftrightarrow H(x) + \alpha + \eta(H(x)) = H(x + \alpha) \\ &\Leftrightarrow U(x + \alpha) - U(x) = \eta(x + U(x)) \end{aligned}$$

A natural approach to build a solution U to the last equation is to look for successive approximations, taking as a first approximation the solution of the linear equation

$$\forall x \in \mathbb{R}, \quad U(x + \alpha) - U(x) = \eta(x) \tag{2.4}$$

As U and η are 1-periodic, it is natural to solve them using Fourier series:

$$U(x) = \sum_{n \in \mathbb{Z}} \hat{U}(n) e^{2i\pi n x}, \quad \eta(x) = \sum_{n \in \mathbb{Z}} \hat{\eta}(n) e^{2i\pi n x}.$$

We find for all $n \in \mathbb{Z}$:

$$(e^{2i\pi n\alpha} - 1)\hat{U}(n) = \hat{\eta}(n). \tag{2.5}$$

Two difficulties arise from this linearized approach. First, in the case $n = 0$, the compatibility condition $\hat{\eta}(0) = 0$ is required. More importantly, one faces a *small divisor problem* when writing

$$\hat{U}(n) = \frac{\hat{\eta}(n)}{e^{2i\pi n\alpha} - 1} \quad \forall n \neq 0.$$

Indeed, as α is irrational, the denominator does not vanish for $n \neq 0$, but it can still take arbitrarily small values, which can prevent the convergence of the Fourier series which is supposed to define U.

To control the smallness of the denominator, it is then natural to put some condition on α, which expresses the fact that α can not be approximated too well by rational numbers. A convenient condition is the following diophantine condition: *there exists* $\nu > 2, K > 0$ *such that*

$$(H_{K,\nu}) \qquad \forall (m,n) \in \mathbb{Z} \times \mathbb{Z}^*, \quad |\alpha - \frac{m}{n}| \ \geq \frac{K}{|n|^\nu}$$

As is well-known, this property is generic: more precisely, one has the following

Lemma 2.2. *For all* $\nu > 2$*, the set of* α *satisfying* $(H_{K,\nu})$ *is of full measure in* $\mathbb{R}$.

Under assumption $(H_{K,\nu})$, the inequality

$$\forall n \neq 0, \quad |e^{2i\pi n\alpha} - 1| \geq \frac{4K}{|n|^{\nu-1}} \tag{2.6}$$

holds. Indeed, for $m \in \mathbb{Z}$ such that $|n\alpha - m| \leq \frac{1}{2}$, one can write

$$|e^{2i\pi n\alpha} - 1| = |e^{2i\pi(n\alpha - m)} - 1| = 2|\sin(\pi(n\alpha - m))| \geq 4|n\alpha - m|$$

using the inequality $\sin\theta \geq \frac{2}{\pi}\theta$, valid for $0 \leq \theta \leq \frac{\pi}{2}$. The inequality follows.

Still, in spite of this control, the resolution of (2.4) exhibits *the phenomenon of loss of derivatives* mentioned in the preamble to these lecture notes. Introducing for instance the periodic Sobolev norm over 1-periodic functions with zero mean:

$$\|F\|_{H^s_{per}} \ = \ \left(\sum_{n \in Z} |n|^{2s} |\hat{F}(n)|^2 \right)^{1/2}, \quad s \geq 0$$

the inequality (2.6) yields

$$\|U\|_{H^s_{per}} \ \leq \frac{1}{4K} \|\eta\|_{H^{s+\nu-1}_{per}}.$$

As explained in the preamble to the lecture notes, this loss is incompatible with a standard resolution through linearization. Up to our knowledge, this kind of small divisor problems was first noticed in the context of celestial mechanics. We will see here how to overcome such problems in the analytic setting.

2.3.2 A reminder about analytic functions

We will consider here functions that are 1-periodic and analytic over the real line. More precisely, we introduce for all $\sigma > 0$:

$$\begin{aligned} &B_\sigma = \{x + iy, \quad x \in \mathbb{R}, |y| < \sigma\}, \\ &X_\sigma = \{\eta \ C^\infty, \text{1-periodic over } \mathbb{R}, \\ &\qquad \text{such that } \eta \text{ has an holomorphic extension over } B_\sigma, \text{ continuous over } \overline{B}_\sigma\}. \end{aligned}$$

Note that the extension of $\eta \in X_\sigma$ is unique by analytique continuation. To lighten notations, we will still denote by $\eta = \eta(z)$ such extension. Note also that

$$\eta(x) = \eta(x+1) \quad \forall x \in \mathbb{R} \;\Rightarrow\; \eta(z) = \eta(z+1) \quad \forall z \in \overline{B}_\sigma \tag{2.7}$$

by analytic continuation. We infer from this property that functions in X_σ are bounded. More precisely,

$$\|\eta\|_\sigma = \max_{z \in \overline{B}_\sigma} |\eta(z)|$$

makes X_σ a Banach space (using Weierstrass theorem).

We now state a version of the Paley-Wiener theorem adapted to our 1-periodic functions:

Lemma 2.3. *Let $\eta \in X_\sigma$. Then, for all $n \in \mathbb{Z}$,*

$$|\hat{\eta}(n)| \le e^{-2\pi\sigma|n|} \|\eta\|_\sigma$$

Proof. We introduce the closed and rectangular oriented contour γ_- defined as the concatenation of the segments $[0,1]$, $[1, 1-i\sigma]$, $[1-i\sigma, -i\sigma]$ and $[-i\sigma, 0]$. By Cauchy's theorem, we have for all $n \ge 0$:

$$\int_{\gamma_-} \eta(z) e^{-2i\pi n z}\, dz = 0.$$

By 1-periodicity of η, see (2.7), the integrals over the two vertical boundaries of γ_- cancel out. We are left with

$$\int_{[0,1]} \eta(z) e^{-2i\pi n z} dz = \int_{[-i\sigma, 1-i\sigma]} \eta(z) e^{-2i\pi n z} dz.$$

The left hand-side is exactly $\hat{\eta}(n)$, while the right-hand side has its modulus bounded by $\|\eta\|_\sigma e^{-2\pi n|\sigma|}$. This gives the desired inequality for all $n \ge 0$. The inequality for $n < 0$ is obtained using the contour γ_+, symmetric of γ_- with respect to the x-axis. $\square$

Remark 2.4. *Conversely, if η is a 1-periodic function satisfying $|\hat{\eta}(n)| \le C\, e^{-2\pi\sigma|n|}$ for all n, it can be shown that it is an element of $X_{\sigma'}$ for all $0 < \sigma' < \sigma$.*

The next lemma will be crucial to overcome the loss of derivatives.

Lemma 2.4. *Let $\eta \in X_\sigma$. For all $0 < \delta < \sigma$, $\eta' \in X_{\sigma-\delta}$, with the estimate*

$$\|\eta'\|_{\sigma-\delta} \le \frac{1}{\delta} \|\eta\|_\sigma \tag{2.8}$$

Proof. This is a consequence of the Cauchy formula. For $z \in \overline{B}_{\sigma-\delta}$, and $\gamma = C(z,\delta)$, the relation

$$\eta'(z) = \frac{1}{2i\pi} \int_\gamma \frac{\eta(\xi)}{(\xi - z)^2} d\xi$$

implies the formula. $\square$

2.3.3 Arnold's theorem

The main theorem of the chapter is

Theorem 2.3. *Let $\nu > 2$, $K > 0$ and α satisfying $(H_{K,\nu})$. Let f be an orientation-preserving homeomorphism with $\rho(f) = \alpha$. For all $\sigma > 0$, there exists $\varepsilon = \varepsilon(K, \nu, \sigma)$ such that if f has a lift F satisfying:*

$$F(x) = x + \alpha + \eta(x) \quad \text{with } \|\eta\|_\sigma \leq \varepsilon$$

then f is conjugated to R_α by an analytic orientation-preserving diffeomorphism.

The next section will be devoted to the detailed proof of the theorem. The general picture is as follows. The aim is to construct the analytic conjugacy h (more precisely its lift H) through an iterative process. One first sets $F_0(x) = F(x) = x + \alpha + \eta(x)$. Then, given $F_n(x) = x + \alpha + \eta_n(x)$, one builds

- The solution U_n of $U_n(x + \alpha) - U_n(x) = \eta_n(x) - \hat{\eta}_n(0)$.
- $H_n(x) = x + U_n(x)$
- $F_{n+1} = H_n^{-1} \circ F_n \circ H_n = (H_1 \circ \cdots \circ H_n)^{-1} \circ F \circ (H_1 \circ \cdots \circ H_n)$.

Roughly, the goal is then to show that $\mathcal{H}_n = H_1 \circ \cdots \circ H_n$ converges to an analytic diffeomorphism $\mathcal{H}$, while F_n converges to the rotation, meaning that the error η_n converges to 0. A keypoint in establishing this convergence is to show an estimate of the type

$$\|\eta_{n+1}\|_{\sigma_n - \delta_{n+1}} \leq \frac{C}{\delta_n^M} \|\eta_n\|_{\sigma_n}^2$$

which encodes both the loss of derivatives through the denominator δ_n^M and the quadratic convergence of the error. Roughly, the idea will be to take $\delta_n = \frac{\delta_0}{n^2}$, $\delta_0 \ll 1$, so that losing $\delta_0 + \delta_1 + ...$ in the regularity index through the iteration process will leave us with some regularity. Meanwhile, we expect as in the Newton's scheme a decay of the error like $\varepsilon_0^{2^n}$. In particular, such fast decay will win over the growth induced by $\frac{1}{\delta_n^M}$. Obviously, the exponent 2 (more generally the fact that the exponent is greater than 1) at the right hand-side is crucial: an estimate like $\|\eta_{n+1}\|_{\sigma_n - \delta_{n+1}} \leq \frac{C}{\delta_n^M} \|\eta_n\|_{\sigma_n}$ has no chance to yield a decay of η_n.

2.4 Proof of Arnold's theorem

The proof presented here follows the lines of paper [25], which is a nice introduction to KAM theory from an analytic viewpoint. We first gather all estimates that will be used at each step of the iteration alluded to above.

Proposition 2.5. *Let $\sigma > 0$, $\eta \in X_\sigma$, $\delta \in (0, \min(\frac{\sigma}{6}, 1))$. Let $F(x) = x + \alpha + \eta(x)$, with $\rho_0(F) = \alpha$.*

1. *The equation*
$$U(\cdot+\alpha)-U=\eta-\hat{\eta}(0)$$
has a unique solution with zero mean in $X_{\sigma-\delta}$, *satisfying*
$$\|U\|_{\sigma-\delta}\leq\frac{C_{\nu,K}}{\delta^{\nu}}\|\eta\|_{\sigma} \tag{2.9}$$

2. *There exists* $c=c_{\nu,K}$ *such that under the assumption* $\|\eta\|_{\sigma}\leq c\delta^{\nu+1}$, *the following statements hold:*
   - $H(z)=z+U(z)$ *is an analytic diffeomorphism from* $B_{\sigma-2\delta}$ *to its range. This range contains* $B_{\sigma-3\delta}$.
   - $V(z)=H^{-1}(z)-(z-U(z))$ *is in* $X_{\sigma-4\delta}$, *with estimate*
   $$\|V\|_{\sigma-4\delta}\leq\frac{C_{\nu,K}}{\delta^{2\nu+1}}\|\eta\|_{\sigma}^{2}. \tag{2.10}$$
   - $\tilde{F}=H^{-1}\circ F\circ H$ *can be written as* $\tilde{F}(x)=x+\alpha+\tilde{\eta}(x)$ *with* $\tilde{\eta}\in X_{\sigma-6\delta}$ *and*
   $$\|\tilde{\eta}\|_{\sigma-6\delta}\leq\frac{C_{\nu,K}}{\delta^{2\nu+1}}\|\eta\|_{\sigma}^{2}. \tag{2.11}$$

Proof.

1. Following the end of subsection 2.3.1, we know that if $U=\sum_{n\in\mathbb{Z}^*}\hat{U}(n)e^{2i\pi nx}$ is a solution of (2.4), then for all $n\in\mathbb{Z}^*$, $\hat{U}(n)=\frac{\hat{\eta}(n)}{e^{2i\pi n\alpha}-1}$. By the lower bound (2.6) on the denominator, and the upper bound on the numerator given by Lemma 2.3, we get, for all $z\in\overline{B}_{\sigma-\delta}$:
$$|\hat{U}(n)e^{2i\pi nz}|\leq\frac{|\hat{\eta}(n)|}{|e^{2i\pi n\alpha}-1|}e^{2\pi|n\mathcal{I}m\,z|}\leq\|\eta\|_{\sigma}\frac{e^{-2\pi|n|\ \delta}}{4K}|n|^{\nu-1}.$$
This shows that the Fourier series converges absolutely over $\overline{B}_{\sigma-\delta}$. It provides a solution to (2.4), with estimate
$$\|U\|_{\sigma-\delta}\leq\frac{\|\eta\|_{\sigma}}{4K}\sum_{n\in\mathbb{Z}^*}e^{-2\pi|n|\ \delta}|n|^{\nu-1}=\frac{\|\eta\|_{\sigma}}{4K\delta^{\nu-1}}\sum_{n\in\mathbb{Z}^*}e^{-2\pi|n|\ \delta}|\delta n|^{\nu-1}$$
One can check that there exists $M=M_{\nu}$ such that the function $g_{\nu}:x\to e^{-2\pi x}x^{\nu-1}$ is increasing for $0\leq x\leq M_{\nu}$ and decreasing for $x\geq M_{\nu}$. We can then split the sum at the right hand-side between indices n with $|n\delta|\leq M_{\nu}$ and those with $|n\delta|\geq M_{\nu}$. The first sum can be bounded through:
$$\sum_{n,\,|n\ \delta|\ \leq M_{\nu}}e^{-2\pi|n|\ \delta}|\delta n|^{\nu-1}\leq\sum_{n,\,|n\ \delta|\ \leq M_{\nu}}g_{\nu}(M_{\nu})\leq\frac{2M_{\nu}}{\delta}g_{\nu}(M_{\nu})$$
while the second sum can be compared to the integral $\int_{|x\delta|\geq M_{\nu}}e^{-2\pi x\delta}|\delta x|^{\nu-1}\,dx=\frac{C_{\nu}}{\delta}$. Estimate (2.9) follows.

2. To prove that $H(z) = z + U(z)$ is an analytic diffeomorphim from $B_{\sigma-2\delta}$ onto its range, it is enough to show that $|U'(z)| < 1$ for all $z \in \overline{B}_{\sigma-2\delta}$: indeed, it follows easily from this property that H is one-to-one, and a local analytic diffeomorphism (by the local inversion theorem in its analytic version). By (2.8),

$$\|U'\|_{\sigma-2\delta} \leq \frac{1}{\delta}\|U\|_{\sigma-\delta} \leq \frac{C_{\nu,K}}{\delta^{\nu+1}}\|\eta\|_\sigma$$

where the last inequality comes from (2.9). If $\|\eta\|_\sigma \leq \frac{\delta^{\nu+1}}{2C_{\nu,K}}$, then $\|U\|_{\sigma-\delta} \leq \frac{\delta}{2}$, and $\|U'\|_{\sigma-2\delta} \leq \frac{1}{2}$.

To show that $B_{\sigma-3\delta} \subset H(B_{\sigma-2\delta})$, it is enough to show that for all $y \in B_{\sigma-3\delta}$, the function $z \to y - U(z)$ has a fixed point in $B_{\sigma-2\delta}$. As $\|U\|_{\sigma-\delta} \leq \frac{\delta}{2}$, this mapping sends $\overline{B}_{\sigma-2\delta}$ to itself, and the inequality $\|U'\|_{\sigma-2\delta} < 1$ implies that it is a contraction over this strip. The result follows from Picard fixed point theorem.

Let now $V(z) = H^{-1}(z) - (z - U(z))$, which is well-defined on $B_{\sigma-3\delta}$ by the previous item. We write

$$z = H^{-1} \circ H(z) = H^{-1}(z + U(z)) = z + U(z) - U(z + U(z)) + V(z + U(z)),$$

so that

$$V(z + U(z)) = U(z + U(z)) - U(z) = \int_0^1 U'(z + sU(z))U(z)ds.$$

Setting $\xi = z + U(z) = H(z)$, we find

$$V(\xi) = \int_0^1 U'\big(H^{-1}(\xi) + sU(H^{-1}(\xi))\big)U(H^{-1}(\xi))ds.$$

Note that, in the same way that $B_{\sigma-3\delta} \subset H(B_{\sigma-2\delta})$, one has $B_{\sigma-4\delta} \subset H(B_{\sigma-3\delta})$, so that the previous formula for $V(\xi)$ holds for all $\xi \in B_{\sigma-4\delta}$. Moreover, $H^{-1}(\xi) \in B_{\sigma-3\delta}$, and $H^{-1}(\xi) + sU(H^{-1}(\xi)) \in B_{\sigma-2\delta}$. Eventually,

$$|V(\xi)| \leq \|U'\|_{\sigma-2\delta}\|U\|_{\sigma-3\delta} \leq \|U'\|_{\sigma-2\delta}\|U\|_{\sigma-\delta} \leq \frac{C'_{\nu,K}}{\delta^{2\nu+1}}\|\eta\|^2_\sigma.$$

The estimate (2.10) follows.

To conclude the proof, we compute

$$
\begin{aligned}
\tilde{F}(x) = H^{-1} \circ F \circ H(x) &= H^{-1}\Big(x + U(x) + \alpha + \eta(x + U(x))\Big) \\
&= x + U(x) + \alpha + \eta(x + U(x)) - U\Big(x + U(x) + \alpha + \eta(x + U(x))\Big) \\
&\quad + V\Big(x + U(x) + \alpha + \eta(x + U(x))\Big) \\
&= x + \alpha + \eta(x) + \Big(U(x) - U(x + \alpha)\Big) \\
&\quad + \Big(\eta(x + U(x)) - \eta(x)\Big) + \Big(U(x + \alpha) - U\Big(x + \alpha + U(x) + \eta(x + U(x))\Big)\Big) \\
&\quad + V\Big(x + U(x) + \alpha + \eta(x + U(x))\Big).
\end{aligned}
$$

Hence, we can write $\tilde{F}(x) = x + \alpha + \tilde{\eta}(x)$, where

$$
\begin{aligned}
\tilde{\eta}(z) &= \hat{\eta}(0) + \Big(\eta(z + U(z)) - \eta(z)\Big) + \Big(U(z + \alpha) - U\Big(z + \alpha + U(z) + \eta(z + U(z))\Big)\Big) \\
&\quad + V\Big(z + U(z) + \alpha + \eta(z + U(z))\Big) \\
&= \hat{\eta}(0) + I(z) + II(z) + III(z).
\end{aligned}
$$

It remains to show that the right-hand side is well-defined for $z \in \overline{B}_{\sigma-6\delta}$, holomorphic in the interior and continuous on the closure, with sup norm bounded by $\frac{C_{\nu,K}}{\delta^{2\nu+1}}\|\eta\|_\sigma^2$.

- We have seen that under the condition $\|\eta\|_\sigma \leq \frac{\delta^{\nu+1}}{2C_{\nu,K}}$, one has $\|U\|_{\sigma-\delta} \leq \frac{\delta}{2}$. It follows easily that I is defined over $\overline{B}_{\sigma-\delta}$.Moreover,

$$
I(z) = \int_0^1 \eta'(z + sU(z))\, U(z)\, ds
$$

  so that

$$
\|I\|_{\sigma-2\delta} \leq \|\eta'\|_{\sigma-\delta}\|U\|_{\sigma-2\delta} \leq \frac{C_{\nu,K}}{\delta^{\nu+1}}\|\eta\|_\sigma^2
$$

  by (2.8) and (2.9).

- Using that $\|U\|_{\sigma-\delta} \leq \frac{\delta}{2}$ and taking $\|\eta\|_\sigma \leq \min\left(\frac{1}{2C_{\nu,K}}, 1\right)\delta^{\nu+1} \leq \delta$, we see easily that II is defined over $\overline{B}_{\sigma-3\delta}$. Moreover,

$$
II(z) = \int_0^1 U'\Big(z + \alpha + s\big(U(z) + \eta(z + U(z))\big)\Big)\,\big(U(z) + \eta(z + U(z))\big)\, ds
$$

  so that

$$
\begin{aligned}
\|II\|_{\sigma-4\delta} &\leq \|U'\|_{\sigma-2\delta}\left(\|U\|_{\sigma-4\delta} + \|\eta\|_{\sigma-3\delta}\right) \\
&\leq \frac{C_{\nu,K}}{\delta^{\nu+1}}\left(\frac{C_{\nu,K}}{\delta^{\nu}} + 1\right)\|\eta\|_\sigma^2 \leq \frac{C'_{\nu,K}}{\delta^{2\nu+1}}\|\eta\|_\sigma^2.
\end{aligned}
$$

- With similar arguments as before, for $z \in \overline{B}_{\sigma-6\delta}$, $z+\alpha+U(z)+\eta(z+U(z)) \in \overline{B}_{\sigma-4\delta}$, so that III is well-defined in $\overline{B}_{\sigma-6\delta}$, and satisfies

$$\|III\|_{\sigma-6\delta} \leq \|V\|_{\sigma-4\delta} \leq \frac{C_{\nu,K}}{\delta^{2\nu+1}}\|\eta\|_\sigma^2$$

by (2.10).

- It remains to evaluate $\hat{\eta}(0)$. To that purpose, we notice that $\rho_0(\tilde{F}) = \rho_0(F) = \alpha$ (see the proof of the last point in Proposition 2.3). By Remark 2.3, there exists x_0 such that $\tilde{\eta}(x_0) = 0$. It follows that

$$\hat{\eta}(0) = -I(x_0) - II(x_0) - III(x_0)$$

and by the previous estimates, $|\hat{\eta}(0)| \leq \frac{C_{\nu,K}}{\delta^{2\nu+1}}\|\eta\|_\sigma^2$.

Putting together all estimates yields (2.11).

□

Thanks to Proposition 2.5, we are now in order to prove Arnold's theorem. As sketched in Paragraph 2.3.3, conjugacy to the rotation is obtained as the limit of an induction process, starting from $F_0 = F$, $\eta_0 = \eta$: $F_0(x) = x+\alpha+\eta_0(x)$. The induction relation is the following: given $F_n(x) = x+\alpha+\eta_n(x)$, we solve

$$U_n(\cdot+\alpha) - U_n = \eta_n - \hat{\eta}_n(0)$$

and then set

$$H_n(x) = x+U_n(x), \quad F_{n+1}(x) = H_n^{-1} \circ F_n \circ H_n(x) = x+\alpha+\eta_{n+1}(x).$$

More precisely, let $\delta_n = \frac{\min(\sigma,1)}{2\pi^2 n^2}$, $n \geq 1$, so that in particular $\sum_{n\geq 1}\delta_n \leq \frac{\min(\sigma,1)}{12}$. Let $\varepsilon_n = \varepsilon_0^{(\frac{3}{2})^n}$, $n \geq 0$, $\varepsilon_0 > 0$. We shall show inductively that for a good choice of ε_0 (depending on ν, K) and a smallness assumption on $\|\eta\|_\sigma$, the following property holds for all $n \geq 0$:

$(\mathcal{P}_n)$ η_n defines an element of X_{σ_n} where $\sigma_n = \sigma - 6\sum_{k=1}^{n}\delta_k$, $\quad \|\eta_n\|_{\sigma_n} \leq \varepsilon_n$, and $\|\eta_n\|_{\sigma_n} \leq c\,\delta_{n+1}^{\nu+1}$ (see Proposition 2.5 for the definition of c).

First, $\mathcal{P}_0$ is clearly satisfied under the smallness assumption $\|\eta\|_\sigma \leq \varepsilon_0$ for any choice of ε_0. Now, assume $\mathcal{P}_n$. To establish $\mathcal{P}_{n+1}$, we first notice that our choice of δ_n ensures that $\delta_{n+1} \leq \min\left(\frac{\sigma_n}{6}, 1\right)$. This, together with inequality $\|\eta_n\|_{\sigma_n} \leq c\,\delta_{n+1}^{\nu+1}$, allows to apply Proposition 2.5 with $\sigma = \sigma_n$ and $\delta = \delta_{n+1}$: it shows that the function η_{n+1} introduced above is well-defined as an element of $X_{\sigma_n - 6\delta_{n+1}} = X_{\sigma_{n+1}}$, and that

$$\|\eta_{n+1}\|_{\sigma_{n+1}} \leq \frac{C_{\nu,K}}{\delta_{n+1}^{2\nu+1}}\|\eta_n\|_{\sigma_n}^2. \tag{2.12}$$

As $\|\eta_n\|_{\sigma_n} \le \varepsilon_n$, this last inequality yields:

$$\|\eta_{n+1}\|_{\sigma_{n+1}} \le \frac{C_{\nu,K}}{\delta_{n+1}^{2\nu+1}}(\varepsilon_n)^2.$$

In particular, $\|\eta_{n+1}\|_{\sigma_{n+1}} \le \varepsilon_{n+1}$ holds whenever $\frac{C_{\nu,K}}{\delta_{n+1}^{2\nu+1}}(\varepsilon_n)^2 \le \varepsilon_{n+1}$, which can be written in the form

$$C'_{\nu,K}(n+1)^{4\nu+2}\varepsilon_0^{\frac{1}{2}(\frac{3}{2})^n} \le 1$$

and is guaranteed for ε_0 small enough (depending only on ν, K). Moreover, still by $(\mathcal{P}_n)$ and (2.12), we have

$$\|\eta_{n+1}\|_{\sigma_{n+1}} \le \frac{C_{\nu,K}}{\delta_{n+1}^{2\nu+1}}\left(c\delta_{n+1}^{\nu+1}\right)\varepsilon_n$$

so that we obtain $\|\eta_{n+1}\|_{\sigma_{n+1}} \le c\,\delta_{n+2}^{\nu+1}$ whenever $\frac{C_{\nu,K}}{\delta_{n+1}^{\nu}}\varepsilon_n \le \delta_{n+2}^{\nu+1}$, which can be again guaranteed for ε_0 small enough (depending on ν, K). This proves $\mathcal{P}_{n+1}$.

Conclusion of the proof of Arnold's theorem.

From Proposition 2.5 and the properties of η_n, one can check that H_n is analytic on $B_{\sigma_n-2\delta_{n+1}}$ with range included in $B_{\sigma_n-\delta_{n+1}} \subset B_{\sigma_{n-1}-2\delta_n}$. Hence, if we set $\mathcal{H}_n = H_1 \circ \dots \circ H_n$,then $\mathcal{H}_n$ is well-defined and analytic over $B_{\sigma_n-2\delta_{n+1}}$, and by construction: $F_{n+1} = \mathcal{H}_n^{-1} \circ F \circ \mathcal{H}_n$. We compute

$$\begin{aligned}\mathcal{H}_n(z) &= H_1 \circ \dots \circ H_{n-1}(z + U_n(z))\\ &= H_1 \circ \dots \circ H_{n-2}(z + U_n(z) + U_{n-1}(z + U_n(z)))\\ &= z + U_n(z) + U_{n-1}(z + U_n(z)) + U_{n-2}(z + U_n(z) + U_{n-1}(z + U_n(z))) + \dots\end{aligned}$$

We deduce from the previous writing that

$$\begin{aligned}\|\mathcal{H}_n - z\|_{\sigma_n-\delta_{n+1}} &\le \sum_{k=0}^{n} \|U_k\|_{\sigma_k-2\delta_{k+1}}\\ &\le C_{\nu,K}\sum_{k=0}^{n}\frac{\|\eta_k\|_{\sigma_k}}{\delta_{k+1}^{\nu}} \le C_{\nu,K}\sum_{k=0}^{n}\frac{\varepsilon_k}{\delta_{k+1}^{\nu}} \le C'_{\nu,K}\varepsilon_0.\end{aligned}$$

Moreover, for all n,

$$\mathcal{H}_{n+1}(z) - \mathcal{H}_n(z) = \mathcal{H}_n \circ H_{n+1}(z) - \mathcal{H}_n(z) = \int_0^1 \mathcal{H}'_n(z + s\,U_{n+1}(z))\,U_{n+1}(z)ds$$

so that

$$\begin{aligned}\|\mathcal{H}_{n+1} - \mathcal{H}_n\|_{\sigma_{n+1}-2\delta_{n+2}} &\le \|\mathcal{H}'_n\|_{\sigma_{n+1}-\delta_{n+2}}\|U_{n+1}\|_{\sigma_{n+1}-2\delta_{n+2}}\\ &\le \|\mathcal{H}'_n\|_{\sigma_n-3\delta_{n+1}}\,\|U_{n+1}\|_{\sigma_{n+1}-\delta_{n+2}}\\ &\le (\|(\mathcal{H}_n - z)'\|_{\sigma_n-3\delta_{n+1}} + 1)\,\|U_{n+1}\|_{\sigma_{n+1}-\delta_{n+2}} \le \left(\frac{C'_{\nu,K}\varepsilon_0}{\delta_{n+1}} + 1\right)\frac{C_{\nu,K}}{\delta_{n+2}^{\nu}}\varepsilon_n.\end{aligned}$$

This implies that $(\mathcal{H}_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in X_{σ_*}, with $\sigma_* = \lim_n \sigma_n$, therefore converging to an element $\mathcal{H}$ in X_{σ_*}. Moreover, the estimate on $\mathcal{H}_n - z$ shows that

$$\|\mathcal{H} - z\|_{\sigma_*} \leq C'_{\nu,K}\varepsilon_0$$

By (2.8), it follows that

$$\|\mathcal{H}' - 1\|_{\sigma_*-\delta_*} \leq \frac{C'_{\nu,K}}{\delta_*}\varepsilon_0,$$

where (for instance) $\delta_* = \frac{\sigma_*}{4}$. Reasoning as in Proposition (2.5) (to establish the properties of H), we can deduce from the two previous inequalities that for ε_0 small enough (depending on ν, K, and also on σ_* therefore on σ), $\mathcal{H}$ is invertible from $B_{\sigma_*-\delta_*}$ onto its range, and that this range contains $B_{\sigma_*-2\delta_*}$. Eventually:

$$F \circ \mathcal{H}(x) = \lim_n F \circ \mathcal{H}_n(x) = \lim_n \mathcal{H}_n \circ F_n(x) = \lim_n \mathcal{H}_n(x + \alpha + \eta_n(x)) = \mathcal{H}(x + \alpha)$$

for all $x \in \mathbb{R}$. This concludes the proof.

Chapter 3

The Nash-Moser theorem

The previous chapter was dedicated to the phenomenon of loss of derivatives in an analytic framework. The main point in solving the problem was to establish an error estimate of the form
$$\|\eta_{n+1}\|_{\sigma-\delta} \leq \frac{C}{\delta^M}\|\eta_n\|_\sigma^2,$$
for arbitrary small δ. The fact that δ could be arbitrarily close to zero was used by taking $\delta = \delta_n$ with $\lim_{n\to+\infty}\delta_n = 0$. In this chapter, we are rather interested to spaces of functions with finite regularity, where the error estimate takes the form

$$\|\eta_{n+1}\|_{s-r} \leq C_{s,r}\|\eta_n\|_s^2, \quad \text{for some fixed } r. \tag{3.1}$$

As one can not play on the index loss r anymore, another strategy is needed to handle the loss of derivatives. We will describe the famous approach introduced by Nash [18] and further developped by Moser [14, 15, 16]. The crude idea behind this approach is to modify the Newton's scheme by adding some *frequency truncation operators* in the iteration formula. The goal of these truncations is to control at iteration n the support of the Fourier transform of η_n. Note that we will consider only 1-periodic functions: the Fourier transform $\hat{f}(n)$, $n \in \mathbb{Z}$, will refer to the Fourier coefficient n of the 1-periodic function f.

To explain a little more the idea, let us assume that, for all n, the error η_n is a 1-periodic function with Fourier support included in a ball of radius N_n:
$$\eta_n(x) = \sum_{|k|\ \leq N_n} \hat{\eta}_n(k)e^{2i\pi k\cdot x}, \quad x \in \mathbb{R}^d.$$

Introducing the Sobolev norm
$$\|\eta\|_s = \left(\sum_{k\in\mathbb{Z}}\langle k\rangle^{2s}|\hat{\eta}(k)|^2\right)^{1/2} < +\infty \qquad (s \in \mathbb{R},\ \langle x\rangle = \sqrt{1+|x|^2})$$
one has easily: $\|\eta_n\|_s \leq \langle N_n\rangle^r\|\eta_n\|_{s-r}$, so that an estimate like (3.1) turns into
$$\|\eta_{n+1}\|_{s-r} \leq C_{s,r}\langle N_n\rangle^{2r}\|\eta_n\|_{s-r}^2.$$

Hence, if the scheme under consideration (itself related to the equation that we try to solve) ensures that the radius N_n grows subquadratically, typically if $N_{n+1} \lesssim N_n^\mu$ for some $\mu < 2$, then the error will converge to zero for η_0 small enough.

3.1 Statement of the Nash-Moser theorem

The Nash-Moser theorem reflects a methodology introduced by Nash and Moser for the resolution of perturbative problems: the goal is to solve an equation of the form $F(u) = f$, with $F(0) = 0$ and f close to zero.

Notation. In what follows, for any two families of norms $\| \|_{s_i}$ and $\| \|_{t_k}$, we shall use the notation

$$\|\mathcal{F}_1(u_1, \dots u_m)\|_{s_1} + \dots + \|\mathcal{F}_k(u_1, \dots u_m)\|_{s_k} \lesssim \|\mathcal{G}_1(v_1, \dots v_n)\|_{t_1} + \dots + \|\mathcal{G}_l(v_1, \dots v_n)\|_{t_l}$$

whenever

$$\|\mathcal{F}_1(u_1, \dots u_m)\|_{s_1} + \dots + \|\mathcal{F}_k(u_1, \dots u_m)\|_{s_k} \le C\Big(\|\mathcal{G}_1(v_1, \dots v_n)\|_{t_1} + \dots + \|\mathcal{G}_l(v_1, \dots v_n)\|_{t_l}\Big)$$

for a constant C that depends only on $s_1, \dots, s_k, t_1, \dots, t_l$.

Functional framework.
We consider a family of Banach spaces $(X_s, \| \|_s)_{s\ge 0}$ with the following properties:

i) $(X_s)_{s\ge 0}$ is decreasing, with $\| \cdot \|_s \le \| \cdot \|_{s'}$ for $s \le s'$.

ii) There is $\delta > 0$, and a sequence $(T_N)_{N\ge 1}$ of linear operators from X_0 to $X_\infty = \cap_{s\ge 0} X_s$ such that:

$$\|T_N u\|_{s'} \lesssim N^{s'-s+\delta} \|u\|_s \quad \forall s' \ge s, \forall u \in X_s,$$
$$\|(Id - T_N)u\|_{s'} \lesssim N^{s'-s+\delta} \|u\|_s \quad \forall s' < s, \forall u \in X_s,$$

Besides $(X_s, || ||_s)$, we consider another family of decreasing Banach spaces $(Y_s, | |_s)_{s\ge 0}$, with $| \cdot |_s \le | \cdot |_{s'}$ for $s \le s'$.

Remark 3.1. *One can consider families (X_s) and (Y_s) either indexed by $s \in \mathbb{R}_+$, or by s in a subset of $\mathbb{R}_+$, typically $\mathbb{N}$.*
Examples of families satisfying i) and ii) are $(H^s(\mathbb{T}^d))_{s\ge 0}$, see the exercise below, or $(H^s(\mathbb{R}^d))_{s\ge 0}$: see chapter 5 for more. In those examples, one can take $\delta = 0$ in i) and ii). The family $(C^s(\mathbb{T}^d))_{s\in\mathbb{N}}$ is also adapted, with for instance $\delta = d$ (exercise).

Assumptions on F.
Let $k \ge 2r \ge 0$, and $F : B_k \to Y_{k-r}$ of class C^1, where B_k is an open ball centered at 0 in X_k. We make the folllowing assumptions

1. ($C^{1,1}$ condition)

   For all $u, v \in X_k$ such that $u, u+v \in B_k$,

$$| \ F(u+v) - F(u) - F'(u)v|_{k-r} \lesssim \|v\|_k^2.$$

2. (C^1 condition)

   For all $u, v \in X_k$ such that $u \in B_k$,

$$| \ F'(u)v|_{k-r} \lesssim \|v\|_k$$

3. (C^0 condition)

   For all $s \geq k$, for all $u \in X_s \cap B_k$,

$$| \ F(u)|_{s-r} \lesssim \|u\|_s$$

4. (right invertibility condition)

   For all $u \in X_\infty \cap B_k$, there exists $L(u)$ such that: for all $s \geq k-r$,

$$L(u) : Y_s \to X_{s-r}, \quad F'(u)L(u)|_{Y_{s+2r}} = Id|_{Y_{s+2r}}$$

   and such that for all $g \in Y_s$,

$$\|L(u)g\|_{s-r} \lesssim | \ g|_s \ + \ | \ g|_{k-r}\|u\|_s$$

Remark 3.2. *Note that conditions 1. and 2. must hold only for the index k, while conditions 3. and 4. hold for all $s \geq k$.*

As will be clear from the proof, the keypoint in conditions 3. and 4. is the fact that the inequalities are subquadratic (and even linear) in the high norm: for instance, there is no term like $\|u\|_s^2$ in 3. or no term like $\|u\|_s| \ g|_s$ in 4. Such estimates, linear in the high norm, are called tame estimates. *We shall come back to this notion in the next chapters.*

Theorem 3.1. *Under the assumptions 1. to 4. above, there exists $k' \geq k-r$, such that for all $f \in Y_{k'}$ with $| \ f|_{k-r}$ small enough (the smallness condition depending on $|f|_{k'}$), there exists a solution u in B_k of $F(u) = f$.*

By *depending on $|f|_{k'}$*, we mean that for all $M > 0$, one can find $\varepsilon = \varepsilon(M) > 0$ such that: for all $f \in Y_{k'}$ with $| \ f|_{k'} \leq M$ and $| \ f|_{k-r} \leq \varepsilon$, there exists a solution.

Remark 3.3. *This statement of the Nash-Moser theorem and its proof follow the article [14] by J. Moser, see also [24]. A refined statement (due to L. Hörmander) can be found in [1]. See also [10].*

Exercice 3.1. (Sobolev spaces of periodic functions)
Let $s \in \mathbb{R}_+$. We define

$$H^s(\mathbb{T}^d) = \{u \in L^2_{loc}(\mathbb{R}^d),\ 1\text{-periodic, such that } \sum_{n\in\mathbb{Z}^d} \langle n\rangle^{2s} |\hat{u}(n)|^2 < +\infty\}$$

where $\langle x\rangle = \sqrt{1+|x|^2}$, *and* $\hat{u}(n) = \int_{[0,1]^d} u(t) e^{-2i\pi n\cdot t} dt$, $\forall n \in \mathbb{Z}^d$.

1. *Show that $H^s(\mathbb{T}^d)$ is a Banach space with the norm* $\|u\|_s = \left(\sum_{n\in\mathbb{Z}^d} \langle n\rangle^{2s} |\hat{u}(n)|^2\right)^{1/2}$.
2. *Show that for $s \in \mathbb{N}$, $H^s(\mathbb{T}^d)$ is the completion of $C^\infty(\mathbb{T}^d)$ for the norm*

$$\|u\|_{H^s} = \left(\sum_{\alpha\in\mathbb{N}^d, |\alpha|\le s} \|\partial^\alpha u\|^2_{L^2}\right).$$

   Show that $\|\cdot\|_s$ and $\|\cdot\|_{H^s}$ are equivalent.

3. *Show that the operators T_N, $N \ge 1$ defined by*

$$T_N f = \sum_{n\in\mathbb{Z}^d, |n|\le N} \hat{u}(n) e^{2i\pi n\cdot x}$$

   satisfy the properties ii) mentioned in the functional framework above, with $\delta = 0$.

4. *Show that for all $s \in \mathbb{N}$ and any $\varepsilon > 0$, $H^{s+\frac{d}{2}+\varepsilon}$ is continuously embedded in $C^s(\mathbb{T}^d)$.*

3.2 Proof of the Nash-Moser theorem

To lighten a bit the notations, we shall only consider the case where the exponent δ in the bounds satisfied by $(T_N)_{N\ge1}$ is zero. We start by collecting a few lemmas:

Lemma 3.1. *Let $\mu \ge 6r$, and $\Lambda > 15r + \frac{9}{2}\mu + 3$. Let $c > 0$, $n_* \in \mathbb{N}\cup\{+\infty\}$, $N_0 > 0$. Let $(u_n)_{n\in\mathbb{N}}$ be a sequence in X_∞ such that for all $n \in [|1, n_*|]$:*

$$\|u_{n+1} - u_n\|_k \le c\left(N_n^{3r}\|u_n - u_{n-1}\|_k^2 + N_n^{5r-\Lambda} N_{n-1}^{\Lambda}\right), \quad N_n = N_0^{(\frac{3}{2})^n}. \tag{3.2}$$

If $N_0 \ge 4c^2$ and $N_1^\mu \|u_1 - u_0\|_k \le \frac{1}{2c}$, then

$$N_{n+1}^\mu \|u_{n+1} - u_n\|_k \le \frac{1}{2c}, \quad \forall n \in [|0, n_*|].$$

Proof. We set $\delta_n = N_n^\mu \|u_n - u_{n-1}\|_k$ for $n \in [|1, n_*|]$. Inequality (3.2) can be written as

$$\delta_{n+1} \le c\left(N_n^{3r-\frac{\mu}{2}}\delta_n^2 + N_{n-1}^{\frac{15}{2}r+\frac{9}{4}\mu-\frac{\Lambda}{2}}\right)$$

Note that by our choice of μ, $3r - \frac{\mu}{2} \leq 0$, and by our choice of Λ, $\frac{15}{2}r + \frac{9}{4}\mu - \frac{\Lambda}{2} < -\frac{3}{2}$. Thus, we obtain

$$\delta_{n+1} \leq c\left(\delta_n^2 + \frac{1}{N_n}\right).$$

Under the assumptions $N_0 \geq 4c^2$ and $\delta_1 < \frac{1}{2c}$, an easy induction shows that $\delta_{n+1} < \frac{1}{2c}$ for all $n \in [1, n_*]$, which concludes the proof.

□

Lemma 3.2. *Let $\Lambda \in \mathbb{N}_*$, $M > 0$, and $f \in Y_{k-r+\Lambda}$, $|f|_{k-r+\Lambda} \leq M$. Then, there exists a positive constant a (depending on Λ and M) such that: for all $N \geq 1$, for all $u \in X_\infty$:*

$$\forall \lambda \in [|0,\Lambda|],\ \|u\|_{k+\lambda} \leq a^{-1}N^\lambda \quad \Rightarrow \quad \forall \lambda \in [|0,\Lambda|],\ \|L(u)(F(u)-f)\|_{k-2r+\lambda} \leq aN^\lambda$$

Proof. Let u as in the statement of the Lemma.Taking $\lambda = 0$ and $a > 1$, we ensure that $u \in B_k$, so that $L(u)$ is well-defined. By the properties 3. and 4. satisfied by F, we get

$$\begin{aligned}\|L(u)(F(u)-f)\|_{k-2r+\lambda} &\lesssim |\ F(u)-f|_{k-r+\lambda} + |\ F(u)-f|_{k-r}\|u\|_{k-r+\lambda} \\ &\lesssim \|u\|_{k+\lambda} + |f|_{k-r+\lambda} + (\|u\|_k + |f|_{k-r})\|u\|_{k+\lambda}.\end{aligned}$$

By the assumption on u, we find that

$$\begin{aligned}\|L(u)(F(u)-f)\|_{k-2r+\lambda} &\leq C_\lambda\left(a^{-1}N^\lambda + |f|_{k-r+\lambda} + a^{-1}\left(a^{-1} + |f|_{k-r}\right)N^\lambda\right) \\ &\leq aN^\lambda\end{aligned}$$

taking a large enough so that:

$$\left(\max_{\lambda \in [|0,\Lambda|]} C_\lambda\right)\left(a^{-1} + M + a^{-1}\left(a^{-1} + M\right)\right) \leq a.$$

□

We now turn to the core of the proof of the Nash-Moser theorem. As mentioned earlier, it consists in a regularized Newton's scheme, involving truncation operators. More precisely, we introduce $N_n = N_0^{(\frac{3}{2})^n}$, N_0 to be specified later, and then consider the following scheme:

$$u_0 = 0, \quad u_{n+1} = u_n - T_{N_{n+1}}L(u_n)\left(f - F(u_n)\right) \tag{3.3}$$

Note that due to the application of the truncation operator, u_n will be in the range of T_{N_n} and therefore in X_∞. The scheme will be well-defined as long as $u_n \in B_k$. We will show that under suitable assumptions on f, and a suitable choice of N_0, it will be globally well-defined and will converge to a solution u of $F(u) = f$.

One main step is to show that the sequence (u_n) satisfies an inequality of the type (3.2). We fix $\Lambda \in \mathbb{N}^*$ as in Lemma 3.1 and then a as in Lemma 3.2. We introduce n^* the biggest index (possibly infinite) such that: for all $n \in [|0, n_*|]$

$$\forall \lambda \in [0,\Lambda],\ \|u_n\|_{k+\lambda} \leq \frac{1}{2}a^{-1}N_n^\lambda.$$

From (3.3), we deduce: for all $n \in [|0, n_*|]$,

$$\begin{aligned} \|u_{n+1} - u_n\|_k = \|T_{N_{n+1}} L(u_n)\,(f - F(u_n))\,\|_k &\lesssim N_{n+1}^{2r} \|L(u_n)\,(f - F(u_n))\,\|_{k-2r} \\ &\lesssim N_{n+1}^{2r} |\ f - F(u_n)|_{k-r} \end{aligned} \tag{3.4}$$

where the last inequality comes from assumption 4. To evaluate $|\ f - F(u_n)|_{k-r}$, we write, for $n \in [|1, n_*|]$

$$\begin{aligned} |\ F(u_n) - f|_{k-r} \leq &|\ F(u_n) - F(u_{n-1}) - F'(u_{n-1})(u_n - u_{n-1})|_{k-r} \\ &+ \ |F(u_{n-1}) + F'(u_{n-1})(u_n - u_{n-1}) - f|_{k-r} \\ &\lesssim \|u_n - u_{n-1}\|_k^2 \ + \ \ |\ F'(u_{n-1})(Id - T_{N_n})L(u_{n-1})(f - F(u_{n-1}))|_{k-r}. \end{aligned}$$

where the first term at the right hand-side of the last inequality comes from assumption 1., while the second term comes from formula (3.3). By assumption 3. we get:

$$|\ F(u_n) - f|_{k-r} \lesssim \|u_n - u_{n-1}\|_k^2 \ + \ \ \|(Id - T_{N_n})L(u_{n-1})(f - F(u_{n-1}))\|_k.$$

To handle the second term, we proceed as follows: taking advantage of the properties of the truncation operator, *cf* property ii), we write

$$\forall \lambda, \quad \|(Id - T_{N_n})L(u_{n-1})(f - F(u_{n-1}))\|_k \lesssim N_n^{2r-\lambda} \|L(u_{n-1})(f - F(u_{n-1}))\|_{k-2r+\lambda}$$

for $\lambda \geq 2r$. By the definition of n_* and Lemma 3.2 we get:

$$\forall \lambda \in [2r, \Lambda], \quad \|(Id - T_{N_n})L(u_{n-1})(f - F(u_{n-1}))\|_k \ \leq \ C_\lambda\, a\, N_n^{2r-\lambda}\, N_{n-1}^{\lambda}.$$

Hence, for all $n \in [|1, n_*|]$, for $\lambda = \Lambda$,

$$\begin{aligned} \|u_{n+1} - u_n\|_k &\leq c\left(N_{n+1}^{2r} \|u_n - u_{n-1}\|_k^2 + N_n^{5r-\Lambda}\, N_{n-1}^{\Lambda}\right) \\ &\leq c\left(N_n^{3r} \|u_n - u_{n-1}\|_k^2 + N_n^{5r-\Lambda}\, N_{n-1}^{\Lambda}\right) \end{aligned}$$

where the constant c depends on Λ and a.

We can apply Lemma 3.1, which yields: if $N_0 \geq 4c^2$, and $N_1^{\mu} \|u_1 - u_0\|_k = N_1^{\mu} \|u_1\|_k \leq \frac{1}{2c}$, then for all $n \in [|0, n^*|]$,

$$\|u_{n+1} - u_n\|_k \leq \frac{1}{2c} N_{n+1}^{-\mu} \tag{3.5}$$

Note that by inequality (3.4) with $n = 0$,

$$\|u_1\|_k \leq C N_1^{2r} |\ f|_{k-r}$$

so that for any N_0, the smallness condition will be satisfied by taking $|\,f|_{k-r}$ is small enough. To conclude we notice that : for all finite $n \in [|0, n_*|]$

$$\begin{aligned}
\|u_{n+1}\|_{k+\lambda} &\leq \sum_{j=0}^{n} \|u_{j+1} - u_j\|_{k+\lambda} \\
&\leq C_\lambda \sum_{j=0}^{n} N_{j+1}^{\lambda} \|u_{j+1} - u_j\|_k \\
&\leq \frac{C_\lambda}{2c} N_{n+1}^{\lambda} \sum_{j=0}^{n} N_{j+1}^{-\mu} \leq a N_{n+1}^{\lambda}
\end{aligned}$$

if we take N_0 large enough so that $\max_\lambda \frac{C_\lambda}{2c} \sum_{j=1}^{n} N_{j+1}^{-\mu} < a$.

This shows that $n_* = +\infty$, and that the estimate (3.5) holds for all n. From there, deducing the convergence of u_n to a solution of $F(u) = f$ is easy, and left to the reader.

Chapter 4

The isometric embedding of riemannian manifolds

We assume here that the reader has a basic knowledge of riemannian geometry (essentially that he knows the classical definitions). If not, he can jump to section 4.2, where the nonlinear PDE (4.1) below is solved thanks to the Nash-Moser theorem, without any reference to geometrical objects.

The general problem is the following: given a smooth manifold M of dimension d, equipped with a riemannian metrics of class C^k, can we find some integer n and a C^k isometric embedding $u : M \to \mathbb{R}^n$?

We remind a few definitions to clarify this statement:

Definition 4.1. (immersion, embedding)
A C^1 mapping $u : M \to R^n$ is an immersion if for any x in M, the differential $u'(x) : T_xM \to \mathbb{R}^n$ is injective. We say that u is an embedding if it is an immersion and if it is an homeomorphism onto its image.

The fact that u is an immersion ensures that u is a local embedding, but not necessarily global. In other words, an embedding excludes self-intersection, while an immersion allows for it. Note that in the case where M is a compact manifold, an injective immersion is an embedding.

Definition 4.2. *For (M, g) and (N, h) two riemannian manifolds, a C^1 mapping $u : M \to N$ is an isometry if $u^*h = g$ (the pullback of h by u is g). In local coordinates, this reads*

$$h\left(\frac{\partial u}{\partial x_i}, \frac{\partial u}{\partial x_j}\right) = g\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right) (= g_{ij}), \quad \forall i \leq j = 1...d$$

Note that when $(N, h) = (\mathbb{R}^n, e)$ with e the usual euclidean scalar product, the previous equation reads

$$\frac{\partial u}{\partial x_i} \cdot \frac{\partial u}{\partial x_j} = g_{ij}, \quad 1 \leq i \leq j \leq d. \tag{4.1}$$

The isometric embedding problem enters a more general framework, that is the correspondence between abstract manifolds and submanifolds of $\mathbb{R}^n$. In this regards, we remind the famous theorem due to Whitney:

Theorem 4.1. *A smooth manifold of dimension d can be smoothly embedded in $\mathbb{R}^{2d}$.*

The exponent $2d$ in the theorem of Whitney can be compared to the dimension suggested by the isometry constraints (4.1). Namely, (4.1) is a system of $s_d = \frac{d(d+1)}{2}$ equations in n unknowns (the components of u), so that it is very natural that $n \geq s_d$.

The isometric embedding problem was formulated by Schläfli in 1871, and has a long history. We refer for instance to the survey [27], or [23]. In these lecture notes, we want to give some insight into the breakthrough result of Nash (1956):

Theorem 4.2. *For $k \geq 3$, and $n \geq 3s_d + 4d$, any smooth compact manifold of dimension d with C^k metrics can be C^k isometrically embedded in $\mathbb{R}^n$.*

The exponent was later improved by Gromov and Rohlin, down to $n \geq s_d + d + \max(d, 5)$, and the case of non-compact manifold was treated as well.

In the rest of the chapter, we will establish a weaker version, not caring about the lower bounds on k and n. *Namely, for a given k, we will show the existence of n and $k' \geq k$ such that for any g in $C^{k'}$, there exists a C^k isometric embedding u from M to $\mathbb{R}^n$.* The first section is dedicated to preliminary steps, that will allow to go from the global embedding problem to the resolution of (4.1) over a torus, in a perturbative regime. In the second section, we will solve this PDE through the use of the Nash-Moser theorem.

4.1 Preliminaries

In all what follows, $k \geq 1$ is fixed, M is smooth and compact, and g is $C^{k'}$ on M, $k' \geq k$.

Step 1: Reduction to the case $M = \mathbb{T}^d$.
We show here that we can always restrict to the case where M is a torus, with a $C^{k'}$ riemannian metrics g. This will allow us to benefit from global coordinates on M. In particular, the equation (4.1) will be set over the whole of M. To restrict to this case, we first invoke Whitney's theorem: by replacing M by $u_W(M)$ and g by its pushforward $(u_W)_* g$, where $u_W : M \to \mathbb{R}^{2d}$ is the Whitney's embedding, we can always consider the case where M is a smooth compact submanifold of $\mathbb{R}^{2d}$. We then rely on

Proposition 4.1. *For any smooth compact submanifold M of $\mathbb{R}^N$, any $C^{k'}$ riemannian metrics g on M, and for any bounded neighborhood U of M, there exists a $C^{k'}$ riemannian metrics G on $\mathbb{R}^N$ with*

- *$G = e$ in $\mathbb{R}^N \setminus U$, e the euclidean scalar product.*
- *$G|_M = g$, in the sense that for all $x \in M$, for all $v_1, v_2 \in T_x M$, $G_x(v_1, v_2) = g_x(v_1, v_2)$.*

This proposition will be proved in Appendix B. As g is constant outside a compact set, we can always consider its restriction to a large cube and periodize it so as to obtain a smooth metrics on a torus, as claimed.

Step 2: From an embedding to an immersion.

We now show that the requirement that u be one-to-one can be removed. In particular, this means that we will only need to find a u satisfying the relation (4.1), as it automatically implies that u is an immersion. This simplification is made possible thanks to the following property:

Lemma 4.1. *Let $u_i : M \to \mathbb{R}^{n_i}$, $g_i = (u_i)^* e_i$, $t_i > 0$, $i = 1, 2$. Then,*

$$u = (t_1 u_1, t_2 u_2) : M \to \mathbb{R}^{n_1+n_2}$$

satisfies $u^* e = t_1^2 g_1 + t_2^2 g_2$.

Here, e_i, resp. e, denotes the scalar product on $\mathbb{R}^{n_i}$, resp. $\mathbb{R}^n$. The proof of this lemma is easy and left to the reader. Now, assume that we know:

i) a smooth embedding $u_1 : M \to \mathbb{R}^{n_1}$ (for instance given by Whithney's theorem).

ii) *for some $k' \geq k$ and any $C^{k'}$ metrics g_2 on M, the existence of $u_2 : M \to \mathbb{R}^{n_2}$ of class C^k such that $g_2 = u_2^* e$.*

Then, given a $C^{k'}$ riemannian metrics g, with k' as in point ii), we set $g_1 = u_1^* e$, $g_2 = g - \lambda^2 g_1$, $\lambda > 0$. As M is compact, for $\lambda > 0$ small enough, g_2 is definite positive, hence a $C^{k'}$ riemannian metrics. By ii), there exists $u_2 : (M, g_2) \to \mathbb{R}^{n_2}$ of class C^k with $g_2 = u_2^* e$. By the lemma, it follows that $u = (\lambda u_1, u_2)$ satisfies $u^* e = g$, and as u_1 is one-to-one, u is one-to-one as well. We get in this way a C^k embedding from (M, g) to some $\mathbb{R}^n$. Therefore, we only need to focus on ii).

Step 3: Density of representable metrics and reduction to a perturbative problem.

We call representable any metrics g on M of the form $u^* e$, for an immersion $u : M \to \mathbb{R}^n$. In a local chart (or even global in the case of a torus), this reads

$$g = (\nabla u)(\nabla u)^t, \quad \nabla u = (\partial_i u_j)_{1 \leq i \leq d, 1 \leq j \leq n}, \quad \text{see (4.1)}.$$

A key ingredient to reduce the problem to a perturbative setting is the following density result of *representable* metrics.

Proposition 4.2. *Let $k' \geq 1$, g a $C^{k'}$ riemannian metrics on M. One can find a family of $C^{k'-1}$ riemannian metrics $(g_\varepsilon)_{\varepsilon > 0}$ such that*

- $\|g_\varepsilon - g\|_{C^{k'-1}} \leq \varepsilon$.
- *$g_\varepsilon = u_\varepsilon^* e$, for some $u_\varepsilon : M \to \mathbb{R}^{n_\varepsilon}$ of class $C^{k'}$.*

The proof of this proposition is postponed to Appendix B. Thanks to this proposition, it is enough to prove the following claim:

Claim: There exists a smooth riemannian metrics g_0 on M, $k' \geq k$ and $\varepsilon > 0$ such that: for all $C^{k'}$ riemannian metrics $\tilde{g}$ with $\|\tilde{g} - g_0\|_{C^{k'}} \leq \varepsilon$, $\tilde{g}$ is representable, i.e. there exists a C^k immersion $u : M \to \mathbb{R}^n$ for some n with $u^* e = \tilde{g}$.

Indeed, assume that this claim holds and let g be an arbitrary $C^{k'+1}$ riemannian metrics on M. We can always assume that $g' = g - g_0$ is also a riemannian metrics: indeed, we can always replace g_0 by $c_0 g_0$ for a small enough positive constant c_0, which amounts to replace ε by $c_0\varepsilon$ in the claim. For the ε of the claim, by the density lemma, there exists a $C^{k'}$ representable metrics g_ε (with a $C^{k'}$ immersion) such that $\|g' - g_\varepsilon\|_{C^{k'}} \leq \varepsilon$. This last inequality reads $\|\tilde{g} - g_0\|_{C^{k'}} \leq \varepsilon$, where $\tilde{g} = g - g_\varepsilon$. It follows from the claim that $g - g_\varepsilon$ is representable (with a C^k immersion), but as g_ε is also representable, Lemma 4.1 implies that the sum is representable (with a C^k immersion). This proves that any regular enough riemannian metric g is representable, as expected. Thus, it remains to prove the claim, which is a perturbative result.

Step 4: Choice of the reference solution.

The last step of these preliminaries consists in selecting a good g_0 for which one is able to solve (4.1) for g near g_0. We take a metrics g_0 that can be written in local coordinates as $g_0 = (\nabla u_0)(\nabla u_0)^t$, where the vectors $\partial_i u_0(x), \partial^2_{ij} u_0(x)$, $1 \leq i \leq j \leq d$ are independent for all $x \in M$. Such immersion will be called free. One can check that the notion of free immersion is independent of the choice of the local chart (exercise). Note that, as we restricted in Step 1 to the case of a torus, we could take global coordinates, but this is not useful here. Obviously, before picking up a free immersion, one must check that free immersions exist ! This can be shown as follows. By Whitney's theorem, we can always assume that M is a submanifold of $\mathbb{R}^{2d}$ of dimension d. Moreover, locally, M can be expressed as a graph: for any $\overline{x} \in M$, after possible reindexing of the coordinates, there exists an open neighborhood U of $\overline{x}$ in $\mathbb{R}^{2d}$ such that all points x in $U \cap M$ take the form

$$x = (x_1, \ldots, x_d, \psi(x_1, \ldots, x_d)), \quad \psi \text{ with values in } \mathbb{R}^d.$$

Let

$$\phi : \mathbb{R}^{2d} \to \mathbb{R}^{2d+d(2d+1)}, \quad \phi(x_1, \ldots, x_{2d}) = (x_i, x_k\, x_l)_{1\leq i\leq 2d, 1\leq k\leq l\leq 2d}.$$

It is a simple verification that $\phi|_M$ is a free immersion of M in $\mathbb{R}^{2d+d(2d+1)}$. Indeed, one can check that

$$\Phi(x_1, \ldots, x_d) = \phi\left(x_1, \ldots, x_d, \psi(x_1, \ldots, x_d)\right)$$

satisfies:

$$\partial_i \Phi(x_1, \ldots, x_d), \partial_k \partial_l \Phi(x_1, \ldots, x_d),\ 1 \leq i \leq d, 1 \leq k \leq l \leq d, \text{ are independent vectors.}$$

Conclusion. After this simplification procedure, it remains to show the following

Theorem 4.3. *Let $M = \mathbb{T}^d$ a d-dimensional torus, $g_0 = (\nabla u_0)(\nabla u_0)^t$ a smooth riemannian metrics given by a smooth and free immersion $u_0 : M \to \mathbb{R}^{n_0}$. There exists $k' \geq k$, $\varepsilon > 0$, such that for any riemannian metrics g with $\|g - g_0\|_{C^{k'}} \leq \varepsilon$, there exists a C^k immersion $u : M \to \mathbb{R}^{n_0}$ (close to u_0) with $u^* e = g$.*

The next section is dedicated to the proof of this result.

4.2 Use of the Nash-Moser theorem

The proof of Theorem 4.3 comes down to the resolution of equation (4.1), for g close to g_0. This resolution can be tackled using the Nash-Moser theorem 3.1. We shall consider the family of spaces $X_s = (C^s(\mathbb{T}^d))^{n_0}$, see Remark 3.1, while $Y_s = (C^s(\mathbb{T}^d))^{\frac{d(d+1)}{2}}$. Setting $f = g - g_0$, $\mathfrak{u} = u - u_0$, we rewrite (4.1) as

$$F(\mathfrak{u}) = f, \quad F_{ij}(\mathfrak{u}) = \frac{\partial u_0}{\partial x_i} \cdot \frac{\partial \mathfrak{u}}{\partial x_j} + \frac{\partial \mathfrak{u}}{\partial x_i} \cdot \frac{\partial u_0}{\partial x_j} + \frac{\partial \mathfrak{u}}{\partial x_i} \cdot \frac{\partial \mathfrak{u}}{\partial x_j}, \quad 1 \leq i \leq j \leq d.$$

We fix an integer $k > 4$, and want to verify the assumptions 1. to 4. in chapter 3. Clearly, F is C^1 from B_k to Y_{k-1} for any ball B_k in X_k, and for all $\mathfrak{u} \in B_k, v \in X_k$:

$$F'(\mathfrak{u})v = \frac{\partial(u_0 + \mathfrak{u})}{\partial x_i} \cdot \frac{\partial v}{\partial x_j} + \frac{\partial v}{\partial x_i} \cdot \frac{\partial(u_0 + \mathfrak{u})}{\partial x_j}$$

while for all $\mathfrak{u}, v \in B_k$

$$F(\mathfrak{u} + v) - F(u) - F'(\mathfrak{u})v = \frac{\partial v}{\partial x_i} \cdot \frac{\partial v}{\partial x_j}$$

Assumptions 1. and 2. hold for any $k \geq r \geq 1$.

To check the so-called *tame estimate 3.*, we rely on the following product estimate

Lemma 4.2. *For all $t \in \mathbb{N}$, and $f, g \in C^t(\mathbb{T}^d)$*

$$\|fg\|_{C^t} \lesssim \|f\|_{C^0} \|g\|_{C^t} + \|f\|_{C^t} \|g\|_{C^0} \tag{4.2}$$

Proof. From the classical formula $\|h'\|_{C^0} \leq \sqrt{2\|h\|_{C^0}\|h''\|_{C^0}}$ valid for any periodic function $h = h(t)$ of one real variable, we deduce that for all $i = 1 \ldots, d$, $\|\partial_i f\|_{C^0} \leq 2\|f\|_{C^0}^{1/2} \|\partial_i^2 f\|_{C^0}^{1/2}$. By a classical induction argument, it follows that for all t, for all $\alpha \in \mathbb{N}^d$ with $|\alpha| \leq t$, we have

$$\|\partial^\alpha f\|_{C^0} \lesssim \|f\|_{C^0}^{1-\frac{|\alpha|}{t}} \|f\|_{C^t}^{\frac{|\alpha|}{t}}.$$

To establish (4.2), it is enough to show by induction that: for all t, for all $\alpha, \beta \in \mathbb{N}^d$ with $|\alpha + \beta| \leq t$,

$$\|\partial^\alpha f\|_{C^0} \|\partial^\beta g\|_{C^0} \lesssim \|f\|_{C^0} \|g\|_{C^t} + \|f\|_t \|g\|_{C^0}$$

The case $t = 0$ is obvious. Let us assume that the inequality holds for all $\alpha, \beta \in \mathbb{N}^d$ with $|\alpha+\beta| \le t$. As $\| \, \|_{C^t} \le \| \, \|_{C^{t+1}}$, the inequality still holds with t replaced by $t+1$ if $|\alpha+\beta| \le t$. For $|\alpha + \beta| = t + 1$, we use the interpolation inequality, that leads to:

$$\|\partial^\alpha f\|_{C^0} \, \|\partial^\beta g\|_{C^0} \lesssim \left(\|f\|_{C^0}^{1-\frac{|\alpha|}{t+1}} \|g\|_{C^{t+1}}^{\frac{|\beta|}{t+1}} \right) \left(\|f\|_{C^{t+1}}^{\frac{|\alpha|}{t+1}} \|g\|_{C^0}^{1-\frac{|\beta|}{t+1}} \right)$$

The desired inequality then folllows from the convexity inequality : $ab \le \frac{1}{p}a^p + \frac{1}{q}b^q$ for $a, b > 0$, $\frac{1}{p} + \frac{1}{q} = 1$, applied with $p = \frac{1}{1-\frac{|\alpha|}{t+1}} = \frac{t+1}{\beta}$ and $q = \frac{1}{1-\frac{|\beta|}{t+1}} = \frac{t+1}{\alpha}$ $\square$

By taking $t = s - 1$ with $s \ge k$, the lemma implies easily assumption 3. for any $r \ge 1$.

The last piece of the proof is to check assumption 4. Here, we face the difficulty that the linearized system $F'(\mathfrak{u})v = g$, which can be written as

$$\frac{\partial(u_0 + \mathfrak{u})}{\partial x_i} \cdot \frac{\partial v}{\partial x_j} + \frac{\partial v}{\partial x_i} \cdot \frac{\partial(u_0 + \mathfrak{u})}{\partial x_j} = g_{ij}, \quad \forall 1 \le i \le j \le d \tag{4.3}$$

is underdetermined. A nice idea of Nash is to add constraints on v, namely

$$\frac{\partial(u_0 + \mathfrak{u})}{\partial x_k} \cdot v = 0, \quad \forall 1 \le k \le d \tag{4.4}$$

By taking $k = i$ (resp. $k = j$) in the previous relation, differentiating with respect to x_j (resp. to x_i), we infer that

$$\frac{\partial^2(u_0 + \mathfrak{u})}{\partial x_i x_j} \cdot v = \frac{\partial(u_0 + \mathfrak{u})}{\partial x_i} \cdot \frac{\partial v}{\partial x_j} = \frac{\partial(u_0 + \mathfrak{u})}{\partial x_j} \cdot \frac{\partial v}{\partial x_i} \tag{4.5}$$

so that (4.3) becomes

$$2 \, \frac{\partial^2(u_0 + \mathfrak{u})}{\partial x_i x_j} \cdot v = g_{ij}, \quad \forall 1 \le i \le j \le d. \tag{4.6}$$

As u_0 is a free immersion, under the assumption that the radius of B_k is small enough, one has that for any $\mathfrak{u} \in X_\infty \cap B_k$, $u_0 + \mathfrak{u}$ is a free immersion as well. Hence, the system (4.4)-(4.6) (which is equivalent to (4.4)-(4.3)) is pointwise solvable as soon as $n_0 > d + \frac{d(d+1)}{2}$. To build a right inverse, we can for instance proceed in the following way. For all $x \in M$, we orthonormalize the family $\frac{\partial^2(u_0+\mathfrak{u})}{\partial x_i x_j}(x)$, $\frac{\partial(u_0+\mathfrak{u})}{\partial x_k}(x)$, resulting in an orthonormal family $e_\ell(x)$, $\ell = 1, \dots, d + \frac{d(d+1)}{2}$, with explicit dependence on the vectors of the former family, and that spans the same vector space. We can then compute from (4.4) and (4.6) the values of $v \cdot e_\ell$ for all ℓ, and set $v = \sum_\ell (v \cdot e_\ell) e_\ell$. Note that the dependence on v is explicit and smooth on $\frac{\partial^2(u_0+\mathfrak{u})}{\partial x_i x_j}(x)$, $\frac{\partial(u_0+\mathfrak{u})}{\partial x_k}(x)$, and it is smooth and linear in the coefficients of the riemannian metrics $g(x)$. Hence, by Lemma 4.2, the right inverse $v = L(\mathfrak{u})g$ satisfies for all t

$$\|L(\mathfrak{u})g\|_t \le \|F\big(\partial^2_{ij}(u_0 + \mathfrak{u}), \partial_k(u_0 + \mathfrak{u})\big)\|_{C^0} \mid g|_t + \|F\big(\partial^2_{ij}(u_0 + \mathfrak{u}), \partial_k(u_0 + \mathfrak{u})\big)\|_{C^t} \mid g|_0. \tag{4.7}$$

where F is a smooth vector valued function of its arguments. We then use the

Lemma 4.3. *Let $G : \Omega \to \mathbb{R}$ a smooth function, Ω an open set of $\mathbb{R}^n$. Then, for all $t \in \mathbb{N}$, for all $U \in C^t(\mathbb{T}^d, K)$ with $K \Subset \Omega$,*

$$\|G(U)\|_{C^t} \lesssim 1 + \mathcal{C}(\|U\|_{L^\infty})\|U\|_t.$$

for an increasing function $\mathcal{C}$.

For brevity, we skip the proof of this lemma, that can be obtained by writing $\partial^\alpha G(u)$ as a linear combination of products of some $G^{(k)}(u)$ and some derivatives of u. These products can then be controlled using Lemma 4.2.

Applying this lemma in (4.7), we see that the assumption 4. of the Nash-Moser theorem is satisfied for $r = 2$. The Nash-Moser theorem therefore provides a solution to the embedding problem.

Chapter 5

Littlewood-Paley decomposition

We have encountered in chapter 3 the notion of *tame estimate*, see Remark 3.2 or Lemma 4.2. We shall see here that tame estimates are very natural, notably in a PDE context: in most functional spaces used in the analysis of PDEs, products and compositions by smooth functions obey tame estimates, similar to those encountered in Lemma 4.2 or Lemma 4.3. To establish these estimates, we shall rely on the so-called *Littlewood-Paley decomposition*, which is an important tool of Fourier analysis. This decomposition, and the related notion of *paraproduct* have revealed efficient in the study of nonlinear PDE's, and sometimes used as a substitute to Nash-Moser techniques ([11, 3]). This will be a bit discussed at the end of the chapter.

For a detailed description of Littlewood Paley analysis and its applications, we refer the reader to the nice monograph [2]. We rather borrow here to the lighter presentation in [1], see also the last chapter in [6]. We start with reminders on temperate distributions.

5.1 The space of temperate distributions $\mathcal{S}'(\mathbb{R}^d)$

5.1.1 Definitions, examples

Definition 5.1. (the Schwartz space $\mathcal{S}(\mathbb{R}^d)$)
The Schwartz space or space of rapidly decreasing functions on $\mathbb{R}^d$ is

$$\mathcal{S}(\mathbb{R}^d) = \left\{\varphi \in C^\infty(\mathbb{R}^d, \mathbb{C}), \quad \forall k \in \mathbb{N},\ \|\varphi\|_k < +\infty\right\}, \quad \|\varphi\|_k = \sup_{x\in\mathbb{R}^d, |\alpha|\le k} (1+|x|)^k |\partial^\alpha \varphi(x)|.$$

Proposition 5.1. *For all $k \in \mathbb{N}$, $\|\varphi\,\|_k = \sup_{x\in\mathbb{R}^d,|\alpha|\le k}(1+|x|)^k|\partial^\alpha\varphi(x)|$ defines a norm on $\mathcal{S}(\mathbb{R}^d)$, and*

$$d_{\mathcal{S}}(\varphi,\psi) = \sum_{k\in\mathbb{N}} \frac{\min(1, \|\varphi-\psi\|_k)}{2^k}$$

defines a distance on $\mathcal{S}(\mathbb{R}^d)$ which makes it a complete metric space.

Remark 5.1. *One can check that $d_{\mathcal{S}}(\varphi_n, \varphi) \to 0$ if and only if $\|\varphi_n - \varphi\|_k \to 0$ for all k.*

Definition 5.2. (the space of temperate distributions $\mathcal{S}(\mathbb{R}^d)$)
The space of temperate distributions $\mathcal{S}'(\mathbb{R}^d)$ is the space of all continuous linear forms on $\mathcal{S}(\mathbb{R}^d)$

Remark 5.2. *By continuous on $\mathcal{S}(\mathbb{R}^d)$, we mean continuous on $(\mathcal{S}(\mathbb{R}^d), d_{\mathcal{S}})$. One can show that $u : \mathcal{S}(\mathbb{R}^d) \to \mathbb{R}$ is continuous if and only if there exists $k \in \mathbb{N}$ and $C_k > 0$ such that*

$$\forall \varphi \in \mathcal{S}(\mathbb{R}^d), \quad |u(\varphi)| \leq C_k \|\varphi\|_k.$$

For $u \in \mathcal{S}'(\mathbb{R}^d)$ and $\varphi \in \mathcal{S}(\mathbb{R}^d)$, we will often note $\langle u, \varphi \rangle$ (duality bracket) for $u(\varphi)$.
We remind here a few examples of temperate distributions:

- If $u \in L^1_{loc}(\mathbb{R}^d)$ grows at most polynomially, *i.e.* there exists $C, k, M > 0$ such that $|u(x)| \leq C(1+|x|)^k$ for a.e. $|x| \geq M$, then u defines an element T_u of $\mathcal{S}'(\mathbb{R}^d)$ through the formula
$$\langle T_u, \varphi \rangle \ = \ \int_{\mathbb{R}^d} u\, \varphi.$$
The same holds if $u \in L^p(\mathbb{R}^d)$ for some $1 \leq p \leq +\infty$. In both cases, the mapping $u \to T_u$ is one-to-one, as a consequence of the following general result: *for any open set $\Omega \subset \mathbb{R}^d$, and any $u \in L^1_{loc}(\Omega)$,*
$$\int_\Omega u\varphi = 0 \quad \forall \varphi \in C_c^\infty(\Omega) \quad \Rightarrow \quad u = 0 \ \text{ a.e. in } \Omega \tag{5.1}$$
(if two functions are equal in the sense of distributions, they are equal almost everywhere). This justifies to identify the function and the distribution, and use notation $\langle u, \varphi \rangle$ instead of $\langle T, \varphi \rangle$.
- If μ is a complex measure on the borelians of $\mathbb{R}^d$, it can be identified to a linear form on $C_b(\mathbb{R}^d)$, and $\mu|_{\mathcal{S}(\mathbb{R}^d)}$ belongs to $\mathcal{S}'(\mathbb{R}^d)$.
- If $\chi \in C^\infty(\mathbb{R}^d)$ is a smooth function that grows at most polynomially (see the first item), and $\mathcal{S}'(\mathbb{R}^d)$ one can define the product $u\chi \in \mathcal{S}'(\mathbb{R}^d)$ (also noted χu) by the formula $\langle u\chi, \varphi \rangle = \langle u, \chi\varphi \rangle$.
- **Derivative of a temperate distribution.** Given $u \in \mathcal{S}'(\mathbb{R}^d)$ and $\alpha \in \mathbb{N}^d$, one defines the element $\partial^\alpha u \in \mathcal{S}'(\mathbb{R}^d)$ by the formula: $< \partial^\alpha u, \varphi >= (-1)^{|\alpha|} \langle u, \varphi \rangle$.

Proposition 5.2. (Support of a distribution)
We say that $u \in S'(\mathbb{R}^d)$ is zero on an open set O of $\mathbb{R}^d$ if $\langle u, \varphi \rangle = 0$ for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$, compactly supported in O. The union of all open sets on which u is zero is still an open set on which u is zero, and we call its support its complementary set.

5.1.2 Link with the Fourier transform

Notation. $\mathcal{F}: L^1(\mathbb{R}^d) \to C_0^0(\mathbb{R}^d)$, $\quad \mathcal{F}f(\xi) = \int_{\mathbb{R}^d} e^{-2i\pi\xi\cdot x} f(x)dx$. One shall also use notation $\hat{f}$ instead of $\mathcal{F}$.

From the algebraic identities:

$$\widehat{\partial^\alpha f}(\xi) = i^{|\alpha|}\xi^\alpha \hat{f}(\xi), \quad \text{and } \widehat{x^\alpha f}(\xi) = i^{|\alpha|}\partial^\alpha \hat{f}(\xi)$$

it is easily seen that $\mathcal{F}$ sends the Schwartz space to itself. Combining this with the Fourier inversion formula, it is easily seen that

Proposition 5.3. *$\mathcal{F}$ is an isomorphism from $\mathcal{S}(\mathbb{R}^d)$ to itself, with inverse given by $\mathcal{F}^{-1}\varphi(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{ix\cdot\xi}\varphi(\xi)d\xi$.*

This proposition allows to define the Fourier transform on the space of temperate distributions, through duality:

Definition 5.3. (Fourier transform of a temperate distribution)
For $u \in \mathcal{S}'(\mathbb{R}^d)$, the Fourier transform $\hat{u}$ of u is given by the formula:

$$\langle \hat{u}, \varphi \rangle = \langle u, \hat{\varphi} \rangle, \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^d)$$

The following can be shown:

Proposition 5.4. *$u \to \hat{u}$ is an isomorphism from $\mathcal{S}'(\mathbb{R}^d)$ to itself.*

Remark 5.3. *Continuity of a function $F : \mathcal{S}'(\mathbb{R}^d) \to X$ with X a topological space is understood as follows: if $u_n \to u$ in $\mathcal{S}'(\mathbb{R}^d)$, in the sense that $\langle u_n, \varphi \rangle \to \langle u, \varphi \rangle$ for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$, then $F(u_n) \to F(u)$.*

Remark 5.4. *For $u \in \mathcal{S}'(\mathbb{R}^d)$, the inverse Fourier transform of u satisfies:*

$$\langle \mathcal{F}^{-1}u, \varphi \rangle \; = \; \langle u, \mathcal{F}^{-1}\varphi \rangle, \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^d)$$

Remark 5.5. *In the case where $u \in \mathcal{S}(\mathbb{R}^d)$, or $u \in L^1(\mathbb{R}^d)$, or $u \in L^2(\mathbb{R}^d)$, the usual definition of $\hat{u}$ coincides with the definition in the sense of $\mathcal{S}'$ (seeing functions as temperate distributions, see the first example before paragraph 5.1.2.*

Definition 5.4. (Fourier multiplier)
Let $\chi \in C^\infty(\mathbb{R}^d, \mathbb{C})$ that grows at most polynomially. We define $\chi(D) : \mathcal{S}'(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ by the formula: $\widehat{\chi(D)u} = \chi\hat{u}$.

In the next paragraph, we will use the convolution between a temperate distribution and a fastly decreasing function, defined as follows:

Definition 5.5. *For $u \in \mathcal{S}'(\mathbb{R}^d)$, $\varphi \in \mathcal{S}(\mathbb{R}^d)$, one can define the convolution $u \star \varphi$ by*

$$u \star \varphi(x) = \langle u, \varphi(x - \cdot) \rangle$$

Note that again, this definition is consistent with the usual ones when u is more than a temperate distribution. In the same spirit, one has $u \star \varphi = \varphi \star u$ when u is such that the right side is well-defined.

Proposition 5.5. *$u \star \varphi$ is a smooth function with at most polynomial growth, and for any $\alpha \in \mathbb{N}^d$, $\partial^\alpha(u \star \varphi) = \partial^\alpha u \star \varphi = u \star \partial^\alpha \varphi$. In particular, $u \star \varphi \in \mathcal{S}'(\mathbb{R}^d)$ (see the first example after paragraph 5.1.2).*

Proposition 5.6. *For $u \in \mathcal{S}'(\mathbb{R}^d)$, $\varphi \in \mathcal{S}(\mathbb{R}^d)$, $\widehat{u \star \varphi} = \hat{u}\,\hat{\varphi}$.*

Note that the right-hand side is well-defined as the product of the temperate distribution $\hat{u}$ and the fastly decreasing function $\hat{\varphi}$ (see the third example after paragraph 5.1.2). Note also that the identity in the proposition implies that *a temperate distribution whose Fourier transform is compactly supported is a smooth function with polynomial growth.* Indeed, one can take $\phi \in \mathcal{S}(\mathbb{R}^d)$ such that $\phi = 1$ near the support of $\hat{u}$. We then have $\hat{u} = \hat{u}\phi$ so that $u = u \star g$ with $\hat{g} = \phi$. The result follows from Proposition 5.5.

5.2 Littlewood-Paley decomposition

The Littlewood-Paley decomposition is a decomposition of functions (or temperate distributions) where each term in the decomposition has a localized frequency support, either in a ball $B(0, R)$ or in an annulus $C(0, R, R')$ (with small radius R, large radius R'). Among its possible interests, such decomposition allows for more precise estimates: the underlying ingredient is the following

Proposition 5.7. (Bernstein inequalities) *Let $r_1, r_2 > 0$. There exists C such that for all $k \geq 1$, for all $\lambda > 0$ $u \in L^p(\mathbb{R}^d)$ ($1 \leq p \leq +\infty$):*

$$\text{supp } \hat{u} \subset B(0, r_1\lambda) \;\Rightarrow\; \sup_{|\alpha|=k} \|\partial^\alpha u\|_{L^p} \leq C^k \lambda^k \|u\|_{L^p} \tag{5.2}$$

$$\text{supp } \hat{u} \subset C(0, r_1\lambda, r_2\lambda) \;\Rightarrow\; C^{-k}\lambda^k \|u\|_{L^p} \leq \sup_{|\alpha|=k} \|\partial^\alpha u\|_{L^p} \leq C^k \lambda^k \|u\|_{L^p} \tag{5.3}$$

Proof. By considering $u_\lambda(x) = u(\lambda x)$ one can restrict to the case $\lambda = 1$. Let $\phi \in C_c^\infty(\mathbb{R}^d)$ such that $\phi = 1$ in a vicinity of $B(0, r_1)$, and $g \in \mathcal{S}(\mathbb{R}^d)$ such that $\hat{g} = \varphi$. If $\hat{u}$ is supported in $B(0, r_1)$, then $\hat{u}\phi = \hat{u}$, which implies that $u = u \star g$, and from there $\partial^\alpha u = u \star (\partial^\alpha g)$. By Young's inequality, we get

$$\begin{aligned}
\|\partial^\alpha u\|_{L^p} &\leq \|\partial^\alpha g\|_{L^1}\, \|u\|_{L^p} \\
&\leq C_d\, \|(1 + |x|^2)^d \partial^\alpha g\|_{L^\infty}\, \|u\|_{L^p} \\
&\leq C_d' \|(1 - \Delta)^d (\xi^\alpha \phi)\|_{L^1}\, \|u\|_{L^p} \leq C^{|\alpha|} \|\ \|u\|_{L^p}
\end{aligned}$$

for C large enough. This shows (5.2), which implies the right-hand side inequality of (5.3).

For the remaining inequality, we shall use the identity

$$|\xi|^{2k} = (\xi_1^2 + \cdots + \xi_d^2)^k = \sum_{|\alpha|=k} A_\alpha \, \xi_1^{2\alpha_1} \ldots \xi_d^{2\alpha_d} = (-1)^k \sum_{|\alpha|=k} A_\alpha (i\xi)^\alpha (i\xi)^\alpha,$$

with $A_\alpha = \frac{k!}{\alpha_1! \ldots \alpha_d!}$. Note that

$$\sum_{|\alpha|=k} A_\alpha = (1 + \cdots + 1)^k = d^k, \quad \text{so that } \; A_\alpha \leq d^k. \tag{5.4}$$

We consider $\phi \in C_c^\infty(\mathbb{R}^d \setminus \{0\})$ such that $\phi = 1$ near $C(0, r_1, r_2)$ and set for all α of length k:

$$\phi_\alpha(\xi) = (-1)^k A_\alpha \frac{(i\xi)^\alpha}{|\xi|^{2k}} \phi(\xi), \quad g_\alpha = \mathcal{F}^{-1}\phi_\alpha.$$

Note that $\phi_\alpha \in \mathcal{S}(\mathbb{R}^d)$: there is in particular no singularity at $\xi = 0$, due to the vanishing of ϕ. By the above algebraic identity, and as $\hat{u}\phi = \hat{u}$, we deduce:

$$u = \sum_{|\alpha|=k} \partial^\alpha u \star g_\alpha, \quad \text{so that } \; \|u\|_{L^p} \leq \sum_{|\alpha|=k} \|g_\alpha\|_{L^1} \|\partial^\alpha u\|_{L^p}$$

One can show as for (5.2) and taking into account (5.4) that $\|g_\alpha\|_{L^1} \leq C^k$ for C large enough, which concludes the proof. $\square$

We now turn to the definition of the Littlewood-Paley decomposition. Let Ψ a smooth and non-increasing function on $\mathbb{R}_+$ such that $\Psi(r) = 1$ for $0 \leq r \; \leq \frac{1}{2}$, $\Psi(r) = 0$ for $r \geq 1$. Let $\psi(x) = \Psi(|x|)$, $\chi(\xi) = \psi(\frac{\xi}{2}) - \psi(\xi)$. From the definitions, it follows that ψ and χ are non-negative, that ψ is supported in $B(0,1)$, that χ is supported in the annulus $\mathcal{C} = C(0, \frac{1}{2}, 2)$, and that for all $\xi \in \mathbb{R}^d$:

$$1 = \psi(\xi) + \sum_{p \geq 0} \chi(2^{-p}\xi) \tag{5.5}$$

Note that the sum above has at most two non-zero terms for each ξ: indeed, $\xi \to \chi(2^{-p}\xi)$ has support in $2^p\mathcal{C} = C(0, 2^{p-1}, 2^{p+1})$, so that $2^p\mathcal{C} \cap 2^q\mathcal{C} = \emptyset$ as soon as $|q - p| \geq 2$. One can then show (exercise) that for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$,

$$\hat{\varphi} = \psi\hat{\varphi} + \sum_{p \geq 0} \chi(2^{-p}\cdot)\hat{\varphi}$$

with convergence of the right-hand side in $\mathcal{S}(\mathbb{R})^d$. It follows easily that

$$\forall u \in \mathcal{S}'(\mathbb{R}^d), \quad u = \psi(D)u + \sum_{p \geq 0} \chi(2^{-p}D)u \tag{5.6}$$

with convergence of the series in $\mathcal{S}'(\mathbb{R}^d)$. We shall note

$$\Delta_{-1}u = \psi(D)u, \quad \Delta_p u = \chi(2^{-p}D)u \quad \text{for all } p \geq 0,$$

and eventually

$$S_p u = \psi(2^{-p}D)u = \sum_{q=-1}^{p-1} \Delta_q u \qquad \text{for all } p \geq 0.$$

Note that

$$u = \lim_{p\to+\infty} S_p u, \quad \forall u \in \mathcal{S}'(\mathbb{R}^d).$$

Remark 5.6. *As $\Delta_p u$ or $S_p u$ has compact frequency support, it is smooth with at most polynomial growth.*

Lemma 5.1. *There exists $C > 0$ such that for all $1 \leq q \leq +\infty$, for all $u \in L^q(\mathbb{R}^d)$,*

$$\sup_{p\geq -1} \|S_p u\|_{L^q} \leq C\|u\|_{L^q}, \quad \sup_{p\geq -1} \|\Delta_p u\|_{L^q} \leq C\|u\|_{L^q}$$

Proof. As $S_p u = \psi(2^{-p}D)u$, it can also be written $2^{pd}(\mathcal{F}^{-1}\psi)(2^p\cdot) \star u$. By Young inequality, it follows that

$$\|S_p u\|_{L^q} \leq \|2^{pd}(\mathcal{F}^{-1}\psi)(2^p\cdot)\|_{L^1}\, \|u\|_{L^q} \leq \|\mathcal{F}^{-1}\psi\|_{L^1}\|u\|_{L^q}$$

The same reasoning applies to $\Delta_p u$, replacing ψ by χ. □

Lemma 5.2. (Almost orthogonality)

$$\forall \xi \in \mathbb{R}^d, \quad \frac{1}{2} \leq \psi^2(\xi) + \sum_{p\geq 0} \chi^2(2^{-p}\xi) \leq 1.$$

and for all $u \in L^2(\mathbb{R}^d)$,

$$\sum_{p\geq -1} \|\Delta_p u\|_{L^2}^2 \leq \|u\|_{L^2}^2 \ \leq\ 2 \sum_{p\geq -1} \|\Delta_p u\|^2$$

Proof. The first double inequality follows from the fact that the sum $1 = \psi(\xi)+\sum_{p\geq 0}\chi(2^{-p}\xi)$ has at most two non-zero terms for each value of ξ. Hence, the result follows from the elementary inequalities $a^2 + b^2 \leq (a+b)^2 \leq 2\,(a^2+b^2)$. The second double inequality is obtained from the first one by multiplying by $\hat{u}(\xi)$, integrating over $\mathbb{R}^d$, and using Plancherel identity. □

5.3 Link with Sobolev and Hölder spaces

Definition 5.6. (Sobolev spaces) *For all $s \in \mathbb{R}$, we define*

$$H^s(\mathbb{R}^d) = \{u \in \mathcal{S}'(\mathbb{R}^d), \quad \xi \to (1+|\xi|^2)^{s/2}\hat{u}(\xi) \in L^2(\mathbb{R}^d)\}$$

Note that for $s \geq 0$, $\hat{u} \in L^2(\mathbb{R}^d)$, so that $u \in L^2(\mathbb{R}^d)$. One can show that it is a Hilbert space, equipped with the norm $\|u\|_{H^s} = \left(\int_{\mathbb{R}^d}(1+|\xi|^2)^s|\hat{u}(\xi)|^2 d\xi\right)^{1/2}$.

In the case when $s \in \mathbb{N}$, there is an alternative definition, in terms of *weak derivatives* in L^2:

Definition 5.7. (**weak derivatives**) *Let Ω an open set of $\mathbb{R}^d$, $1 \leq p \leq +\infty$ and $\alpha \in \mathbb{N}^d$. For $u \in L^1_{loc}(\Omega)$, we say that $\partial^\alpha u \in L^p(\Omega)$ if there exists $v_\alpha \in L^p(\Omega)$ satisfying: for all $\varphi \in C_c^\infty(\Omega)$,*

$$\int_\Omega u\, \partial^\alpha \varphi = (-1)^{|\alpha|} \int_\Omega v_\alpha\, \varphi. \tag{5.7}$$

Remark 5.7. *One can show that this v_α is unique, and we denote it $\partial^\alpha u$. In terms of classical distributions, it means that the derivative $\partial^\alpha u \in \mathcal{D}'(\Omega)$ is representable by a function in $L^q(\Omega)$.*

Definition 5.8. *For $s \in \mathbb{N}$, $1 \leq p \leq +\infty$, Ω an open set of $\mathbb{R}^d$:*

$$W^{s,p}(\Omega) = \{u \in L^p(\Omega), \partial^\alpha u \in L^p(\Omega), \quad \forall |\alpha| \ \leq s\}$$

It is a Banach space, equipped with the norm $\|u\|_{W^{s,p}} = \left(\sum_{|\alpha|\leq s} \|u\|^2_{L^p}\right)^{1/2}$ (and a Hilbert space in the case $p = 2$).

Proposition 5.8. *$H^s(\mathbb{R}^d) = W^{s,2}(\mathbb{R}^d)$, with equivalent norms.*

Proof. Here are a few ideas of the proof. If $u \in H^s(\mathbb{R}^d)$, for all α with $|\alpha| \leq s$, the function $(i\xi)^\alpha \hat{u}$ belongs to L^2, so that $v_\alpha = \mathcal{F}^{-1}\left((i\xi)^\alpha \hat{u}\right) \in L^2(\mathbb{R}^d)$. One then shows that v_α is a weak derivative of u, *cf* identity (5.7). This identity can be easily proved thanks to the Plancherel formula. Conversely, if $u \in W^{s,2}(\mathbb{R}^d)$, and $v_\alpha = \partial^\alpha u$ denoting the weak derivative, the point is to show that $\widehat{v_\alpha} = (i\xi)^\alpha \hat{u}$. In other words, one has to show that this algebraic identity, valid for usual derivatives, extends to the weak derivatives. This again relies on identity (5.7), through Plancherel formula. □

Definition 5.9. (Hölder spaces) *For $\alpha > 0$, $\alpha \notin \mathbb{N}$, $C^\alpha(\mathbb{R}^d)$ is the space of functions $f \in C_b^k(\mathbb{R}^d)$, $k = E(\alpha)$, such that*

$$\forall \beta, \quad |\beta| \ \leq k, \quad \sup_{x \neq y} \frac{|\partial^\beta f(x) - \partial^\beta f(y)|}{|x-y|^{\alpha-k}} < +\infty$$

$C^\alpha(\mathbb{R}^d)$ is a Hilbert space equipped with the norm

$$\|u\|_{\tilde{C}^\alpha} = \sup_{|\beta|\leq k} \|\partial^\beta u\|_{L^\infty} \ + \ \sup_{|\beta|=k} \sup_{x\neq y} \frac{|\partial^\beta u(x) - \partial^\beta u(y)|}{|x-y|^{\alpha-k}}$$

We now characterize these functional spaces in terms of properties of the Littlewood Paley decomposition.

Proposition 5.9. (characterization of Sobolev spaces)
For all $s \in \mathbb{R}$, for all $u \in \mathcal{S}'(\mathbb{R}^d)$, $u \in H^s(\mathbb{R}^d)$ if and only if $\sum_{p \geq -1} 2^{2ps} \|\Delta_p u\|_{L^2}^2 < +\infty$. Moreover, there exists $C > 0$ such that for all $u \in H^s(\mathbb{R}^d)$

$$\frac{1}{C} \sum_{p \geq -1} 2^{2ps} \|\Delta_p u\|_{L^2}^2 \leq \|u\|_{H^s}^2 \leq C \sum_{p \geq -1} 2^{2ps} \|\Delta_p u\|_{L^2}^2$$

Proof. If $u \in H^s(\mathbb{R}^d)$, $\|u\|_s^2 = \|\langle D \rangle^s u\|_{L^2}$, where $\langle \xi \rangle = \sqrt{1 + |\xi|^2}$. By the second property of Lemma 5.2, we deduce

$$\sum \|\Delta_p \langle D \rangle^s u\|_{L^2}^2 \leq \|u\|_{H^s}^2 \leq 2 \sum \|\Delta_p \langle D \rangle^s u\|_{L^2}^2.$$

Then, using Plancherel formula and the definition of Δ_p, we have easily that there exists $C > 0$ such that for all $p \geq -1$

$$\frac{1}{C} 2^{ps} \|\Delta_p u\|_{L^2} \leq \|\Delta_p \langle D \rangle^s u\|_{L^2} \leq C 2^{ps} \|\Delta_p u\|_{L^2}$$

so that for some $C > 0$

$$\frac{1}{C} \sum_{p \geq -1} 2^{2ps} \|\Delta_p u\|_{L^2}^2 \leq \|u\|_s^2 \leq C \sum_{p \geq -1} 2^{2ps} \|\Delta_p u\|_{L^2}^2 \tag{5.8}$$

The left inequality yields in particular that $\sum_{p \geq -1} 2^{2ps} \|\Delta_p u\|_{L^2}^2 < +\infty$.

Conversely, let $u \in \mathcal{S}'(\mathbb{R}^d)$ such that $\sum_{p \geq -1} 2^{2ps} \|\Delta_p u\|_{L^2}^2 < +\infty$. It implies that for all $p \geq -1$, $\Delta_p u \in L^2(\mathbb{R}^d)$, so that for any $N \geq 0$, $S_N u$ belongs to $L^2(\mathbb{R}^d)$, and even $H^s(\mathbb{R}^d)$ as it is compactly supported in frequency. We can apply the reasoning above replacing u by $S_N u$, and the right inequality in (5.8) yields

$$\|S_N u\|_s^2 \leq C \sum_{p \geq -1} 2^{2ps} \|\Delta_p S_N u\|_{L^2}^2 \leq \sum_{p \geq -1} 2^{2ps} \|\Delta_p u\|_{L^2}^2$$

Hence, the sequence $(S_N u)_N$ is bounded in $H^s(\mathbb{R}^d)$, so that it has a subsequence weakly converging to $\tilde{u} \in H^s(\mathbb{R}^d)$. But it also converges to $u \in S'(\mathbb{R}^d)$, so that $u = \tilde{u} \in H^s(\mathbb{R}^d)$. The equivalence of the norms is then given by (5.8). □

Proposition 5.10. (characterization of Hölder spaces)
For all $\alpha \in \mathbb{R}_+ \setminus \mathbb{N}$, for all $u \in \mathcal{S}'(\mathbb{R}^d)$, $u \in C^\alpha(\mathbb{R}^d)$ if and only if $\sup_{p \geq -1} 2^{p\alpha} \|\Delta_p u\|_{L^\infty} < \infty$. Moreover, there exists $C > 0$ such that for all $u \in C^\alpha(\mathbb{R}^d)$

$$\frac{1}{C} \sup_{p \geq -1} 2^{p\alpha} \|\Delta_p u\|_{L^\infty} \leq \|u\|_{C^\alpha}^{\sim 2} \leq C \sup_{p \geq -1} 2^{p\alpha} \|\Delta_p u\|_{L^\infty}$$

Proof. We focus on the first part of the equivalence. Let $u \in C^\alpha(\mathbb{R}^d)$, and $k = E(\alpha)$. By Lemma 5.1, $\|\Delta_{-1}u\|_{L^\infty} \leq \|u\|_{L^\infty} \leq \|u\|_\alpha$. By Bernstein inequality (5.3), we find that for all $p \geq 0$,

$$\begin{aligned}
2^{p\alpha}\|\Delta_p u\|_{L^\infty} &= 2^{p(\alpha-k)}2^{pk}\|\Delta_p u\|_{L^\infty} \\
&\leq C^k 2^{p(\alpha-k)} \sup_{|\beta|=k} \|\partial^\beta \Delta_p u\|_{L^\infty} = C^k 2^{p(\alpha-k)} \sup_{|\beta|=k} \|\Delta_p \partial^\beta u\|_{L^\infty}
\end{aligned}$$

As $\partial^\beta u \in C^{\alpha-k}(\mathbb{R}^d)$, the problem reduces to proving the first part of the equivalence for $\alpha \in (0,1)$. In this case, we proceed as follows. We introduce $h_\chi = 2^{pd}\mathcal{F}^{-1}\chi(2^p\cdot)$, so that $\Delta_p u = h_\chi \star u$. As $\chi(0) = 0$, $\int_{\mathbb{R}^d} h_\chi = 0$, which implies

$$\forall x, \quad \Delta_p u(x) = \Delta_p u(x) - \left(\int_{\mathbb{R}^d} h_\chi\right) u(x) = \int_{\mathbb{R}^d} h_\chi(y)(u(x-y) - u(x))dy.$$

We deduce :

$$\begin{aligned}
|\ \Delta_p u(x)| &\leq \int_{\mathbb{R}^d} h_\chi(y)|y|^\alpha dy\, \|u\|_{C^\alpha} \leq\ 2^{-p\alpha}\|y \to h_\chi(y)2^p|y|^\alpha\|_{L^1}\|u\|_{C^\alpha} \\
&\leq 2^{-p\alpha}\|y \to\ \mathcal{F}^{-1}\chi(y)\,|y|^\alpha\|_{L^1}\,\|u\|_{C^\alpha}
\end{aligned}$$

Hence, $2^{p\alpha}\|\Delta_p u\|_{L^\infty}$ is bounded uniformly in p by $C\|u\|_{C^\alpha}$.

Conversely, let us assume that $M_\alpha = \sup_{p\geq -1} 2^{p\alpha}\|\Delta_p u\|_{L^\infty} < +\infty$. We set again $k = E(\alpha)$. We write, for any $|\beta| \leq k$:

$$\begin{aligned}
\|\partial^\beta S_N u - \partial^\beta S_M u\|_{L^\infty} &\leq \sum_{p=M}^{N-1} \|\partial^\beta\ \Delta_p u\|_{L^\infty} \leq C^k \sum_{p=M}^{N-1} 2^{pk}\|\Delta_p u\|_{L^\infty} \\
&\leq C^k\, M_\alpha \sum_{p=M}^{N-1} 2^{p(k-\alpha)}
\end{aligned}$$

It follows that $(S_N u)_N$ is a Cauchy sequence in $C^k_b(\mathbb{R}^d)$, converging to some $\tilde{u}$. But as $(S_N u)$ converges to u in $\mathcal{S}'(\mathbb{R}^d)$, $u = \tilde{u} \in C^k_b(\mathbb{R}^d)$. By taking $S_N u$ instead of $(S_N - S_M)u$ in the inequalities above, and letting $N \to +\infty$, we get:

$$\|\partial^\beta u\|_{L^\infty} \leq C \sup_{p\geq -1} 2^{p\alpha}\|\Delta_p u\|_{L^\infty}.$$

It remains to show that for all β with $|\beta| = k$,

$$\|\partial^\beta u\|_{C^{\alpha-k}} \leq C \sup_p 2^{p\alpha}\|\Delta_p u\|_{L^\infty}$$

By Bernstein inequality, $\|\partial^\beta \Delta_p u\|_{L^\infty} \leq C^k 2^{pk}\|\Delta_p u\|_{L^\infty}$, so that it is enough to show that

$$\|\partial^\beta u\|_{C^{\alpha-k}} \leq C \sup_p 2^{p(\alpha-k)}\|\partial^\beta \Delta_p u\|_{L^\infty} = C \sup_p 2^{p(\alpha-k)}\|\Delta_p\ \partial^\beta u\|_{L^\infty}.$$

Setting $v = \partial^\beta u$, we see that we can restrict to the case $\beta = 0$, $\alpha \in (0,1)$ (and u continuous). We write

$$u(x) = S_p u(x) + R_p u(x), \quad R_p u = \sum_{q \geq p} \Delta_q u.$$

We find

$$\|R_p u\|_{L^\infty} \leq \sum_{q \geq p} \|\Delta_q u\|_{L^\infty} \leq M_\alpha \sum_{q \geq p} 2^{-q\alpha} \leq C M_\alpha 2^{-p\alpha}.$$

Also,

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

But by (5.2), $\|\nabla \Delta_q u\|_{L^\infty} \leq C 2^q \|\Delta_q u\|_{L^\infty} \leq C' M_\alpha 2^{q(1-\alpha)}$. Collecting the previous bounds, we obtain:

$$|u(x) - u(y)| \leq C M_\alpha \left(|x-y| 2^{p(1-\alpha)} + 2^{-p\alpha} \right)$$

One can then optimize in p, taking for p the biggest value such that $|x-y|2^p \leq 1$. It follows that

$$|u(x) - u(y)| \leq C' M_\alpha |x-y|^\alpha$$

which concludes the proof. □

This last proposition suggests to extend the definition of the Hölder spaces through the following

Definition 5.10. *For any $\alpha \in \mathbb{R}$, we denote*

$$C^\alpha(\mathbb{R}^d) = \{u \in \mathcal{S}'(\mathbb{R}^d), \quad \sup_{p \geq -1} 2^{p\alpha} \|\Delta_p u\|_{L^\infty} < +\infty\}$$

which is a Banach space under the norm $\|u\|_{C^\alpha} = \sup_{p \geq -1} 2^{p\alpha} \|\Delta_p u\|_{L^\infty}$.

Remark 5.8. *Warning ! One can show that for $\alpha \in \mathbb{N}$, $C^\alpha(\mathbb{R}^d)$ does not coincide with the space $C_b^\alpha(\mathbb{R}^d)$. To emphasize the difference, it is customary to denote $C_*^\alpha(\mathbb{R}^d)$ the space $C^\alpha(\mathbb{R}^d)$ in the special case where $\alpha \in \mathbb{N}$.*

Exercice 5.1. (convexity inequalities)

1. *Show that if $s = \theta s_0 + (1-\theta) s_1$, with $0 \leq \theta \leq 1$, $s_0, s_1 \in \mathbb{R}$,*

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

2. *Show that If $\alpha = \theta \alpha_0 + (1-\theta)\alpha_1$, with $0 \leq \theta \leq 1$, $\alpha_0, \alpha_1 \in \mathbb{R}$,*

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

Proposition 5.11. (first Sobolev imbedding)
For all $s \in \mathbb{R}$, $H^s(\mathbb{R}^d)$ is embedded continuously in $C^{s-\frac{d}{2}}(\mathbb{R}^d)$.

Proof. Let $u \in H^s(\mathbb{R}^d)$. As a byproduct of Proposition 5.9, we know that $\Delta_p u \in L^2(\mathbb{R}^d)$. Thus, its Fourier transform is in $L^2(\mathbb{R}^d)$ and also compactly supported, so that it is in $L^1(\mathbb{R}^d)$. We write the Fourier inversion formula: for $p \geq -1$, $\Delta_p u(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{ix\cdot\xi}\widehat{\Delta_p u}(\xi)d\xi$, which implies

$$\begin{aligned}\|\Delta_p u\|_{L^\infty} &\leq \|\Delta_p u\|_{L^2}\,|B(0,C2^p)|^{1/2} \leq C' 2^{p(\frac{d}{2}-s)} \sup_p 2^{ps}\|\Delta_p u\|_{L^2}\\ &\leq C''\, 2^{p(\frac{d}{2}-s)}(\sum_p 2^{ps}\|\Delta_p u\|_{L^2}^2)^{1/2}\end{aligned}$$

By Propositions 5.9 and 5.10, we get that $\|u\|_{C^{s-\frac{d}{2}}} \leq C\|u\|_{H^s}$. □

Corollary 5.1. *For $s > \frac{d}{2}$, $H^s(\mathbb{R}^d)$ is embedded continuously in $C_b(\mathbb{R}^d)$.*

Proposition 5.12. (second Sobolev imbedding)
For $0 < s < \frac{d}{2}$, $H^s(\mathbb{R}^d)$ is continuously embedded in $L^{\frac{2d}{d-2s}}(\mathbb{R}^d)$. More precisely, there exists $C > 0$ such that

$$\|u\|_{L^{\frac{2d}{d-2s}}} \leq C\|u\|_{\dot{H}^s}, \quad \|u\|_{\dot{H}^s} = \left(\int_{\mathbb{R}^d} |\xi|^{2s}|\hat{u}(\xi)|^2\,d\xi\right)^{1/2}.$$

Remark 5.9. *The exponent $\frac{2d}{d-2s}$ can be deduced from a homogeneity argument. Indeed, suppose that we have an inequality of the type*

$$\|u\|_{L^p} \leq C\left(\int_{\mathbb{R}^d} |\xi|^{2s}|\hat{u}(\xi)|^2\,d\xi\right)^{1/2}$$

for some p. Given some non-zero u, we consider the family $u_\lambda(x) = u(\lambda x)$, $\lambda > 0$. A simple calculation shows that $\|u_\lambda\|_{L^p} = \lambda^{-d/p}\|u\|_{L^p}$, while, as $\widehat{u_\lambda} = \lambda^{-d}\,\hat{u}(\lambda^{-1}\cdot)$, we get $\|u_\lambda\|_{\dot{H}^s} = \lambda^{s-\frac{d}{2}}\|u\|_{\dot{H}^s}$. Taking u_λ in the previous inequality, we deduce

$$\lambda^{-d/p}\|u\|_{L^p} \leq C\lambda^{s-\frac{d}{2}}\|u\|_{\dot{H}^s}.$$

Sending λ to 0, resp. λ to $+\infty$, we find $-\frac{d}{p} - s + \frac{d}{2} \geq 0$, resp. $-\frac{d}{p} - s + \frac{d}{2} \leq 0$. We get $p = \frac{2d}{d-2s}$.

Remark 5.10. *In the case $s \in \mathbb{N}^*$, with the identification $H^s(\mathbb{R}^d) = W^{s,2}(\mathbb{R}^d)$ (see Proposition 5.8 and its proof), the inequality can be written equivalently as:*

$$\|u\|_{L^{\frac{2d}{d-2s}}} \leq C\sup_{|\alpha|=s}\|\partial^\alpha u\|_{L^2}. \tag{5.9}$$

Proof. [1] We denote $p = \frac{2d}{d-2s}$. We start from a famous identity for the L^p norm, namely

$$\|f\|_{L^p}^p = p\int_0^{+\infty} \lambda^{p-1} Leb(\{|f| > \lambda\})\,d\lambda.$$

[1]This proof is borrowed from the lecture notes [6], and close in spirit to the proof of the famous Marcinkiewicz interpolation theorem, see [26, 187-194].

We decompose f in low and high frequencies, setting

$$f = f_{1,A} + f_{2,A} \quad f_{1,A} = \mathbf{1}_{B(0,A)}(D)f, \quad f_{2,A} = \mathbf{1}_{B(0,A)^c}(D)f.$$

If $s < \frac{d}{2}$, $\xi \to |\xi|^{-s}\mathbf{1}_{B(0,A)}(\xi)$ belongs to L^2, so that $\widehat{f_{1,A}}(\xi) = |\xi|^{-s}\mathbf{1}_{B(0,A)}(\xi)|\xi|^s\hat{f}$ belongs to L^1, and

$$\|f_{1,A}\|_{L^\infty} \leq \frac{1}{(2\pi)^d}\|\widehat{f_{1,A}}\|_{L^1} \leq \frac{1}{(2\pi)^d}\left(\int_{B(0,A)} |\xi|^{-2s}d\xi\right)^{1/2} \|f\|_{\dot{H}_s} \leq C_s A^{\frac{d}{2}-s}\|f\|_{\dot{H}_s}.$$

Now, we notice that

$$\{|f| > \lambda\} \subset \{|f_{1,A}| > \frac{\lambda}{2}\} \cup \{|f_{2,A}| > \frac{\lambda}{2}\}$$

which implies

$$Leb(\{|f| > \lambda\}) \ \leq \ Leb(\{|f_{1,A}| > \frac{\lambda}{2}\}) + Leb(\{|f_{2,A}| > \frac{\lambda}{2}\})$$

We take $A = A_\lambda = \left(\frac{\lambda}{4C_s\|f\|_{\dot{H}_s}}\right)^{\frac{p}{d}}$, so that $Leb(\{|f_{1,A}| > \frac{\lambda}{2}\}) = 0$. Hence,

$$\|f\|_{L^p}^p = p\int_0^{+\infty} \lambda^{p-1} Leb(\{|f_{2,A}| > \frac{\lambda}{2}\})d\lambda.$$

For the high-frequency part, we use the Tchebytchev inequality, and then Plancherel identity, which leads to

$$Leb(\{|f_{2,A}| > \frac{\lambda}{2}\}) \leq 4\frac{\|f_{2,A}\|_{L^2}^2}{\lambda^2} \leq \frac{C}{\lambda^2}\|\widehat{f_{2,A}}\|_{L^2}^2$$

We finally get

$$\|f\|_{L^p}^p \leq C\int_0^{+\infty}\int_{\mathbb{R}^d} \lambda^{p-3}1_{\{|\xi|\geq A_\lambda\}}|\hat{f}(\xi)|^2 d\xi$$

Note that, by definition of A_λ, $|\xi| \geq A_\lambda$ is equivalent to $\lambda \leq 4C_s\|f\|_{\dot{H}_s}|\xi|^{\frac{d}{p}} := C_f|\xi|^{\frac{d}{p}}$. By Fubini's theorem:

$$\|f\|_{L^p}^p \leq C\int_{\mathbb{R}^d} |\hat{f}(\xi)|^2 \int_0^{C_f|\xi|^{\frac{d}{p}}} \lambda^{p-3}d\lambda d\xi \leq C' C_f^{p-2}\int_{\mathbb{R}^d} |\xi|^{\frac{d(p-2)}{p}}|\hat{f}(\xi)|^2 d\xi = C''\|f\|_{\dot{H}_s}^p.$$

□

Corollary 5.2. *$H^{d/2}(\mathbb{R}^d)$ is continuously embedded in $L^q(\mathbb{R}^d)$ for all $q \geq 2$.*

Proof. Let $s_n = \frac{d}{2} - \frac{1}{n}$. By the previous proposition H^{s_n} is continuously embedded in $L^{p_n}(\mathbb{R}^d)$ $(\cap L^2(\mathbb{R}^d))$, with $p_n = \frac{2d}{d-2s_n}$. As $p_n \to +\infty$ as $n \to +\infty$, the result follows. □

We conclude this paragraph with three lemma, that express some robustness with respect to the choice of the frequency decomposition in the Littlewood-Paley approach. The first two are concerned with decompositions of the form $u = \sum_p u_p$, with the spectrum of u_p localized in either a ball of radius $O(2^p)$, or an annulus of typical radius $O(2^p)$. We show that the Sobolev and Hölder characterizations in terms of the family $(\Delta_p u)_p$ extend at least partially to the family $(u_p)_p$. The third lemma treats the case of a decomposition $u = \sum u_p$ where u_p mimics a localization frequency, in the sense that one derivative costs $O(2^p)$.

Lemma 5.3. *Let $(u_p)_{p\geq -1}$ such that supp $\hat{u}_p \subset B(0, R2^p)$, for some $R > 0$, for all $p \geq -1$.*

- *If $\sup_p 2^{p\alpha}\|u_p\|_{L^\infty} < +\infty$ (with $\alpha > 0$), then $u = \sum u_p$ belongs to $C^\alpha(\mathbb{R}^d)$, and*

$$\|u\|_{C^\alpha} \leq C \sup_p 2^{p\alpha}\|u_p\|_{L^\infty}.$$

- *If $\sum_p 2^{2ps}\|u_p\|_{L^2}^2 < +\infty$ (with $s > 0$), then $u = \sum u_p$ belongs to $H^s(\mathbb{R}^d)$, and*

$$\|u\|_{H^s}^2 \leq C \sum_p 2^{2ps}\|u_p\|_{L^2}^2.$$

Remark 5.11. *Note the limitation $s > 0$ in the last statement.*

Proof. We focus on the second item, the first one is very similar and slighty easier. Let $q \geq -1$. By the support condition on (u_p), there exists $N > 0$ such that for all $q \geq -1$,

$$\Delta_q u = \sum_{p \geq q-N} \Delta_q u_p. \tag{5.10}$$

It follows that

$$2^{qs}\|\Delta_q u\|_{L^2} \leq 2^{qs} \sum_{p\geq q-N} \|\Delta_q u_p\|_{L^2} \leq 2^{qs} \sum_{p \geq q-N} \|u_p\|_{L^2} \quad \leq \sum_{p\geq q-N} 2^{ps}\|u_p\|_{L^2} 2^{(q-p)s} \leq \sum_{p\in\mathbb{Z}} a_p\, b_{q-p}$$

where:

- $a_p = 2^{ps}\|u_p\|_{L^2}$ for $p \geq -1$, $a_p = 0$ otherwise.
- $b_r = 2^{rs}$ for $r \leq N$, $b_r = 0$ otherwise.

As $s > 0$, $(b_r)_{r\in\mathbb{Z}}$ is sommable, the discrete Young convolution inequality yields

$$\|2^{ps}\|\Delta_p u\|_{L^2}\|_{l^2} \leq \|a_p\|_{l^2}\, \|b_p\|_{l^1}$$

so that

$$\sum_{p\geq -1} 2^{2ps}\|\Delta_p u\|_{L^2}^2 \leq C \sum_{p\geq -1} 2^{2ps}\|u_p\|_{L^2}^2$$

and we conclude thanks to Proposition 5.9. $\square$

Lemma 5.4. *Let $(u_p)_{p\geq 0}$ such that supp $\hat{u}_p \subset C(0, c2^p, C2^p)$, for some $C > c > 0$, for all $p \geq -1$.*

- *If $\sup_p 2^{p\alpha}\|u_p\|_{L^\infty} < +\infty$ (with $\alpha \in \mathbb{R}$), then $u = \sum u_p$ belongs to $C^\alpha(\mathbb{R}^d)$, and*

$$\|u\|_{C^\alpha} \leq C \sup_p 2^{p\alpha}\|u_p\|_{L^\infty}.$$

- *If $\sum_p 2^{2ps}\|u_p\|_{L^2}^2 < +\infty$ (with $s \in \mathbb{R}$), then $u = \sum u_p$ belongs to $H^s(\mathbb{R}^d)$, and*

$$\|u\|_s^2 \leq C \sum_p 2^{2ps}\|u_p\|_{L^2}^2.$$

Proof. The proof is very similar to the proof of the previous statement. The difference is that the relation (5.10) is replaced by

$$\Delta_q u = \sum_{|p-q|\leq N} \Delta_q u_p.$$

We arrive this time at the inequality

$$2^{qs}\|\Delta_q u\|_{L^2} \leq \sum_{p\in\mathbb{Z}} a_p\, b_{q-p}$$

where:

- $a_p = 2^{ps}\|u_p\|_{L^2}$ for $p \geq -1$, $a_p = 0$ otherwise.
- $b_r = 2^{rs}$ for $|r| \leq N$, $b_r = 0$ otherwise.

so that the sequence $(b_r)_{r\in\mathbb{Z}}$ is summable without constraint on the sign of s (or α). $\square$

Lemma 5.5. *Let $s > 0$, $n \in \mathbb{N}$, $n > s$. There exists $C > 0$ such that for all family u_p ($p \in \mathbb{Z}$) in $H^n(\mathbb{R}^d)$, with*

$$\|\partial^\alpha u_k\|_{L^2} \leq 2^{k(|\alpha|-s)}\varepsilon_k, \quad \textit{with } (\varepsilon_k)_{k\in\mathbb{Z}} \in l^2(\mathbb{Z}),$$

the sum $u = \sum u_k \in H^s(\mathbb{R}^d)$, and $\|u\|_{H^s}^2 \leq C\sum \varepsilon_p^2$.

Proof. We first note that $\sum f_j$ converges in L^2, so that f is well-defined as an element of L^2. We write

$$2^{js}\|\Delta_j f\|_{L^2} \leq \sum_{k\geq j} 2^{js}\|\Delta_j f_k\|_{L^2} + \sum_{k<j} 2^{js}\|\Delta_j f_k\|_{L^2}.$$

To bound the first sum, we shall use the Bernstein inequality and the assumption of the lemma, which yields

$$\|\Delta_j f_k\|_{L^2} \leq C\|f_k\|_{L^2} \leq C2^{-ks}\varepsilon_k, \quad \forall j \geq -1$$

while to bound the second sum, we shall use

$$\|\Delta_j f_k\|_{L^2} \le C2^{-nj} \sum_{|\alpha|=n} \|\partial^\alpha \Delta_j f_k\|_{L^2} \le C2^{-nj}\|f_k\|_{H^n} \le C2^{-(j-k)n}2^{-ks}\varepsilon_k, \quad \forall j \ge 0.$$

We find

$$2^{js}\|\Delta_j f\|_{L^2} \le C\sum_{k\ge j} 2^{(j-k)s}\varepsilon_k + C\sum_{k<j} 2^{(j-k)(s-n)}\varepsilon_k = C(a\star\varepsilon)_j$$

where $\star$ refers to a discrete convolution, $\varepsilon = (\varepsilon_k)_{k\in\mathbb{Z}}$, and $a = (a_l)_{l\in\mathbb{Z}}$ is defined by

$$a_l = 2^{ls} \text{ for } \ l \le 0, \quad a_l = 2^{l(s-n)} \text{ for } \ l > 0.$$

As $a \in l^1(\mathbb{Z})$, the result follows from Young's inequality.

□

5.4 Paraproduct and tame estimates

We will now use the Littlewood Paley tools to obtain a refined decomposition of a product of two functions, due to Bony [4]. Formally, the idea is to write

$$\begin{aligned} uv &= \sum_{k,j} \Delta_k u\, \Delta_j v = \sum_{j\ge 2} S_{j-2}u\Delta_j v \ + \sum_{k\ge 2} S_{k-2}v\Delta_k u + \sum_{|k-j|\le 2} \Delta_k u\, \Delta_j v \\ &= T_u v + T_v u + R(u,v) \end{aligned}$$

where

- $T_u v$ describes the part of the interaction in which the frequencies of u are much lower than the frequencies of v.
- $T_v u$ describes the part of the interaction in which the frequencies of v are much lower than the frequencies of u.
- $R(u,v)$ describes the part of the interaction in which the frequencies of u and v are comparable.

The term $T_u v$ (resp. $T_v u$) is called the paraproduct of v by u (resp. of u by v), while $R(u,v)$ is the remainder.

Remark 5.12. *One can show that, for any $u, v \in \mathcal{S}'(\mathbb{R}^d)$, the series defining $T_u v$ or $T_v u$ are converging in $\mathcal{S}'(\mathbb{R}^d)$ (see the exercise below). The third series, that has to be understood as*

$$R(u,v) = \sum_{j\ge 1}\left(\sum_{k\in\{j,j\pm 2,j\pm 1\}} \Delta_k u \Delta_j v\right)$$

is not always converging (which is consistent with the fact that the product uv is not always well-defined for two temperate distributions u and v). We shall give sufficient conditions below that ensure the convergence of the three series.

Exercice 5.2. *Let $u, v \in \mathcal{S}'(\mathbb{R}^d)$.*

1. *Show that for all $\xi \in C(0, 2^{j-1}, 2^{j+1})$, for all $\xi' \in B(0, 2^{j-2})$, for all $\alpha \in \mathbb{N}^d$, for all $K > 0$, for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$, for all $j \geq 2$,*
$$|\partial_\xi^\alpha \hat{\varphi}(\xi - \xi')| \leq C_{K,\alpha}\, 2^{-jK}.$$
2. *Deduce from the previous question that for all $u \in \mathcal{S}'(\mathbb{R}^d)$, for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$, the series $\sum_{j\geq 2} \Delta_j\,(S_{j-2} u \varphi)$ converges in $\mathcal{S}(\mathbb{R}^d)$.*
3. *Deduce that the paraproduct $T_u v$ is well-defined for all $u, v \in \mathcal{S}'(\mathbb{R}^d)$.*

The main proposition of this section is

Proposition 5.13. (estimates for paraproduct and remainder)

- *For all $s \in \mathbb{R}$, $u \in L^\infty$, $v \in H^s$,*
$$\|T_u v\|_s \leq C_s \|u\|_{L^\infty} \|v\|_s.$$
- *For all $\alpha \in \mathbb{R}$, $u \in L^\infty$, $v \in C^\alpha$,*
$$\|T_u v\|_{C^\alpha} \leq C_\alpha \|u\|_{L^\infty} \|v\|_{C^\alpha}.$$
- *For all r, s with $r + s > 0$, $u \in C^r$, $v \in H^s$,*
$$\|R(u, v)\|_{H^{r+s}} \leq C_{r,s} \|u\|_{C^r}\, \|v\|_{H^s}.$$
- *For all α, α' with $\alpha + \alpha' > 0$, $u \in C^\alpha$, $v \in C^{\alpha'}$,*
$$\|R(u, v)\|_{C^{\alpha+\alpha'}} \leq C_{\alpha,\alpha'} \|u\|_{C^\alpha} \|v\|_{C^{\alpha'}}.$$

Remark 5.13. *The bottom line of this proposition is that the regularity of $T_u v$ is mainly determined by v, while the regularities of u and v add to provide the regularity of $R(u, v)$.*

Proof. We remark that $T_u v$ has a decomposition that obeys the assumption of Lemma 5.4. Indeed, $T_u v = \sum_{p\geq 2} S_{p-2} u\, \Delta_p v$, and the Fourier transform of $S_{p-2} u \Delta_p v$ is included in $B(0, 2^{p-2}) + C(0, 2^{p-1}, 2^{p+1}) \subset C(0, c2^p, C2^p)$ for some $C > c > 0$. As regards the first item, we write:
$$\begin{aligned}\sum_p 2^{2ps} \|S_{p-2} u \Delta_p v\|_{L^2}^2 \leq \sum_p 2^{2ps} \|S_{p-2} u\|_{L^\infty}^2 \|\Delta_p v\|_{L^2}^2 \leq C \|u\|_{L^\infty}^2 \sum_p 2^{2ps} \|\Delta_p v\|_{L^2}^2 \\ \leq C' \|u\|_{L^\infty}^2 \|u\|_{H^s}^2.\end{aligned}$$
where the last inequality comes from Proposition 5.9. The inequality $\|T_u v\|_s \leq C_s \|u\|_{L^\infty} \|v\|_s$ is then deduced from the first item of Lemma 5.4. The second item of the proposition can be handled along the same lines, using this time the second item of Lemma 5.4. Eventually, the third and fourth items rely on the same kind of arguments together with Lemma 5.3. We leave the details to the reader. $\square$

Corollary 5.3. (Tame estimates for the product)

- *For any* $s > 0$*, for all* $u, v \in L^\infty \cap H^s(\mathbb{R}^d)$*,*

$$\|uv\|_s \leq C\left(\|u\|_{L^\infty}\|v\|_s + \ \|v\|_{L^\infty}\|u\|_s\right)$$

- *For any* $\alpha \in \mathbb{R}_+ \setminus \mathbb{N}$*, for all* $u, v \in C^\alpha(\mathbb{R}^d)$*,*

$$\|uv\|_{C^\alpha} \leq C\left(\|u\|_{L^\infty}\|v\|_{C^\alpha} + \ \|v\|_{L^\infty}\|u\|_{C^\alpha}\right)$$

Proof. This is a simple consequence of the estimates on the paraproduct and the remainder, taking into account that L^∞ is included in C^0_*. □

Proposition 5.14. (Tame estimate for the composition)
Let F *a* C^∞ *function on* $\mathbb{R}$*, with* $F(0) = 0$*. If* $u \in H^s(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ *with* $s \geq 0$*, then* $F(u) \in H^s(\mathbb{R}^d)$ *and*

$$\|F(u)\|_{H^s} \leq C_{s,F,\|u\|_{L^\infty}}\|u\|_{H^s}.$$

Proof. For $s = 0$, it is enough to write $F(u) = G(u)\,u$ with $G \in C^\infty(\mathbb{R})$: it follows that $\|F(u)\|_{L^2} \leq \sup\{\ |G(t)|, t \in [-\|u\|_{L^\infty}, \|u\|_{L^\infty}]\}\,\|u\|_{L^2}$.

For $s > 0$, we want to write

$$F(u) = F(S_0 u) + \sum_{p=0}^{+\infty} F(S_{p+1}u) - F(S_p u).$$

Let us show that this series converges in L^2, that is $F(S_n u) \to F(u)$ in L^2 as $n \to +\infty$. We know that $S_n u \to u$ in L^2, using for instance the almost orthogonality of the family $(\Delta_k u)$, see Lemma 5.2. Moreover, $\|S_p u\|_{L^\infty} \leq C\|u\|_{L^\infty}$. It follows that

$$\|F(S_n u) - F(u)\|_{L^2} \leq C \sup_{t\in[0,1]} \|F'(tS_n u + (1-t)u)\|_{L^\infty}\|S_n u - u\|_{L^2} \to 0.$$

Now, we write this sum as

$$F(u) = \sum_{p\geq -1} m_p\,\Delta_p u, \quad m_p = \int_0^1 F'(S_p u + t\Delta_p u)dt$$

with the convention that $S_{-1}u = 0$. For any $\beta \in \mathbb{N}^d$,

$$\|\partial^\beta(S_p u + t\Delta_p u)\|_{L^\infty} \leq C_\beta 2^{|\beta|p}\|u\|_{L^\infty}.$$

It follows by Leibnitz rule that for all $\alpha \in \mathbb{N}^d$,

$$\|\partial^\alpha m_p\|_{L^\infty} \leq C_{\alpha,F,\|u\|_{L^\infty}} 2^{|\alpha|p}$$

Using that $u \in H^s$, we also have for all α

$$\|\partial^\alpha\Delta_p u\|_{L^2} \leq C_\alpha 2^{|\alpha|(p-s)}\varepsilon_p, \quad \sum \varepsilon_p^2 = \|u\|^2_{H^s}.$$

Combining previous inequalities, we find for all $t\alpha$,

$$\|\partial^\alpha(m_p\Delta_p v)\|L^2 \leq C_{\alpha,F,\|u\|_{L^\infty}} 2^{|\alpha|(p-s)}\varepsilon_p.$$

We conclude by applying Lemma 5.5 □

5.5 Connection to the Nash-Moser methodology

This last paragraph is a very informal discussion on how the paraproducts (and more generally paradifferential calculus) can substitute to the Nash-Moser technique. We borrow here from the article [11]. For much more on paradifferential calculus, we refer to [4] and to the nice monograph [12].

As seen from Proposition 5.13, the paraproduct $T_u v$ has essentially the regularity of v: the derivatives almost commute to the paraproduct. Hence, in PDE problems with loss of derivatives, when the loss comes from the coefficients of the right inverse $g \to L(u)g$ of the linearized operator $v \to F'(u)v$, it can be very useful to substitute to a linearization a *paralinearization*. We note that this approach, that we will discuss here in the context of the isometric embedding problem, relies on the paraproduct, that is on removing certain high frequencies. To this respect, it relates to the frequency truncation in the Nash-Moser approach.

The starting point of the paralinearization approach is the following theorem, due to Bony:

Theorem 5.1. (Paralinearization Theorem)
Let $F = F(x,u) \in C^\infty(\mathbb{R}^d \times \mathbb{R}^m)$, F and its derivatives bounded on $\mathbb{R}^d \times K$ for all bounded $K \subset \mathbb{R}^m$. Assume $F(x,0) = 0$. Denote F'_u the fonction $x \to \frac{\partial}{\partial u} F(x, u(x))$.

- *If $u \in H^s(\mathbb{R}^d)$ with $\rho := s - \frac{d}{2} > 0$, then*

$$F(\cdot, u) - T_{F'_u} u \in H^{s+\rho}(\mathbb{R}^d).$$

- *If $u \in C^r(\mathbb{R}^d)$ with $r > 0$, then*

$$F(\cdot, u) - T_{F'_u} u \in C^{2r}(\mathbb{R}^d).$$

See [4] or [12] for a proof. Another property of the paraproduct that is often used is

Proposition 5.15. *Let $r > 0$. Let $u, v \in C^r(\mathbb{R}^d)$. Then,*

$$T_u T_v - T_{uv} \in L(H^s(\mathbb{R}^d), H^{s+r}(\mathbb{R}^d)) \quad \forall s \in \mathbb{R},$$
$$T_u T_v - T_{uv} \in L(C^s(/R^d), C^{s+r}(\mathbb{R}^d)) \quad \forall s \in \mathbb{R}.$$

See again [4] or [12] for a proof.

We can now introduce the notion of paradifferential operator, in the restricted context of classical differential operators. For a much wider framework, *cf* [12].

Definition 5.11. *For a differential operator of order m over $\mathbb{R}^d$, $A = \sum_{|\alpha| \le m} a_\alpha(x) \partial^\alpha$, the paradifferential operator associated to A is the operator $T_A = \sum_{|\alpha| \le m} T_{a_\alpha} \partial^\alpha$.*

For a general nonlinear partial differential operator $F(x,(\partial^\alpha u)_{|\alpha|\le m})$, with $F = F(x,U)$ satisfying the assumptions of Theorem 5.1, the application of the theorem to the function $F(x,U(x))$, $U(x) = (\partial^\alpha u)_{|\alpha|\le m}(x)$, shows the following: *for all* $u \in C^{r+m}(\mathbb{R}^d)$,

$$F(\cdot,(\partial^\alpha u)_{|\alpha|\le m})) - T_{\mathbf{F}'_u} u \in C^{2r}(\mathbb{R}^d),$$

where this time $T_{\mathbf{F}'_u}$ *is the paradifferential operator associated to*

$$\mathbf{F}'_u = \sum_{|\alpha|\le M} \frac{\partial}{\partial U_\alpha} F(x,U(x))\partial^\alpha.$$

Similarly, one can from Proposition 5.15 deduce the following result: *for two differential operators* $A = \sum_{\alpha\le m} a_\alpha \partial^\alpha$ *and* $B = \sum_{|\beta|\le m'} b_\beta \partial^\beta$*, if the coefficients* a_α, b_β *belong to* $C^r(\mathbb{R}^d)$ *with* $r > m$*, then*

$$T_A T_B - T_{AB} \in L(C^{m+m'+\sigma}(\mathbb{R}^d), C^{\sigma+r}(\mathbb{R}^d)), \quad \forall \sigma \in \mathbb{R}.$$

To apply these results and illustrate the paralinearization approach, we now come back to the isometric problem from Chapter 4. We will give indications on how to solve it, based on the article [11]. We refer to this article for all necessary additional details. As seen in Chapter 4, the heart of the proof is to be able to solve an equation of the type

$$\Phi(u_0 + u) = \Phi(u_0) + f, \quad x \in \mathbb{T}^d, \quad \Phi(u) = \left(\frac{\partial u}{\partial x_i} \cdot \frac{\partial u}{\partial x_j}\right)_{1\le i\le j\le d}. \tag{5.11}$$

Here u_0 is a free immersion, C^∞, with values in $\mathbb{R}^n$. Ideally, one would like to find for $f \in C^{r+1}(\mathbb{T}^d, \mathbb{R}^{\frac{d(d+1)}{2}})$ small enough, a solution $u \in C^{r+1}(\mathbb{T}^d, \mathbb{R}^n)$ (we will explain how to do so for any $r > 2$). As u_0 is a free immersion, one can build for $\|u\|_{C^{r+1}}$ small enough a right-inverse $g \to L(u_0+u)g$ to the linearized operator $v \to \Phi'(u_0+u)v$. We remind that $L(u_0+u)$ is a zero order operator (matrix product), with coefficients involving second order derivatives of $u_0 + u$, which is where the strongest loss of derivatives occurs.

If L was not involving any loss of derivatives, one could proceed as follows to solve (5.11). One could rewrite the equation (5.11) as

$$\Phi'(u_0)u + R(u) = f,$$

look for u under the form $u = u(g) = L(u_0)g$, and solve the equation with unknown g:

$$\Phi'(u_0)u(g) + R(u(g)) = f, \quad \text{that is } g + R(u(g)) = f.$$

Solvability would be possible for small f by the implicit function theorem.

Here, to treat the loss of derivatives, the idea is the following. We start by rewriting (5.11) as

$$T'_\Phi(u_0 + u) + R(u) = f \tag{5.12}$$

where for $u \in C^{r+1}$,

$$R(u) = \Phi(u_0 + u) - \Phi(u_0) - T'_\Phi(u_0 + u) \in C^{2r}(\mathbb{T}^d)$$

by the paralinearization result for nonlinear partial differential operators (here, the operator is first order). Note that all notions and results of this chapter were stated in $\mathbb{R}^d$, but straightforward analogues are available on $\mathbb{T}^d$. In particular, for $2r \geq r+1$, that is $r \geq 1$, the functional R sends C^{r+1} to C^{r+1}. Moreover, one can show that this functional is smooth, see [11] for details.

The next step is to look at the equation

$$u - T_{L(u_0+u)}g - L(u_0)g + T_{L(u_0)}g = 0. \tag{5.13}$$

of unknown u, with data $g \in C^{r+1}$. Using Proposition 5.13, one can check that as u_0 is C^∞, $-L(u_0)g + T_{L(u_0)}g$ is in C^σ for all σ. By the same proposition, $T_{L(u_0+u)}g$ belongs to C^{r+1}. Eventually, one can show that this equation is of the form $G(u, g) = 0$ with a smooth $G : C^{r+1} \times C^{r+1} \to C^{r+1}$, satisfying $\partial_u G(0, 0) = \mathrm{Id}$. It follows from the usual implicit function theorem that for $\|g\|_{C^{r+1}}$ small enough, (5.13) has a unique solution $u = u(g) \in C^{r+1}$.

The last step is to solve the equation

$$\Phi(u_0 + u(g)) = \Phi(u_0) + f$$

with unknown $g \in C^{r+1}$. Taking into account (5.12) and (5.13), we can write it

$$T_{\Phi'(u_0+u(g))}T_{L(u_0+u(g))}g + T_{\Phi'(u_0+u(g))}\left(L(u_0)g - T_{L(u_0)}g\right) + R(u(g)) = f$$

The second and last terms at the left-hand side belong to C^{r+1}. As regards the first term, we use the result on the commutation of paradifferential operators: as

$$\Phi'(u_0 + u)\, L(u_0 + u) = \mathrm{Id},$$

taking into account that $\Phi'(u_0 + u)$ and $L(u_0 + u)$ are respectively first order and zero order operators, and that their coefficients are respectively in C^r and C^{r-1}, we get that, for $r - 1 > 1$,

$$T_{\Phi'(u_0+u)}T_{L(u_0+u)} \in L(C^{\sigma+1}, C^{\sigma+r-1}).$$

Taking $\sigma = r > 2$, we find that $T_{\Phi'(u_0+u(g))}T_{L(u_0+u(g))}g$ belongs to C^{r+1}. Eventually, one can check that the function

$$g \to \Phi(u_0 + u(g)) - \Phi(u_0)$$

(which sends C^{r+1} to C^{r+1} as we just established), is smooth and that its differential at 0 is 0. We conclude by the implicit function theorem.

Chapter 6

The regularity theorem of De Giorgi for scalar elliptic equations

We will present here a celebrated result on the regularity of weak solutions to scalar elliptic equations. This result was established independently by De Giorgi [7] and Nash [19], and revisited later by Moser [17]. We shall present here the approach of De Giorgi. Our main references for this chapter are [20] and [5]. We shall emphasize some proximity with the arguments encountered in Chapter 2, used to overcome losses of derivatives.

The PDEs under consideration are diffusion equations of the form:

$$-\text{div } (A\nabla u) = f, \tag{6.1}$$

or in coordinates

$$-\sum_{i,j} \partial_i(a_{ij}\partial_j u) = f.$$

Such equations are considered in a domain (that is an open and connected set) Ω of $\mathbb{R}^d$, $d \geq 2$:

$$u, f : \Omega \to \mathbb{R}, \quad A = (a_{ij})_{1\leq i,j\leq d} : \Omega \to M_d(\mathbb{R}).$$

The diffusion matrix satisfies the following fundamental assumptions:

(H1) (L^∞ bound): $a_{ij} \in L^\infty(\Omega)$ for all $1 \leq i, j \leq d$.

(H2) (uniform ellipticity): There exists $\lambda > 0$ such that for a.e. $x \in \Omega$, and all $\xi \in \mathbb{R}^d$,

$$A(x)\xi \cdot \xi \geq \lambda \xi \cdot \xi.$$

The analysis of equation (6.1) has a long history, starting from the study of harmonic functions $\Delta u = 0$ and of the Poisson equation $-\Delta u = f$. As regards harmonic functions, it is a well-known result that if a distribution $u \in \mathcal{D}'(\Omega)$ satisfies $\Delta u = 0$ in Ω, it is in fact smooth (and even analytic). This in turn allows to establish various regularity results for the Poisson equation. As an example, we state

Proposition 6.1. (Hölder regularity for Poisson equation)
Let $\alpha \in (0,1)$, Ω an open set of $\mathbb{R}^d$. If f is of class C^α on Ω, any distributional solution u of $-\Delta u = f$ in Ω is of class $C^{2+\alpha}$ on Ω.

Proof. Let $x_0 \in \Omega$, and $\chi \in C_c^\infty(\Omega)$ such that $\chi = 1$ near x_0. Let u_χ the solution in $\mathcal{S}'(\mathbb{R}^d)$ of

$$-\Delta u_\chi = \chi f$$

which is obtained as the convolution of the fundamental solution and χf (seen as a function over $\mathbb{R}^d$ after extension by zero outside Ω). We note that $u - u_\chi$ is harmonic near x_0, so that it is smooth near x_0. As $\Delta_{-1} u_\chi$ is also smooth, it is then enough to show that $\sum_{p\ge 0} \Delta_p u_\chi$ belongs to $C^{2+\alpha}(\mathbb{R}^d)$, given that $\chi f \in C^\alpha(\mathbb{R}^d)$ (after extension by zero outside Ω). As

$$-\Delta \Delta_p u_\chi = -\Delta_p \Delta u_\chi = -\Delta_p(\chi f)$$

we deduce from (5.3) that for all $p \ge 0$,

$$2^{2p} \|\Delta_p u_\chi\|_{L^\infty} \ \le\ C \|\Delta_p(\chi f)\|_{L^\infty} \le C' 2^{-p\alpha} \|\chi f\|_{C^\alpha}$$

The conclusion follows from Lemma 5.3 □

How is the situation modified in the case of heterogeneous diffusions (that is of non-constant matrices) ? *When the coefficients a_{ij} have $C^{0,\alpha}$ regularity*, one can show through a perturbative approach that the conclusion for constant coefficients equations remains: if the data has Hölder continuity C^α, the solution of (6.1) has regularity $C^{2+\alpha}$. Such result, together with a complete existence theory and set of estimates, was established by J. Schauder in the 1930's. For a description of *Schauder estimates*, we refer the reader to the textbook [9, chapter 6]. Still, with regards to applications, a Hölder continuity assumption on the coefficients is restrictive, and one would like to have some existence and regularity theory under the mere assumptions (H1) and (H2). We remind in the next chapter the standard notion of weak solutions, before turning to the De Giorgi-Nash-Moser theorem on the Hölder regularity of weak solutions.

6.1 Weak solutions of elliptic equations

In this whole section, Ω denotes a domain of $\mathbb{R}^d$ (open connected set). The Sobolev space $W^{s,p}(\Omega)$ was introduced in Definition 5.8. For an extensive study of Sobolev spaces (approximation by smooth functions, extension to the whole space, traces, compactness...), we refer to [8, chapter 5]. We shall focus here on the space $H^1(\Omega) = W^{1,2}(\Omega)$. We shall notably use

Theorem 6.1. (Rellich compactness theorem)
If Ω is of class C^1 and bounded, the embedding of $H^1(\Omega)$ into $L^2(\Omega)$ is compact: any bounded sequence in $H^1(\Omega)$ has a subsequence that converges in $L^2(\Omega)$.

Remark 6.1. *Let (u_n) be a bounded sequence in $H^1(\Omega)$, strongly converging in $L^2(\Omega)$ to some u. This sequence being bounded in the Hilbert space $H^1(\Omega)$, it also has a subsequence that converges weakly in $H^1(\Omega)$ to some $\tilde{u}$. By comparison of the limits in the distributional sense, we obtain $\tilde{u} = u$. Hence, combining Rellich theorem and this argument, we see that one can extract from a bounded sequence in $H^1(\Omega)$ a subsequence that converges strongly in L^2 and weakly in H^1 to some u.*

For later use, we also single out the following composition result (see [8, chapter 5, problem 16]):

Proposition 6.2. *Let $G \in C^1(\mathbb{R})$ such that $G(0) = 0$ and G' bounded. For all $u \in H^1(\Omega)$, $G(u) \in H^1(\Omega)$, with for all $1 \le i \le d$, $\partial_i G(u) = G'(u)\partial_i u$.*

Exercice 6.1. ([8, chapter 5, problem 17])
Let $u \in H^1(\Omega)$, $u^+ = \max(u, 0)$, $u^- = \max(-u, 0)$. Show that $u^\pm \in H^1(\Omega)$,

$$\nabla u^+ = \mathbf{1}_{\{u>0\}} \nabla u\,, \quad \nabla u^- = -\mathbf{1}_{\{u<0\}} \nabla u.$$

Hint: use the previous proposition with the function $F^\varepsilon(z) = (z^2 + \varepsilon^2)^{1/2} - \varepsilon$ for $z > 0$, $F^\varepsilon(z) = 0$ for $z < 0$.

In the resolution of elliptic equations, a crucial role is played by a subspace of $H^1(\Omega)$. Namely, one denotes by $H^1_0(\Omega)$ the closure of $C^\infty_c(\Omega)$ in $H^1(\Omega)$. A key property of the space $H^1_0(\Omega)$ is the so-called

Proposition 6.3. (Poincaré inequality) *If Ω is bounded, there exists $C > 0$ (that can be taken proportional to the diameter of Ω) such that for all $u \in H^1_0(\Omega)$:*

$$\|u\|_{L^2} \le C \|\nabla u\|_{L^2}.$$

Hence, $\|u\|_{H^1_0(\Omega)} := \left(\int_\Omega |\nabla u|^2\right)^{1/2}$ is a norm on $H^1_0(\Omega)$, equivalent to the usual one.

Remark 6.2. *When Ω is C^1 and bounded (so that Rellich theorem applies), one can prove the inequality*

$$\|u\|_{L^2} \le C \|\nabla u\|_{L^2}, \quad C = C(\Omega), \forall u \in V$$

for V any closed subspace of $H^1(\Omega)$ such that the only constant function in V is zero. This applies for instance to the functions of $H^1(\Omega)$ that have zero mean. See also Lemma 6.3

It is easily seen that, given $u \in H^1_0(\Omega)$, its extension $\overline{u}$ by 0 outside Ω belongs to $H^1(\mathbb{R}^d)$. This allows to apply the results of Chapter 5, notably the results on the Sobolev embeddings. We find (combined with Poincaré inequality)

Proposition 6.4. (Sobolev embeddings)

- *If $d > 2$, the space $H_0^1(\Omega)$ embedds continuously in $L^{\frac{2d}{d-2}}(\Omega)$, with*
$$\|u\|_{L^{\frac{2d}{d-2}}} \le C\|\nabla u\|_{L^2}$$
for some C that depends on d and Ω (and can be taken proportional to the diameter of Ω).
- *If $d = 2$, $H_0^1(\Omega)$ embedds continuously in $L^q(\Omega)$ for all $q \in [2, +\infty)$ ($q \in [1, +\infty)$ for bounded Ω), with*
$$\|u\|_{L^q} \le C\|\nabla u\|_{L^2}$$

Remark 6.3. *In the whole space case, the second embedding involves the non-homogeneous H^1 norm of u, not only the L^2 norm of the gradient. But in the case of $H_0^1(\Omega)$, the Poincaré inequality allows to replace the H^1 norm by the H_0^1 norm.*

Remark 6.4. *These embeddings still hold when $H_0^1(\Omega)$ is replaced wth $H^1(\Omega)$, under the assumption that Ω is of class C^1, and if the H_0^1 norm at the right-hand side is replaced by the (inhomogeneous) H^1 norm.*

The space $H_0^1(\Omega)$ arises naturally in the notion of weak solutions for Poisson type systems:

$$-\mathrm{div}\ (A\nabla u) = f \quad \text{in } \Omega, \quad u|_{\partial\Omega} = 0. \tag{6.2}$$

Definition 6.1. (weak solution of the Poisson equation)
For any $f \in H^{-1}(\Omega) = (H_0^1(\Omega))'$, we say that $u \in H_0^1(\Omega)$ is a weak solution of (6.2) *if:*

$$\int_\Omega \nabla u \cdot \nabla \varphi = f(\varphi), \quad \forall \varphi \in H_0^1(\Omega). \tag{6.3}$$

Note that if Ω is bounded and if $f \in L^q(\Omega)$, with $q \ge \frac{2d}{d+2}$ if $d > 2$ or $q > 1$ if $d = 2$, then f can be seen as an element of $H^{-1}(\Omega)$ through the mapping $\varphi \to \int_\Omega f\varphi$. For instance, in the case $d > 2$:

$$|\int_\Omega f\varphi| \le \|f\|_{L^{\frac{2d}{d+2}}(\Omega)}\|\varphi\|_{L^{\frac{2d}{d-2}}(\Omega)} \le C\|f\|_{L^{\frac{2d}{d+2}}(\Omega)}\|f\|_{H_0^1(\Omega)} \le C'\|f\|_{L^q(\Omega)}\|f\|_{H_0^1(\Omega)}$$

where we have used Proposition 6.4 for the second inequality.

The notion of weak solution is extremely convenient, as it provides an easy framework for unique solvability of (6.2). This is the sense of

Proposition 6.5. (Existence and uniqueness of weak solutions)
Let Ω be a bounded domain. For any $f \in H^{-1}(\Omega)$, there is a unique weak solution of (6.2).

Proof. This is a consequence of the following abstract result:

Lemma 6.1. (Lax-Milgram Lemma) *Let H be a Hilbert space, $a : H \times H \to \mathbb{R}$ bilinear, continuous, and coercive in the sense that $a(u, u) \ge \alpha u$ for some $\alpha > 0$, for all $u \in H$. Then, for all $f \in H'$, there exists a unique solution $u \in H$ of $a(u, v) = f(v)$.*

For a proof of this lemma, see [8, chapitre 6]. The theorem follows, with $H = H_0^1(\Omega)$, $a(u, v) = \int_\Omega A(x)\nabla u(x) \cdot \nabla v(x)dx$. ☐

6.2 Regularity of the weak solution

A natural question is: how regular is the weak solution provided by Theorem 6.5 ? We focus here on interior regularity (regularity up to the boundary would require more assumptions on Ω). The main goal of this chapter is proving Hölder regularity. We introduce, for an open set $\tilde{\Omega}$, the norm

$$\|u\|_{C^\alpha(\tilde{\Omega})} = \|u\|_{L^\infty(\tilde{\Omega})} + \sup_{x \neq y \in \tilde{\Omega}} \frac{|u(x) - u(y)|}{|x - y|^\alpha}.$$

Theorem 6.2. [7, 19]
Suppose that $A = (a_{ij})$ satisfies assumptions (H1)-(H2). Let $q > \frac{d}{2}$, $f \in L^q(\Omega)$. Suppose that $u \in H^1(\Omega)$ satisfies

$$\int_\Omega A\nabla u \cdot \nabla\varphi = \int_\Omega f\varphi, \quad \forall \varphi \in H^1_0(\Omega).$$

Then, there exists $\alpha = \alpha(d, \lambda, \|A\|_{L^\infty}, q) \in (0,1)$ such that for all domain $\tilde{\Omega} \Subset \Omega$,

$$\|u\|_{C^\alpha(\tilde{\Omega})} \leq C\left(\|f\|_{L^q(\Omega)} + \|u\|_{H^1(\Omega)}\right), \tag{6.4}$$

where C depends on $d, \lambda, \|A\|_{L^\infty}, q$ and $\tilde{\Omega}$.

Note that by a covering argument, it is enough to establish the bound

$$\|u\|_{C^\alpha(B_r(x_0))} \leq C\left(\|f\|_{L^q(B_{4r}(x_0))} + \|u\|_{H^1(B_{4r}(x_0))}\right),$$

for any $x_0 \in \Omega$ and $r > 0$ such that $B(x_0, 4r) \Subset \Omega$, where C is allowed to depend on $d, \lambda, \Lambda, q, r$. Furthermore, by dilation and translation, we can always assume $x_0 = 0$, $r = \frac{1}{4}$. Eventually, it is enough to prove the theorem in the special case $\Omega = B(0,1)$, $\tilde{\Omega} = B(0, 1/4)$.

6.2.1 First step: from L^2 to L^∞.

Proposition 6.6. *Suppose that $A = (a_{ij})$ is bounded measurable in $B(0,1)$ and satisfies for some $\lambda > 0$:*

$$\lambda\xi \cdot \xi \leq A(x)\xi \cdot \xi, \quad \textit{for almost every } x \in B(0,1), \quad \forall \xi \in \mathbb{R}^d.$$

Let $q > \frac{d}{2}$, $f \in L^q(B(0,1))$ and $u \in H^1(B(0,1))$ a subsolution of the Poisson equation in B_1, namely,

$$\int_{B(0,1)} A\nabla u \cdot \nabla\varphi \leq \int_\Omega f\varphi \quad \forall \varphi \in H^1_0(B(0,1)),\ \varphi \geq 0 \ \textit{ in } \ B(0,1)$$

Then $u^+ \in L^\infty(B(0,1/2))$, with

$$\|u^+\|_{L^\infty(B(0,1/2))} \leq C\left(\|f\|_{L^q(B(0,1))} + \|u^+\|_{L^2(B(0,1))}\right)$$

for a constant C that depends only on $d, q, \lambda, \|A\|_{L^\infty}$.

Proof. We follow here the approach by De Giorgi and assume $d \geq 3$ (the case $d = 2$ resuires very minor modifications). We consider

- a sequence of balls $B_k = B(0, r_k)$ where $r_0 = 1$ and r_k is decreasing towards $\frac{1}{2}$.
- a sequence of smooth functions $(\chi_k)_{k\geq 1}$, with values in $(0,1)$, such that $\chi_k = 1$ in B_k, $\chi_k = 0$ in B_{k-1}^c, with $|\nabla \chi_k| \leq \frac{C}{r_k - r_{k-1}}$.
- a sequence of increasing constants C_k going to 1.

The goal is to show that there exists $\delta > 0$ such that

$$\int_{B(0,1/2)} ((u-1)^+)^2 = \lim_{k\to+\infty} \int_{B(0,1)} \chi_k((u - C_k)^+)^2 = 0$$

under the assumption $\|f\|_{L^q(B(0,1))} + \|u^+\|_{L^2(B(0,1))} \leq \delta$. Then the general case follows by considering $(u_\delta, f_\delta) = \delta \frac{(u,f)}{\|f\|_{L^q(B(0,1))} + \|u^+\|_{L^2(B(0,1))}}$.

We denote $u_k = (u - C_k)^+$. The first ingredient of the proof is sometimes called Cacciopoli inequality. Taking $\varphi = \chi_{k+1}^2 u_{k+1}$ as a test function, we find

$$\int_{B(0,1)} A\nabla u \cdot \nabla \left(\chi_{k+1}^2 u_{k+1}\right) \leq \int_\Omega f \chi_{k+1}^2 u_{k+1}$$

Expanding the gradient at the left-hand side, we find

$$\int_{B(0,1)} \chi_{k+1}^2 A\nabla u \cdot \nabla u_{k+1} \leq \int_\Omega f \chi_{k+1}^2 u_{k+1} - 2\int_{B(0,1)} (A\nabla u \cdot \nabla \chi_{k+1})\chi_{k+1} u_{k+1}$$

We note that $A\nabla u \cdot \nabla u_{k+1} = A\nabla u_{k+1} \cdot \nabla u_{k+1} \geq \lambda |\nabla u_{k+1}|^2$, and that

$$\begin{aligned}
&-2\int_{B(0,1)} (A\nabla u \cdot \nabla \chi_{k+1})\chi_{k+1} u_{k+1} = -2\int_{B(0,1)} (A\nabla u_{k+1} \cdot \nabla \chi_{k+1})\chi_{k+1} u_{k+1} \\
&\leq \frac{\lambda}{2}\int_{B(0,1)} \chi_{k+1}^2 |\nabla u_{k+1}|^2 + C\int_{B(0,1)} |\nabla \chi_{k+1}|^2 |u_{k+1}|^2 \\
&\leq \frac{\lambda}{2}\int_{B(0,1)} \chi_{k+1}^2 |\nabla u_{k+1}|^2 + \frac{C'}{(r_k - r_{k+1})^2}\int_{B(0,1)} |\chi_k u_k|^2
\end{aligned}$$

where C, C' depend on λ and $\Lambda = \|A\|_{L^\infty}$. We have used here that $\chi_k = 1$ on the support of $\nabla \chi_{k+1}$ and that $u_{k+1} \leq u_k$. Hence,

$$\int_{B(0,1)} \chi_{k+1}^2 |\nabla u_{k+1}|^2 \leq C\left(\frac{1}{(r_k - r_{k+1})^2}\int_{B(0,1)} |\chi_k u_k|^2 + \int_\Omega f \chi_{k+1}^2 u_{k+1}\right).$$

This implies

$$\int_{B(0,1)} |\nabla(\chi_{k+1} u_{k+1})|^2 \leq C'\left(\frac{1}{(r_k - r_{k+1})^2}\int_{B(0,1)} |\chi_k u_k|^2 + \int_\Omega f \chi_{k+1}^2 u_{k+1}\right).$$

The second ingredient of the proof is the Sobolev imbedding. More precisely, we write, for $p = \frac{2d}{d-2}$:

$$\begin{aligned}
\int_\Omega f\chi_{k+1}^2 u_{k+1} \;&\leq\; \|f\|_{L^q}\, Leb(\{\chi_{k+1}u_{k+1} > 0\})^{1-\frac{1}{p}-\frac{1}{q}}\, \|\chi_{k+1}u_{k+1}\|_{L^p} \\
&\leq\; C\,\|f\|_{L^q(B(0,1))}\, Leb(\{\chi_{k+1}u_{k+1} > 0\})^{1-\frac{1}{p}-\frac{1}{q}}\, \|\nabla(\chi_{k+1}u_{k+1})\|_{L^2} \\
&\leq\; \delta\,\|\nabla(\chi_{k+1}u_{k+1})\|_{L^2}^2 \;+\; C_\delta\,\|f\|_{L^q(B(0,1))}^2\, Leb(\{\chi_{k+1}u_{k+1} > 0\})^{2-\frac{2}{p}-\frac{2}{q}}
\end{aligned}$$

We also have

$$\begin{aligned}
\int_{B(0,1)} |\chi_{k+1}u_{k+1}|^2 &\leq C \left(\int_{B(0,1)} |\chi_{k+1}u_{k+1}|^p \right)^{2/p} Leb(\{\chi_{k+1}u_{k+1} > 0\})^{2/d} \\
&\leq C \left(\int_{B(0,1)} |\nabla(\chi_{k+1}u_{k+1})|^2 \right) Leb(\{\chi_{k+1}u_{k+1} > 0\})^{2/d}
\end{aligned}$$

Combining, we get

$$\begin{aligned}
\int_{B(0,1)} |\chi_{k+1}u_{k+1}|^2 \leq C\Big(&\frac{1}{(r_k - r_{k+1})^2} \Big(\int_{B(0,1)} |\chi_k u_k|^2 \Big)\, Leb(\{\chi_{k+1}u_{k+1} > 0\})^{2/d} \\
&+ \|f\|_{L^q}^2\, Leb(\{\chi_{k+1}u_{k+1} > 0\})^{1+\varepsilon} \Big)
\end{aligned}$$

where $\varepsilon = 1 - \frac{2}{p} - \frac{2}{q} + \frac{2}{d} > 0$. To evaluate the Lebesgue measure of the set $\{\chi_{k+1}u_{k+1} > 0\}$, we use the fact that

$$\chi_{k+1}u_{k+1} > 0 \quad \Rightarrow \quad \chi_k u_k > C_{k+1} - C_k$$

which by Bienaimé-Tchebytchev yields

$$Leb(\{\chi_{k+1}u_{k+1} > 0\}) \leq \frac{1}{(C_{k+1} - C_k)^2} \int_{B(0,1)} |\chi_k u_k|^2.$$

Setting $U_k = \int_{B(0,1)} |\chi_k u_k|^2$, we end up with the inequality

$$U_{k+1} \leq C\left(\frac{1}{(r_k - r_{k+1})^2 (C_{k+1} - C_k)^{2/d}} U_k^{1+2/d} \;+\; \frac{1}{(C_{k+1} - C_k)^{2(1+\varepsilon)}} \|f\|_{L^q} U_k^{1+2\varepsilon} \right) \tag{6.5}$$

Taking $r_k = \frac{1}{2} + \frac{1}{2^{k+1}}$, $C_k = 1 - \frac{1}{2^k}$, we get for some fixed N:

$$U_{k+1} \leq C 2^{Nk} (U_k^{1+2/d} + \|f\|_{L^q}^2 U_k^{1+2\varepsilon}) \tag{6.6}$$

From there, one can show by induction that for $\|f\|_{L^q} \leq 1$, for δ and $\varepsilon > 0$ small enough and $U_0 = \int_{B(0,1)} |u^+|^2 \leq \delta$, one has $U_k \leq \delta^{(1+\varepsilon)^k}$ for all k. This implies that $U_k \to 0$ as $k \to +\infty$, which ends the proof. $\square$

Corollary 6.1. *Under the same assumptions on A and f as in the previous proposition, if $u \in H^1(B(0,1))$ satisfies*

$$\int_{B(0,1)} A \cdot \nabla u \cdot \nabla \varphi = \int_{B(0,1)} f\varphi, \quad \forall \varphi \in H_0^1(B(0,1))$$

then $u \in L^\infty(B(0,1/2))$ and

$$\|u\|_{L^\infty(B(0,1/2))} \leq C \left(\|f\|_{L^q(B(0,1))} + \|u\|_{L^2(B(0,1))}\right)$$

for a constant C that depends only on $d, q, \lambda, \|A\|_{L^\infty}$.

Remark 6.5. *There is a strong analogy between the method used to prove Proposition 6.6 and the one encountered in Chapter 2. In the former, the loss of derivatives was resulting in negative powers of δ, where δ corresponded to a loss in the radius of analyticity. Here, the loss is due to the shrinking of the domain on which regularity estimates are obtained (negative powers of $r_k - r_{k+1}$), and on negative powers of $C_{k+1} - C_k$. In both contexts, what compensates this loss is the nonlinear nature of the iterative process. Note that in the proof of Proposition 6.6, this nonlinearity is made possible thanks to the use of the Sobolev imbedding, despite the linear nature of the equation.*

6.2.2 Hölder regularity with no source

We now turn to the proof of De Giorgi-Nash-Moser theorem. We first focus on the homogeneous case, namely on solution of

$$-\text{div } (A\nabla u) = 0. \tag{6.7}$$

Definition 6.2. *Let $\Omega \subset \mathbb{R}^d$ an open set. $u \in H^1(\Omega)$ is a subsolution (resp. supersolution) of* (6.7) *in Ω if*

$$\int_\Omega A\nabla u \cdot \nabla \varphi \leq 0 \ \ (\textit{resp.} \geq 0) \ \ \forall \varphi \in H_0^1(\Omega), \quad \varphi \geq 0 \ \ \textit{in} \ \ \Omega$$

Lemma 6.2. *Let $\phi \in C^1(\mathbb{R})$ a convex function with $\phi(0) = 0$ and ϕ' bounded. Let $u \in H^1(B(0,1))$. Then,*

i) If u is a subsolution and $\phi' \geq 0$, then $v = \phi(u)$ is also a subsolution.

ii) If u is a supersolution and $\phi' \leq 0$, then $v = \phi(u)$ is a subsolution.

Remark 6.6. *One could replace $B(0,1)$ by any smooth bounded domain.*

Proof. Note that by Proposition 6.2, $\phi(u) \in H^1(B(0,1))$. We prove the first point, in the case $\phi \in C_b^2(\mathbb{R})$: the general case follows by considering $\phi_\varepsilon = \phi \star \rho_\varepsilon$ with ρ_ε a standard mollifier and passing to the limit into the relation

$$\int_{B(0,1)} A\nabla \phi_\varepsilon(u) \cdot \nabla \varphi = \int_{B(0,1)} \phi_\varepsilon'(u) \nabla u \cdot \nabla \varphi \leq 0.$$

Let $\varphi \in C_c^\infty(B(0,1))$ with $\varphi \geq 0$. We compute

$$\begin{aligned}\int_{B(0,1)} ADv \cdot D\varphi &= \int_{B(0,1)} \phi'(u)\, ADu \cdot D\varphi \\ &= \int_{B(0,1)} ADu \cdot D\left(\phi'(u)\varphi\right) - \int_{B(0,1)} ADu \cdot Du\phi''(u)\varphi \leq 0\end{aligned}$$

where the first term is non-positive as u is a subsolution, and the second term is non-positive by the convexity of Φ. Note that the test function $\phi'(u)\varphi$ in the second integral belongs to $H_0^1(B_1)$, as $\phi'(u) \in H^1(B(0,1))$ (see Proposition 6.2) and $\varphi \in C_c^\infty(B(0,1))$.

To conclude, we leave to the reader to show that any non-negative function φ in $H_0^1(B(0,1))$ is the limit in H^1 of a sequence (φ_n) of non-negative functions in $C_c^\infty(B(0,1))$. By the previous lines, for all n, $\int_{B(0,1)} ADv \cdot D\varphi_n \leq 0$, so that $\int_{B(0,1)} ADv \cdot D\varphi \leq 0$. □

We now state a variation of Poincaré inequality:

Lemma 6.3. *For any $\varepsilon > 0$, there exists $C = C(\varepsilon, d)$ such that, for all $u \in H^1(B(0,1))$ with $Leb(\{u = 0\}) \geq \varepsilon$, we have*

$$\int_{B(0,1)} |u|^2 \leq C \int_{B(0,1)} |\nabla u|^2$$

Remark 6.7. *One could replace $B(0,1)$ by any smooth bounded domain.*

Proof. If the statement does not hold, one can find $\varepsilon > 0$ and a sequence (u_n) in H^1 such that

$$Leb(\{u_n = 0\}) \geq \varepsilon, \quad \|u_n\|_{L^2} = 1, \quad \|\nabla u_n\|_{L^2} = O(1/n).$$

By the Rellich theorem, see [8], one can up to a subsequence assume that u_n converges weakly in H^1 and strongly in L^2 to some $u \in H^1$. By the gradient bound, we find that $\nabla u = 0$, so that $u = \overline{u}$ is constant over $B(0,1)$. As $\|u\|_{L^2} = \lim_n \|u_n\|_{L^2} = 1$, the constant $\overline{u}$ is non-zero. Eventually, we get

$$\varepsilon \leq Leb(\{u_n = 0\}) \leq Leb(\{|u_n - \overline{u}| \geq |\overline{u}|\}) \leq \frac{1}{|\overline{u}|^2} \int_{B(0,1)} |u_n - \overline{u}|^2$$

where the last inequality is the Bienaimé-Tchebytchev inequality. As the right-hand side goes to zero when n goes to infinity, we reach a contradiction. □

We now prove:

Proposition 6.7. *Suppose that u is a non-negative supersolution in $B(0,2)$, and*

$$Leb(\{x \in B(0,1),\ u \geq 1\}) \geq \varepsilon, \quad \varepsilon > 0.$$

Then there exists a constant $c > 0$ depending on ε, d, λ, $\Lambda = \|A\|_{L^\infty}$ such that

$$\inf_{B(0,1/2)} u \geq c.$$

Proof. We may assume that $u \geq \delta > 0$, then the result for any non-negative u can be obtained by considering $u_\delta = u + \delta$ and letting δ go to 0^+.

We then consider $v = (\ln(u))^- = (-\ln u)^+$. We claim that v a non-negative subsolution of (6.7) in $B(0,2)$. This can be shown by considering the function $F_\varepsilon \circ (-\ln u)$ where F_ε was defined in Exercice 6.1, applying Proposition 6.2 and sending ε to 0. Moreover, we have

$$Leb(\{x \in B(0,1),\ u \geq 1\}) = Leb(\{x \in B(0,1),\ v = 0\})$$

From the previous proposition, we can apply the Poincaré inequality:

$$\int_{B(0,1)} |v|^2 \leq C \int_{B(0,1)} |\nabla v|^2$$

Moreover, we can also apply Proposition 6.6 to get

$$\|v\|_{L^\infty(B(0,1/2))} \leq C \|v\|_{L^2(B(0,1))}.$$

Combining both, we get

$$\|v\|_{L^\infty(B(0,1/2))} \leq C \int_{B(0,1)} |\nabla v|^2.$$

To prove that the right-hand side is bounded uniformly in v, we take $\varphi = \frac{\chi^2}{u}$, χ smooth and compactly supported in $B(0,2)$ in the definition of the supersolution. We find:

$$0 \leq \int_{B(0,2)} A\nabla u \cdot \nabla \left(\frac{\chi^2}{u}\right) = -\int_{B(0,1)} \frac{\chi^2}{u^2} A\nabla u \cdot \nabla u + 2\int_{B(0,2)} A\nabla u \cdot \nabla \chi\, \chi\, u$$

which implies after a few manipulations that

$$\int_{B(0,2)} \chi^2 |\nabla \ln(u)|^2 \leq C \int_{B(0,2)} |\nabla \chi|^2$$

Taking $\chi = 1$ in a vicinity of $B(0,1)$, we deduce

$$\int_{B(0,1)} |\nabla \ln(u)|^2 \leq C$$

Eventually, $\|v\|_{L^\infty(B(0,1/2))} \leq C$, so that $\inf_{B(0,1/2)} u \geq c = e^{-C}$. $\square$

Proposition 6.8. (Control of the oscillations)
Assume that $u \in H^1(B(0,2)) \cap L^\infty(B(0,2))$ satisfies

$$\int_{B(0,2)} A\nabla u \cdot \nabla \varphi = 0, \quad \forall \varphi \in H^1_0(B(0,2)).$$

Then, there exists a $\gamma \in (0,1)$ depending on d, λ, $\|A\|_{L^\infty}$ such that

$$osc_{B(0,1/2)} u \leq \gamma\, osc_{B(0,2)} u$$

Proof. We set $\alpha = \inf_{B(0,2)} u$, $\beta = \sup_{B(0,2)} u$, and $\alpha' = \inf_{B(0,1/2)} u$, $\beta' = \sup_{B(0,1/2)} u$. Note the following equivalences:

$$u \geq \frac{1}{2}(\alpha+\beta) \Leftrightarrow 2\frac{u-\alpha}{\beta-\alpha} \geq 1$$
$$u \leq \frac{1}{2}(\alpha+\beta) \Leftrightarrow 2\frac{\beta-u}{\beta-\alpha} \geq 1$$

Hence, there are two cases:

1. $Leb(\{x \in B(0,1), 2\frac{u-\alpha}{\beta-\alpha} \geq 1\} \geq \frac{1}{2}Leb(B(0,1))$. Then we apply the previous proposition to $2\frac{u-\alpha}{\beta-\alpha}$ which is a non-negative supersolution in $B(0,2)$. It follows that: for some $C > 1$,
$$\inf_{B(0,1/2)} 2\frac{u-\alpha}{\beta-\alpha} > \frac{1}{C},$$
which amounts to
$$\alpha' \geq \alpha + \frac{1}{2C}(\beta-\alpha)$$
so that $\beta' - \alpha' \leq (1 - \frac{1}{2C})(\beta-\alpha)$.

2. $Leb(\{x \in B(0,1), 2\frac{\beta-u}{\beta-\alpha} \geq 1\}) \geq \frac{1}{2}Leb(B(0,1))$. This case can be handled similarly.

□

Corollary 6.2. *Theorem 6.2 holds in the case $f = 0$.*

Proof. As explained after the statement of Theorem 6.2, it is enough to consider the case $\tilde{\Omega} = B(0,1/4)$, $\Omega = B(0,1)$. By Corollary (6.1), we already have the L^∞ control (remind that $f = 0$ here):
$$\|u\|_{L^\infty(B(0,1/2))} \leq C\|u\|_{L^2(B(0,1))}.$$
We then consider, for an arbitrary $x_0 \in B(0,1/4)$:

$$u_1(y) = u(x_0 + y/8), \quad u_{n+1}(y) = u_n(y/8) = u(x_0 + y/8^{n+1}).$$

It is easy to see that for all $n \geq 1$, $u_n \in H^1(B(0,2)) \cap L^\infty(B(0,2))$ (u being bounded in $B(0,1/2)$) and that it satisfies

$$\int_{B(0,2)} A_n \nabla u_n \cdot \nabla\varphi = 0, \quad \forall \varphi \in H_0^1(B(0,2))$$

where $A_n(y) = A(x_0 + \frac{y}{8^n})$ satisfies the same assumptions as the original A. We can apply the previous proposition, that yields

$$\mathrm{osc}_{B(0,1/2)} u_{n+1} \leq \gamma \mathrm{osc}_{B(0,2)} u_{n+1} \leq \gamma \mathrm{osc}_{B(0,1/4)} u_n \leq \gamma \mathrm{osc}_{B(0,1/2)} u_n.$$

This leads to

$$\mathrm{osc}_{B(0,1/2)} u_n \leq \gamma^n \mathrm{osc}_{B(0,1/2)} u_1$$

which implies

$$\sup_{|x-x_0|\leq \frac{1}{8^n}} |u(x) - u(x_0)| \leq 2\gamma^n \|u\|_{L^\infty(B(0,1/2))}.$$

We claim that this relation implies that there exists $C > 0$ such that for all $x \in B(0, 1/4)$,

$$|u(x) - u(x_0)| \ \leq C|x - x_0|^\alpha \|u\|_{L^\infty(B(0,1/2))}, \qquad \text{with } \ \alpha = -\frac{\ln \gamma}{\ln 8}.$$

Indeed, it is enough to show that there exists $C > 0$ such that this inequality holds for $C|x - x_0|^\alpha \leq 2$, otherwise no matter the value of C the inequality is obvious. Taking C large enough so that $(2/C)^{1/\alpha} \leq 1/8$, there is a biggest n for which $|x - x_0| \leq \frac{1}{8^n}$, which implies by the previous bound:

$$|u(x) - u(x_0)| \leq 2\gamma^n \|u\|_{L^\infty(B(0,1/2))}$$

One can then use the inequality $\frac{1}{8^{n+1}} \leq |x - x_0|$ to have an upper bound on n in terms of $|x - x_0|$, and show the desired inequality for a good C. We leave the last details to the reader. □

6.2.3 Hölder regularity with source

After the treatment of the homogeneous equation (6.7) performed in the previous paragraph, we now consider the equation (6.1) with general source term f. We will use a very useful characterization of Hölder spaces due to Campanato, in terms of the growth of local integrals.

For Ω a bounded domain of $\mathbb{R}^d$, $u \in L^1(\Omega)$, and x, r such that $B(x, r) \subset \Omega$, we define

$$u_{x,r} = \frac{1}{Leb(B(x, r))} \int_{B(x,r)} u.$$

Theorem 6.3. (Campanato's characterization of Hölder spaces)
Suppose that $u \in L^2(\Omega)$ satisfies for all $B(x, r) \subset \Omega$:

$$\int_{B(x,r)} |u - u_{x,r}|^2 \leq M^2 r^{d+2\alpha}$$

for some $\alpha \in (0, 1)$. Then u is of class C^α in Ω, and for all $\tilde{\Omega} \Subset \Omega$,

$$\|u\|_{C^\alpha(\tilde{\Omega})} \ \leq \ C(M + \|u\|_{L^2(\Omega)})$$

for some C depending on d, α, Ω, Ω'.

Exercice 6.2. *1. Show that for all $c \in \mathbb{R}$, for all $u \in L^2(\Omega)$ and all $B(x, r) \subset \Omega$,*

$$\int_{B(x,r)} |u - u_{x,r}|^2 \leq \int_{B(x,r)} |u - c|^2.$$

2. Show that if u is of class C^α in a vicinity $\tilde{\Omega}$ of $B(x,r)$,

$$\int_{B(x,r)} |u - u_{x,r}|^2 \leq \|u\|_{C^\alpha(\tilde{\Omega})} r^{d+2\alpha}.$$

Proof. Denote $R_0 = dist(\tilde{\Omega}, \Omega^c)$. For any $x_0 \in \tilde{\Omega}$ and any $0 < r_1 < r_2 < R_0$,

$$|u_{x_0,r_1} - u_{x_0,r_2}|^2 \leq 2\left(|u(x) - u_{x_0,r_1}|^2 + |u(x) - u_{x_0,r_2}|^2\right)$$

so that after integration over $B(x_0, r_1)$:

$$|u_{x_0,r_1} - u_{x_0,r_2}|^2 \leq \frac{2}{\omega_d r_1^d}\left(\int_{B(x_0,r_1)} |u - u_{x_0,r_1}|^2 + \int_{B(x_0,r_2)} |u - u_{x_0,r_2}|^2\right).$$

Here, ω_d is the volume of the unit ball in $\mathbb{R}^d$. Hence,

$$|u_{x_0,r_1} - u_{x_0,r_2}|^2 \leq CM^2 r_1^{-d}\left(r_1^{d+2\alpha} + r_2^{d+2\alpha}\right). \tag{6.8}$$

For any $r < R_0$, set $r_1 = \frac{r}{2^{i+1}}$, $r_2 = \frac{r}{2^i}$, we obtain

$$|u_{x_0,2^{-(i+1)}r} - u_{x_0,2^{-i}r}| \leq C' M 2^{-(i+1)\alpha} r^\alpha.$$

Summing from $i = m-1$ to n (with $m < n$), we find easily

$$|u_{x_0,2^{-m}r} - u_{x_0,2^{-n}r}| \leq \frac{C''}{2^{m\alpha}} M r^\alpha$$

so that $(u_{x_0,2^{-i}r})_i$ is a Cauchy sequence, hence converging, for any $r > 0$. This limit is independent of r, as can be seen easily from taking $r_1 = 2^{-i}r$ and $r_2 = 2^{-i}r'$, for $0 < r \neq r' \leq R_0$. We denote by $\overline{u}(x_0)$ this limit. Taking $m = 0$, $n = +\infty$, we find

$$|u_{x_0,r} - \overline{u}(x_0)| \leq C'' M r^\alpha, \quad \forall r \leq R_0.$$

In particular, the convergence as $r \to 0$ is uniform over $\tilde{\Omega}$, and as $x_0 \to u_{x0,r}$ is continuous, $\overline{u}$ is continuous as well.

Now, invoking Lebesgue's differentiation theorem, we know that for a.e. x_0, $u_{x_0,r} \to u(x_0)$ as $r \to 0$, which means that $\overline{u} = u$ a.e. Hence, u is continuous (in the sense that it has a continuous representative), and from the previous inequality:

$$|u_{x,r} - u(x)| \leq C'' M r^\alpha, \quad \forall r \leq R_0. \tag{6.9}$$

With $r = R_0$, we obtain:

$$\sup_{\tilde{\Omega}} |u| \leq C'' M R_0^\alpha + \sup_{x \in \Omega'} u_{x,R_0} \leq c\left(M + \|u\|_{L^2(\Omega)}\right) \tag{6.10}$$

As regards the Hölder continuity, let $x \neq y \in \tilde{\Omega}$, and set $r = |x-y|$. If $r < \frac{R_0}{2}$, we write

$$|u(x) - u(y)| \leq |u(x) - u_{x,2r}| + |u(y) - u_{y,2r}| + |u_{x,2r} - u_{y,2r}|.$$

The first two terms are controlled by (6.9). For the last one, we split again:

$$|u_{x,2r} - u_{y,2r}|2 \leq 2\left(|u(z) - u_{x,2r}|^2 + |u(z) - u_{y,2r}|^2\right)$$

and integrate with respect to $z \in B(x,2r) \cap B(y,2r) \supset B(x,r)$. Thus,

$$|u_{x,2r} - u_{y,2r}|^2 \leq \frac{2}{Leb(B(x,r))}\left(\int_{B(x,2r)} |u - u_{x,2r}|^2 \; + \int_{B(y,2r)} |u - u_{y,2r}|^2\right) \leq CM^2 r^{2\alpha}$$

We get, if $r = |x-y| \leq \frac{R_0}{2}$,

$$|u(x) - u(y)| \leq CM|x-y|^\alpha.$$

On the other hand, if $|x-y| \geq \frac{R_0}{2}$

$$|u(x) - u(y)| \leq 2\sup_{\tilde{\Omega}} |u| \leq \frac{2^{\alpha+1}}{R_0^\alpha}|x-y|^\alpha \leq C(M + \|u\|_{L^2(\Omega)})|x-y|^\alpha$$

by (6.10). $\square$

Corollary 6.3. *Assume $u \in H^1_{loc}(\Omega)$ satisfies*

$$\int_{B(x,r)} |\nabla u|^2 \leq M^2 r^{d-2+2\alpha} \quad \text{for any } \; B(x,r) \subset \Omega$$

for some $\alpha \in (0,1)$. Then u is of class C^α in Ω, and for all $\tilde{\Omega} \Subset \Omega$,

$$\|u\|_{C^\alpha(\tilde{\Omega})} \; \leq \; C(M + \|u\|_{L^2(\Omega)})$$

for some C depending on d, α, Ω, Ω'.

Proof. By the Poincaré inequality for functions with zero mean, see Remark 6.2, we obtain

$$\int_{B(x,r)} |u - u_{x,r}|^2 \leq Cr^2 \int_{B(x,r)} |\nabla u|^2 \leq C' M^2 r^{d+2\alpha}$$

and the result follows from the previous theorem. $\square$

We will now prove Theorem 6.2 in the case $f \neq 0$. $x_0 \in \tilde{\Omega}$, and ρ, r such that $0 < r < R_0$, $R_0 = dist(\tilde{\Omega}, \Omega^c)$. We can with no loss of generality assume that $R_0 < 1$. We decompose $u = v + w$, where $v \in H^1_0(B(x,r))$ satisfies the homogeneous equation:

$$\int_\Omega A\nabla v \cdot \nabla \varphi = \int_\Omega f\varphi, \quad \forall \varphi \in H^1_0(B(x,r))$$

(and $w = u - v$). Taking $\varphi = v$, we get

$$\int_{B(x,r)} |\nabla v|^2 \;\le\; C \int_{B(x,r)} |f\,v| \le C\|f\|_{L^{\frac{2d}{d+2}}(B(x,r))}\,\|v\,\|_{L^{\frac{2d}{d-2}}(B(x,r))}$$
$$\le\; C'\|f\|_{L^{\frac{2d}{d+2}}(B(x,r))}\,\|\nabla v\|_{L^2(B(x,r))} \;\le\; C''\|f\|_{L^q(\Omega)} r^{\frac{1}{2}(d+2-\frac{2d}{q})}\|\nabla v\|_{L^2(B(x,r))}$$

where the second inequality comes from the Sobolev embedding, while the third comes from Hölder inequality. We end up with

(6.11)
$$\int_{B(x,r)} |\nabla v|^2 \le C\|f\|^2_{L^q(B(0,1))} r^{d+2-\frac{2d}{q}}\|\nabla v\|_{L^2(B(x,r))} = C\;\|f\|^2_{L^q(B(0,1))} r^{d-2+2\alpha_1}\|\nabla v\|_{L^2(B(x,r))}$$

with $\alpha_1 = 2 - \frac{d}{q} > 0$.

As regards the estimate for w, we state:

Lemma 6.4. *Let $w \in H^1(B(x_0,r))$ satisfying*

$$\int_{B(x_0,r)} A\nabla w \cdot \nabla\varphi = 0, \quad \forall \varphi \in H^1_0(B(x_0,r)).$$

There exists $\alpha_2 \in (0,1)$ such that for any $\rho < r$ we have:

$$\int_{B(x_0,\rho)} |\nabla w|^2 \le C\left(\frac{\rho}{r}\right)^{d-2+2\alpha_2} \int_{B(x_0,r)} |\nabla w|^2 \tag{6.12}$$

where C only depends on d, λ, $\|A\|_{L^\infty}$.

Let us postpone the proof of this lemma. On the basis of estimates (6.11) and (6.12), we can write

$$\begin{aligned}
\int_{B(x_0,\rho)} |\nabla u|^2 &\le 2\left(\int_{B(x_0,\rho)} |\nabla w|^2 + \int_{B(x_0,\rho)} |\nabla v|^2\right)\\
&\le C\left(\left(\frac{\rho}{r}\right)^{d-2+2\alpha_2}\int_{B(x_0,r)} |\nabla w|^2 + \int_{B(x_0,r)} |\nabla v|^2\right)\\
&\le C'\left(\left(\frac{\rho}{r}\right)^{d-2+2\alpha_2}\int_{B(x_0,r)} |\nabla u|^2 + \int_{B(x_0,r)} |\nabla v|^2\right)\\
&\le C''\left(\left(\frac{\rho}{r}\right)^{d-2+2\alpha_2}\int_{B(x_0,r)} |\nabla u|^2 + \|f\|^2_{L^q(B(0,1))} r^{d-2+2\alpha_1}\right).
\end{aligned}$$

We then state another lemma:

Lemma 6.5. *Let ϕ a non-negative and non-decreasing function on $[0,R]$. Assume that*

$$\phi(\rho) \le A\left(\frac{\rho}{r}\right)^{\gamma}\phi(r) + Br^{\beta},$$

for any $0 < \rho \leq r \leq R$, with $A, B, \gamma, \beta \geq 0$, and $\beta < \alpha$. Then, for any $\delta \in (\beta, \gamma)$, there exists c depending on d, β, γ, δ such that

$$\phi(r) \leq c \left(\frac{\Phi(R)}{R^\delta} r^\delta + B r^\beta \right).$$

Applying this lemma with $R = \frac{R_0}{2}$, $\gamma = d-2+\alpha_2$, $B = \|f\|_{L^q(B(0,1))}$, $\beta = \min(d-2+\alpha_1, d-2+2\alpha_2)$, and $\phi(r) = \int_{B(x_0,r)} |\nabla u|^2$, we obtain the existence of $\alpha > 0$ such that

$$\int_{B(x_0,r)} |\nabla u|^2 \leq C_0 \left(\int_{B(x_0,\frac{R_0}{2})} |\nabla u|^2 + \|f\|^2_{L^q(\Omega)} \right) r^{d-2+2\alpha}$$

As a last step, one must establish that

$$\int_{B(x_0,\frac{R_0}{2})} |\nabla u|^2 \leq C \left(\|u\|^2_{L^2(\Omega)} + \|f\|^2_{L^q(\Omega)} \right)$$

which is classical: one can for instance test the weak formulation against $\chi^2 u$, where χ is a smooth truncation that is compactly supported in Ω, and satisfying $\chi = 1$ near $B(x_0, \frac{R_0}{2})$. See the proof of Proposition 6.6 for similar (but more involved) estimates. Eventually, we get

$$\int_{B(x_0,r)} |\nabla u|^2 \leq C \left(\|u\|^2_{L^2(\Omega)} + \|f\|^2_{L^q(\Omega)} \right) r^{d-2+2\alpha}$$

and we can conclude by application of Corollary 6.3.

We still have to prove the last two lemmas.

Proof of Lemma 6.4.

By translation and dilation, one can restrict to $x_0 = 0, r = 2$. One can also restrict to $\rho \leq \frac{1}{4}$, otherwise the statement is trivial. Furthermore, we can assume that u has zero mean over $B(0,2)$, because the statement of the lemma is insensitive to the addition of a constant. By Poincaré inequality for functions with zero mean, see Remark 6.2:

$$\int_{B(0,2)} w^2 \leq C \int_{B(0,2)} |\nabla w|^2.$$

Hence, Corollary 6.2 yields: for all $x \leq \frac{1}{2}$,

$$|w(x) - w(0)|^2 \ \leq C|x|^{2\alpha} \int_{B(0,2)} |\nabla w|^2.$$

We introduce χ_ρ a smooth truncation function that is compactly supported in $B(0, 2\rho)$, $\chi_\rho = 1$ in $B(0,\rho)$, $|\nabla \chi_\rho| \leq \frac{2}{\rho}$. We then take as a test function: $\varphi = \chi_\rho^2 (w - w(0))$. Straightforward manipulations yield

$$\int_{B(0,\rho)} |Dw|^2 \leq \int \chi_\rho^2 |Dw|^2 \leq \frac{C}{\rho^2} \int_{B(0,2\rho)} |w - w(0)|^2 \leq C' \rho^{d-2+2\alpha} \int_{B(0,2)} |Dw|^2$$

see the proof of Proposition 6.6 for similar (but more involved) estimates. We end up with

$$\int_{B(0,\rho)} |Dw|^2 \leq C' \rho^{d-2+2\alpha} \int_{B(0,2)} |Dw|^2$$

Proof of Lemma 6.5.
Let $\delta \in (\beta, \gamma)$, $r < R$. Choose $\tau \in (0,1)$ such that $A\tau^\gamma = \tau^\delta$. We find

$$\phi(\tau r) \leq \tau^\gamma \phi(r) + Br^\beta.$$

For all integers $k > 0$,

$$\begin{aligned} \phi(\tau^{k+1} r) &\leq \tau^\gamma \phi(\tau^k r) + B\tau^{k\beta} r^\beta \\ &\tau^{(k+1)\gamma} + B\tau^{k\beta} r^\beta \sum_{j=0}^{k} \tau^{j(\gamma-\beta)} \\ &\tau^{(k+1)\gamma} + \frac{B\tau^{k\beta} r^\beta}{1-\tau^{\gamma-\beta}} \end{aligned}$$

Choosing k such that $\tau^{k+2} r < \rho \leq \tau^{k+1} r$, this inequality gives

$$\phi(\rho) \leq \frac{1}{\tau^\gamma} \left(\frac{\rho}{r}\right)^\gamma \phi(r) + \frac{B\rho^\beta}{\tau^{2\beta}(1-\tau^{\gamma-\beta}}.$$

Appendix A

Complement to the proof of the local inversion theorem

We show first the following lemma, which easily implies the second part of Lemma 1.1.

Lemma A.1. *Let $F : X \to Y$ a bi-Lipschitz map. Let U be an open set of X, such that F is C^1 on U and $F'(u)$ is an isomorphism from X to Y for all $u \in U$. Then, F is a C^1 diffeomorphism from U to $F(U)$.*

Proof. As F is an homeomorphism from X to Y, $F(U)$ is an open set. To prove that F^{-1} is C^1 over $F(U)$, it is enough to show that F^{-1} is differentiable: indeed, one can then differentiate the identity $F \circ F^{-1} = Id$ in $F(U)$, and see that necessarily $(F^{-1})'(F(u)) = F'(u)^{-1}$ and is therefore continuous in u.

To prove the differentiability of F^{-1} at any point $f_0 = F(u_0) \in F(U)$, we first write for any $f = F(u) \in F(U)$:

$$f - f_0 = F(u) - F(u_0) = F'(u_0)(u - u_0) + \|u - u_0\|\varepsilon(u), \quad \lim_{u \to u_0} \varepsilon(u) = 0$$

We apply $F'(u_0)^{-1}$ to this identity, and replace u by $F^{-1}(f)$, u_0 by $F^{-1}(f_0)$ to get

$$F^{-1}(f) - F^{-1}(f_0) \ = \ F'(u_0)^{-1}(f - f_0) - \|F^{-1}(f) - F^{-1}(f_0)\|\varepsilon'(f)$$

where $\varepsilon'(f) = F'(u_0)^{-1}\varepsilon(F^{-1}(f))$ goes to zero as f goes to f_0 (by the continuity of F^{-1} at f_0). Also, as F^{-1} is Lipschitz, we can write

$$\|F^{-1}(f) - F^{-1}(f_0)\| = M(f)\|f - f_0\|$$

where the function M is bounded by $\mathrm{Lip}(F^{-1})$. Eventually,

$$F^{-1}(f) - F^{-1}(f_0) \ = \ F'(u_0)^{-1}(f - f_0) + \|f - f_0\|\varepsilon''(f)$$

where $\varepsilon'' = M\,\varepsilon'$ goes to zero as f goes to f_0. This proves the differentiability. $\square$

[7, 19]

We then give the proof of:

Lemma 1.2. Let $r > 0$, $\rho_r : x \to \overline{B}_r$ such that

$$\rho_r(x) = x, \quad x \in \overline{B}_r, \quad \rho_r(x) = r\frac{x}{\|x\|}, \quad x \notin \overline{B}_r$$

Then, ρ_r is Lipschitz, with $\mathrm{Lip}(\rho_r) \leq 2$.

Proof. We split following the values of $\|x\|$ and $\|y\|$.

If $\|x\| \leq r, \|y\| \leq r$, $\|\rho_r(x) - \rho_r(y)\| = \|x - y\|$.

If $\|x\| \leq r \leq \|y\|$,

$$\begin{aligned}\|\rho_r(x) - \rho_r(y)\| = \|x - r\frac{y}{\|y\|}\| = \frac{1}{\|y\|}\| \, \|y\|x - ry\| = \frac{1}{\|y\|}\| \, (\|y\|x - \|y\| \, y) + (\|y\| \, y - ry) \, \| \\ \leq \|x - y\| + \|y\| - r \leq \|x - y\| + \|y\| - \|x\| \leq 2\|x - y\|.\end{aligned}$$

If $\|x\| \geq r$, $\|y\| \geq r$,

$$\begin{aligned}\|\rho_r(x) - \rho_r(y)\| &= r\|\frac{x}{\|x\|} - \frac{y}{\|y\|}\| \leq \frac{r}{\|x\| \, \|y\|} \, \| \, \|y\|x - \|x\|y \, \| \\ &= \frac{r}{\|x\| \, \|y\|}\|(\|y\|x - \|y\|y) + (\|y\|y - \|x\|y)\| \\ &\leq \frac{r}{\|x\|}\|x - y\| \; + \; \frac{r}{\|x\|}| \, \|y\| - \|x\|| \leq 2\|x - y\|.\end{aligned}$$

□

Appendix B

Complement to the proof of the isometric embedding

Proposition 4.1.
For any smooth compact submanifold M of $\mathbb{R}^N$, any $C^{k'}$ riemannian metrics g on M, and for any bounded neighborhood U of M, there exists a $C^{k'}$ riemannian metrics G on $\mathbb{R}^N$ with

- $G = e$ *in* $\mathbb{R}^N \setminus U$, *e the euclidean scalar product.*
- $G|_M = g$, *in the sense that for all $x \in M$, for all $v_1, v_2 \in T_xM$, $G_x(v_1, v_2) = g_x(v_1, v_2)$.*

Proof. Let $(U_j)_{j\in J}$ be a covering of a neighborhood of M by local charts, meaning that there are smooth diffeomorphisms φ_j from $U_j \subset \mathbb{R}^N$ to open sets $\Omega_j \subset \mathbb{R}^N$ such that

$$\varphi_j(U_j \cap M) = \Omega_j \cap \mathbb{R}^d \times \{0\}$$

where d is the dimension of M. We can always assume that $\cup_j U_j \subset U$. Let $(\psi_j)_{j\in J}$ a partition of unity such that the support of ψ_j is included in U_j, and $\sum_j \psi_j = 1$ in an open neighborhood $\mathcal{U}$ of M. Let $g_j = g|_{U_j\cap M}$, and $h_j = (\varphi_j)_* g_j$, that is the push forward of g_j by φ_j. It defines a riemannian metrics on $\Omega_j \cap \mathbb{R}^d \times \{0\}$, with for all $y = (y', 0) \in \Omega_j \cap \mathbb{R}^d \times \{0\}$, $h_j|_y$ is a scalar product on $\mathbb{R}^d \times \{0\}$. We extend it as a metrics over Ω_j by setting: for all $y = (y', y'') \in \Omega_j$, for all $v = (v', v'') \in \mathbb{R}^N$,

$$\tilde{h}_j|_y(v_1, v_2) = h_j|_{(y',0)}\left((v_1', 0), (v_2', 0)\right) + v_1'' \cdot v_2''$$

where the second term at the right-hand side refers to the scalar product in $\mathbb{R}^{N-d}$. Eventually, we define for all $y \in \Omega_j$:

$$H_j|_y = \chi(y'')\tilde{h}_j|_{(y',0)} \;+\; (1 - \chi(y''))(\varphi_j)_* e$$

where e is the euclidean scalar product on $\mathbb{R}^N$, and χ is a smooth function with values in $(0, 1)$, $\chi = 1$ near 0, $\chi = 0$ outside some open neighbourhood U'' of 0. More precisely, we take U'' so that

$$\Omega_j \cap (\mathbb{R}^d \times \overline{U''}) \subset \varphi(U_j \cap \mathcal{U}).$$

Eventually, we define $G_j = (\varphi_j)^* H_j$ in U_j, $G_j = e$ outside U_j, and set

$$G = \sum \psi_j G_j \text{ in } \mathcal{U}, \quad G = e \text{ outside } \mathcal{U}.$$

$\square$

Proposition 4.2.
Let $k' \geq 1$, g a $C^{k'}$ riemannian metrics on M. One can find a family of $C^{k'-1}$ riemannian metrics $(g_\varepsilon)_{\varepsilon>0}$ such that

- $\|g_\varepsilon - g\|_{C^{k'-1}} \leq \varepsilon$.
- $g_\varepsilon = u_\varepsilon^* e$, for some $u_\varepsilon : M \to \mathbb{R}^{n_\varepsilon}$ of class $C^{k'}$.

Proof. First, we prove that there exists a radial, smooth, and compactly supported function χ such that

$$\int_{\mathbb{R}^d} (\nabla\chi)(\nabla\chi)^t(x)\, dx = Id \tag{B.1}$$

Taking $\chi(x) = \psi(|x|)$, we compute:

$$\int_{\mathbb{R}^d} (\nabla\chi)(\nabla\chi)^t(x) dx = \int_0^{+\infty} \psi'(r) r^{n-1} \int_{\mathbb{S}^{d-1}} e_r \times e_r d\sigma dr.$$

We need to show that $A = \int_{\mathbb{S}^{d-1}} e_r \times e_r \, d\sigma$ is proportional to the identity matrix: indeed, afterwards, one can take $\psi = c\underline{\psi}$ for $\underline{\psi}$ a smooth compactly supported function such that $\int_0^{+\infty} \underline{\psi}'(r) r^{n-1} dr \neq 0$ and c a proper constant. Therefore, we notice that

$$A = \int_{B(0,1)} \left(\frac{x}{|x|} \otimes \frac{x}{|x|} \right) dx$$

which implies easily that $Ax \cdot x$ is constant on the unit sphere of $\mathbb{R}^d$. From there, we deduce that Ax is proportional to x for all x (Lagrange multiplier theorem), and a classical argument implies that A is a scalar endomorphism.

Now, for $a \in C_c^\infty(\mathbb{R}^d)$, $A \in C^\infty(\mathbb{R}^d, GL_d(\mathbb{R}))$, and $y \in \mathbb{R}^d$, we define $u_{\delta,y} : \mathbb{R}^d \to \mathbb{R}$ by

$$u_{\delta,y} = |\det(A(x))|^{1/2} \delta^{1-d/2} \chi \left(A(x) \frac{x-y}{\delta} \right) a(x).$$

A direct calculation shows that

$$\nabla_x u_{\delta,y} = |\det(A(x))|^{1/2} \delta^{1-d/2} a(x) \left(\frac{A^t}{\delta} + \frac{x-y}{\delta} \cdot \nabla A^t(x) \right) \nabla\chi \left(A(x) \frac{x-y}{\delta} \right) + O(\delta^{1-d/2})$$

which implies (after the change of variable $z = A(x)\frac{x-y}{\delta}$):

$$\int_{\mathbb{R}^d} (\nabla_x u_{\delta,y})(\nabla_x u_{\delta,y})^t \, dy = a(x)^2 A(x)^t \left(\int_{\mathbb{R}^d} (\nabla\chi(z))(\nabla\chi(z))^t dz \right) A(x) \; + \; O(\delta)$$
$$= a(x)^2 A(x)^t A(x) + O(\delta)$$

where the $O(\delta)$ is measured in $C^{k'-1}$ if A and a are $C^{k'}$.

To conclude the proof of the proposition, we introduce a partition of unity $(\psi_j)_{j=1\dots J}$ associated to a covering $(U_j)_{j=1\dots J}$ of the manifold M by local charts. We can then decompose $g = \sum_j \psi_j g$ and remark that, expressed in local coordinates, each $\psi_j g$ is of the form $a(x)A(x)$ where $a \in C_c^\infty(\mathbb{R}^d)$, while A is $C^{k'}$ and has values in the set of symmetric definite positive matrices. Hence, as above, one can construct $u^j_{\delta,y}$ defined on M such that

$$\int_{\mathbb{R}^d} (\nabla u^j_{\delta,y})(\nabla u^j_{\delta,y})^t \, dy \; \to \psi_j g \quad \text{in } C^{k'-1}, \qquad \text{as } \delta \to 0.$$

By approximating the integral in y by a discrete sum of the form

$$\sum_{i=1}^{l} \lambda_i^2 (\nabla u^j_{\delta,y_i})(\nabla u^j_{\delta,y_i})^t = \sum_{i=1}^{l} (\nabla \lambda_i \nabla u^j_{\delta,y_i})(\nabla \lambda_i u^j_{\delta,y_i})^t$$

we find eventually that for each $\varepsilon > 0$, for each j, there exists $\mathfrak{u}^j_1, \dots, \mathfrak{u}^j_l$, $l = l(\varepsilon, j)$, of class $C^{k'}$, supported in the local chart U_j, such that

$$\| \sum_{i=1}^{k(\varepsilon,j)} (\nabla \mathfrak{u}^j_i)(\nabla \mathfrak{u}^j_i)^t \; - \; \psi_j g \|_{C^{k'-1}} \leq \frac{\varepsilon}{J}$$

so that

$$\| \sum_{j=1}^{J} \sum_{i=1}^{k(\varepsilon,j)} (\nabla \mathfrak{u}^j_i)(\nabla \mathfrak{u}^j_i)^t \; - \; \psi_j g \|_{C^{k'-1}} \leq \varepsilon.$$

By Lemma 4.1, we deduce that for all $\varepsilon > 0$, there exists $n = n(\varepsilon)$ and $u_\varepsilon = (\mathfrak{u}^j_i)$ of class $C^{k'}$ from M to $\mathbb{R}^n$ such that $\|u_\varepsilon^* \, e - g\|_{C^{k'-1}} \leq \varepsilon$. □

Bibliography

[1] S. Alinhac, P. Gérard. Pseudo-differential operators and the Nash-Moser theorem. Graduate Studies in Mathematics, 82. *American Mathematical Society, Providence, RI*, 2007.

[2] H. Bahouri, J.-Y. Chemin, R. Danchin. Fourier analysis and nonlinear partial differential equations. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 343. *Springer, Heidelberg*, 2011.

[3] Jacob Bedrossian, Nader Masmoudi, Clément Mouhot. Landau damping: paraproducts and Gevrey regularity. *Ann. PDE* 2 2016 71 pages.

[4] J.-M. Bony. Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires. *Ann. Sci. Ec. Norm. Sup.* 14 (1981), 209-246.

[5] L. Caffarelli, A. Vasseur. The De Giorgi method for regularity of solutions of elliptic equations and its applications to fluid dynamics. *Discrete and Continuous Dynamical Systems- S*, 2010, 3 (3) : 409-427.

[6] R. Danchin. Analyse non-linéaire, notes de cours disponibles sur http://perso-math.univ-mlv.fr/users/danchin.raphael/cours/ananonlin09.pdf, 2010.

[7] E. De Giorgi. Sulla differenziabilitá e l'analiticitá delle estremaili degli integrali multipli regolari. *Mem. Accad. Sci. Torino. Cl. Sci. Fis. Math. Nat.*, 3 (1957), 25–43.

[8] L.-C. Evans. Partial Differential Equations. Graduate Studies in Mathematics, AMS, 2010.

[9] D. Gilbarg, N. Trudinger. Elliptic Partial Differential Equations of Second Order. Classics in Mathematics, *Springer-Verlag Berlin Heidelberg*, 2001.

[10] R.S. Hamilton. The inverse function theorem of Nash and Moser. *Bull. Amer. Math. Soc. (N.S.)* Volume 7, Number 1 (1982), 65-222.

[11] Lars Hörmander. The Nash-Moser theorem and paradifferential operators. *Analysis et cetera*, 429-449, *Academic Press, Boston, MA* 1990.

[12] Guy Métivier. Para-differential calculus and applications to the Cauchy problem for nonlinear systems. *Centro di Ricerca Matematica Ennio De Giorgi (CRM) Series*, 5. Edizioni della Normale, Pisa, 2008.

[13] Clément Mouhot, Cédric Villani. On Landau damping. *Acta Math.* 207 2011 29-201.

[14] Jürgen Moser. A new technique for the construction of solutions of nonlinear differential equations. *Proc. Nat. Acad. Sci. U.S.A* 47 1961 1824-1831.

[15] Jürgen Moser. A rapidly convergent iteration method and non-linear partial differential equations - I. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* 20 1966 265-315.

[16] Jürgen Moser. A rapidly convergent iteration method and non-linear partial differential equations - II. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* 20 1966 499-535.

[17] Jürgen Moser. On Harnack's theorem for elliptic differential equations. *Comm. Pure Appl. Math.* 14 1961 577-591.

[18] John Nash. The embedding problem for Riemannian manifolds. *Ann. of Math.* 63 1956 20-63.

[19] J. Nash. Continuity of solutions of parabolic and elliptic equations. *American Journal of Mathematics*, Vol. 80, No. 4. (1958).

[20] Qing Han and Fanghua Lin. Elliptic Partial Differential Equations. Courant Lecture Notes. *American Mathematical Society, Providence, Rhode Island, USA*, 2000.

[21] Q. Rahman, G. Schmeisser. Analytic Theory of Polynomials. London Mathematical Society Monographs. *Oxford University Press, USA*, 2002

[22] Ya. G. Sinai. Topics in Ergodic Theory. Princeton Mathematical Series 44. *Princeton University Press, Princeton, NJ*, 1994.

[23] C. De Lellis, L. Szekelyhidi. High-dimensionality and h-principle in PDE. *Bull. Ams*, 2017.

[24] Terry Tao. The Nash-Moser iteration scheme. Available at http://www.math.ucla.edu/ tao/preprints/Expository/

[25] C. E. Wayne. An introduction to KAM theory. *Dynamical systems and probabilistic methods in partial differential equations* (Berkeley, CA, 1994), 3–29, Lectures in Appl. Math., 31, *Amer. Math. Soc., Providence, RI*, 1996.

[26] M. Willem. Analyse harmonique réelle. Collection Méthodes. *Hermann, Paris* (1995).

[27] A. Zeghib. Histoire des immersions isométriques. *Géométrie au XX e siècle*. Editeurs : J. Kouneiher , D. Flament , P. Nabonnand , J. Szczeciniarz, Hermann, Paris '2005).