The Dissertation Committee for John Andrew Vano
certifies that this is the approved version of the following dissertation:

A Nash-Moser Implicit Function Theorem with Whitney Regularity and Applications

Committee:

Rafael de la Llave, Supervisor

Hans Koch

Lorenzo Sadun

Charles Radin

Philip Morrison

A Nash-Moser Implicit Function Theorem with Whitney Regularity and Applications

by

John Andrew Vano, B.A.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2002

To my parents

Who taught me the value of understanding

To my wife

Who makes sure I eat and sleep somewhat regularly

And to all my cats

Who occasionally inspire mathematical insight

Acknowledgments

I would like to acknowledge the support of all who helped make this endeavor possible. I would specifically like to thank my undergraduate thesis adviser Dr. Jeff Rosoff at Gustavus Adolphus College, my "prelim option" adviser Dr. Hans Koch at University of Texas at Austin, all my graduate student cohorts including Nikola Petrov, and my Ph.D. thesis adviser Dr. Rafael de la Llave also at University of Texas at Austin. These individuals have served as teachers, mentors and friends and I am deeply grateful for their help and encouragement. I am also deeply grateful for help and patience of the members of my committee. Finally, on a personal note I would like to thank my parents Andrew and Sally Vano and my wife Jennie Vano for all their love and support over the years.

Thesis available for download at http://www.ma.utexas.edu/mp_arc/

A Nash-Moser Implicit Function Theorem with Whitney Regularity and Applications

Publication No. ______________

John Andrew Vano, Ph.D.
The University of Texas at Austin, 2002

Supervisor: Rafael de la Llave

This dissertation establishes the Whitney regularity with respect to parameters of implicit functions obtained from a Nash-Moser implicit function theorem. As an application of this result, we study the problem of wave propagation in resonating cavities.

Using a modification of the general setup in [Zeh75], we consider functionals $\mathcal{F} : U \times V \to Z$ which have an approximate right inverse $R : \mathcal{C} \times V \to L(Z, Y)$. Here $U \subseteq X$ and $V \subseteq Y$ are open sets of scales of Banach spaces (scale parameters are suppressed here for brevity) and $\mathcal{C} \subseteq U$ is an arbitrary set of parameters (in applications $\mathcal{C}$ is often a Cantor set). Under appropriate hypothesis on $\mathcal{F}$, which are natural extensions of [Zeh75], we show that given (x_0, y_0) with $\mathcal{F}(x_0, y_0) = 0$ for $x \in \mathcal{C}$ near x_0 there exists a function $g(x)$, Whitney regular with respect to x, which satisfies $\mathcal{F}(x, g(x)) = 0$.

The problem of wave propagation in a cavity with (quasi-periodically) moving boundary can be reduced to the study of a family of torus maps. Because of their extremely degenerate nature, this family is not covered by known versions of KAM theory. However, our implicit function theorem approach allows us to overcome these problems and prove a degenerate KAM theory. Our approach can also be applied to other problems of current interest.

Table of Contents

Chapter 1

Introduction

In this dissertation we present a Nash-Moser implicit function theorem and establish the Whitney regularity of the resulting implicit function. This is the thrust of Chapters 2-7 where we extend the abstract formulation of hard implicit function theorems (in particular [Zeh75]) to include smooth dependence on parameters, even when the parameters range over Cantor sets. As an application of this Nash-Moser implicit function theorem with Whitney regularity in Chapters 8 and 9 we establish a KAM theory for a family of torus maps that arise in the study of wave propagation in a domain with a quasi-periodically moving boundary. This family is extremely degenerate since the frequencies available lie in a one-dimensional space. Moreover the dependence on parameters turns out to have critical points in the region of interest. We introduce the method of "borrowing of parameters" which allows us to prove versions of the KAM theorem which apply to such degenerate situations.

Recall that implicit function theorems allow one to solve equations provided the function defining the equation satisfy some non-degeneracy conditions. An important prototype is the classical implicit function theorem that allows one to solve the equation $F(x, y) = 0$ for y in terms of x in a neighborhood of (x_0, y_0) with $F(x_0, y_0) = 0$ provided the operator $D_2F(x_0, y_0)$ has a (bounded) inverse. It is well known that this theorem remains valid when x and y range over a general Banach spaces (see e.g.[Die69]) which makes the implicit function theorem one of the basic tools of nonlinear functional analysis.

Hard implicit function theorems cover cases where the assumption of boundedness for the inverse of $D_2F(x_0, y_0)$ is weakened. In these settings, one usually considers functionals that map between Banach spaces in which one can separate out one parameter families of Banach subspaces at various "scales" (for example one might have a functional acting on the space of C^γ functions which has, for $\alpha > \gamma$, the Banach subspaces C^α). In this setting one usually assumes that while $D_2F(x_0, y_0)$ may not have a bounded inverse when viewed under the "scales" which make $D_2F(x_0, y_0)$ a bounded operator, it does have a bounded inverse when taking one "scale" into a bigger space at another "scale" (by analogy with the C^γ spaces, smaller scale parameter correspond to larger spaces). The manner in which the inverse of $D_2F(x_0, y_0)$ becomes bounded by changing "scale" must also satisfy certain quantitative estimates (tameness). Several versions of hard implicit function theorems have been developed to serve various problems, see [Ham82], [Hör76], [Hör85], [Hör90] , [Sch60], [Ser72] or [Ser73]. The closest to our point of view is [Zeh75].

One important motivation for the development of such hard implicit function theorems has been the study of persistence of quasi-periodic solutions in Hamiltonian systems. A class of problems related to the persistence of tori are conjugacy problems (see e.g. Example 4.3.1). Such problems are studied by KAM theory, named in honor of Kolmogorov, Arnold and Moser who originated and developed the theory in the late 50's and early 60's. The connection of these problems with hard implicit functions theorems appears because, when one writes down the equation for invariance, the resulting functional equations involve small divisors (see (4.25) in Example 4.3.1). To obtain boundedness of the inverse of $D_2F(x_0, y_0)$ for such functionals, the small divisors require one to "change scales" in order to obtain estimates (this "change of scales" is often referred to as a loss of smoothness/regularity or, when thinking of analytic functions, as a loss of domain). Furthermore, to obtain boundedness not only must one "change scales," but certain number theoretic (Diophantine) properties of the quasi-periodic frequency are also necessary to obtain quantitative estimates (tameness). See Section 8.5 and Definition 8.5.1 for more details on this matter.

When considering problems of the above type, it is very natural to consider the dependence of the results on the frequency, that is, to view the frequency as a parameter. Since the set of vectors satisfying Diophantine conditions has empty interior, the appropriate concept for regularity is Whitney regularity (see Definition 3.1.1). The study of the dependence of solutions on frequency parameters is interesting on several grounds. For example, the dependence of the frequency leads immediately to geometric properties of the set of tori which are observed. The abundance and geometry of the set of tori plays an important role in applications and is a subject of current theoretical and experimental interest (see [TLRF02, Las93]).

Perhaps more importantly, as we will show in Chapter 9, study of the dependence on the frequency allows one to obtain, rather quickly, results for systems whose map are very degenerate ([Rüs90], [CS94], [Rüs01]). See also [BHS96b], [BHS96a], [Sev99], [Sev96]. Such degenerate systems appear often in practice due in part to the abundance of symmetry (around certain points) in applied problems. In particular, we note that the most famous problem in mechanics, the planetary system, is degenerate because the Kepler solutions present only one frequency (they are periodic) while one would expect three independent frequencies in a system with three degrees of freedom. Other examples with extreme degeneracy occur in chemical systems where degeneracy occurs due to the fact that all the particles of the same species have the same mass and other mechanical properties. The weakest assumptions on non-degeneracy that presently allow for the proof of KAM theory are the so called Rüssmann non-degeneracy conditions (see [Pös01], [Rüs01]). These conditions can be obtained as a corollary of our methods. Our method of "borrowing of parameters" also can be applied to examples which do not satisfy the conditions of Rüssmann (see Chapter 9).

The key to the development of our results are constructive implicit function theorems (Theorem 6.1.1 and Theorem 7.1.1). Informally, these constructive theorems state that given an object which approximately satisfies the functional equation, there exists a true solution which is close (in appropriate norm) to the approximate object. Such constructive theorems are useful in numerical analysis, where they go under the name of a posteriori estimates. A numerical algorithm, if correctly implemented, produces objects which approximately satisfy the desired equation to a very high accuracy. If one has such constructive theorems or a posteriori estimates then the computed approximate solutions have true solutions nearby. These constructive theorems can also be used to validate approximate solutions obtained from other methods, e.g. through formal expansions. With a constructive implicit function theorem we do not need to analyze or justify the procedure used to obtain our approximate solutions. To obtain the existence of similar (i.e. nearby) true solutions we only need to verify that our approximate solutions satisfies the equation approximately.

The constructive hard implicit function theorems we present (Theorem 6.1.1 and Theorem 7.1.1) are patterned after that of [Zeh75] but we have paid attention to some quantitative issues and incorporated the more modern Brjuno-Rüssmann small divisor condition. Using these constructive hard implicit function theorems, we establish Whitney regularity with respect to parameters in two different settings (see Theorem 6.2.1 and Theorem 7.2.1 and Theorem 6.2.3 and Theorem 7.2.2)

Our first approach to obtaining Whitney regularity is to apply our constructive hard implicit function theorem (Theorem 6.1.1 or Theorem 7.1.1) in the context of Banach spaces of Whitney differentiable functions (a similar approach to obtaining differentiability on parameters was used in [dlLO99] for functions depending on parameters on manifolds). This approach has the advantage that implicit solutions of the functional need not be unique. This non-uniqueness occurs, for example, in the isometric embedding problem. On the other hand, we need to assume there is a consistent way to obtain an approximate right inverse which depends smoothly on parameters. Such approximate right inverse can be obtained if the functional has some type of group structure, as described in [Zeh75].

Our second approach to obtaining Whitney regularity requires uniqueness for solutions to the functional equation. If this is the case, we can use the formal expansions of the implicit function to directly verify the Whitney regularity of the implicit function. The terms of this expansion play the role of the Whitney derivatives. These formal expansions are a natural abstraction of the Lindstedt expansions of solutions in terms of their frequencies. Note that this approach provides some validation for the formal expansions which appear in the study of KAM problems of mechanics.

The layout of our exposition is as follows:

- **Chapter 2** presents some basic results about polynomials, asymptotic polynomials and formal power series. Of particular interest is the behavior of polynomials and asymptotic polynomials under composition. The coefficients that arise from the composition of polynomials are identical to derivatives of the composition. This will be used in the following chapter when we consider the composition of Whitney differentiable functions. In particular, we will use it to determine the Whitney derivatives of such a composition.

- **Chapter 3** introduces notion of Whitney Regularity. The definition and some basic consequences are presented in Section 3.1. Of particular interest is Theorem 3.1.8 which proves that the composition of two Whitney differentiable functions produces a function which is again Whitney differentiable. Section 3.2 explores the issues of the uniqueness of Whitney derivatives. The Whitney Extension theorems, which makes Whitney Regularity a very useful concept, appear in Section 3.3.

- In **Chapter 4** we present the abstract setting in which we work. Section 4.1 describes the one parameter families of Banach spaces X_σ along with the corresponding accumulation spaces X_0^q and C^ω smoothing. Section 4.2 presents the Brjuno-Rüssmann condition which is exactly the quantitative estimates (tameness) needed to obtain our results. A list of the various sets of hypotheses we use to obtain results in these various settings appears in Section 4.3 (broken down into: hypotheses for polynomial approximate solutions in Section 4.3.1, hypotheses for solutions in the analytic spaces X_σ in Section 4.3.2 and hypotheses for solutions in the smooth spaces X_0^q in Section 4.3.3).

- **Chapter 5** begins the development our results by establish the existence of polynomial approximate solutions akin to the Lindstedt expansions in mechanics.

- **Chapter 6** presents the development of solutions in the analytic spaces X_σ with Section 6.1 containing the "constructive" implicit function theorem (Theorem 6.1.1), Section 6.2 establishing the Whitney regularity and Section 6.3 presenting one approach to establishing uniqueness.

- **Chapter 7** mirrors the development of Chapter 6 but with results in the smooth spaces X_0^q.

- In **Chapter 8** we study maps of the torus. This develops the framework for the following chapter.

- **Chapter 9** presents an application of our Nash-Moser implicit function with Whitney regularity. Here we establish a degenerate version of KAM theory which applies to a families of torus maps that arise in the study of wave propagation in a domain with a quasi-periodically moving boundary.

Chapter 2

Polynomial Preliminaries

It is useful to begin by developing some notation and results about polynomials. For us, a polynomial is a finite sum of symmetric multi-linear operators (see Definition 2.1.1 below). The notation we define in this chapter for expressing polynomials will be used extensively in Chapter 3 and will appear throughout the rest of the dissertation (using the one parameter families of Banach spaces defined in Section 4.1). A detailed study of polynomial algebras, etc. can be found in [Gla58].

2.1 Polynomials

Let X, Y be Banach spaces and let $\mathrm{Sym}_n(X, Y)$ denote the space of continuous symmetric n-linear forms from X^n to Y (for ease of notation we take $\mathrm{Sym}_0(X, Y) = Y$). For $a \in \mathrm{Sym}_i(X, Y)$ define the operator norm

$$\|a\|_{\mathrm{Sym}_i(X,Y)} = \sup\{\|a[v_1, \ldots, v_i]\|_Y : v_j \in X, \|v_j\|_X \leq 1, 1 \leq j \leq i\}$$

Definition 2.1.1. *Given $a_i \in \mathrm{Sym}_i(X, Y)$, $0 \leq i \leq k$, for $n, \ell \geq 0$ with $n + \ell \leq k$ we define the polynomials $a_n^{\leq \ell} : X \to \mathrm{Sym}_n(X, Y)$ by*

$$a_n^{\leq \ell}(\Delta)[-]^{\otimes n} = \sum_{i=0}^{\ell} \frac{1}{i!} a_{n+i}[\Delta^{\otimes i}, [-]^{\otimes n}] \tag{2.1}$$

Here $\Delta \in X$ and $[-]$ is used as a placeholder for terms from X used which are inserted when applying this an element of $\mathrm{Sym}_n(X,Y)$. For the polynomials $a_0^{\leq k}(\Delta)$, i.e. (2.1) for $n = 0$ and $\ell = k$, we write $a(\Delta)$ or, to emphasize the degree, $a^{\leq k}(\Delta)$.

Let $\mathbb{P}_k[X;Y]$ denote the set of all $a^{\leq k}$ and define

$$\left\|a^{\leq k}\right\|_{\mathbb{P}_k} = \max\left\{\|a_i\|_{\mathrm{Sym}_i(X,Y)} : 0 \leq i \leq k\right\} \tag{2.2}$$

■

Definition 2.1.1, in particular (2.1), is motivated by the computation of derivatives. See Remark 2.1.5.

Remark 2.1.2. *For $n < m$, given $a \in \mathbb{P}_n[X;Y]$ by taking $a_i = 0$ for $n < i \leq m$ one can view $a \in \mathbb{P}_m[X;Y]$. This gives a natural inclusion of $\mathbb{P}_n[X;Y]$ into $\mathbb{P}_m[X;Y]$ Conversely, for $n < m$, given $a \in \mathbb{P}_m[X;Y]$ the truncation $a^{\leq n} \in \mathbb{P}_n[X;Y]$. This gives a natural projection of $\mathbb{P}_m[X;Y]$ onto $\mathbb{P}_n[X;Y]$.*

There are several useful variations of $\mathbb{P}_k[X;Y]$ which we now define.

Definition 2.1.3. *Define $\breve{\mathbb{P}}_k[X;Y]$ to be the subset of $\mathbb{P}_k[X;Y]$ of polynomials $a^{\leq k}$ with $a_0 = 0$. Furthermore, for $a^{\leq k} \in \mathbb{P}_k[X;Y]$, given $n, \ell \geq 0$ with $n + \ell \leq k$, we define*

$$\breve{a}_n^{\leq \ell}(\Delta)[-]^{\otimes n} = \sum_{i=1}^{\ell} \frac{1}{i!} a_{n+i}[\Delta^{\otimes i}, [-]^{\otimes n}] \tag{2.3}$$

so that $a_n^{\leq \ell}(\Delta) = a_n + \breve{a}_n^{\leq \ell}(\Delta)$.

As with the polynomials $a_0^{\leq k}(\Delta)$*, we use* $\breve{a}(\Delta)$ *or, to emphasize the degree,* $\breve{a}^{\leq k}(\Delta)$ *to express* $\breve{a}_0^{\leq k}(\Delta)$*, i.e.* (2.3) *for* $n = 0$ *and* $\ell = k$*.* ■

Next, we define polynomials whose coefficients a_i depend on a variable $p \in M$.

Definition 2.1.4. *Given functions* $g_i : M \to \mathrm{Sym}_i(X, Y)$ *for* $0 \leq i \leq k$*, for* $n, \ell \geq 0$ *with* $n + \ell \leq k$*, we define the variable coefficient polynomials*

$$g_n^{\leq \ell} : M \times X \to \mathrm{Sym}_n(X, Y)$$

by

$$g_n^{\leq \ell}(p; \Delta)[-]^{\otimes n} = \sum_{i=0}^{\ell} \frac{1}{i!} g_{n+i}(p)[\Delta^{\otimes i}, [-]^{\otimes n}] \tag{2.4}$$

Here $p \in M$ *and again we take* $\Delta \in X$ *and* $[-]$ *represents a placeholder for elements of* X*. Let* $\mathbb{P}_k[M, X; Y]$ *be the set of all variable coefficient polynomials* $g_0^{\leq k}$*.*

As in the constant coefficient case, we use $g(p; \Delta)$ *or, to emphasize the degree,* $g^{\leq k}(p; \Delta)$ *to express* $g_0^{\leq \ell}(p; \Delta)$*, i.e.* (2.4) *for* $n = 0$ *and* $\ell = k$*. We also define* $\breve{\mathbb{P}}_k[M, X; Y]$ *to be the subset* $\mathbb{P}_k[M, X; Y]$ *with* $g_0(p) = 0$*. Given* $g_0^{\leq k} \in \mathbb{P}_k[M, X; Y]$ *and* $n, \ell \geq 0$ *with* $n + \ell \leq k$ *define* $\breve{g}_n^{\leq \ell} \in \breve{\mathbb{P}}_\ell[M, X; \mathrm{Sym}_n(X, Y)]$ *by* $g_n^{\leq \ell}(p; \Delta) = g_n(p) + \breve{g}_n^{\leq \ell}(p; \Delta)$ *and use* $\breve{g}(p; \Delta)$ *or* $\breve{g}^{\leq k}(p; \Delta)$ *to express* $\breve{g}_0^{\leq k}(p; \Delta)$*.*
■

Remark 2.1.5. *Note that with the factorial normalization in the coefficients of* $a_n^{\leq \ell}(\Delta)$ *and* $g_n^{\leq \ell}(p; \Delta)$*, for all* $m \leq \ell$ *we have*

$$D_\Delta^m[a_n^{\leq \ell}(\Delta)] = a_{n+m}^{\leq \ell - m}(\Delta) \quad \text{and} \quad D_\Delta^m[g_n^{\leq \ell}(p; \Delta)] = g_{n+m}^{\leq \ell - m}(p; \Delta)$$

This is one of the motivations behind our choice of notation. Also observe that $a_n^{\leq \ell}(0) = a_n$ *and* $g_n^{\leq \ell}(p;0) = g_n(p)$ *while* $\breve{a}_n^{\leq \ell}(0) = 0$ *and* $\breve{g}_n^{\leq \ell}(p;0) = 0$.

Remark 2.1.6. *An element* $g \in \mathbb{P}_k[M, X; Y]$ *can be viewed both as a mapping* $g : M \times X \to Y$ *and as a mapping* $g : M \to \mathbb{P}_k[X; Y]$.

If M *is a Banach space and the variable coefficients are themselves polynomials with* $g_i \in \mathbb{P}_{k-i}[M; \mathrm{Sym}_i(X, Y)]$ *then the mapping* $g^{\leq k} : M \times X \to Y$ *is a constant coefficient polynomial, i.e. there is* $a^{\leq k} \in \mathbb{P}_k[M \times X; Y]$ *with* $a^{\leq k}((p, x)) = g^{\leq k}(p; x)$. *Conversely, any polynomial* $a^{\leq k} \in \mathbb{P}_k[M \times X; Y]$ *can be thought of as a polynomial in* Y *with variable coefficient depending on* M*, i.e. there is* $g^{\leq k} \in \mathbb{P}_k[M, X; Y]$ *with variable coefficients* $g_i \in \mathbb{P}_{k-i}[M; \mathrm{Sym}_i(X, Y)]$ *such that* $a^{\leq k}((p, x)) = g^{\leq k}(p; x)$. *Going between these two viewpoints is is useful when we consider the composition of polynomials. In particular, see Lemma 2.1.12, Lemma 2.1.13 and Theorem 2.1.14.*

Proposition 2.1.7. *Under the norm* $\|-\|_{\mathbb{P}_k}$, $\mathbb{P}_k[X; Y]$, $\mathbb{P}_k[M, X; Y]$, $\breve{\mathbb{P}}_k[X; Y]$, *and* $\breve{\mathbb{P}}_k[M, X; Y]$ *are all Banach spaces.*

If Y *is a Banach Algebra,* $\bigcup_{0 \leq k} \mathbb{P}_k[X; Y]$, $\bigcup_{0 \leq k} \breve{\mathbb{P}}_k[X; Y]$, $\bigcup_{0 \leq k} \mathbb{P}_k[M, X; Y]$ *and* $\bigcup_{0 \leq k} \breve{\mathbb{P}}_k[M, X; Y]$ *are normed algebras, however they are not complete.*

Finally, the natural inclusions and projections described in Remark 2.1.2 are bounded linear operators with operator norms of 1.

Proof. Straightforward. □

Proposition 2.1.8. (Polynomial Composition)

Let X, Y and Z be Banach spaces. If $a \in \mathbb{P}_n[Y;Z]$ and $b \in \mathbb{P}_m[X;Y]$ we denote by $a \circ b$ the polynomial in $\mathbb{P}_{nm}[X;Z]$ defined by $a \circ b(\Delta) = a(b(\Delta))$. Letting $c = a \circ b$, with $c = c^{\leq nm}$ as in (2.1) of Definition 2.1.1, we have

$$c_0 = a_0^{\leq n}(b_0) = P_0(a_0, \dots, a_n; b_0) \tag{2.5}$$

and, using the convention $a_j = 0$ for $j > n$ and $b_j = 0$ for $j > m$, for $i > 0$

$$c_i = P_i(a_1, \dots, a_n; b_0, \dots, b_i) \tag{2.6}$$

where P_i is a polynomial in b_i with coefficients a_i. The expression of polynomials P_i is independent of the spaces X, Y and Z (see Remark 2.1.10).

Furthermore, there exists a constants $M_{n,m} \geq 1$ such that given any $a \in \mathbb{P}_n[Y;Z]$ and $b \in \mathbb{P}_m[X;Y]$ we have

$$\|a \circ b\|_{\mathbb{P}_{nm}} \leq M_{n,m} \|a\|_{\mathbb{P}_n} (1 + \|b\|_{\mathbb{P}_m}^m) \tag{2.7}$$

and for any $a \in \mathbb{P}_n[Y;Z]$ and $e, f \in \mathbb{P}_m[X;Y]$

$$\|a \circ e - a \circ f\|_{\mathbb{P}_{nm}} \leq C(e,f) \|a\|_{\mathbb{P}_n} \|e - f\|_{\mathbb{P}_m} \tag{2.8}$$

where $C(e,f) = M_{n,m}(1 + \max(\|e\|_{\mathbb{P}_m}, \|f\|_{\mathbb{P}_m})^{m-1})$ and

$$\left\|a \circ e - a \circ f - (a_1^{\leq n-1} \circ f)[e - f]\right\|_{\mathbb{P}_{nm^2}} \leq D(e,f) \|a\|_{\mathbb{P}_n} \|e - f\|_{\mathbb{P}_m}^2 \tag{2.9}$$

where $D(e,f) = M_{n,m}(1 + \max(\|e\|_{\mathbb{P}_m}, \|f\|_{\mathbb{P}_m})^{m-2})$.

Proof. Equation (2.6) follows from the definitions.

To establish (2.7) note that the norm defined in (2.2) is equivalent to the following norm

$$\left\|a^{\leq k}\right\|_{\sup} = \max\left\{\left\|a^{\leq k}(\Delta)\right\|_Y : \|\Delta\|_X \leq 1\right\} \tag{2.10}$$

i.e. there exists a constant $C_k \geq 1$ such that

$$(1/C_k)\left\|a^{\leq k}\right\|_{\sup} \leq \left\|a^{\leq k}\right\|_{\mathbb{P}_k} \leq C_k\left\|a^{\leq k}\right\|_{\sup}$$

For $\lambda \geq 1$, we have the scaling property

$$\left\|a^{\leq k}\circ(\lambda \mathrm{Id})\right\|_{\sup} \leq \lambda^k\left\|a^{\leq k}\right\|_{\mathbb{P}_k}$$

and thus for $\lambda \in \mathbb{R}$

$$\left\|a^{\leq k}\circ(\lambda \mathrm{Id})\right\|_{\sup} \leq (1 + |\lambda|^k)\left\|a^{\leq k}\right\|_{\mathbb{P}_k}$$

From this it is clear

$$\|a\circ b\|_{\sup} \leq \|a\|_{\mathbb{P}_k}(1 + \|b\|_{\sup}^m)$$

By the equivalence of norms, (2.7) follows.

Note that (2.8) follows from (2.9). To establish (2.9), note that for fixed $v, w \in Y$ with $\|v\|_Y, \|w\|_Y \leq R$ we have

$$\left\|a(v) - a(w) - a_1^{\leq n-1}(w)[v - w]\right\|_Z \leq M_n(1 + R^{m-2})\|a\|_{\mathbb{P}_n}\|v - w\|_Y^2$$

Replacing v, w with e, f we get

$$\left\|a\circ e - a\circ f - (a_1^{\leq n-1}\circ f)[e - f]\right\|_{\sup} \leq D(e, f)\|e - f\|_{\sup}^2$$

with $D(e,f) = M_{n,m}(1+\max(\|e\|_{\mathbb{P}_m}, \|f\|_{\mathbb{P}_m})^{m-2})$. By the equivalence of norms, (2.8) follows. $\square$

Remark 2.1.9. *Note* (2.8) *in Proposition 2.1.8 establishes the continuity of the map* $a_* : \mathbb{P}_m[X;Y] \to \mathbb{P}_{nm}[X;Z]$ *defined by* $a_*(b) = a \circ b$ *while* (2.9) *proves that it is differentiable with derivative* $Da_*(b)[\Delta b] = a_1^{n-1} \circ b[\Delta b]$.

Remark 2.1.10. *The polynomials* P_i *defined in* (2.6) *(and* Q_i *defined in* (4) *of Proposition 2.1.11) have the same form independent of the choice of* X, Y, Z, *and* m. *Furthermore, their dependence in* n *can be understood by taking* $a_n = \ldots = a_{n-k} = 0$. *Writing* $\tilde{P}_i$ *for coefficients that arise when composing polynomials of degree* k *with degree* m *and* $\tilde{P}_i$ *for the coefficients which arise when composing polynomials of degree* k *with degree* m*, we clearly have*

$$\tilde{P}_i(a_1, \ldots, a_{n-k}; b_0, \ldots, b_i) = P_i(a_1, \ldots, a_n; b_0, \ldots, b_i)$$

Writing $a_j p_j(b_0, \ldots, b_j)$ *for* $P_j(0, \ldots, 0, a_j, 0 \ldots, 0; b_0, \ldots, b_j)$*, by the linearity of* P_i *described in* (2) *of Proposition 2.1.11 we have*

$$P_i(a_1, \ldots, a_n; b_0, \ldots, b_i) = \sum_{j=1}^{n} a_j p_j(b_0, \ldots, b_j)$$

Using the notion of formal power series we can think of P_i *as being independent of* n *(see Remark 2.3.2).*

Finally, note that when $X = Y = Z = \mathbb{R}$*, the fact that* $a_i b_j = b_j a_i$*, allows one to simplify the formulas for* P_i*. In this setting, the Faá di Bruno's formula (see [AR67]) gives an explicit formula for the derivative in* (1) *of*

Proposition 2.1.11 and thus P_i (the Faá di Bruno's formula actually can be expressed in this arbitrary setting but care must be taken since one does not have commutativity).

Proposition 2.1.11. *The polynomials P_i defined in* (2.6) *have the following useful properties:*

1. *(Computing via differentiation)*

$$P_i(a_1, \ldots, a_n; b_0, \ldots, b_i) = D^i_\Delta \left[c_0^{\leq nm}(\Delta)\right]_{\Delta=0} = D^i_\Delta \left[a_0^{\leq n}(b_0^{\leq m}(\Delta))\right]_{\Delta=0}$$

2. *(Linearity)*

$$P_i(\alpha a_1 + \beta b_1, \ldots, \alpha a_n + \beta b_n; c_0, \ldots, c_i) =$$
$$= \alpha P_i(a_1, \ldots, a_n; c_0, \ldots, c_i) + \beta P_i(b_1, \ldots, b_n; c_0, \ldots, c_i)$$

3. *(Explicit $a_{i+1}, \ldots, a_n$ independence)*

$$P_i(a_1, \ldots, a_n; 0, b_1, \ldots, b_i) = P_i(a_1, \ldots, a_i; 0, b_1, \ldots, b_i)$$

4. *(Explicit b_i dependence)*

$$(i!)P_i(a_1, \ldots, a_n; b_0, \ldots, b_i) = a_1[b_i] + Q_i(a_1, \ldots, a_n; b_0, \ldots, b_{i-1})$$

for Q_i a polynomial with coefficients $a_1, \ldots, a_n$ depending on $b_0, \ldots, b_{i-1}$.

Proof. Straightforward. □

Lemma 2.1.12. *Given $a \in \mathbb{P}_n[Y;Z]$ let $f \in \mathbb{P}_n[Y,Y;Z]$ with coefficients $f_i \in \mathbb{P}_{n-i}[Y;\mathrm{Sym}_i(Y,Z)]$, defined by $f_0(x) = a(x)$ and, for $0 < i \leq n$,*

$$f_i(x) = P_i(a_1, \ldots, a_n; x, 0, \ldots, 0) \tag{2.11}$$

Then

$$f(x;\Delta) = a(x + \Delta) \tag{2.12}$$

Conversely, given $a \in \mathbb{P}_n[Y;Z]$ and $f \in \mathbb{P}_n[Y,Y;Z]$ satisfying (2.12) *the coefficients of f must satisfy* (2.11).

Proof. Apply Proposition 2.1.11. □

Lemma 2.1.13. *If $a \in \mathbb{P}_n[Y;Z]$, $b \in \mathbb{P}_m[X;Y]$ and $f \in \mathbb{P}_n[Y,Y;Z]$ with coefficients $f_i \in \mathbb{P}_{n-i}[Y;\mathrm{Sym}_i(Y,Z)]$ defined by $f_0(x) = a(x)$ and, for $0 < i \leq n$,*

$$f_i(x) = P_i(a_1, \ldots, a_n; x, 0, \ldots, 0)$$

then

$$P_i(a_1, \ldots, a_n; b_0, \ldots, b_i) = P_i(f_1(b_0), \ldots, f_i(b_0); 0, b_1, \ldots, b_i)$$

Proof. Note by definition of f_i we from Lemma 2.1.12 that (2.12) holds. Note that one has

$$a(b(\Delta)) = a(b_0 + \breve{b}(\Delta)) = f(b_0; \breve{b}(\Delta))$$

Applying Proposition 2.1.8 and 3 from Proposition 2.1.11 the result follows. □

We now present a fundamental and very useful relationship among the polynomials P_i. This relationship arises and is easy to establish when we consider the derivatives of the composition of polynomials. However, it also arises when we consider the composition of asymptotic polynomials in Section 2.2 and again when we consider the composition of Whitney differentiable functions (Theorem 3.1.8) in Section 3.1.

Theorem 2.1.14. *Given polynomials* $a \in \mathbb{P}_n[Y;Z]$ *and* $b \in \mathbb{P}_m[X;Y]$, *let* $f \in \mathbb{P}_n[Y,Y;Z]$, $g \in \mathbb{P}_m[X,X;Y]$ *and* $h \in \mathbb{P}_{nm}[X,X;Z]$ *be variable coefficient polynomials with coefficients* $f_i \in \mathbb{P}_{n-i}[Y;\mathrm{Sym}_i(Y,Z)]$, $g_i \in \mathbb{P}_{m-i}[X;\mathrm{Sym}_i(X,Y)]$ *and* $h_i \in \mathbb{P}_{nm-i}[X;\mathrm{Sym}_i(X,Z)]$ *defined by*

$$f_i(y) = P_i(a_1, \ldots, a_n; y, 0, \ldots, 0)$$

$$g_i(x) = P_i(b_1, \ldots, b_m; x, 0, \ldots, 0)$$

$$h_0(x) = a_0^{\leq n}(b_0^{\leq m}(x)) = f_0^{\leq n}(g_0^{\leq m}(x)) \tag{2.13}$$

and

$$h_i(x) = P_i(f_1(g_0(x)), \ldots, f_i(g_0(x)); 0, g_1(x), \ldots, g_i(x)) \tag{2.14}$$

Then one has the property

$$h(x;\Delta) = a(b(x+\Delta)) \tag{2.15}$$

and, for $0 \leq i \leq nm$, *the coefficients* $h_i \in \mathbb{P}_{nm-i}[X;\mathrm{Sym}_i(X,Z)]$ *satisfy*

$$h_{i+1}(x) = D_x h_i(x) \tag{2.16}$$

Proof. By Lemma 2.1.13, note

$$f(y;\Delta) = a(y+\Delta) \quad \text{and} \quad g(x;\Delta) = b(x+\Delta)$$

Hence

$$a(b(x+\Delta)) = a(g(x;\Delta)) = f(g_0(x);\breve{g}(x;\Delta))$$

and applying Proposition 2.1.8 gives (2.15) with coefficients (2.13) and (2.14).

To prove (2.16), we use induction on i. Note that the case $i=0$ can be established by differentiating (2.15) with respect to Δ and evaluating at $\Delta = 0$. Assume (2.16) holds for $i \leq k$ and note that differentiating (2.15) $k+1$ times with respect to Δ and evaluate at $\Delta = 0$ we have

$$h_{k+1}(x) = D_\Delta^{k+1}[h_0(x+\Delta)]_{\Delta=0} = D_x^{k+1}h_0(x) = D_x h_k(x)$$

which completes the induction. □

2.2 Asymptotic polynomials

Definition 2.2.1. *Let X, Y be Banach spaces, $\gamma > 1$ a real number, $k < \gamma \leq k+1$ with $k \in \mathbb{Z}^+$, and A an arbitrary subset of X with $0 \in A$. Define "big-O notation" as follows. The symbol $O(x^\gamma)$ is used to denote any function $f : A \to Y$ with the property*

$$\|f(x)\|_Y \leq M\|x\|_X^\gamma \quad \forall x \in A$$

To emphasize the constant M, we write $O_M(x^\gamma)$. ■

Proposition 2.2.2. *Given k, γ and $A \subseteq X$ bounded (i.e. $A \subseteq B(0,R)$ for some $R \geq 1$), there exists a constant $C_{\gamma,k,R} \geq 1$ such that for any $a \in \mathbb{P}_k[X;Y]$ and any $g : A \to X$ with $g(\Delta) = O_{N_g}(\Delta^\gamma)$*

$$a^{\leq k}(\Delta + g(\Delta)) = a^{\leq k}(\Delta) + O_{N_a}(\Delta^\gamma)$$

with $N_a \leq C_{\gamma,k,R} \|a\|_{\mathbb{P}_k} N_g$.

Proof. Fix x and let $f \in \mathbb{P}_k[X,X;Y]$ with $f_i \in \mathbb{P}_{k-i}[X;\mathrm{Sym}_i(X,Y)]$ defined by $f_0(\Delta) = a(\Delta)$ and, for $0 < i \leq k$

$$f_i(\Delta) = P_i(a_1, \ldots, a_k; \Delta, 0, \ldots, 0)$$

so that

$$a^{\leq k}(\Delta + g(x)) = f^{\leq k}(\Delta; g(x)) = a^{\leq k}(\Delta) + \breve{f}^{\leq k}(\Delta; g(x)) \tag{2.17}$$

Note

$$\left\| \breve{f}^{\leq k}(\Delta; g(x)) \right\|_Y \leq \sum_{i=1}^{k} \frac{1}{i!} \|f_i(\Delta)\|_Y \|g(x)\|_Y^i \tag{2.18}$$

Furthermore using the linearity of P_i in a_j (see (2) in Proposition 2.1.11) we can factor out $\|a\|_{\mathbb{P}_k}$ leaving P_i with coefficients of operator norm ≤ 1 and since $\|\Delta\|_X \leq R$ we get $\|f_i(\Delta)\|_Y \leq \|a\|_{\mathbb{P}_k} N_{i,R}$ for constants $N_{i,R}$ depending only on P_i and R. Substituting into (2.18), we get

$$\left\| \breve{f}^{\leq k}(\Delta; g(x)) \right\|_Y \leq \underbrace{\left(\sum_{i=1}^{k} \frac{1}{i!} N_{i,R} R^{\gamma(i-1)} \right)}_{C_{\gamma,k,R}} \|a\|_{\mathbb{P}_k} N_g \|x\|_X^\gamma$$

Taking $\Delta = x$ and combining with (2.17) the result follows. ☐

We now extend the polynomial spaces defined in the previous section ($\mathbb{P}_n[X;Y]$, $\breve{\mathbb{P}}_n[X;Y]$, $\mathbb{P}_n[M,X;Y]$ and $\breve{\mathbb{P}}_n[M,X;Y]$) by adding $O(x^\gamma)$ terms. We will refer to these objects as *asymptotic polynomials*.

Definition 2.2.3. *Let X, Y be Banach spaces, A an arbitrary subset of X with $0 \in A \subseteq X$ and γ with $k < \gamma \leq k+1$. Define*

$$\mathbb{P}^\gamma[A;Y] = \left\{a^{\leq k} + a^{\geq\gamma} : a^{\leq k} \in \mathbb{P}_k[X;Y],\ a^{\geq\gamma} : A \to Y,\ a^{\geq\gamma}(\Delta) = O(\Delta^\gamma)\right\}$$

and

$$\|a\|_{\mathbb{P}^\gamma} = \sup\left\{M : \|a_i\|_{\mathrm{Sym}_i(X,Y)} \leq M, \left\|a^{\geq\gamma}(\Delta)\right\|_Y \leq M\|\Delta\|_X^\gamma\right\} \tag{2.19}$$

The spaces $\mathbb{P}^\gamma[M,A;Y]$, $\breve{\mathbb{P}}^\gamma[A;Y]$ and $\breve{\mathbb{P}}^\gamma[M,A;Y]$ are defined analogously. ■

Proposition 2.2.4. *Under the norm $\|-\|_{\mathbb{P}^\gamma}$, $\mathbb{P}^\gamma[A;Y]$, $\mathbb{P}^\gamma[M,A;Y]$, $\breve{\mathbb{P}}^\gamma[A;Y]$ and $\breve{\mathbb{P}}^\gamma[M,A;Y]$ are all Banach spaces.*

If Y is a Banach algebra then $\mathbb{P}^\gamma[A;Y]$, $\mathbb{P}^\gamma[M,A;Y]$, $\breve{\mathbb{P}}^\gamma[A;Y]$ and $\breve{\mathbb{P}}^\gamma[M,A;Y]$ are also Banach algebras.

Proof. Straightforward. □

Proposition 2.2.5. *Let X, Y and Z be Banach spaces and $k < \gamma \leq k+1$ be given. If $a = \breve{a}^{\leq k} + a^{\geq\gamma} \in \breve{\mathbb{P}}^\gamma[A;Y]$ and $b = \breve{b}^{\leq k} + b^{\geq\gamma} \in \breve{\mathbb{P}}^\gamma[B;Z]$ then, defining $C = B \cap a^{-1}(A)$, the composition $a \circ b : C \to Z$ lies in $\breve{\mathbb{P}}^\gamma[C;Z]$. Denoting*

$$a \circ b = c = \breve{c}^{\leq k} + c^{\geq\gamma}$$

one has, for $0 < i \leq k$,

$$c_i = P_i(a_1, \dots, a_i, 0, b_1, \dots, b_i) \tag{2.20}$$

with P_i *defined in* (2.6) *of Proposition 2.1.8.*

Furthermore, there exists a constant $M_\gamma \geq 1$ *such that the following inequalities hold:*

(i) For any $a \in \breve{\mathbb{P}}^\gamma[A; Z]$ *and* $b \in \breve{\mathbb{P}}^\gamma[B; Y]$

$$\|a \circ b\|_{\mathbb{P}^\gamma} \leq M_\gamma \|a\|_{\mathbb{P}^\gamma}(1 + \|b\|_{\mathbb{P}^\gamma}^\gamma) \tag{2.21}$$

(ii) For any $a \in \breve{\mathbb{P}}^\gamma[A; Z]$ *with* $a^{\geq\gamma}$ *satisfying*

$$\left\|a^{\geq\gamma}(v) - a^{\geq\gamma}(w)\right\|_Z \leq M_\gamma \|a\|_{\mathbb{P}^\gamma} \|v - w\|_Y \tag{2.22}$$

for all $v, w \in A$*, then for any* $e, f \in \breve{\mathbb{P}}^\gamma[B; Y]$

$$\|a \circ e - a \circ f\|_{\mathbb{P}^\gamma} \leq C(e, f) \|a\|_{\mathbb{P}^\gamma} \|b - d\|_{\mathbb{P}^\gamma} \tag{2.23}$$

with $C(e, f) = M_\gamma(1 + \max(\|e\|_{\mathbb{P}^\gamma}, \|f\|_{\mathbb{P}^\gamma})^{\gamma+1})$

(iii) For any $a \in \breve{\mathbb{P}}^\gamma[A; Z]$ *with* $a^{\geq\gamma}$ *having the property that there exists* $Da^{\geq\gamma} : A \to L(X, Y)$ *such that*

$$\|a(v) - a(w) - Da(w)[v - w]\|_Z \leq M_\gamma \|a\|_{\mathbb{P}^\gamma} \|v - w\|_Y^2 \tag{2.24}$$

for all $v, w \in A$*, then for any* $e, f \in \breve{\mathbb{P}}^\gamma[B; Y]$

$$\|a \circ e - a \circ f - (Da \circ f)[e - f]\|_{\mathbb{P}^\gamma} \leq D(e, f) \|a\|_{\mathbb{P}^\gamma} \|b - d\|_{\mathbb{P}^\gamma}^2 \tag{2.25}$$

with $D(e, f) = M_\gamma(1 + \max(\|e\|_{\mathbb{P}^\gamma}, \|f\|_{\mathbb{P}^\gamma})^{\gamma+2})$

Proof. Note that for any fixed x, applying Proposition 2.1.8 we get

$$\breve{a}^{\leq k}(\breve{b}^{\leq k}(\Delta) + b^{\geq\gamma}(x)) = h_0^{\leq k^2}(x;\Delta)$$

with

$$h_i(x) = P_i(a_0, \ldots, a_n; b^{\geq\gamma}(x), b_1, \ldots, b_i)$$

Using Proposition 2.2.2 we have

$$h_i(x) = \underbrace{P_i(a_0, \ldots, a_n; 0, b_1, \ldots, b_i)}_{=c_i} + h_i^{\geq\gamma}(x)$$

with $h_i^{\geq\gamma}(x) = O(x^\gamma)$. Setting $\Delta = x$, (2.20) follows with

$$c^{\geq\gamma}(x) = a^{\geq\gamma}\circ b + \sum_{i=0}^{k} \frac{1}{i!} h_i^{\geq\gamma}(x) + \sum_{i=k+1}^{k^2} \frac{1}{i!} h_i(x) x^{\otimes i}$$

To obtain (2.21), (2.23) and (2.25) note that by the definition of the norm in (2.19) it suffices to establish (2.21), (2.23) and (2.25) for each coefficient of the composition and then for the remaining $O(x^\gamma)$ term of the composition. Furthermore, estimates (2.21), (2.23) and (2.25) on $c^{\leq k}$ follow directly from (2.7), (2.8) and (2.9) in Proposition 2.1.8. Thus, we need to consider is the $c^{\geq\gamma}$ terms of the composition.

To obtain (2.21), note that Proposition 2.2.2 gives us $h_i^{\geq\gamma}(x) = O_{N_i}(x^\gamma)$ with $N_i \leq C_{\gamma,i,R}\|a\|_{\mathbb{P}_k}\|b\|_{\mathbb{P}^\gamma}$. Estimating the remaining terms of $c^{\geq\gamma} = O_M(x^\gamma)$ we get $M \leq M_{\gamma,k,R}\|a\|_{\mathbb{P}_k}(1 + \|b\|_{\mathbb{P}^\gamma}^\gamma)$ which establishes (2.21).

To obtain (2.23), note that by (2.22), we have

$$\begin{aligned}\left\|(a^{\geq\gamma}\circ e - a^{\geq\gamma}\circ f)(x)\right\|_Z &\leq C(e,f)\|a\|_{\mathbb{P}^\gamma}\|e(x) - f(x)\|_Y \\ &\leq \left(C(e,f)\|a\|_{\mathbb{P}^\gamma}\|e - f\|_{\mathbb{P}^\gamma}\right)\|x\|_X^\gamma\end{aligned}$$

Estimating the remaining terms of $c^{\geq\gamma}$, (2.23) follows.

Similarly, to obtain (2.25), note that by (2.24), we have

$$\begin{aligned}\left\|\left(a^{\geq\gamma}\circ e - a^{\geq\gamma}\circ f(Da\circ e)[e-w]\right)(x)\right\|_Z &\leq D(e,f)\|a\|_{\mathbb{P}^\gamma}\|e(x)-f(x)\|_Y^2 \\ &\leq \left(D(e,f)\|a\|_{\mathbb{P}^\gamma}\|e-f\|_{\mathbb{P}^\gamma}^2\right)\|x\|_X^{2\gamma}\end{aligned}$$

Estimating the remaining terms of $c^{\geq\gamma}$, (2.25) follows. □

Using polynomials to approximate functions, as in Taylor Theorem (Theorem 3, p.7 in [Nel69]), Proposition 2.2.5 gives us the following:

Corollary 2.2.6. *Let X, Y and Z be Banach spaces with $U \subseteq X$ and $V \subseteq Y$ open. Given functions $f : V \to Z$ and $g : U \to V$ which are C^γ, $k < \gamma \leq k+1$, the composition $f \circ g : U \to Z$ is again C^γ. Denoting the Frechet derivatives of $f(y)$ at $g(x_0)$ by $f_i(g(x_0))$ and of $g(x)$ at x_0 by $g_i(x_0)$, the derivatives of $f \circ g$ at x_0 have the form of (2.6) with $b_0 = 0$ and $a_i = f_i(g(x_0))$ and $b_i = g_i(x)$ for $0 < i \leq k$.*

Proof. Use Taylor Theorem on f and g and apply Proposition 2.2.5. □

Corollary 2.2.7. *Let X, Y and Z be Banach spaces with $U \subseteq X$ and $V \subseteq Y$ open. Given:*

(a) $F : U \times V \to Z$ *and* $g : U \to V$ *both* C^γ *with* $k < \gamma \leq k+1$.

(b) $(x_0, y_0) \in U \times V$ *with* $g(x_0) = y_0$

define $G : B(0, \epsilon) \to U \times V$ *by*

$$G(\Delta) = \begin{pmatrix} x_0 + \Delta \\ g(x_0 + \Delta) \end{pmatrix}$$

The composition $F \circ G$ *is Frechet differentiable to order* k *and, denoting the Frechet derivatives of* F *at* (x_0, y_0) *by* $F_i(x_0, y_0)$ *(or* F_i *when space is tight) and the Frechet derivatives of* g *at* x_0 *by* $g_i(x_0)$*, the derivatives* $D^i_\Delta \left[F(G(\Delta))\right]_{\Delta=0}$ *(or* $D^i \left[F \circ G\right]$ *when space is tight) have the form of* (2.6) *with* $b_0 = 0$ *and* $a_i = F_i(x_0, y_0)$ *and* $b_i = G_i$*, for* $0 < i \leq k$*, where* $G_1 = \begin{pmatrix} \mathrm{Id} \\ g_1(x_0) \end{pmatrix}$ *and* $G_i = \begin{pmatrix} 0 \\ g_i(x_0) \end{pmatrix}$ *for* $1 < i \leq k$*. In particular, for* $0 < i \leq k$ *we have*

$$D^i \left[F \circ G\right] = D_2 F(x_0, y_0)[g_i(x_0)] + Q_i(F_1, \dots, F_i; 0, G_1, \dots, G_{i-1}) \tag{2.26}$$

Proof. Follows from Corollary 2.2.6 and (4) of Proposition 2.1.11. □

Proposition 2.2.8. *Using* (2.26) *from Corollary 2.2.7, define the polynomials* $Q_i^F(x_0, y_0; b_1, \dots, b_{i-1})$ *so that*

$$Q_i^F(x_0, y_0; g_1, \dots, g_{i-1}) = Q_i(F_1, \dots, F_i; 0, G_1, \dots, G_{i-1}) \tag{2.27}$$

Note F_j*, for* $0 < j \leq i$*, are the coefficients of* Q_i^F *and we have*

$$\begin{aligned} Q_{n+1}^F = {} & D_1 D_2 F[b_n] + D_2^2 F[b_1, b_n] \\ & + D_x Q_n^F + D_y Q_n^F[b_1] + \nabla_b Q_n^F \cdot (b_2, \dots, b_n) \end{aligned} \tag{2.28}$$

Proof. Using $g^{\leq k} = b^{\leq k}$ and differentiating (2.26) the result follows. □

2.3 Formal power series

To round out our discussion of polynomials we give the following:

Definition 2.3.1. *A Formal Power Series (FPS) is a collection*

$$\{a_i \in \mathrm{Sym}_i(X,Y)\}_{i=0}^{\infty}$$

which is formally written as

$$a^{\leq\infty}(\Delta) \equiv \sum_{i=0}^{\infty} \frac{1}{i!} a_i \Delta^{\otimes i} \tag{2.29}$$

Let $\mathbb{P}_\infty[X;Y]$ denote the set of all $a^{\leq\infty}$ and define

$$\left\| a^{\leq\infty} \right\|_{\mathbb{P}_\infty} = \sup \left\{ \|a_i\|_{\mathrm{Sym}_i(X,Y)} : 0 \leq i < \infty \right\}$$

■

Remark 2.3.2. *Because no assertion is made as to the convergence of this power series, one must treat* (2.29) *as a formal object. Since we can only "evaluate" $y_0^\infty(\Delta)$ for $\Delta = 0$ the ℓ_∞ type norm $\left\| a^{\leq\infty} \right\|_{\mathbb{P}_\infty}$ no longer corresponds to a supremum norms as with $\mathbb{P}_{\leq n}[X;Y]$ and $\mathbb{P}^\gamma[X;Y]$. Also, even on the formal level the composition of formal power series results in infinite sums for finite order coefficients and thus cannot be defined.*

However, formal power series provide a useful formalism to describe sequences of polynomials of the form $p_n(\Delta) = \sum_{i=0}^{n} y_i \Delta^{\otimes i}$. Examples include P_i, defined in (2.6)*, Q_i, defined in* (4) *of Proposition 2.1.11, Q_i^F, defined Proposition 2.2.8, and h_i, defined in* (2.14)*, (all viewed as a formal power series in n, not i, see Remark 2.1.10).*

Chapter 3

Whitney Regularity

Informally, a function $f : A \subseteq X \to Y$ is C^γ Whitney differentiable, $k < \gamma \leq k+1$, if one can find suitable substitutes $\{f_i\}_{i=0}^k$ for the derivatives of f so that the estimates of the classical C^γ Taylor's theorem, which make sense on arbitrary domains, are satisfied. One can think of this substitute $\{f_i\}_{i=0}^k$ for the derivatives as prescribing the k-jet of f.

While many of the classical notions of differentiability continue to hold for Whitney differentiable functions, e.g. the product rule and chain rule, when moving to arbitrary domains such simple results as the uniqueness of the derivatives or the k-jet of f need not hold.

While classically, the i-th derivatives of a function automatically satisfy the estimates of the classical $C^{\gamma-i}$ Taylor's theorem, for Whitney differentiability, this condition must be imposed as additional conditions on the k-jet of f. To establish the Whitney regularity of a function, it is often relatively easy to obtain the C^γ Taylor estimates on the function while the $C^{\gamma-i}$ Taylor estimates for the Whitney derivatives are often more difficult to establish. To this end, we present two "Whitney Verification Lemmas" which establish the $C^{\gamma-i}$ Taylor estimates on the derivatives from the C^γ estimate of the function provided one have either certain additional relationships among the $\{f_i\}_{i=1}^k$ (Lemma 3.1.10) or additional conditions on the set A (Lemma 3.2.6).

Section 3.1 presents the definition of $C^{\gamma}_{Wht}(A, Y)$, the set of C^{γ} Whitney differentiable functions. Here we explore some of the basic consequences of this definition. In particular, restricting this definition to the interior of A one recovers the classical notion of C^{γ} regularity for functions defined on open sets (see Proposition 3.1.5). We also establish that the composition of two C^{γ} Whitney differentiable functions is again C^{γ} Whitney (see Theorem 3.1.8). In Section 3.2 under some reasonable conditions on A we can ensure the uniqueness of Whitney derivatives can be established (see Proposition 3.2.3 and Remark 3.2.4). Finally, Section 3.3 presents the Whitney Extension theorem, which extends C^{γ} Whitney regular functions in finite dimensions (i.e. $\mathbb{R}^n$) to classically C^{γ} regular functions.

3.1 The definition and some consequences

The following definition generalizes the spaces $\mathrm{Lip}(\gamma, A)$ as defined on p. 176 in [Ste70] to functions with domain and range in arbitrary (infinite dimensional) Banach spaces. (Of course, theorems using the standard notion of Whitney regularity, e.g. the Whitney Extension theorem discuss in Section 3.3, will apply if we restrict A to be a closed finite dimensional subset.)

Definition 3.1.1. *Let X, Y be Banach spaces, $A \subseteq X$ arbitrary, and $\gamma > 1$ with $k < \gamma \leq k + 1$ for $k \geq 1$ a positive integer. Define $C^{\gamma}_{Wht}(A, Y)$ to be the collection of functions $f : A \to Y$ with the property that for some choice* *of functions $f_i : A \to \mathrm{Sym}_i(X,Y)$ for $0 \le i \le k$ with $f = f_0$ there exists a positive constant M such that for every $0 \le i \le k$*

$$\|f_i(x)\|_{\mathrm{Sym}_i(X,Y)} \le M \tag{3.1}$$

and for every $0 \le i \le k$ and every $x, x+\Delta \in A$

$$\left\| f_i(x+\Delta) - f_i^{\le k-i}(x;\Delta) \right\|_{\mathrm{Sym}_i(X,Y)} \le M \|\Delta\|_X^{\gamma - i} \tag{3.2}$$

Define

$$\|f\|_{C^\gamma_{Wht}} = \inf \{M : (3.1) \textit{ and } (3.2) \textit{ hold}\}$$

Given $f \in C^\gamma_{Wht}(A,Y)$ we say f is C^γ Whitney in A with Whitney derivatives $\{f_i\}_{i=0}^k$. We also refer to the collection $\{f_i\}_{i=0}^k$ as the k-jet of f.

Define

$$C^\infty_{Wht}(A,Y) \equiv \bigcap_{\gamma > 0} C^\gamma_{Wht}(A,Y)$$

Given $f \in C^\infty_{Wht}(A,Y)$ we say the function f is C^∞ Whitney in A with Whitney derivatives $\{f_i\}_{i=0}^\infty$. We refer to the collection $\{f_i\}_{i=0}^k$ as the ∞-jet of f. ■

Remark 3.1.2. (Whitney's formulation) *The original definition given by Whitney in [Whi34] for "functions of class C^m in A" took $m = \gamma \in \mathbb{Z}^+ \cup \{\infty\}$, $X = \mathbb{R}^n$, $Y = \mathbb{R}$ and A closed. Furthermore, conditions (3.1) and (3.2) were replaced by the condition that for any $x' \in X$ and $\epsilon > 0$ there exists a $\delta > 0$ such that for any $x, x+\Delta \in A \cap B(x',\delta)$ one has*

$$\left\| f_i(x+\Delta) - f_i^{\le m-i}(x;\Delta) \right\|_{\mathrm{Sym}_i(X,Y)} \le \epsilon \|\Delta\|_X^{m-i} \tag{3.3}$$

for all $i \le m$.

Note that with $m < \delta \leq m+1$, all the functions in $Lip(\delta, A)$ as defined by Stein are of "class C^m in A" by Whitney's definition. Similarly, if A is compact then any function which is of "class C^m in A" a la Whitney is also $Lip(\eta, A)$ for all $\eta \leq m$.

Remark 3.1.3. (Big-O notation) *Whitney and Stein both write*

$$f_i(x+\Delta) = f_i^{\leq k-i}(x;\Delta) + R_i(x, x+\Delta) \tag{3.4}$$

and use $R_i(x, x+\Delta)$ to express (3.2) *or* (3.3)*. Using "big-O" notation (see Definition 2.2.1) developed in Chapter 2 we could write* (3.4) *as*

$$f_i(x+\Delta) = f_i^{\leq k-i}(x;\Delta) + O_M(\|\Delta\|_X^{\gamma-i}) \tag{3.5}$$

thus expressing condition (3.2)*. In a similar manner one can use a modified "little-o" notation to express Whitney's original notion of C^m in A.*

Remark 3.1.4. (A arbitrary and infinite) *Note in Definition 3.1.1 the set A need not be closed. Furthermore, the linear spaces X and Y need not be finite dimensional. Only when we consider the Whitney extension theorem (Theorem 3.3.1) in Section 3.3 will we require A to be closed and X to be finite dimensional.*

In applications, working with Whitney regularity in infinite dimensions on arbitrary sets is useful since one can establish the Whitney regularity of the implicit function in this setting and then restrict to a finite dimensional closed set of parameters and use the Whitney extension theorem (Theorem 3.3.1) to obtain measure estimates.

We now present some of the basis consequences that follow from our definition of $C^{\gamma}_{Wht}(A, Y)$.

Proposition 3.1.5. *Let* $\operatorname{int}(A)$ *denote the interior of the set* A*. For any* $f \in C^{\gamma}_{Wht}(A, Y)$ *we have the following:*

1. *If* $\gamma \notin \mathbb{Z}$*, the function* $f|_{\operatorname{int}(A)}$ *is* C^{γ} *in the classical sense.*

2. *If* $\gamma = k + 1$ *then* $f|_{\operatorname{int}(A)}$ *is* C^k *with Lipschitz continuous derivatives (often denoted* $C^{k,1}$*).*

Furthermore, on $\operatorname{int}(A)$ *the estimates* (3.2) *for* $0 < i \leq k$ *follow from* (3.2) *for* $i = 0$.

Proof. Apply the converse Taylor theorem (see e.g. p. 6 [AR67]). □

Proposition 3.1.6. *The set* $C^{\gamma}_{Wht}(A, Y)$ *with* $\|-\|_{C^{\gamma}_{Wht}}$ *is a Banach space. If* Y *is a Banach Algebra then so is* $C^{\gamma}_{Wht}(A, Y)$.

Proof. Straightforward (see, e.g. p. 176 in [Ste70]). □

Proposition 3.1.7. *If* f *is* $C^{\gamma}_{Wht}(A, Y)$ *with Whitney derivatives* $\{f_i\}_{i=0}^{k}$ *then its Whitney derivatives* f_n *are* $C^{\gamma-n}_{Wht}(A, \operatorname{Sym}_n(X, Y))$ *with Whitney derivatives* $\{f_{n+i}\}_{i=0}^{k-n}$.

The converse is not true. Namely, there exists $f \in C^{\gamma}_{Wht}(A, Y)$ *whose* k*-jets* $\{f_i\}_{i=1}^{k}$ *are* $C^{\infty}_{Wht}(A, \operatorname{Sym}_n(X, Y))$ *but* f *is not* $C^{\eta}_{Wht}(A, Y)$ *for any* $\eta > \gamma$*. (See Proposition 3.1.11 for conditions under which the converse does hold.)*

Proof. The first assertion follows from the definitions. For an example of the second, see [Whi34]. □

Theorem 3.1.8. (Whitney Composition)

Given $k < \gamma \leq k+1$, X, Y, Z be linear spaces, $A \subseteq X$ and $B \subseteq Y$ let

$$g : A \to Y \quad \text{and} \quad f : B \to Z$$

and define $C \equiv A \cap g^{-1}(B) \subseteq X$ and

$$h \equiv f \circ g : C \to Z$$

If $g \in C^{\gamma}_{Wht}(A, Y)$ with Whitney derivatives g_i and $f \in C^{\gamma}_{Wht}(B, Z)$ with Whitney derivatives f_i then $h \in C^{\gamma}_{Wht}(C, Z)$ with Whitney derivatives $\{h_i\}_{i=0}^{k}$ given by $h_0(x) = f_0(g_0(x))$ and, for $0 < i \leq k$,

$$h_i(x) = P_i(f_1(g(x)), \ldots, f_i(g(x)); 0, g_1(x), \ldots, g_i(x)) \tag{3.6}$$

with P_i as defined in (2.5) and (2.6) in Proposition 2.1.8.

Furthermore, there exists a constant $M_\gamma \geq 1$ such that the following inequalities hold:

(i) For any $f \in C^{\gamma}_{Wht}(B; Z)$ and $g \in C^{\gamma}_{Wht}[A; Y]$

$$\|f \circ g\|_{C^{\gamma}_{Wht}} \leq M_\gamma \|f\|_{C^{\gamma}_{Wht}} (1 + \|g\|^{\gamma}_{C^{\gamma}_{Wht}}) \tag{3.7}$$

(ii) For any $f \in C^{\gamma+1}_{Wht}[B; Z]$, $g_1, g_2 \in C^{\gamma}_{Wht}[A; Y]$

$$\|f \circ g_1 - f \circ g_2\|_{C^{\gamma}_{Wht}} \leq C(e, f) \|f\|_{C^{\gamma+1}_{Wht}} \|g_1 - g_2\|_{C^{\gamma}_{Wht}} \tag{3.8}$$

with $C(g_1, g_2) = M_\gamma (1 + \max(\|g_1\|_{C^{\gamma}_{Wht}}, \|g_2\|_{C^{\gamma}_{Wht}})^{\gamma})$

(iii) For any $f \in C^{\gamma+1}_{Wht}[B;Z]$ with y $g_1, g_2 \in C^{\gamma}_{Wht}[A;Y]$

$$\|f \circ g_1 - f \circ g_2 - (\mathcal{WD} f \circ g_2)[g_1 - g_2]\|_{C^{\gamma}_{Wht}} \leq C(e,f)\|f\|_{\mathbb{P}^{\gamma+1}}\|g_1 - g_2\|^2_{C^{\gamma}_{Wht}} \tag{3.9}$$

with $C(e,f)$ as above.

Remark 3.1.9. *See [dlLO99] for a more detailed discussion of the regularity of the composition functional (although not done for Whitney differentiability).*

If we work in with closed sets in finite dimensions, the extension theorems presented in Section 3.3 would allow us to trivially conclude that the composition of two C^{γ}_{Wht} functions is again C^{γ}_{Wht}. However, as mentioned in Remark 3.1.4, we can use the definition of $C^{\gamma}_{Wht}(A,Y)$ to establish this in a more general setting. In the proof of the general case of Theorem 3.1.8, we use the following:

Lemma 3.1.10. (Whitney Verification Lemma I)
Let $n < \eta \leq n+1$, $A \subseteq X$, $U \subseteq Y$ and $g_i \in C^{\eta-i}_{Wht}(A \times U, Y)$ for $0 < i \leq n$ be given with

$$g_i : A \times U \to U$$

Given $f : A \to U$, define $f_0(x) = f(x)$ and $f_i(x) = g_i(x, f(x))$. If, for all $(x,y) = (x, f(x))$, one has

$$g_{i+1}(x,y) = \mathcal{WD}_x(g_i(x,y)) + \mathcal{WD}_y(g_i(x,y))[-, g_1(x,y)] \tag{3.10}$$

and

$$\left\|f_0(x+\Delta) - f_0^{\leq n}(x;\Delta)\right\|_Y \leq M\|\Delta\|^{\eta}_X \tag{3.11}$$

for all $x, x+\Delta \in A$*, then* $f \in C^{\eta}_{Wht}(A, Y)$ *with Whitney derivatives* $f_i(x) = g_i(x, f(x))$*.*

This lemma is useful independent of Theorem 3.1.8. It is used in Section 6.2 and Section 7.2 to establish the Whitney regularity of the implicit function. Also, note that this lemma establishes conditions under which the converse for Proposition 3.1.7 is true, i.e.:

Proposition 3.1.11. *Given* $n < \eta \leq n+1$ *and a collection of functions* $\{f_i\}_{i=0}^{n}$ *with* $f_i \in C^{\eta-i}_{Wht}(A, \mathrm{Sym}_i(X, Y))$ *for* $0 < i \leq n$ *provided* $f_{i+1}(x) = \mathcal{WD}_x f_i(x)$ *for* $0 < i \leq n$*, and* (3.11) *holds, then* $f \in C^{\eta}_{Wht}(A, Y)$*.*

Proof. If take $g_i(x, y) = f_i(x)$ and apply Lemma 3.1.10. □

The proofs of Lemma 3.1.10 and Theorem 3.1.8 are related in the following sense. To prove Theorem 3.1.8 for some $k < \gamma \leq k+1$ we use Lemma 3.1.10 for $\eta = k$. Likewise to prove Lemma 3.1.10 for $n < \eta \leq n+1$ we use Theorem 3.1.8 for $\gamma = \eta - 1$. Thus, we give the proofs of Theorem 3.1.8 and Lemma 3.1.10 simultaneously.

Proof of Theorem 3.1.8 and Lemma 3.1.10

We establish both results by induction. In the base cases $k = 0$ in Theorem 3.1.8 and $n = 0$ in Lemma 3.1.10 both results are immediate.

Assume that Theorem 3.1.8 and Lemma 3.1.10 hold for $n, k \leq N$. We will first establish Lemma 3.1.10 for $N+1 < \eta \leq N+2$ and then use this to establish Theorem 3.1.8 for $N+1 < \gamma \leq N+2$.

Establishing Lemma 3.1.10 for $N + 1 < \eta \leq N + 2$:

Let $f : A \to U$ and $g_i \in C^{\eta-i}_{Wht}(A \times U, Y)$, $0 < i \leq N + 1$ be given as in Lemma 3.1.10 for $N + 1 < \eta \leq N + 2$. Note that the nature of the hypothesis in Lemma 3.1.10 allow us to use it to conclude that $f \in C^{N+1}_{Wht}(A, Y)$ with Whitney derivatives $f_i(x) = g_i(x, f(x))$ for $0 < i \leq N$. For $0 < i \leq N + 1$, since $g_i \in C^{\eta-i}_{Wht}$ and $\eta - i \leq N + 2 - i \leq N + 1$, by Theorem 3.1.8, we have that $f_i(x) = g_i(x, f(x))$ is $C^{\eta-i}_{Wht}$ for all $0 < i \leq N + 1$. Furthermore, for $0 < i < N+1$ note the first Whitney derivative of $f_i(x) = g_i(x, f(x))$ will have the form

$$\mathcal{WD}_x f_i(x) = \mathcal{WD}_x g_i(x, f(x)) + \mathcal{WD}_y g_i(x, g(x))[\mathcal{WD}_x f(x)]$$

Since $\mathcal{WD}_x f(x) = g_1(x, f(x))$, applying (3.10) we get, for $0 < i < N + 1$

$$\mathcal{WD}_x f_i(x) = f_{i+1}(x) \tag{3.12}$$

By assumption we have (3.2) for $i = 0$ For $0 < i < N + 1$ since $f_i(x) = g_i(x, f(x))$ is $C^{\eta-i}_{Wht}$ with (3.12) the estimates (3.2) for $f_i \in C^{\eta-i}_{Wht}$ are exactly the estimates for f_0 we need to establish (3.2) for $0 < i < N + 1$. Finally, the fact that $f_{N+1}(x) = g_{N+1}(x, f(x))$ is $C^{\eta-(N+1)}_{Wht}$ is sufficient to establish (3.2) for $i = N + 1$. This establishes Lemma 3.1.10 for $N + 1 < \eta \leq N + 2$. $\square$

Establishing Theorem 3.1.8 for $N + 1 < \gamma \leq N + 2$:

As in Theorem 3.1.8 with $N+1 < \gamma \leq N+2$, let $g \in C^{\gamma}_{Wht}(A, Y)$ with $(N+1)$-jet $\{g_i\}_{i=0}^{N+1}$ and $f \in C^{\gamma}_{Wht}(B, Z)$ with $(N + 1)$-jet $\{f_i\}_{i=0}^{N+1}$ be given. For fixed

x, applying Proposition 2.2.5 we have $f \circ g = h = h_0^{\leq k} + h^{\geq \gamma} \in P^\gamma[C; Z]$ with (2.20) giving us $h_i(x)$ as in (3.6). Estimates (3.7), (3.8) and (3.9) follow directly from (2.21), (2.23) and (2.25). Since $h^{\geq \gamma}(x) = O(x^\gamma)$, we also have (3.2) for $i = 0$. Furthermore, since

$$f_i \in C^{\gamma-i}_{Wht}(B, \mathrm{Sym}_i(Y, Z)) \quad \text{and} \quad g_i \in C^{\gamma-i}_{Wht}(C, \mathrm{Sym}_i(X, Y))$$

by Theorem 3.1.8 we, for $0 < i \leq N$, the functions h_i defined in (3.6) have $h_i \in C^{\gamma-i}_{Wht}(C, \mathrm{Sym}_i(X, Z))$. To compute $\mathcal{WD}_x h_i \in C^{\gamma-(i+1)}_{Wht}(C, \mathrm{Sym}_{i+1}(X, Z))$ for $0 \leq i \leq N$, note that the Whitney derivatives computed by Theorem 3.1.8 will have the same form as in Theorem 2.1.14, and thus $\mathcal{WD}_x h_i(x) = h_{i+1}(x)$. Applying Lemma 3.1.10 with $N + 1 < \eta = \gamma \leq N + 2$, we conclude $h = f \circ g \in C^\gamma_{Wht}(C, Z))$. □

This establishes Lemma 3.1.10 for $N + 1 < \eta \leq N + 2$ and Theorem 3.1.8 for $N + 1 < \gamma \leq N + 2$. Hence by induction we have Lemma 3.1.10 for every $n < \eta \leq n + 1$ and Theorem 3.1.8 for every $k < \gamma \leq k + 1$. □

3.2 Conditions for uniqueness of Whitney derivatives

Remark 3.2.1. *Note that the Whitney derivatives, $\{f_i\}_{i=0}^k$, need not be unique (for example, if $f : A = \{(x, 0)\} \subseteq \mathbb{R}^2 \to \mathbb{R}$ then the Whitney partial derivative in the y direction, f_y, is completely arbitrary). To avoid ambiguity, when speaking of $f \in C^\gamma_{Wht}(A, Y)$ we will usually consider a specific k-jet $\{f_i\}_{i=0}^k$. In this context, we use $\mathcal{WD}^i_x f(x)$ to refer to $f_i(x)$.*

Remark 3.2.2. *In some sense, Proposition 3.1.5 and the example given in Remark 3.2.1 represent the extremes of density a about a point x in A. Proposition 3.1.5 illustrates that given any $x \in A$, with enough points in A close to x the f_i are unique. Furthermore, in this case one can obtain estimates* (3.2) *for $0 < i \leq k$ from* (3.2) *with $i = 0$. For intermediate cases of density a about a point x in A one can still obtain uniqueness (e.g. Proposition 3.2.3) as well as estimates* (3.2) *for $0 < i \leq k$ from* (3.2) *with $i = 0$ (e.g. Lemma 3.2.6).*

Proposition 3.2.3. *Given $f \in C^{\gamma}_{Wht}(A, Y)$, for any point $x \in A$ and $v \in X$, if for some $\sigma(t)$ with $\sigma(t)/t \to 0$ as $t \to 0$ the set*

$$\{t : x + tv + w(t) \in A, \|w(t)\|_X \leq \sigma(t)\}$$

has 0 *as an accumulation point then $f_1(x)[v]$ are unique.*

Proof. If f_1 and $\tilde{f}_1$ are possible Whitney derivatives of f, note using (3.2) with $i = 0$ we have

$$f_1(x)[v] - \tilde{f}_1(x)[v] = \frac{\tilde{f}_1(x)[w(t)] - f_1(x)[w(t)] + O(\|tv + w(t)\|_X^{\min(2,\gamma-1)})}{t}$$

For small values of t the RHS is arbitrarily small. Since the LHS does not depend on t we have $f_1(x)[v] = \tilde{f}_1(x)[v]$. □

Remark 3.2.4. *Around a given point $x \in A$ one can formulate "higher order" density conditions on the set A, similar to those given in Proposition 3.2.3, which ensure additional uniqueness of $f_i(x)$. These "higher order" density conditions are related to the conditions for obtaining* (3.2) *for $0 < i \leq k$ from* (3.2) *for $i = 0$ (see Definition 3.2.5 and Lemma 3.2.6).*

Definition 3.2.5. *We say that a point* $x \in A$ *has the* γ density property, $k < \gamma \leq k+1$ *if* x *is a limit point of* A *and there exists positive constants* ϵ, M *and* $\lambda_1, \ldots, \lambda_k$ *distinct such that for any* $\|\Delta\|_X \leq \epsilon$ *with* $x + \Delta \in A$ *and any* z *with* $\|z\|_X = \|\Delta\|_X$ *one can find* $w_i \in X$, $i = 1, \ldots, k$, *with* $\|w_i\|_X \leq M\|\Delta\|_X^\gamma$ *such that* $x + \Delta + \lambda_i z + w_i \in A$ *(or* $x + \Delta - \lambda_i z + w_i \in A$*).*

We say the set A *has the* γ density property *if each* $x \in A$ *has the* γ *density property for the same choice of* M *and* $\lambda_1, \ldots, \lambda_k$. ■

Lemma 3.2.6. (Whitney Verification Lemma II)

Let $k < \gamma \leq k+1$, *and assume* $A \subseteq X$ *has the* γ *density property (see Definition 3.2.5 in Section 2.2). Let* $f : A \to Y$ *and*

$$f_i : A \to \mathrm{Sym}_i(X, Y) \quad \textit{for} \quad 0 \leq i \leq k$$

with $f = f_0$ *and a positive constant* M *such that for* $0 \leq i \leq k$

$$\|f_i(x)\|_{\mathrm{Sym}_i(X,Y)} \leq M$$

If, for every $x, x + \Delta \in A$, *one has*

$$\left\| f_0(x + \Delta) - f_0^{\leq k}(x; \Delta) \right\|_Y \leq M\|\Delta\|_X^\gamma$$

then in fact one has (3.2) *for* $0 < i \leq k$ *and hence* f *is* C^γ *Whitney in* A *with Whitney derivatives* $\{f_i\}_{i=0}^k$, *i.e.* $f \in C^\gamma_{Wht}(A, Y)$.

Proof of Lemma 3.2.6

Following Converse Taylor Theorem (Theorem 3, p.7 in [Nel69]), we proceed by induction on i. The base case of (3.2) for $i = 0$ is assumed.

Assume that we have established (3.2) for $i \le m-1 < k$. To establish (3.2) for $i = m$, note that since A has the γ density property for any x, Δ and z with $x, x+\Delta \in A$ and $\|z\| = \|\Delta\|$ we can find w_i with $\|w_i\|_X \le M\|\Delta\|_X^\gamma$ such that $x + \Delta + \lambda_i z + w_i \in A$ for $i = 1, \dots, k$ (or $x + \Delta - \lambda_i z + w_i \in A$ for $i = 1, \dots, k$, which can be though of as as a special case of the first with $-z$ in place of z). Using that (3.2) holds for $m-1$, note

$$f_{m-1}(x + \Delta + \lambda_i z + w_i) = f_{m-1}^{k-m+1}(x + \Delta; \lambda_i z) + O(\|\Delta\|_X^{\gamma-m+1})$$

and

$$f_{m-1}(x + \Delta + \lambda_i z + w_i) = f_{m-1}^{k-m+1}(x; \Delta + \lambda_i z) + O(\|\Delta\|_X^{\gamma-m+1})$$

hence after subtracting the two expressions and collecting the coefficients of z, we have

$$g_{m-1}(\Delta) + \cdots + \lambda_i^{k-m+1} g_k(\Delta)[z]^{\otimes k-m+1} = O(\|\Delta\|^{\gamma-m+1}) \tag{3.13}$$

with

$$g_j(\Delta) \equiv \frac{f_j(x+\Delta) - f_j^{\le k-j}(x;\Delta)}{j!} \tag{3.14}$$

Putting together (3.13) with the various λ_i we have

$$\begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \dots & \lambda_1^{k-m+1} \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^{k-m+1} \\ 1 & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_{k-m} & \lambda_{k-m}^2 & \dots & \lambda_{k-m}^{k-m+1} \end{pmatrix} \begin{pmatrix} g_{m-1}(\Delta) \\ g_m(\Delta)[z] \\ \vdots \\ g_k(\Delta)[z]^{\otimes k-m+1} \end{pmatrix} = \begin{pmatrix} O(\|\Delta\|^{\gamma-m+1}) \\ O(\|\Delta\|^{\gamma-m+1}) \\ \vdots \\ O(\|\Delta\|^{\gamma-m+1}) \end{pmatrix}$$

This matrix is a Vandermonde matrix and since the $\lambda_1, \lambda_2, \dots, \lambda_{k-m}$ are distinct, it can be inverted this matrix. Furthermore, since λ_i are fixed, the norm of the inverse matrix is bounded. From this, we obtain

$$g_m(\Delta)[z] = O(\|\Delta\|^{\gamma-m+1})$$

and since this hold for every z with $\|z\|_X = \|\Delta\|_X$ we get

$$f_m(x+\Delta) - f_m^{\leq k-j}(x;\Delta) = O(\|\Delta\|^{\gamma-m})$$

This holds for any any $x, x+\Delta \in A$, hence we have established (3.2) for $i = m$. This completes the induction and establishes (3.2) for $0 \leq i \leq k$. □

3.3 Extension theorems

Proposition 3.1.5 illustrates that, restricted to the interior of A, the notion of Whitney regularity coincides with the classical notions of regularity. If A is finite dimensional, this correspondence with the classical notions of regularity extends to a neighborhood of A. In particular, we have the following powerful and important result:

Theorem 3.3.1. (Whitney Extension Theorem)
Let $k \in \mathbb{Z}^+$, $k < \gamma \leq k+1$ and A a closed subset of $\mathbb{R}^n$. Then there is a continuous linear mapping

$$\mathcal{E}_k : C^\gamma_{Wht}(A,Y) \to C^\gamma_{Wht}(X,Y) = C^\gamma(X,Y) = \Lambda_\gamma(X,Y)$$

such that:

(i) $[\mathcal{E}_k f](x) = f(x)$ for all $x \in A$.

(ii) $D^i[\mathcal{E}_k f](x) = f_i(x)$ *for* $i \leq k$ *and* $x \in A$

(iii) *The operator norm of* $\mathcal{E}_k$ *is independent of the set* A*.*

Alternatively, using the extension described by Whitney in [Whi34], we have the following. Given $f \in C^\gamma_{Wht}(A, Y)$ *(or* f *of class* C^k *in* A*, see Remark 3.1.2; here* $k = \infty$ *is also permissible) we can find a function* $F(x)$ *of class* C^k *in* $\mathbb{R}^n$ *such that:*

(I) $F(x) = f(x)$ *for all* $x \in A$*.*

(II) $D_i F(x) = f_i(x)$ *for all* $i \leq k$ *and* $x \in A$*.*

(III) $F(x)$ *is analytic for* $x \in \mathbb{R}^n - A$*.*

Proof. For the case $Y = \mathbb{R}$ see Theorem 4 on p. 177 of [Ste70] (note *(ii)* follows from 2.3.2 a' on p. 187) and Theorem I in [Whi34]. Given Y any other Banach space, these same constructions and estimates can be followed. □

Remark 3.3.2. *The key ingredient of the proof of the extension theorems in [Ste70] and [Whi34] is to obtain a decomposition of* $X - A$ *into cubes whose size is comparable with the distance to the boundary of* A*. This is accomplished in finite dimensions using a Calderon-Zygmund decomposition. However, in an infinite dimensional Banach space it is not clear when such a good decomposition exists (see Chapter V of [KM97] for a discussion of Whitney Extension Theorems in more general settings).*

Chapter 4

Abstract Setup

In this chapter, we describe the general setup in which we will work. As in [Zeh75], we consider two types of one parameter families of Banach spaces, X_σ, which are abstractions of spaces of analytic functions, and X_0^ℓ, which are abstractions of the usual spaces of C^r functions for $\ell = r \notin \mathbb{Z}$. In this abstract setting, the "smooth" X_0^q spaces are obtained as subsets of X_0 described by their approximation properties in X_σ (see Definition 4.1.4). Smoothing operators (see Definition 4.1.9 and 4.1.13) which can be sued to explicitly construct approximations in X_σ also play an important role. The complete presentation of the one-parameter families of Banach spaces, the construction of these "approximation spaces" and the definition of the "smoothing operators" can be found in Section 4.1.

Using these one parameter families of Banach spaces, as in [Zeh75] we consider functionals $\mathcal{F} = \mathcal{F}(x, y)$ of two variables, x (which we think of as the independent variable) and y (which we think of as the dependent variable). To solve the implicit equation $\mathcal{F}(x, g(x)) = 0$ locally near some (x_0, y_0) with $\mathcal{F}(x_0, y_0) = 0$, we require the functional $\mathcal{F}$ satisfy several hypothesis. Informally, we will assume:

1. $\mathcal{F}$ is continuous in x and y and differentiable with respect to y (see conditions (F.A0) and (F.A1) in Section 4.3).

2. On some set $\mathcal{C} \times V$ the differential of $\mathcal{F}$ with respect to y has an approximate right inverse (see condition (F.A2) in Section 4.3). As in [Zeh75], this approximate right inverse, while not bounded when viewed as a mapping between spaces at the same scale, becomes bounded when viewed as a mapping between spaces at different scale (see Remark 4.0.3). These bounds must satisfy certain quantitative estimates (namely the Brjuno-Rüssmann condition discuss in Section 4.2).

Under these hypothesis (which are described in detail in Section 4.3), the local existence of a solution g to the implicit equation $\mathcal{F}(x, g(x)) = 0$ follows for $x \in \mathcal{C}$ near any (x_0, y_0) with $\mathcal{F}(x_0, y_0) = 0$ (see Corollary 6.1.2 and Corollary 7.1.3). Under some additional hypothesis on $\mathcal{F}$ and R (see (F.W1), (F.W2) in Section 4.3) we can further establish the Whitney regularity with respect to parameters of the implicit function g (see Remark 4.0.3 for the terminology "regularity with respect to parameters,"; for the results see Theorems 6.2.1, 6.2.3, 7.2.1 and 7.2.2).

The main differences between the hypotheses used in this work and those found in [Zeh75] are:

1. We only require an approximate right inverse when the x variable ranges over a not-necessarily-open-set (see (F.A2) in Section 4.3).

2. We require weaker quantitative estimates on the bounds for the quadratic remainder and approximate right inverse (in applications, this in turn allows us to consider weaker "Diophantine" conditions). Specifically, we require Ω_Q, Ω_R and Ω_A, satisfy the Brjuno-Rüssmann condition discuss in Section 4.2 (see (F.A1) and (F.A2) in Section 4.3).

3. In the smooth setting (i.e. $X_0^q \times Y_0^q$, see Definition 4.1.4) we eliminate the requirement that approximate solutions are analytic (i.e. they lie in $X_\sigma \times Y_\sigma$). Instead, we require an additional compatibility condition between the functional and the smoothing operators (see (F.S4) in Section 4.3 and Theorem 7.1.1).

Remark 4.0.3. *Some remarks about terminology are in order.*

In the applications we consider, the one parameter families X_σ and X_0^q are often spaces of functions with the scale parameters σ and q measuring their regularity (for analytic functions the σ measures the domain of analyticity). Furthermore, in these applications various linear operators (such as differentials of the functional and the corresponding approximate inverses) have the property that they are bounded when mapping a space at one scale into a space at a different scale. That is, if L is the linear operator under consideration, we would have $L : X_\sigma \to Y_{\sigma'}$ bounded only for $0 \leq \sigma' < \sigma \leq 1$. Due to this association of the scale parameter with regularity, this phenomena is referred to as a "loss of regularity/smoothness" or, in the analytic case, as "loss of domain."

While thinking of σ as a regularity parameter is natural, it can become very confusing since we also consider the regularity of functionals acting between these spaces of functions. To illustrate, consider $f : X_\sigma \to Y_\tau$ for some fixed σ and τ. We can think of σ and τ as measuring the regularity of the functions $x \in X_\sigma$ and $f(x) \in Y_\tau$, but we also want to consider the regularity of the functional f as a map between the Banach spaces X_σ and Y_τ.

When referring to a particular Banach space in the one parameter family X_σ or X_0^q, i.e. fixing σ or q, we will avoid referring to σ or q in terms of regularity and speak of the space at a given "scale." Whenever referring to regularity of a functional acting between one parameter family of Banach spaces, such as the regularity of the functional f as a map between the Banach spaces X_σ and Y_τ above, we will speak of "regularity with respect to parameters."

Remark 4.0.4. *When working in one parameter families of Banach spaces, we often are able to gain desirable properties, such as continuity, differentiability or inverses (see (FA.0), (F.A1), (F.A2) in section 4.3) by sacrificing some arbitrary amount of scale. In addition to the semantic issues described in Remark 4.0.3, this arbitrary loss of scale can cause a fair amount of difficulty with overly burdensome notation.*

For example, in Chapter 5, we use an iterative definition to obtain the coefficients of a polynomial approximate solution. Each step in the iteration uses an (unbounded) inverse (see (F.P2) in section 4.3) and thus at each step we have to loose an arbitrary amount of scale. The overall domain loss can still be arbitrary. Provided the iterative process was only repeated a finite number *of times, when we compute estimates we can simply assume the loss at each step was* $(\sigma - \sigma')/n$ *with* $(\sigma - \sigma')$ *being the overall loss. However, when we have an infinite number of steps (such as in the modified Newton method used in the proof of Theorem 6.1.1) significantly more care must be taken with the domain loss and estimates at each stage.*

4.1 Scales of spaces, the X_0^q spaces and C^ω smoothing

Following [Zeh75], let X_σ, Y_σ and Z_σ be three one parameter families of Banach spaces indexed by σ with $0 \leq \sigma \leq 1$, such that for $0 \leq \sigma' < \sigma \leq 1$ one has

$$X_0 \supseteq X_{\sigma'} \supseteq X_\sigma \supseteq X_1 \tag{4.1}$$

and the inclusion of X_σ into $X_{\sigma'}$ is a bounded linear operator with operator norm ≤ 1, i.e.

$$\|x\|_{X_{\sigma'}} \leq \|x\|_{X_\sigma} \tag{4.2}$$

for all $x \in X_\sigma$ (analogously for Y_σ and Z_σ).

Remark 4.1.1. *Note that re-parameterizing the scale parameter* σ*, i.e. taking* ϕ *to be an increasing function with* $\phi(0) = 0$ *and* $\phi(1) = 1$*, the one parameter family of Banach spaces* $\mathbf{X}_\sigma = X_{\phi(\sigma)}$ *also satisfies* (4.1) *and* (4.2)*. (To keep the approximation spaces* X_0^q*, described in Definition 4.1.4, from (drastically) changing, we will require the re-parameterization* ϕ *to be sufficiently "tame," e.g. there exists* $\epsilon > 0$ *so that* $\epsilon s \leq \phi(s) \leq s/\epsilon$ *for* s *sufficiently small. See Remark 4.1.6.)*

Given any interval $[a,b]$ re-parameterizing σ in the same manner, i.e. ϕ increasing with $\phi(a)=0$ and $\phi(b)=1$, the one parameter family of Banach spaces $\mathbf{X}_\sigma = X_{\phi(\sigma)}$ again satisfies (4.1) *and* (4.2). *We consider $[0,1]$ simply to keep our notation from becoming overly complicated.*

Remarks 4.1.6, 4.1.11, 4.2.3, and 4.3.11 discuss re-parameterizing the scale parameter. In particular, affine re-parameterizations have little effect.

Remark 4.1.2. *The interested reader is invited to compare this setup with the "tame Frechet space" of [Ham82]. In particular, how does the completion of the tame semi-norms of Hamilton differ from the one parameter families X_σ or the approximation space X_0^q discuss in Definition 4.1.4 (see Question 1 in Appendix A)?*

Example 4.1.3. *In Section 8.2, we define the one parameter family, $X_\sigma = A(r\sigma, C^m)$, of real holomorphic functions on complex neighborhoods of $\mathbb{T}^n$. This is an important examples of a one parameter Banach space satisfying* (4.1) *and* (4.2) *and they play a key role in the study of torus diffeomorphisms.*

While the Banach spaces X_σ for $\sigma > 0$ often consist of analytic functions, X_0 may consist of functions with finite differentiability (e.g. C^m). The transition from analytic functions to finitely differentiable functions overlooks a large continuum of intermediate scales (e.g. spaces of functions with higher (finite) regularity). Some of the intermediate scales can be recovered by constructing an intermediate one parameter family of Banach spaces, which we will denoted by X_0^q for $q > 0$, that lies between X_0 and X_σ.

Definition 4.1.4. *Define the approximation space* X_0^q *for* $q > 0$ *as follows:* $x \in X_0$ *lies in* X_0^q *if there exists a sequence* $x_j \in X_{2^{-j}}$ *with* $x_0 = 0$*,*

$$[(x_j)] \equiv \sup_j \{2^{qj} \|x_j - x_{j-1}\|_{X_{2^{-j}}}\} < \infty$$

and $x_j \to x$ *in* X_0*. Taking*

$$\|x\|_{X_0^q} \equiv \inf \left\{ [(x_j)] \;\middle|\; \begin{array}{c} x_j \in X_{2^{-j}}, x_0 = 0 \\ \text{and } x_j \to x \text{ in } X_0 \end{array} \right\}$$

gives one a norm on X_0^q *which makes* X_0^q *into a Banach space (for proof see Lemma 1.1 in [Zeh75]).* ■

Remark 4.1.5. *As with* X_σ*, the norms* $\|\cdot\|_{X_0^q}$ *satisfy* $\|x\|_{X_0^{q'}} \leq \|x\|_{X_0^q}$ *for* $0 < q' < q < \infty$ *so with* $0 < \sigma \leq 1$ *one has*

$$X_0 \supseteq X_0^{q'} \supseteq X_0^q \supseteq X_0^\infty \equiv \left(\bigcap_{s>0} X_0^s \right) \supseteq X_\sigma \supseteq X_1$$

[An interesting question is if one has an abstract version of the Arzela-Ascoli in X_0^q*, i.e. is the embedding of* X_0^{q+m} *into* X_0^q *is compact? See 2 in Appendix A.]*

Remark 4.1.6. *Note that if* ϕ*, an increasing function with* $\phi(0) = 0$ *and* $\phi(1) = 1$*, is used to re-parameterize* $\mathbf{X}_\sigma = X_{\phi(\sigma)}$ *as described in Remark 4.1.1, then given some "tameness" conditions on* ϕ*, e.g. there exists* $\epsilon > 0$ *so that* $\epsilon s \leq \phi(s) \leq s/\epsilon$ *for* s *sufficiently small, we have* $\mathbf{X}_0^q = X_0^q$*.*

Remark 4.1.7. *If the Banach spaces* X_σ *are all Banach algebras under multiplication, so that* $\|ab\|_{X_\sigma} \leq \|a\|_{X_\sigma} \|b\|_{X_\sigma}$*, then* X_0^q *will also be a Banach algebra under multiplication.*

Example 4.1.8. *For $X_\sigma = A(r\sigma, C^m)$ as in Example 4.1.3 the spaces X_0^q can be explicitly computed. In particular $X_0^q = C^q$ for $q \notin \mathbb{Z}$ while $X_0^q = \hat{C}^q$ for $q \notin \mathbb{Z}$ where C^q are the usual spaces of Hölder functions and $\hat{C}^q$ (also denoted Λ_q) are functions satisfying a Zygmund condition. See Section 8.2 for definitions of C^q, $\hat{C}^q$, $A(r\sigma, C^m)$ and other details.*

The spaces X_0^q are defined as subspaces of X_0 through approximation properties and thus it is natural to define an operator which allows one to approximate any element of X_0^q by elements in X_σ.

Definition 4.1.9. *Let X_σ, $0 \le \sigma \le 1$, and X_0^q, $0 < q$, be two one parameter families of Banach spaces such that:*

1. *For $0 \le \sigma' \le \sigma \le 1$ one has $X_{\sigma'} \supseteq X_\sigma$ with $\|x\|_{X_{\sigma'}} \le \|x\|_{X_\sigma}$*
2. *For $0 \le q' \le q$ one has $X_0^{q'} \supseteq X_0^q$ with $\|x\|_{X_0^{q'}} \le \|x\|_{X_0^q}$.*

An analytic smoothing in the family X_σ with respect to X_0^q *is a family $\{S_t\}_{t\ge 0}$ of linear operators $S_t : X_0 \to X_1$ together with constants $k(q) > 0$ for every $0 < q < \infty$ satisfying the following three conditions:*

$$\lim_{t\to\infty} \|(S_t - I)[v]\|_{X_0} = 0 \qquad \text{for } v \in X_0 \tag{4.3}$$

$$\|S_t[v]\|_{X_{t^{-1}}} \le k(q)\|v\|_{X_0^q} \qquad \text{for } v \in X_0^q, t \ge 1 \tag{4.4}$$

$$\|(S_\tau - S_t)[v]\|_{X_{\tau^{-1}}} \le t^{-q}k(q)\|v\|_{X_0^q} \qquad \text{for } v \in X_0^q, \tau \ge t \ge 1 \tag{4.5}$$

■

Remark 4.1.10. *In Definition 4.1.9, the one parameter family X_0^q does not need to be the approximation spaces of X_σ as described in Definition 4.1.4. This said, throughout the rest of this paper, unless explicitly stated, X_0^q will always represent the approximation spaces of X_σ as described in Definition 4.1.4. But note that even when X_0^q is an approximation space of X_σ as described in Definition 4.1.4,* **analytic smoothing in X_σ with respect to X_0^q is not guaranteed**. *One must* **explicitly exhibit** *such smoothing.*

Remark 4.1.11. *Note given ϕ increasing with $\phi(a) = 0$ and $\phi(b) = 1$ is used to re-parameterize the scale parameter σ, if ϕ is sufficiently "tame," e.g. there exists $\epsilon > 0$ so that $\epsilon s \leq \phi(s) \leq s/\epsilon$ for s sufficiently small, then analytic smoothing in X_σ carries over to analytic smoothing in $\mathbf{X}_\sigma = X_{\phi(\sigma)}$.*

Example 4.1.12. *For $X_\sigma = A(r\sigma, C^m)$ as in Example 4.1.3 and $X_0^q = C^q$ or $\hat{C}^q$ as in Example 4.1.8 there exists an analytic smoothing S_t in X_σ with respect to X_0^q. The smoothing operator S_t is a convolution operator with $S_t u = s_t * u$, $s_t(z) = ts(tz)$ and $s(\cdot)$ an entire real holomorphic function. See Section 8.6 for details.*

The smoothing given in Definition 4.1.9 intertwines two one parameter family of Banach spaces. There are simpler types of smoothing operators, e.g. C^∞ smoothing described below, which are defined for a single one parameter family of Banach spaces.

Definition 4.1.13. *Let X_0^q be a one parameter family of Banach spaces with*

$0 \le q < \infty$ such that for $0 \le q' \le q < \infty$ one has

$$X_0 \equiv X_0 \supseteq X_0^{q'} \supseteq X_0^q \supseteq X_0^\infty \equiv \bigcap_{s>0} X_0^s$$

with $\|x\|_{X_0^{q'}} \le \|x\|_{X_0^q}$ for all $x \in X_0^q$. A C^∞-smoothing in the family X_0^q is a family $\{S_t\}_{t\ge0}$ of linear operators $S_t : X_0 \to X_0^\infty$ together with constants $C(q,m) > 0$ for every $0 < q, m < \infty$ satisfying the following three conditions:

$$\lim_{t\to\infty} \|(S_t - I)[v]\|_{X_0} = 0 \qquad \text{for } v \in X_0 \tag{4.6}$$

$$\|S_t[v]\|_{X_0^m} \le t^{(m-q)} C(q,m) \|v\|_{X_0^q} \qquad \text{for } v \in X_0^q, 0 \le q \le m, t \ge 1 \tag{4.7}$$

$$\|(S_t - I)[v]\|_{X_0^q} \le t^{-(m-q)} C(q,m) \|v\|_{X_0^m} \qquad \text{for } v \in X_0^m, 0 \le q \le m, t \ge 1 \tag{4.8}$$

■

Example 4.1.14. *The analytic smoothing S_t in $X_\sigma = A(r\sigma, C^m)$ with respect to $X_0^q = C^q$ or $\hat{C}^q$ discuss in Example 4.1.12 is also C^∞ smoothing when restricted to $X_0^q = C^q$ or $\hat{C}^q$. This concrete smoothing also satisfies a number of other useful estimates, see Sections 8.6 and 8.7 in Chapter 8.*

Remark 4.1.15. *The interested reader is invited to consider question 3 in Appendix A which asks if it is true in the abstract setting if the restriction of analytic smoothing to the family X_0^q always gives a C^∞ smoothing.*

Remark 4.1.16. *Analytic smoothing and C^∞ smoothing have several useful consequences.*

1. *Conditions (4.3) and (4.6) imply X_1 and X_0^∞ are dense in X_0.*

2. *The following interpolation inequalities hold: for $0 \leq r \leq t$ and $0 < \mu < 1$, with $s = \mu r + (1-\mu)t$, there exists a positive constant $M_{r,s,t}$ such that*

$$\|x\|_{X_0^s} \leq M_{r,s,t} \|x\|_{X_0^r}^{\mu} \|x\|_{X_0^t}^{(1-\mu)}$$

   for every $x \in X_0^t$.

3. *In the case of analytic smoothing, the interpolation inequalities correspond to the "three line theorem."*

4. *For certain one parameter families of Banach spaces it is know that interpolation estimates do not exists and as a result these one parameter families of Banach spaces do not have C^∞ smoothing (see [dlLO99] for a further discussion).*

5. *See [Zeh75], [Had98] and [Kol49] for these and other results.*

It is also useful to consider how the smoothing operator S_t acts on certain subsets $\mathcal{C}_0 \subseteq X_0$.

Definition 4.1.17. *Given analytic smoothing S_t in X_σ with respect to X_0^q and a subset $\mathcal{C}_0 \subseteq X_0$ we say that S_t is* $\mathcal{C}_0$-invariant *if for every $\bar{x} \in \mathcal{C}_0$ there exists positive constants r and T_0 such that for all $x \in \mathcal{C}_0$ with $\|x - \bar{x}\|_{X_0} < r$ and all $t \geq T_0$ one has $S_t[x] \in \mathcal{C}_1 \equiv \mathcal{C}_0 \cap X_1$.* ■

Example 4.1.18. *The motivating example for Definition 4.1.17 is when the set $\mathcal{C}_0$ has the form $A \cap B$ where A is invariant under S_t, i.e. $S_t A \subseteq A$, and B is* *an open set. In this case, given* $\bar{x} \in \mathcal{C}_0$ *there is an* $\delta > 0$ *so that* $B_0(\bar{x}, \delta) \subseteq B$*. Given* $\delta^* < \delta$ *and taking* $r < \delta^*/k(0)$*, for all* x *with* $\|x - \bar{x}\|_{X_0} < r$ *one has*

$$\|S_t[x - \bar{x}]\|_{X_0} \leq k(0)\|x - \bar{x}\|_{X_0} < \delta^*$$

Also, since $S_t[\bar{x}] \to \bar{x}$ *in* X_0 *there is a* T_0 *such that for all* $t \geq T_0$*,*

$$\|S_t[\bar{x}] - \bar{x}\|_{X_0} \leq (\delta - \delta^*)$$

and thus

$$\|S_t[x] - \bar{x}\|_{X_0} \leq \delta^* + (\delta - \delta^*) = \delta$$

so $S_t[x] \in B_0(\bar{x}, \delta) \subseteq B$ *and since* $S_t[x^*] \in A$*, we have* $S_t[x] \in A \cap B = \mathcal{C}_0$ *for all* $t \geq T_0$*.*

4.2 The Brjuno-Rüssmann condition

In this section we define the Brjuno-Rüssmann condition and explore some of its consequences. Informally, the Brjuno-Rüssmann condition is, in some sense, the optimal condition for obtaining convergence of the modified Newton iteration scheme introduced in Section 6.1. Motivated by [Rüs75] and especially [Rüs80] (and related to conditions obtained by different methods with different motivation in [Brj71] and [Brj72]) we define the Brjuno-Rüssmann growth condition as follows:

Definition 4.2.1. *Let* $\Omega : (0, 1] \to [1, \infty)$ *be a decreasing function. The function* Ω *satisfies the Brjuno-Rüssmann condition provided there exists a* *sequence* $\{\delta_n\}_{n=0}^{\infty}$ *of positive numbers less than* 1 *with* $\sum_{i=0}^{\infty} \delta_i < \infty$ *so that*

$$\sum_{i=0}^{\infty} 2^{-(i+1)} \log(\Omega(\delta_i)) < \infty \tag{4.9}$$

■

Several observations are in order.

Proposition 4.2.2. *Given a finite collection of functions each satisfying the Brjuno-Rüssmann condition, without loss of generality one can use the same sequence* $\{\delta_n\}_{n=0}^{\infty}$ *for condition* (4.9).

Proof. Note that if Ω_α and Ω_β satisfy condition (4.9) on the sequences $\{(\delta_\alpha)_n\}$ and $\{(\delta_\beta)_n\}$ respectively, then they also satisfy condition (4.9) on the sequence $\{\delta_n\} = \{\max((\delta_\alpha)_n, (\delta_\beta)_n)\}$. □

Remark 4.2.3. *Note given* ϕ *a re-parameterization as described in Remarks 4.1.1, 4.1.6, 4.1.11 and 4.2.3, if* ϕ *has the property that* $\sum_{i=0}^{\infty} \delta_n < \infty$ *if and only if* $\sum_{i=0}^{\infty} \phi(\delta_n) < \infty$*, e.g. there exists* $\epsilon > 0$ *so that* $\epsilon s \leq \phi(s) \leq s/\epsilon$ *for* s *sufficiently small, then the Brjuno-Rüssmann condition is invariant under this re-parameterization, i.e.* Ω *satisfies the Brjuno-Rüssmann condition if and only if* $\Omega \circ \phi$ *satisfies the Brjuno-Rüssmann condition.*

The terms δ_n in the sequence $\{\delta_n\}_{n=0}^{\infty}$ arising in Definition 4.2.1 are related to the loss of smoothness/domain at each step of the modified Newton method introduced in Section 6.1 and thus the sum $\sum_{i=0}^{\infty} \delta_i$ is related to the total loss of smoothness/domain.

Proposition 4.2.4. *Given any $\epsilon > 0$ one can assume that the value of the sum $\sum_{i=0}^{\infty} \delta_i$ arising in Definition 4.2.1 is less than ϵ.*

Proof. Discarding the first k terms from the sequence $\{\delta_n\}_{n=0}^{\infty}$ and re-indexing ensures $\sum_{i=0}^{\infty} \delta_i < \epsilon$ while $\sum_{i=0}^{\infty} 2^{-(i+1)} \log(\Omega(\delta_i))$ will increase by a factor of 2^k but remain finite. □

In Section 7.1, we use approximation to obtain smooth (i.e. $X_0^q \times Y_0^q$) existence. The interplay between the sum $\sum_{i=0}^{\infty} \delta_i$ (i.e. the domain loss) and the sum in (4.9) plays a key role. Motivated by this we make the following:

Definition 4.2.5. *Given $\Omega : (0, 1] \to [1, \infty)$ a decreasing function satisfying the Brjuno-Rüssmann condition define $\Psi_\Omega(\epsilon)$ to be any function such that*

$$\min \left\{ \sum_{i=0}^{\infty} 2^{-(i+1)} \log(\Omega(\delta_i)) \;\middle|\; \sum_{i=0}^{\infty} \delta_i < \epsilon \right\} \leq \log(\Psi_\Omega(\epsilon)) \tag{4.10}$$

■

Note that Proposition 4.2.4 guarantees one can choose $\Psi_\Omega(\epsilon) < \infty$.

We now give two important examples of functions which satisfy the Brjuno-Rüssmann condition given in Definition 4.2.1 above.

Definition 4.2.6. *Let $\Upsilon : (1, \infty) \to (1, \infty)$ be an increasing function. If Υ satisfies*

i) $\frac{1}{n} \log \Upsilon(n)$ *is decreasing (or* $\log \Upsilon(n)$ *is convex)*

ii) $\sum_{n=1}^{\infty} \frac{1}{n^2} \log \Upsilon(n) < \infty$

ii') $\sum_{n=0}^{\infty} \frac{1}{2^n} \log \Upsilon(2^n) < \infty$

we say that Υ is a Rüssmann Modulus*.* ■

Note that given i), conditions ii) and ii') are equivalent (by the Cauchy condensation theorem).

Example 4.2.7. *Note that for $c > 0$, $\nu > 0$ the function $\Upsilon(r) = cr^{\nu}$ is a Rüssmann Modulus.*

Example 4.2.8. *Given Υ a Rüssmann Modulus, define*

$$\Omega_{\Upsilon}(s) = \sum_{n=1}^{\infty} \Upsilon(n) e^{-ns} \tag{4.11}$$

and

$$\delta_n = \frac{1}{2^n} \log(2\Upsilon(2^n))$$

Note that

$$\sigma = \sum_n \delta_n < \infty \tag{4.12}$$

and

$$\sum_{i=0}^{\infty} 2^{-(i+1)} \log(\Omega_{\Upsilon}(\delta_i)) < \infty \tag{4.13}$$

so the function Ω_{Υ} satisfies the Brjuno-Rüssmann given in Definition 4.2.1. For proof of (4.12) *and* (4.13) *see Lemma 1 in [Rüs80]. Also see Remark 4.2.10.*

When considering the small divisor problems (see Section 8.5) that arises when constructing the approximate right inverse R, we can impose a

Diophantine type condition in which the usual power law is replaced with a Rüssmann Modulus Υ. The resulting operator norm of the small divisor operator as the form of Ω_Υ as defined in Example 4.2.8. If Υ has the same form as Example 4.2.7, then Ω_Υ has the following simple form:

Example 4.2.9. *Let A, α and σ be positive constants with $A \geq 1$ and $0 < \sigma \leq 1$. The function $\Omega(s) \equiv As^{-\alpha}$ satisfies the Brjuno-Rüssmann condition on $\delta_n \equiv 2^{-n}$. Furthermore, there exists a positive constant C such that $\Psi_\Omega(\epsilon)$ as defined in (4.10) satisfies $\Psi_\Omega(\epsilon) \leq C\epsilon^{-\alpha}$.*

Remark 4.2.10. *A key property of Example 4.2.9 above is the fact that:*

$$\text{There exists } \alpha > 0 \text{ such that } \Psi_\Omega(s) \leq Cs^{-\alpha}. \tag{4.14}$$

This plays a key role in obtaining existence in the smooth case (i.e. $X_0^q \times Y_0^q$) in Section 7.1. Question 4 in Appendix A asks what reasonable hypothesis can be placed on Υ to guarantee (4.14) for Ψ_{Ω_Υ} as defined in Example 4.2.8.

In the proof of Theorem 6.1.1 in Section 6.1 we obtain estimates of the form

$$\epsilon_{n+1} \leq C(n)\epsilon_n^2 \tag{4.15}$$

where $C(n) \geq 1$ is built of from Ω_Q, Ω_R and Ω_A (see 6.14). Iterating (4.15) one obtains

$$\epsilon_{n+1} \leq C(n)C(n-1)^2 \cdots C(0)^{2^n} \epsilon_0^{2^{n+1}} \tag{4.16}$$

as the sharp upper bound for sequences $\{\epsilon_n\}$ satisfying (4.15). The motivation for the definition of the Brjuno-Rüssmann condition is to ensure that the growth of the corresponding $C(n)$ is slow enough so that (4.15) can be used to grantee $\epsilon_n \to 0$ as $n \to \infty$. To this end, consider the following property:

(C1) Assume $C(n)$ is a sequence of positive numbers with $C(n) \geq 1$ such that

$$\sum_{i=0}^{\infty} 2^{-(i+1)} \log(C(i)) \leq \log(M_C) < \infty$$

for some constant $M_C \geq 1$

Lemma 4.2.11. *Let $C(n)$ be a sequence satisfying property (C1). Given $\{\epsilon_n\}$ a sequence of positive numbers satisfying* (4.15)*, on has the estimate*

$$\epsilon_n \leq (\epsilon_0 M_C)^{2^n} \tag{4.17}$$

Proof. Iterating (4.15) repeatedly, one gets (4.16). Taking the logarithm of both sides of (4.16), one has

$$\begin{aligned}\log \epsilon_{n+1} &\leq 2^{n+1} \left(\log(\epsilon_0) + \sum_{i=0}^{n} 2^{-(i+1)} \log(C(i)) \right) \\ &\leq 2^{n+1} \left(\log(\epsilon_0) + \log(M_C) \right) \\ &\leq 2^{n+1} \log(\epsilon_0 M_C)\end{aligned}$$

Exponentiating the above gives estimate (4.17). □

Corollary 4.2.12. *Given $\sum_{i=0}^{n} 2^{-(i+1)} \log(C(i)) \leq \log(D(n+1))$ for some sequence $D(n)$ then $\epsilon_n \leq (\epsilon_0 D(n))^{2^n}$.*

The following two propositions are consequences of property (C1):

Proposition 4.2.13. *Given two sequences $C_1(n)$ and $C_2(n)$, both satisfying property (C1), the sequences defined by $C_3(n) \equiv C_1(n) + C_2(n)$, $C_4(n) \equiv C_1(n)C_2(n)$ and $C_5(n) \equiv C_1(n + n_0)$ also satisfy property (C1).*

Proof. Straightforward. □

Proposition 4.2.14. *Given any sequence $C(n)$ satisfying property (C1), there exists a constant $R_C > 1$ such that*

$$C(n) \leq (R_C)^{2^n} \quad \forall n \geq 1 \tag{4.18}$$

Proof. Since the terms $2^{-(n+1)} \log(C(n))$ are summable they tend to 0 as $n \to \infty$ and hence are bounded for all n by some constant, $\log(R_C)$, and hence $2^{-(i+1)} \log(C(i)) \leq \log(R_C)$. (If one only considers (4.18) for n large then R_C can be made arbitrarily small.) Exponentiating we get (4.18). □

Remark 4.2.15. *Proposition 4.2.14 .*

Related to the functions in Example 4.2.9 we have the following important class of sequences which satisfy (C1):

Example 4.2.16. *Let A and B be positive constants and let $C(n)$ be any sequence with $C(n) \leq AB^n$. Note the sequence $C(n)$ satisfies condition (C1). In fact, one has*

$$\sum_{i=0}^{n} 2^{-(i+1)} \log(C(i)) \leq \log(A) \underbrace{\left(\sum_{i=0}^{n} 2^{-(i+1)}\right)}_{=1-\left(\frac{1}{2}\right)^{(n+1)}} + \log(B) \underbrace{\left(\sum_{i=0}^{n} i2^{-(i+1)}\right)}_{=1-(n+2)\left(\frac{1}{2}\right)^{(n+1)}} \tag{4.19}$$

$$= \log\left(A^{\left(1-\left(\frac{1}{2}\right)^{(n+1)}\right)} B^{\left(1-(n+2)\left(\frac{1}{2}\right)^{(n+1)}\right)}\right)$$

so applying Corollary 4.2.12 with

$$D(n) = A^{\left(1-\left(\frac{1}{2}\right)^n\right)} B^{\left(1-(n+1)\left(\frac{1}{2}\right)^n\right)} = \frac{AB}{\left(AB^{(n+1)}\right)^{(1/2)^n}} \tag{4.20}$$

one has

$$\epsilon_n \leq \frac{(\epsilon_0 AB)^{2^n}}{AB^{(n+1)}} \tag{4.21}$$

Remark 4.2.17. *Proposition 4.2.13 and Lemma 4.2.11 guarantee that if the sequences $\Omega_Q(\delta_n)$, $\Omega_R(\delta_n)$ and $\Omega_A(2\delta_n)$ all satisfy property (C1) then, provided ϵ_0 is sufficiently small, super-exponential estimates can be made on the decay of ϵ_n. Note $\Omega_Q(\delta_n)$, $\Omega_R(\delta_n)$ and $\Omega_A(2\delta_n)$ all have the form of $\Omega : (0,1] \to [1,\infty)$ evaluated on the points of a summable sequence $\{\delta_n\}_{n=0}^{\infty}$. This is the motivation behind the Brjuno-Rüssmann condition (Definition 4.2.1), i.e. a function Ω satisfies the Brjuno-Rüssmann condition provided there exists a summable sequence δ_n such that the sequence defined by $C(n) \equiv \Omega(\delta_n)$ satisfies property (C1).*

4.3 Hypothesis for the functional $\mathcal{F}$

We consider functionals acting between one parameter families of Banach spaces as follows:

(F0) Let X_σ, Y_σ and Z_σ be one parameter families of Banach spaces with $0 \leq \sigma \leq 1$ as discussed in Section 4.1 (at this point X_σ, Y_σ and Z_σ are not assumed to have analytic smoothing). Assume the functional

$$\mathcal{F} : U_0 \times V_0 \to Z_0$$

is given with $U_0 \subseteq X_0$ and $V_0 \subseteq Y_0$. Let $U_\sigma \equiv U_0 \cap X_\sigma$ and $V_\sigma \equiv V_0 \cap Y_\sigma$.

In the following sections, we present additional hypotheses for the functional $\mathcal{F}$ defined in (F0) that are sufficient to allow us, around various (x_0, y_0) with $\mathcal{F}(x_0, y_0) = 0$, to establish the existence, regularity and uniqueness of an implicit function g which solves $\mathcal{F}(x, g(x)) = 0$. Before discussing these additional hypotheses, we describe an example which will motivate much of our development.

Example 4.3.1. *Given a family F_λ of torus maps $F_\lambda = \mathrm{Id} + f_\lambda : \mathbb{T}^d \to \mathbb{T}^d$ (see Chapter 8) and a vector $\omega \in \mathbb{R}^d$ we want to find vectors $a \in \mathbb{R}^d$ and torus maps $H = \mathrm{Id} + h : \mathbb{T}^d \to \mathbb{T}^d$ so that*

$$(F_\lambda + a) \circ H(\theta) - H(\theta + \omega) = 0 \tag{4.22}$$

With the variables $x = (f_\lambda, \omega)$ and $y = (h, a)$, using the functional

$$\mathcal{F}(x, y) = (\mathrm{Id} + f_\lambda + a) \circ (\mathrm{Id} + h)(\theta) - (\mathrm{Id} + h)(\theta + \omega) \tag{4.23}$$

equation (4.22) can be expressed as $\mathcal{F}(x, y) = 0$. Taking $x_0 = (\omega_0, \omega_0)$ and $y_0 = (0, 0)$, it is easy to check $\mathcal{F}(x_0, y_0) = 0$. Also, note that – at least formally – we have

$$D\mathcal{F}(x, y)[\Delta h, \Delta a](\theta) = \Delta h(\theta) - \Delta h(\theta + \omega_0) + \Delta a + D_\theta f_\lambda(H(\theta))[\Delta h(\theta)] \tag{4.24}$$

in particular

$$D\mathcal{F}(x_0, y_0)[\Delta h, \Delta a](\theta) = \Delta h(\theta) - \Delta h(\theta + \omega_0) + \Delta a \tag{4.25}$$

(see Lemma 6 in [Mey75] for a proof of this calculation). In Fourier space (4.25) *is diagonal and can formally be inverted provided* $\omega \cdot k \neq 0$ *for all* $k \in \mathbb{Z}^d \setminus \{0\}$.

To ensure the the formal inverse is "meaningful," one needs to restrict ω *so that the "small divisors"* $1/(e^{2\pi i\omega\cdot k}-1)$ *do not grow to rapidly with* k. *This leads to the fact that the inverse (or approximate inverse) for the derivative is only defined in a set of* ω *which is totally disconnected. It is precisely to deal with problems of this kind that we introduce the Whitney regularity of the dependence.*

Additional details for to this example can be found in Chapter 9.

With this example in mind, we now describe the additional hypotheses we use in the following settings:

- In Section 4.3.1 we describe the hypotheses used to obtain polynomial approximate solutions in Chapter 5.
- In Section 4.3.2 we describe the hypotheses used to obtain analytic solutions in Chapter 6.
- In Section 4.3.3 we describe the hypotheses used to obtain smooth solutions in Chapter 7.

4.3.1 Hypothesis for polynomial approximate solutions

Provided $\mathcal{F}$ is C^γ for some $k < \gamma \le k+1$, given a (x_0, y_0) with $\mathcal{F}(x_0, y_0) = 0$, motivated by Lindstedt series in mechanics, we consider the problem of find polynomials $g^{\le k}(x_0, y_0; \Delta)$ which act as approximate solutions to $\mathcal{F}(x, g(x)) = 0$ around (x_0, y_0), i.e. $\mathcal{F}(x_0 + \Delta, g^{\le k}(x_0, g_0; \Delta)) = O(\Delta^\gamma)$. This can be done provided one:

(F.P1) Assume $\mathcal{F}$ as in (F0). Let $\gamma > 1$ with $k < \gamma \le k+1$ and assume for every $0 \le \sigma' < \sigma \le 1$ the map

$$\mathcal{F} : U_\sigma \times V_\sigma \to Z_{\sigma'}$$

is C^γ in x and y (in particular assume that U_σ and V_σ are open) and let $\Omega_F : (0, 1] \to [1, \infty)$ be a decreasing function such that

$$\|F\|_{C^\gamma(X_\sigma \times Y_\sigma, Z_{\sigma'})} \le \Omega_F(\sigma - \sigma') \tag{4.26}$$

For $i, j \ge 0$ with $i + j \le k$, denote the Frechet derivatives of $\mathcal{F}$ at $(x, y) \in U_\sigma \times V_\sigma$ by $D_1^i D_2^j \mathcal{F}(x, y)$ where

$$D_1^i D_2^j \mathcal{F} : U_\sigma \times V_\sigma \to \mathrm{Sym}_{i,j}(X_\sigma, Y_\sigma; Z_{\sigma'})$$

Here $\mathrm{Sym}_{i,j}(X_\sigma, Y_\sigma; Z_{\sigma'})$ denotes continuous $(i+j)$-linear operators with i symmetric terms in X_σ and j symmetric terms in Y_σ, equivalently denoted $\mathrm{Sym}_i(X_\sigma, \mathrm{Sym}_j(Y_\sigma, Z_{\sigma'}))$ or $\mathrm{Sym}_j(Y_\sigma, \mathrm{Sym}_i(X_\sigma, Z_{\sigma'}))$.

(F.P2) Let $(x_0, y_0) \in U_\sigma \times V_\sigma$ with $\mathcal{F}(x_0, y_0) = 0$ be given. Assume there exists an (unbounded) right inverse $R(x_0, y_0)$ such that, for all σ' with

$0 \le \sigma' < \sigma \le 1$, $R(x_0, y_0) \in L(Y_\sigma, Z_{\sigma'})$ and let $\Omega_R : (0, 1] \to [1, \infty)$ be a decreasing function such that

$$\|R(x_0, y_0)[v]\|_{Y_{\sigma'}} \le \Omega_R(\sigma - \sigma')\|v\|_{Z_\sigma} \tag{4.27}$$

In addition, assume that

$$\mathrm{Id} - D_2\mathcal{F}(x, y)R(x, y)][v] = 0 \tag{4.28}$$

where here Id actually represents the inclusion of some Y_σ into $Y_{\sigma'}$.

For any (x_0, y_0) with $\mathcal{F}(x_0, y_0) = 0$, Hypotheses (F.P1) and (F.P2) are sufficient to obtain a polynomials $g^{\le k}(x_0, y_0; \Delta)$ which are approximate solution to the functional equation in the sense that $\mathcal{F}(x_0 + \Delta, g^{\le k}(\Delta)) = O(\Delta^\gamma)$.

(F.PU) Given $(x_0, y_0) \in U_\sigma \times V_\sigma$ with $\mathcal{F}(x_0, y_0) = 0$, assume there exists an (unbounded) left inverse $R(x_0, y_0)$ such that, for all σ' with $0 \le \sigma' < \sigma \le 1$, $L(x_0, y_0) \in L(Y_\sigma, Z_{\sigma'})$ and

$$\mathrm{Id} - D_2\mathcal{F}(x, y)R(x, y)][v] = 0 \tag{4.29}$$

4.3.2 Hypothesis for analytic solutions

To obtain analytic results (i.e. results in $X_\sigma \times Y_\sigma$) in, e.g. Theorem 6.1.1 and Theorem 6.2.1, we require:

(F.A0) Assume $\mathcal{F}$ defined in (F0) has the property that for every $0 \le \sigma' < \sigma \le 1$

$$\mathcal{F} : U_\sigma \times V_\sigma \to Z_{\sigma'}$$

is continuous. Here $U_\sigma \equiv U_0 \cap X_\sigma$ and $V_\sigma \equiv V_0 \cap Y_\sigma$.

(F.A1) Assume $\mathcal{F}$ defined in (F0) has the property that for every $0 < \sigma' < \sigma \leq 1$

$$\mathcal{F} : U_\sigma \times V_\sigma \to Z_{\sigma'}$$

is differentiable with respect to its second argument (in particular assume U_σ and V_σ are open). Denote its Frechet derivative at $(x, y) \in U_\sigma \times V_\sigma$ by $D_2\mathcal{F}(x, y)$ with

$$D_2\mathcal{F} : U_\sigma \times V_\sigma \to L(Y_\sigma, Z_{\sigma'})$$

For any $x \in U_\sigma$ and $y_1, y_2 \in V_\sigma$, define the quadratic remainder

$$Q(x; y_1, y_2) \equiv \mathcal{F}(x, y_1) - \mathcal{F}(x, y_2) - D_2\mathcal{F}(x, y_2)[y_1 - y_2] \qquad (4.30)$$

and assume that

$$\|Q(x; y_1, y_2)\|_{Z_{\sigma'}} \leq \Omega_Q(\sigma - \sigma')\|y_1 - y_2\|^2_{Y_\sigma} \qquad (4.31)$$

with $\Omega_Q : (0, 1] \to [1, \infty)$ satisfying the Brjuno-Rüssmann growth condition (see Definition 4.2.1 in Section 4.2).

(F.A2) Assume there exists a subset $\mathcal{C}_0 \subseteq U_0$ (with $\mathcal{C}_\sigma \equiv \mathcal{C}_0 \cap X_\sigma$) such that for all $(x, y) \in \mathcal{C}_\sigma \times V_\sigma$ there exists an (unbounded) approximate right inverse $R(x, y)$ such that for all σ, σ' with $0 \leq \sigma' < \sigma \leq 1$

$$R : \mathcal{C}_\sigma \times V_\sigma \to L(Z_\sigma, Y_{\sigma'})$$

satisfies

$$\|R(x, y)[v]\|_{Y_{\sigma'}} \leq \Omega_R(\sigma - \sigma')\|v\|_{Z_\sigma} \qquad (4.32)$$

and

$$\| \, [\mathrm{Id} - D_2\mathcal{F}(x,y)R(x,y)][v] \, \|_{Z_{\sigma'}} \le \Omega_A(\sigma-\sigma')\|\mathcal{F}(x,y)\|_{Z_\sigma}\|v\|_{Z_\sigma} \quad (4.33)$$

with $\Omega_R, \Omega_A : (0,1] \to [1,\infty)$ satisfying the Brjuno-Rüssmann growth condition (see Definition 4.2.1 in Section 4.2)).

Remark 4.3.2. *Note (F.A2) generalizes the unbounded inverse R defined in (F.P2) to (x, y) with $\mathcal{F}(x, y) \neq 0$.*

Remark 4.3.3. *Note that in* (4.33) *to compute*

$$\| \, [I - D_2\mathcal{F}(x,y)R(x,y)][v] \, \|_{Z_{\sigma'}} \quad (4.34)$$

given $v \in Z_\sigma$, one must choose σ'' with $\sigma' < \sigma'' < \sigma$ and first compute $R(x, y)[v] \in Y_{\sigma''}$ and then $D_2\mathcal{F}(x, y)R(x, y)[v] \in Z_{\sigma'}$. One of the consequences of condition (4.33) *is that this choice of intermediate scale σ'' does not affect* (4.34)*.*

Using the notion of C^γ Whitney regularity presented in Chapter 3, the following hypotheses can be used to establish the Whitney differentiable of g:

(F.W1) Let $\gamma > 1$ with $k < \gamma \le k + 1$ and assume that $\mathcal{F}$ satisfies (F.P1). In addition assume $\Omega_F : (0, 1] \to [1,\infty)$ satisfies the Brjuno-Rüssmann growth condition (see Definition 4.2.1 in Section 4.2).

(F.W2) Let $\gamma > 1$ with $k < \gamma \le k + 1$ and let R be given as in (F.A2). Assume for all σ, σ' with $0 \le \sigma' < \sigma \le 1$, we have

$$R \in C^\gamma_{Wht}(\mathcal{C}_\sigma \times V_\sigma, L(Z_\sigma, Y_{\sigma'}))$$

with

$$\|R\|_{C^\gamma} \leq \Omega_R(\sigma - \sigma') \tag{4.35}$$

with $\Omega_R(s)$, as in (F.A2), satisfying the Brjuno-Rüssmann growth condition (see Definition 4.2.1 in Section 4.2)).

(F.W3) Let R be given as in (F.W2). In addition, for $0 \leq \sigma'' < \sigma' < \sigma \leq 1$, assume

$$\begin{aligned} &\| \, [\mathrm{Id} - D_2\mathcal{F}(-,y)R(-,y)][v(-)] \, \|_{C^\gamma_{Wht}(\mathcal{C}_\sigma;Z_{\sigma''})} \\ &\qquad \leq \Omega_A(\sigma - \sigma')\|\mathcal{F}(-,y)\|_{C^\gamma_{Wht}(\mathcal{C}_\sigma;Z_{\sigma'})}\|v\|_{C^\gamma_{Wht}(\mathcal{C}_\sigma;Z_{\sigma'})} \end{aligned} \tag{4.36}$$

with $\Omega_A(s)$, as in (F.A2), satisfying the Brjuno-Rüssmann growth condition (see Definition 4.2.1 in Section 4.2)).

(F.W4) Let R be given as in (F.W2). In addition, assume that the Whitney derivatives of R, which we denote by $\mathcal{W}\mathcal{D}_x R_(x,y)$ and $\mathcal{W}\mathcal{D}_y R(x,y)$, satisfy

$$\mathcal{W}\mathcal{D}_x R(x,y)[v,w] = -R(x,y)[D_1D_2\mathcal{F}(x,y)[R(x,y)[v],w]]$$

$$\mathcal{W}\mathcal{D}_y R(x,y)[v,w] = -R(x,y)[D_2^2\mathcal{F}(x,y)[R(x,y)[v],w]]$$

for all (x,y) with $\mathcal{F}(x,y) = 0$.

Remark 4.3.4. *Provided $\gamma \geq 2$, Hypothesis (F.A1) follows from Hypothesis (F.W1) with $\Omega_Q = \Omega_F$.*

Remark 4.3.5. *Note, taking $\mathcal{F}(x,y) = 0$, Hypothesis (F.W3) implies Hypothesis (F.W4).*

Remark 4.3.6. *Informally, we can understand (F.W4) (and F.W3)) by noting that to even establish the existence of g we need Hypothesis (F.A2) which requires the approximate right inverse R to be a right inverse up to zeroth order in* $\|\mathcal{F}(x,y)\|$, *i.e. evaluating* (4.33) *when* $\mathcal{F}(x,y)=0$ *implies*

$$I - D_2\mathcal{F}(x,y)R(x,y) = 0 \tag{4.37}$$

Condition (F.W2) simply requires that the approximate right inverse R be a right inverse up to first order. That is, the approximate right inverse R is Whitney Differentiable and the Whitney partial derivatives of R satisfy the equations we get by implicitly differentiating (4.37), *namely*

$$D_1D_2\mathcal{F}(x,y)[R(x,y)[v],w] + D_2\mathcal{F}(x,y)[\mathcal{WD}_1R(x,y)[v,w]] = 0$$

$$D_2^2\mathcal{F}(x,y)[R(x,y)[v],w] + D_2\mathcal{F}(x,y)[\mathcal{WD}_2R(x,y)[v,w]] = 0$$

To obtain local uniqueness for the zeros of implicit function in $X_\sigma \times Y_\sigma$ the following hypothesis is sufficient (see Section 6.3):

(F.AU) Assume that for all $x \in \mathcal{C}_\sigma \subseteq U_\sigma$ there is an approximate left inverse $L(x,y)$ such that for all σ, σ' with $0 \le \sigma' < \sigma \le 1$

$$L : \mathcal{C}_\sigma \times V_\sigma \to L(Z_\sigma, Y_{\sigma'})$$

satisfies

$$\|L(x,y)[v]\|_{Y_{\sigma'}} \le \Omega_L(\sigma - \sigma')\|v\|_{Z_\sigma} \tag{4.38}$$

with $\Omega_L : (0,1] \to [1,\infty)$ satisfying the Brjuno-Rüssmann growth condition (see Definition 4.2.1 in Section 4.2) and

$$L(x,y)D_2\mathcal{F}(x,y) = \mathrm{Id} \quad \text{when} \quad \mathcal{F}(x,y) = 0 \tag{4.39}$$

Remark 4.3.7. *There are several possible variations one can make in the above hypothesis and obtain the same or similar results. We list here a few such "improvements"*

(F.V1) *In place of* (4.31), *it suffices that the quadratic remainder* Q *defined in* (4.30) *satisfy*

$$\|Q(x; y_1, y_2)\|_{Z_{\sigma'}} \leq \Omega_Q(\sigma - \sigma')\|y_1 - y_2\|_{Y_\sigma}^{(1+\alpha)} \tag{4.40}$$

for some $\alpha > 0$ *with* $\Omega_Q : (0, 1] \to [1, \infty)$ *again satisfying the Brjuno-Rüssmann growth condition.*

(F.V2) *To model the method of Arnold, we can replace the single approximate right inverse* R *satisfying* (4.33) (4.32) *with a sequence of operators* R_j *satisfying*

$$\|R_j(x, y)[v]\|_{Y_{\sigma'}} \leq C^{2^j}\Omega_R(\sigma - \sigma')\|v\|_{Z_\sigma} \tag{4.41}$$

and

$$\begin{aligned} \| \ [I - D_2\mathcal{F}(x, y)R_j(x, y)][v] \ \|_{Z_{\sigma'}} \\ \leq C^{2^j}\Omega_A(\sigma - \sigma')(\|\mathcal{F}(x, y)\|_{Z_\sigma} + C^{-2^j})\|v\|_{Z_\sigma} \end{aligned} \tag{4.42}$$

for some constant $C \geq 1$ *with* $\Omega_R, \Omega_A : (0, 1] \to [1, \infty)$ *again satisfying the Brjuno-Rüssmann growth condition.*

The key property that we maintain, even with the above modifications, is that the iteration of the modified Newton method still have super-exponential convergence.

4.3.3 Hypothesis for smooth solutions

To obtain smooth results (i.e. results in $X_0^q \times Y_0^q$) in, e.g. Theorem 7.1.1 and Theorem 7.2.1 we will use analytic smoothing on approximate solutions.

(XYZ.S1) Assume X_σ, Y_σ and Z_σ have analytic smoothing with respect to X_0^q , Y_0^q and Z_0^q (see Definitions 4.1.4 and 4.1.9).

(XYZ.S2) Assume that analytic smoothing in X_σ is both U_0 and $\mathcal{C}_0$-invariant and the analytic smoothing in Y_σ is V_0-invariant (see Definition 4.1.17).

The hypotheses on $\mathcal{F}$ are essentially the same as in the analytic setting with the terms satisfying the Brjuno-Rüssmann growth condition further restricted to have the same form as Example 4.2.9 (actually, we only need the corresponding Ψ_Ω as described in Definition 4.2.5 satisfy $\Psi_\Omega(s) \leq Cs^{-\alpha}$ for some α, see Question 4 in Appendix A).

(F.S0) Same as Hypothesis (F.A0)

(F.S1) Same as Hypothesis (F.A1) with $\Omega_Q(s) \leq C_Q s^{-\alpha}$ (as noted above we only need the corresponding Ψ_Ω satisfy $\Psi_\Omega(s) \leq Cs^{-\alpha}$ for some α).

(F.S2) Same as Hypothesis (F.A2) with $\Omega_R(s) \leq C_R s^{-\beta}$ and $\Omega_A(s) \leq C_A s^{-\gamma}$ (as noted above we only need the corresponding Ψ_Ω satisfy $\Psi_\Omega(s) \leq Cs^{-\alpha}$ for some α).

(F.S3) Assume that there is a constant $M_3 > 0$ so that, for every $0 < \sigma' < \sigma \leq 1$, $\mathcal{F} : U_\sigma \times V_\sigma \to Z_{\sigma'}$ is uniformly Lipschitz with respect to its first argument, i.e.

$$\|\mathcal{F}(x_1, y) - \mathcal{F}(x_2, y)\|_{Z_{\sigma'}} \leq M_3 \|x_1 - x_2\|_{X_\sigma} \tag{4.43}$$

In addition, to control the process of smoothing an approximate solution, we assume $\mathcal{F}$ interacts with smoothing in a natural way. Specifically:

(F.S4) Assume that for q sufficiently large there exists positive constants $q*$ and $M_4(q) > 0$ such that for any $(x, y) \in U_0^q \times V_0^q$ there exists a $t_0 > 0$ such that for all $t \geq t_0$ one has $(S_t x, S_t y) \in U_1 \times V_1$ and the functional $\mathcal{F} : U_\sigma \times V_\sigma \to Z_{\sigma'}$ satisfies the estimate

$$\|\mathcal{F}(S_t x, S_t y) - S_t \mathcal{F}(x, y)\|_{Z_{t^{-1}}} \leq M_4(q) t^{-q+q*} \tag{4.44}$$

Remark 4.3.8. *Note that we take the approximate solutions to be in $X_0^q \times Y_0^q$ (as compared to Zehnder [Zeh75] and Poschel [Pös82] who require an approximate solution in $X_\sigma \times Y_\sigma$). Hypothesis (F.S4) allows us to obtain analytic approximate solutions by applying analytic smoothing to smooth approximate solutions. In particular, given $(x, y) \in U_0^q \times V_0^q$ combining Hypothesis (F.S4) with the standard smoothing estimates we get*

$$\begin{aligned}\|\mathcal{F}(S_t x, S_t y)\|_{Z_{t^{-1}}} &\leq \|\mathcal{F}(S_t x, S_t y) - S_t \mathcal{F}(x, y)\|_{Z_{t^{-1}}} + \|S_t \mathcal{F}(x, y)\|_{Z_{t^{-1}}} \\ &\leq M_4(q) t^{-q+q*} + k(q) \|\mathcal{F}(x, y)\|_{Z_0^q}\end{aligned}$$

Remark 4.3.9. *The exponent* $-q+q*$ *in* (4.44) *and the exponent* β *of* $\Omega_R(s) \leq C_R s^{-\beta}$ *in (F.S2) combine to give us the resulting loss in the smooth case, e.g. given* $(x,y) \in U_0^q \times V_0^q$ *with* $\mathcal{F}(x,y)$ *sufficiently small we get* $y_\infty \in V_0^{q-(q*+\beta)}$ *with* $\mathcal{F}(x,y_\infty) = 0$.

Example 4.3.10. *For the space of analytic torus diffeomorphisms, the composition functional* $\mathcal{F}(f,g) = f \circ g$*, which appears often in KAM theory, satisfies property (F.S4). See Section 8.7, Lemma 8.7.2.*

Remark 4.3.11. *Note given* ϕ *increasing with* $\phi(a) = 0$ *and* $\phi(b) = 1$ *is used to re-parameterize the scale parameter* σ*, if* ϕ *is sufficiently "tame," e.g. there exists* $\epsilon > 0$ *so that* $\epsilon s \leq \phi(s) \leq s/\epsilon$ *for* s *sufficiently small, then as discuss in Remarks 4.1.11 condition (F.S4) will be invariant under the change of scales, i.e. (F.S4) will hold in* $\mathbf{X}_\sigma = X_{\phi(\sigma)}$.

To obtain Whitney regularity in the smooth setting we have the follow hypotheses:

(F.SW1) Same as (F.AW1) with the additional assumption that $\Omega_F(s) \leq C_F s^\alpha$ (as noted above we only need the corresponding Ψ_Ω satisfy $\Psi_\Omega(s) \leq Cs^{-\alpha}$ for some α).

(F.SW2) Same as (F.W2) and (F.W3) with the additional assumption that as in (F.S2), we have $\Omega_R(s) \leq C_R s^{-\beta}$ and $\Omega_A(s) \leq C_A s^{-\gamma}$ (as noted above we only need the corresponding Ψ_Ω satisfy $\Psi_\Omega(s) \leq Cs^{-\alpha}$ for some α).

Finally, to obtain local uniqueness for the zeros of implicit function in $X_0^q \times Y_0^q$ the following hypothesis is sufficient (see Section 7.3):

(F.SU) Same as (F.AU) with the additional assumption that $\Omega_L(s) \leq C_L s^{\beta *}$, (as noted above we only need the corresponding Ψ_Ω satisfy $\Psi_\Omega(s) \leq Cs^{-\alpha}$ for some α).

Chapter 5

Polynomial Approximate Solutions

In this chapter, we consider a functional $\mathcal{F}$ satisfying the various hypotheses described in Section 4.3.1. For $(x_0, y_0) \in U_\sigma \times V_\sigma$ with $\mathcal{F}(x_0, y_0) = 0$, we develop a polynomial $g^{\leq k}(x_0, y_0; \Delta)$ (i.e. the coefficients of Δ depend on x_0, y_0) such that $\mathcal{F}(x_0 + \Delta, g^{\leq k}(x_0, y_0; \Delta))$ vanish at $\Delta = 0$ to order γ, i.e. $\mathcal{F}(x_0 + \Delta, g^{\leq k}(x_0, y_0; \Delta)) = O(\Delta^k)$ (see Chapter 2 for notation and definitions, also for the time being we suppress the specifics of the various scales at which terms occur).

These polynomials are not "exact" solutions to $\mathcal{F}(x, g(x)) = 0$. We think of them as "approximate" solutions with $\left\|\mathcal{F}(x + \Delta, g^k(x_0, y_0; \Delta))\right\|_Z$ measuring the "error" of $g^k(x_0, y_0; \Delta)$. Their construction is, in some sense, easier then obtaining an exact solution g to $\mathcal{F}(x, g(x)) = 0$ for $x \in \mathcal{C}$ around a given (x_0, y_0) with $\mathcal{F}(x_0, y_0) = 0$ (in Chapters 6 and 7 we will obtain such exact solutions, see Corollaries 6.1.2 and 7.1.3). Also, since the coefficients of $g^k(x_0, y_0; \Delta)$, which we denote by $g_i(x_0, y_0)$, are independent of k we can think of $g^k(x_0, y_0; \Delta)$ as the truncations of a formal power series (see Remark 2.3.2). This formal power series is related to, and motivated by, the Lindstedt series in mechanics.

Polynomial approximate solutions are very useful for computational problems. They are related to asymptotic expansions and given local uniqueness to $\mathcal{F}(x, y) = 0$, the coefficients of the polynomial approximate solutions uniquely determine the Whitney derivative of the implicit function solutions obtained in Corollary 6.1.2 or Corollary 6.1.2. Finally, unlike the exact solution g to $\mathcal{F}(x, g(x)) = 0$ we obtain in Chapters 6 and 7 which can only be evaluated for $x \in \mathcal{C}$, polynomial approximate solutions $g^k(x_0, y_0; \Delta)$ can evaluated at any point $x \in U_0$, i.e. $\mathcal{F}(x, g^k(x_0, y_0; x - x_0))$.

Theorem 5.0.12. (Existence of polynomial approximate solutions) *Let $\mathcal{F}$ be as defined in (F0) and assume that $\mathcal{F}$ satisfies Hypothesis (F.P1) for some $k < \gamma \leq k + 1$. Choose $0 \leq \sigma'' < \sigma' < \sigma \leq 1$.*

Let $(x_0, y_0) \in U_\sigma \times V_\sigma$ with $\mathcal{F}(x_0, y_0) = 0$ and assume (F.P2) holds. Choosing intermediate scales as in Remark 5.0.13 inductively define

$$g_i(x_0, y_0) \in \mathrm{Sym}_i(X_\sigma, Y_{\sigma'}) \quad \textit{for} \quad 1 \leq i \leq k \tag{5.1}$$

by the recurrence

$$g_i(x_0, y_0)[-]^{\otimes i} \equiv -R(x_0, y_0)[Q_i^{\mathcal{F}}(x_0, y_0; g_1, \ldots, g_{i-1})] \tag{5.2}$$

Here the $Q_i^{\mathcal{F}}$ are the polynomials described in Proposition 2.2.8 and g_j are used to denote $g_j(x_0, y_0)$ for $1 \leq j \leq i - 1$.

Taking $\epsilon > 0$ sufficiently small one has $g^{\leq k} : B_\sigma(0, \epsilon) \to V_{\sigma'}$. and viewing $\mathcal{F} : U_{\sigma'} \times V_{\sigma'} \to Z_{\sigma''}$ one has $\mathcal{F}(x_0 + \Delta, g^{\leq k}(x_0, y_0; \Delta)) = O(\Delta^\gamma)$, i.e.

there exists $M > 0$ *so that*

$$\left\|\mathcal{F}(x_0 + \Delta, g^{\leq k}(x_0, y_0; \Delta))\right\|_{Z_{\sigma''}} \leq M\|\Delta\|_{X_\sigma}^{\gamma} \tag{5.3}$$

with

$$M = M_\gamma \Omega_F((\sigma - \sigma')/(2k)) \left(1 + \left(\Omega_R((\sigma - \sigma')/(2k))\right)^{\gamma}\right)$$

If, $D_1^i D_2^j \mathcal{F} : U_0^q \times V_0^q \to \sum_{i,j}(X_0^q \times Y_0^q; Z_0^{q-\alpha})$ *(alternatively, as in (F.SW1) we can take* $\Omega_F(s) \leq C_F s^{-\alpha}$*) and* $\Omega_R(s) \leq C_R s^{-\beta}$ *then, provided* $q \geq (\alpha + \beta)k$*, we can use* (5.2) *to define*

$$g_i(x_0, y_0) \in \mathrm{Sym}_i(X_0^q, Y_0^{q-(\alpha+\beta)i})$$

so that

$$\left\|\mathcal{F}(x_0 + \Delta, g^{\leq k}(x_0, y_0; \Delta))\right\|_{Z_0^q} \leq M\|\Delta\|_{X_0^{q'}}^{\gamma} \tag{5.4}$$

with $q' < q - (\alpha + \beta)$.

Finally, if we have R (or an appropriate generalization, e.g. (F.A2)), defined on a larger set of (x, y)*, including* (x, y) *for which* $\mathcal{F}(x, y) \neq 0$*, provided we still have* (4.28) *whenever* (x_0, y_0) *is such that* $\mathcal{F}(x_0, y_0) = 0$*,* (5.2) *can still be used to define polynomials* $g^{\leq k}(x, y; \Delta)$ *with* (5.3) *holding for every* (x_0, y_0) *is such that* $\mathcal{F}(x_0, y_0) = 0$.

Remark 5.0.13. *To determine* (5.1) *using the recurrence formulas in* (5.2) *we need to incur some loss of scale when applying R. Furthermore, we need to do this without going below* σ'*. To this end, choose intermediate scales* σ_i *and* τ_i *with*

$$0 \leq \sigma' < \sigma_k < \tau_k < \sigma_{k-1} < \cdots < \tau_2 < \sigma_1 < \tau_1 < \sigma \leq 1$$

by taking $\sigma_n - \sigma_{n+1} = (\sigma - \sigma')/k$ *and* $\tau_{n+1} = (\sigma_{n+1} + \sigma_n)/2$. *Given the form of the* $Q_i^{\mathcal{F}}$, *we can view it as follows*

$$Q_i^{\mathcal{F}} : (X_\sigma \times Y_\sigma) \times Y_{\sigma_1} \times \cdots \times Y_{\sigma_{i-1}} \to Z_{\tau_i}$$

Taking $R : Z_{\tau_i} \to Y_{\sigma_i}$ *we can thus apply* (5.2) *and obtain*

$$g_i(x_0, y_0) \in \mathrm{Sym}_i(X_\sigma, Y_{\sigma_i})$$

for $1 \leq i \leq k$. *Once the* $g_i(x_0, y_0)$ *are defined, using the inclusion of* Y_{σ_i} *into* $Y_{\sigma'}$ *gives us* (5.1).

Proof of Theorem 5.0.12

Applying Corollary 2.2.7 note that the composition

$$D_2\mathcal{F}(x_0, y_0)[R(x_0, y_0)] : Z_\sigma \to Z_{\sigma'}$$

is the identity (or rather the inclusion of Z_σ into $Z_{\sigma'}$). Using $D^i\,[F \circ G]$ to represent $D^i_\Delta\left[F(x_0 + \Delta, g^{\leq k}(x_0, y_0; \Delta))\right]_{\Delta=0}$, for $0 < i \leq k$ we have,

$$D^i\,[F \circ G] = D_2\mathcal{F}(x_0, y_0)[g_i(x_0, y_0)] + Q_i^{\mathcal{F}}(x_0, y_0; g_1, \ldots, g_{i-1}) = 0$$

Applying Taylor's Theorem with integral remainder (see e.g. Theorem 6 in [Nel69]) we obtains (5.3). The form of M follows from the form of the integral remainder and Proposition 2.2.5.

The generalizations to $X_0^q \times Y_0^q$ and to arbitrary sets of (x, y) are straightforward. $\square$

In Section 6.2 and 7.2 we will use the polynomial approximate solutions obtained in Theorem 5.0.12 with either of the Whitney Verification Lemmas (Lemma 3.1.10 or Lemma 3.2.6) to establish the Whitney differentiability of the implicit function. Anticipating applying Lemma 3.1.10 later, we now establish the following:

Proposition 5.0.14. *If R satisfies Hypothesis (F.W2) and (F.W4), then $g_i(x, y)$ satisfies* (3.10) *in the Whitney Verification Lemma I (Lemma 3.1.10), i.e.*

$$g_{i+1}(x, y) = \mathcal{WD}_x(g_i(x, y)) + \mathcal{WD}_y(g_i(x, y))[g_1(x, y)]$$

Proof. To prove $g_i(x, y)$ satisfy (3.10), we proceed by induction on i. Although (3.10) is only taken $0 < i \leq k$, with $g_0(x, y) = y$ note that $\mathcal{WD}_x(g_0(x, y)) = 0$ and $\mathcal{WD}_y(g_0(x, y)) = Id$ so in fact (3.10) also holds for $i = 0$ and we use this as the base case for our induction.

Inductively assume that (3.10) holds for all $i \leq n - 1 \leq k$. Using Hypotheses (F.W2), (F.W4) and Proposition 2.2.8, note that for (x, y) with

$\mathcal{F}(x, y) = 0$, suppressing (x, y) for compactness of notation, we have

$$\begin{aligned}
&\mathcal{WD}_x(g_n) + \mathcal{WD}_y(g_n)[g_1] = -\mathcal{WD}_x(R[Q_n^{\mathcal{F}}]) - \mathcal{WD}_y(R[Q_n^{\mathcal{F}}])[g_1] \\
&= -(\mathcal{WD}_x R)[Q_n^{\mathcal{F}}] - (\mathcal{WD}_y R)[Q_n^{\mathcal{F}}, g_1] - R[\mathcal{WD}_x Q_n^{\mathcal{F}} + \mathcal{WD}_y Q_n^{\mathcal{F}}[g_1]] \\
&= -R\Big[\ D_1 D_2 \mathcal{F}[g_n] + D_2^2 \mathcal{F}[g_1, g_n] + \\
&\qquad \mathcal{WD}_x Q_n^{\mathcal{F}} + \nabla_b Q_n^F \cdot (\mathcal{WD}_x g_1, \ldots, \mathcal{WD}_x g_{n-1}) + \\
&\qquad \mathcal{WD}_y Q_n^{\mathcal{F}}[g_1] + \nabla_b Q_n^F \cdot (\mathcal{WD}_y g_1, \ldots, \mathcal{WD}_y g_{n-1})[g_1]\ \Big] \\
&= -R\Big[\ D_1 D_2 \mathcal{F}[g_n] + D_2^2 \mathcal{F}[g_1, g_n] + \\
&\qquad \mathcal{WD}_x Q_n^{\mathcal{F}} + \mathcal{WD}_y Q_n^{\mathcal{F}}[g_1] + \nabla_b Q_n^F \cdot (g_2, \ldots, g_n)\ \Big] \\
&= -R[Q_{n+1}^{\mathcal{F}}] = g_{n+1}
\end{aligned}$$

This completes the induction. ☐

Theorem 5.0.15. (Uniqueness of polynomial approximate solutions) *Let $\mathcal{F}$ as in (F0) satisfying Hypotheses (F.A0) and (F.P1) for some $\gamma > 1$ with $k < \gamma \leq k+1$. Choose $0 \leq \sigma'' < \sigma' < \sigma \leq 1$. Given any (x_0, y_0) satisfying (F.PU) with $\mathcal{F}(x_0, y_0) = 0$ and*

$$g_i(x_0, y_0) \in \mathrm{Sym}_i(X_\sigma, Y_{\sigma'}) \quad \textit{for} \quad 1 \leq i \leq k$$

such that for all $\Delta \in B_\sigma(0, \epsilon)$ one has (5.3) then the g_i must satisfy

$$g_i(x_0, y_0) + L(x_0, y_0)[Q_i^{\mathcal{F}}(x_0, y_0; g_1, \ldots, g_{i-1})] = 0 \tag{5.5}$$

where the $Q_i^{\mathcal{F}}$ are polynomial as described in Proposition 2.2.8 and g_j are used to denote $g_j(x_0, y_0)$ for $1 \leq j \leq i-1$.

Furthermore, if $\mathcal{F}$ also satisfies (F.P2) then the $g_i(x_0, y_0)$ are unique and hence have the same form as described in (5.2) *of Theorem 5.0.12.*

Finally, as in Theorem 5.0.12, if $D_1^i D_2^j \mathcal{F} : U_0^\ell \times V_0^\ell \to \sum_{i,j}(X_0^q \times Y_0^q; Z_0^{q-\alpha})$ (or $\Omega_F(s) \leq C_F s^{-\alpha}$) and $\Omega_L(s) \leq C_L s^{-\beta*}$ then, provided $q \geq (\alpha + \beta*)k$, (5.5) *also holds in $X_0^q \times Y_0^q$.*

Proof. Equation (5.5), along with the $X_0^q \times Y_0^q$ case, follows directly by applying L to (2.26).

To demonstrate that (5.5) implies (5.2), note that up to loss of scale, if $R[v] = w$ then $v = \Delta_2 \mathcal{F}[R[v]] = \Delta_2 \mathcal{F}[w]$. Hence $L[v] = L[\Delta_2 \mathcal{F}[w]] = w$, i.e. $R[v] = L[v]$, so (5.5) and (5.2) are equivalent. □

Remark 5.0.16. (Formal Power Series "Solutions")
To obtain a formal power series solution around (x_0, y_0) in the analytic setting, it is sufficient to have condition (F.P1) for all $\gamma \geq 0$. Then, for an infinite choice of decreasing scales, (5.2) *can be used to define the coefficients of a formal power series $g^{\leq\infty}$ (see Definition 2.3.1) which formally solves*

$$\mathcal{F}(x + \Delta, g^{\leq\infty}(\Delta)) = 0 \tag{5.6}$$

i.e. for any $k < \gamma \leq k+1$ the truncated power series (i.e. polynomial) $g^{\leq\leq k}(\Delta)$ solve (5.6) *to order γ.*

Obtaining a FPS in $X_0^q \times Y_0^q$ is generally not possible due to the fact that each use of (5.2) *requires a loose of fixed amount in the q scale.*

Finally, one generally cannot establish anything about the convergence of the FPS (especially since determining the coefficients of $g^{\leq\infty}$ *requires an infinite choice of decreasing scales).*

Chapter 6

Solutions in Analytic Spaces

We now turn from the development of (polynomial, formal power series) approximate solutions to $\mathcal{F}(x, g(x)) = 0$ around $\mathcal{F}(x_0, y_0) = 0$ to establishing exact solutions. To show the existence of such an implicit function g defined for $x \in \mathcal{C}$ near x_0, we establish (in Theorem 6.1.1) that for any y with $\mathcal{F}(x, y)$ sufficiently small (in appropriate norm) there existence y_∞ satisfying $\mathcal{F}(x, y_\infty) = 0$. Theorem 6.1.1 is our "constructive" theorem and, in some sense, it is the key to everything. The basic idea is to balance the rapid convergence of our Newton like iteration scheme against the domain loss at each stage of the iteration in such a way to maintain control throughout this process and obtain convergence. Using this theorem in conjunction with the continuity of $\mathcal{F}$ it is a simple matter to build the implicit function g from individual solutions y_∞ (see Corollary 6.1.2).

In Chapter 7 we use the analytic smoothing discussed in Section 4.1 to extend the results of this chapter to the smooth case (i.e. $X_0^q \times Y_0^q$).

6.1 Existence in analytic space via a modified Newton method

Theorem 6.1.1. *Given $\mathcal{F}$ as in (F0) satisfying Hypothesis (F.A0),(F.A1) and (F.A2), there exists positive constants δ and N, depending only on Ω_Q, Ω_R, Ω_A and τ, such that for any $0 < \tau < \tau' \leq 1$ and $(x, y) \in \mathcal{C}_{\tau'} \times V_{\tau'}$ with*

$$\|\mathcal{F}(x,y)\|_{Z_\tau} \leq \delta \min(1, \operatorname{dist}(y, V^c_{\tau/2})) \tag{6.1}$$

there exists a $y_\infty = y_\infty(x, y) \in V_{\tau/2}$ with

$$\mathcal{F}(x, y_\infty) = 0 \qquad \text{and} \qquad \|y - y_\infty\|_{Y_{\tau/2}} \leq N\|\mathcal{F}(x,y)\|_{Z_\tau} \tag{6.2}$$

Moreover:

(A) Writing $\Omega_Q(s) \equiv C_Q\Psi_Q(s)$, $\Omega_R(s) \equiv C_R\Psi_R(s)$ and $\Omega_A(s) \equiv C_A\Psi_A(s)$, where C_Q, C_R, C_A are constants and $\Psi_Q, \Psi_R, \Psi_A : (0, 1] \to [1, \infty)$ are functions which "carry the shape" of Ω_Q, Ω_R, Ω_A, the constants δ and N can be chosen as follows

$$\delta = M_\delta \frac{1}{C_R \max(C_Q C_R^2, C_A)} \tag{6.3}$$

$$N = M_N C_R \max(C_Q C_R^2, C_A) \tag{6.4}$$

where M_δ and M_N are constants which depend only on Ψ_Q, Ψ_R and Ψ_A and τ. Furthermore, choosing M_δ sufficiently small, one can make $M_\delta M_N = N\delta$ arbitrarily small.

(B) If $\Omega_Q(s) \le C_Q s^{-\alpha}$, $\Omega_R(s) \le C_R s^{-\beta}$ and $\Omega_A(s) \le C_A s^{-\gamma}$, as in (F.S1) and (F.S2), then the constants δ and N can be chosen as follows

$$\delta = M_\delta \frac{\tau^{\max(\alpha+2\beta,\gamma)}}{\max(C_Q C_R^2, C_A)} \tag{6.5}$$

$$N = M_N C_R \tau^{-\beta} \tag{6.6}$$

where M_δ and M_N are constants which depend only on α, β and γ. Furthermore, for any $\eta > 0$, M_δ can be chosen so that $N\delta \le \eta\tau^{\alpha+\beta}$.

A useful application of this point-wise existence is the following:

Corollary 6.1.2. *Given $\mathcal{F}$ as in (F0) satisfying Hypothesis (F.A0),(F.A1) and (F.A2), for any $(\bar{x}, \bar{y}) \in \mathcal{C}_{\tau'} \times V_{\tau'}$ with $\mathcal{F}(\bar{x}, \bar{y}) = 0$, there exists a positive constant ϵ and a function*

$$g : \mathcal{C}_{\tau'} \cap B_{\tau'}(\bar{x}, \epsilon) \to V_{\tau/2}$$

with

$$\mathcal{F}(x, g(x)) = 0 \tag{6.7}$$

Proof. Note for $y = \bar{y}$ fixed, the RHS of (6.1) is a constant. By continuity of $\mathcal{F} : U_{\tau'} \times V_{\tau'} \to Z_\tau$, for $0 \le \tau < \tau' \le 1$, since $\mathcal{F}(\bar{x}, \bar{y}) = 0$ there exists a constant $\epsilon > 0$ so that, for $x \in B_{\tau'}(\bar{x}, \epsilon)$ condition (6.1) is satisfied. Applying Theorem 6.1.1 to the approximate solution $(x, \bar{y})$ with $x \in \mathcal{C}$, we get a $y_\infty \in Y_{\tau/2}$ with $\mathcal{F}(x, y_\infty(x, \bar{y})) = 0$. Define $g(x) = y_\infty$. □

Remark 6.1.3. *In the proof of Theorem 6.1.1, the convergence of sequence y_n obtained from the modified Newton method is uniform in $\|\mathcal{F}(x,y)\|_Z$. Furthermore, all the estimates that appear in the proof of Theorem 6.1.1 are uniform in x and only require $x \in \mathcal{C}$ in order that the right inverse R to exist.*

Remark 6.1.3 gives us the following corollary:

Corollary 6.1.4. *If, for any $0 \le \sigma < \sigma' \le 1$, the functions $\mathcal{F} : U_{\sigma'} \times V_{\sigma'} \to Z_\sigma$ and $R : \mathcal{C}_{\sigma'} \times V_{\sigma'} \to L(Z_{\sigma'}, Y_\sigma)$ are uniformly continuous then the implicit function*

$$g : \mathcal{C}_{\tau'} \cap B_{\tau'}(\bar{x}, \epsilon) \to V_{\tau/2}$$

obtained in Corollary 6.1.2 is uniformly continuous.

Proof. If $\mathcal{F}$ and R are uniformly continuous then the Newton map defined by

$$\mathcal{N}(f)(x) = f(x) - R(x, f(x))[\mathcal{F}(x, f(x))]$$

maps any uniformly continuous function $f : \mathcal{C}_{\tau'} \cap B_{\tau'}(\bar{x}, \epsilon) \to V_{\sigma'}$ to a uniformly continuous function $N(f) : \mathcal{C}_{\tau'} \cap B_{\tau'}(\bar{x}, \epsilon) \to V_\sigma$ for $0 \le \sigma < \sigma' \le \tau' \le 1$. In light of Remark 6.1.3, since $\|\mathcal{F}(x,y)\|_Z \le \epsilon$ viewing the sequence y_n as a sequence of functions $g_n(x)$ with $g_0(x) = \bar{y}$ we have a uniformly convergent sequence of uniformly continuous functions and thus the limit $g(x) = g_\infty(x)$ will be uniformly continuous. □

Remark 6.1.5. *The choice of the $\tau/2$ scale is arbitrary. It is done primarily to keep notation clean (it also is convenient for generating approximations in* *the smooth case). By rescaling, or choosing different scales in the proof of Theorem 6.1.1, one can obtain* $y_\infty \in Y_{\tau''}$ *for any* $0 \le \tau'' < \tau < \tau' \le 1$. *The trade-off for choosing* $\tau'' > \tau/2$ *is the corresponding* δ *is smaller.*

Remark 6.1.6. *Extension (A) of Theorem 6.1.1 is useful for KAM theory for computing the measure of the KAM tori. Informally, the functions* Ω_R *and* Ω_A *and the set* $\mathcal{C} \subseteq X$ *are related to the choice of Diophantine conditions used for the frequency vectors (see for example* (8.10) *in Section 8.5). One wants to understand how scaling the Diophantine conditions to increase the size of* $\mathcal{C}$ *(which in turn increases* Ω_R *and* Ω_A*) effects* N *and, in particular,* δ *– which corresponds to the size of the perturbations considered. In [Pös82], while using very different methods, Poschel employs this idea of trading the size of the perturbation for the size of the Cantor set and obtains sharp estimates on the measure of KAM tori. Also see [Neĭ81]*

Remark 6.1.7. *Extension (B) of Theorem 6.1.1 reflects the growth conditions originally in [Zeh75]. As in [Zeh75], to establish the smooth (i.e.* $X_0^q \times Y_0^q$*) existences in Theorem 7.1.1 we repeatedly apply Theorem 6.1.1 to generate a sequence* $y_n \in V_{\tau_n}$ *with* $\tau_n = (2^n T)^{-1}$. *The comparison of* y_{n+1} *to* y_n *uses* (6.2), *so at each step extension (B) gives us* $\delta_n N_n \le \eta 2^{n(\alpha+\beta)}$. *In Theorem 6.1.1, this is combined with a certain smoothing of* $x \in X_0^q$ *to guarantee (using Definition 4.1.4) the sequence* y_n *converges to some* y_∞ *in* $Y_0^{q-(q*+\beta)}$.

It is possible to establish smooth (i.e. $X_0^q \times Y_0^q$*) existence results under what could be slightly more general conditions by tracking the* $\Psi_\Omega(\epsilon)$ *functions* *(see Definition 4.2.5) arising in Ω_Q, Ω_R and Ω_A and requiring a certain combination not exceed $C\epsilon^{-\alpha}$ for some α. However, this condition is overly awkward and is left to Question 4 in Appendix A.*

Remark 6.1.8. *Section 6.3 addresses the question of uniqueness of y_∞ (especially uniqueness for different y).*

Proof of Theorem 6.1.1

Let $\mathcal{F}$ as in (F0) satisfying Hypothesis (F.A0),(F.A1) and (F.A2) be given. The solution $y_\infty = y_\infty(x, y)$ to $\mathcal{F}(x, y_\infty) = 0$ is constructed by establishing the convergence of a "modified" Newton sequence $\{y_n\}$, defined inductively using the recurrence

$$y_{n+1} \equiv y_n - R(x, y_n)[\mathcal{F}(x, y_n)] \tag{6.8}$$

First, we develop "a priori" estimates for sequences satisfying (6.8). As discussed in Section 4.2 (see Remark 4.2.17) the definition of the Brjuno-Rüssmann condition is motivated primarily by these estimates. The second step of our proof is to use these "a priori" estimates and show that, provided $\|\mathcal{F}(x, y)\|_{Z_\tau}$ is sufficiently small, taking $y_0 \equiv y \in V_\tau$ and using (6.8), the sequence $\{y_n\}_{n=0}^{\infty}$ not only remains in V_0 (in fact $\|y_n - y\|_{Y_{\tau/2}} \leq N\|\mathcal{F}(x, y)\|_{Z_\tau}$) but in fact converges to some y_∞ in $V_{\tau/2}$ with $\mathcal{F}(x, y_\infty) = 0$ (and $\|y_\infty - y\|_{Y_{\tau/2}} \leq N\|\mathcal{F}(x, y)\|_{Z_\tau}$).

Given y_n and y_{n+1} satisfying (6.8), we have that $\mathcal{F} : U_\sigma \times V_\sigma \to Z_{\sigma'}$ and $R : \mathcal{C}_\sigma \times V_\sigma \to L(Z_\sigma, Y_{\sigma'})$ for $0 \leq \sigma' < \sigma \leq 1$, if $y_n \in V_{\sigma_n}$ then one must consider $R(x,y_n)[\mathcal{F}(x,y_n)]$, and thus y_{n+1}, in $Y_{\sigma_{n+1}}$ for $0 \le \sigma_{n+1} < \sigma_n \le 1$.

By Proposition 4.2.2, we can use the same sequence $\{\delta_n\}_{n=0}^{\infty}$ for condition (4.9) in the definition for the Brjuno-Rüssmann conditions for Ω_Q, Ω_R and Ω_A. Furthermore, by Proposition 4.2.4 one can assume $\sum_{n=1}^{\infty} \delta_n < (\tau/4)$. Define σ_n by

$$\sigma_n \equiv \tau - 2\left(\sum_{i=1}^{n} \delta_{i-1}\right) \tag{6.9}$$

Let $\tau_n = \sigma_n - \delta_n$ and note

$$(\tau/2) < \cdots < \sigma_{n+1} < \tau_n < \sigma_n < \cdots < \sigma_0 = \tau$$

With these scales, we will consider

$$y_n \in Y_{\sigma_n}, \quad \mathcal{F}(x,y_n) \in Z_{\tau_n}, \quad \text{and} \quad R(x,y_n)[\mathcal{F}(x,y_n)] \in Y_{\sigma_n - 2\delta_n}$$

To establish a priori bounds on y_n, we establish estimates of $\|y_{n+1} - y_n\|_{X_{\sigma_{n+1}}}$. Since

$$y_{n+1} - y_n = -R(x,y_n)[\mathcal{F}(x,y_n)]$$

using (4.32) it suffices to estimate $\|\mathcal{F}(x,y_n)\|_{X_{\tau_n}}$. Note that one has the identity

$$\begin{aligned}\mathcal{F}(x,y_{n+1}) = & \underbrace{\mathcal{F}(x,y_{n+1}) - \mathcal{F}(x,y_n) - D_2\mathcal{F}(x,y_n)[y_{n+1}-y_n]}_{(i)} \\ & + \underbrace{\mathcal{F}(x,y_n) + D_2\mathcal{F}(x,y_n)[y_{n+1}-y_n]}_{(ii)}\end{aligned} \tag{6.10}$$

We note (i) has the form of the quadratic remainder Q defined in (4.30) and since $y_{n+1} - y_n = -R(x, y_n)[\mathcal{F}(x, y_n)]$, applying (4.31) and (4.32) we get

$$\begin{aligned} \|(i)\|_{Z_{\tau_{n+1}}} &= \|Q(x; y_{n+1}, y_n)\|_{Z_{\sigma_{n+1}-\delta_{n+1}}} \\ &\leq \Omega_Q(\delta_{n+1}) \|R(x, y_n)[\mathcal{F}(x, y_n)]\|^2_{Y_{\sigma_{n+1}}} \\ &\leq \Omega_Q(\delta_{n+1}) \left(\Omega_R(\delta_n)\right)^2 \|\mathcal{F}(x, y_n)\|^2_{Y_{\tau_n}} \end{aligned} \tag{6.11}$$

Similarly, since $y_{n+1} - y_n = -R(x, y_n)[\mathcal{F}(x, y_n)]$ using (4.33) we get

$$\begin{aligned} \|(ii)\|_{Y_{\tau_{n+1}}} &= \|\mathcal{F}(x, y_n) - D_2\mathcal{F}(x, y_n)[R(x, y_n)[\mathcal{F}(x, y_n)]]\|_{Z_{\tau_n-\delta_n-\delta_{n+1}}} \\ &\leq \Omega_A(\delta_n + \delta_{n+1}) \|\mathcal{F}(x, y_n)\|^2_{Z_{\tau_n}} \end{aligned} \tag{6.12}$$

Combining (6.11) and (6.12) yields

$$\|\mathcal{F}(x, y_{n+1})\|_{Z_{\tau_{n+1}}} \leq \underbrace{\left(\Omega_Q(\delta_{n+1}) \left(\Omega_R(\delta_n)\right)^2 + \Omega_A(\delta_n + \delta_{n+1})\right)}_{=C(n)} \|\mathcal{F}(x, y_n)\|^2_{Z_{\tau_n}} \tag{6.13}$$

Defining

$$C(n) \equiv \left(\Omega_Q(\delta_{n+1}) \left(\Omega_R(\delta_n)\right)^2 + \Omega_A(\delta_n + \delta_{n+1})\right) \tag{6.14}$$

and

$$\epsilon_n = \|\mathcal{F}(x, y_n)\|_{Z_{\tau_n}}$$

estimate (6.13) has the same form as (4.15) in Section 4.2 and, as noted in Remark 4.2.17, using Proposition 4.2.13 the Brjuno-Rüssmann conditions for Ω_Q, Ω_R and Ω_A guarantee that $C(n)$ as defined in (6.14) will satisfy property (C1).

We are now ready to apply our "a priori" estimates to establish that y_n defined inductively by (6.8) not only remains in $V_{\tau/2}$ but in fact converges in $V_{\tau/2}$ to some y_∞ satisfying the desired properties. We begin by determining N and, more importantly, δ.

As noted above, the sequence $C(n)$ defined in (6.14) satisfies property (C1), i.e. there is a constant $M_C > 1$ with

$$\sum_{i=0}^{\infty} 2^{-(i+1)} \log(C(i)) \leq \log(M_C) < \infty \tag{6.15}$$

Since Ω_R satisfies the Brjuno-Rüssmann condition on $\{\delta_n\}$, by Remark 4.2.17 the sequence $\Omega_R(\delta_n)$ will also satisfy (C1). Hence, by Proposition 4.2.14, there is a constant $R_\Omega > 1$ so that

$$\Omega_R(\delta_n) \leq (R_\Omega)^{2^n} \tag{6.16}$$

Using $M_C \geq 1$ and $R_\Omega \geq 1$ choose positive constants δ and N satisfying

$$\delta \leq \frac{1}{3R_\Omega M_C} \tag{6.17}$$

$$N = \frac{3R_\Omega M_C}{2} \tag{6.18}$$

Given $(x, y) \in \mathcal{C}_\tau \times V_\tau$ satisfying (6.1), i.e.

$$\|\mathcal{F}(x,y)\|_{Z_\tau} < \delta \min(1, \operatorname{dist}(y, V^c_{\tau/2}))$$

define $y_0 = y \in V_{\sigma_0}$. Provided $y_n \in V_{\sigma_n}$, use (6.8) to inductively define y_{n+1} in terms of y_n. One can apply the a priori estimate (6.13) and Lemma 4.2.11, with $\epsilon_n = \|\mathcal{F}(x, y_n)\|_{Z_{\tau_n}}$, to get

$$\|\mathcal{F}(x, y_n)\|_{Z_{\tau_n}} = \epsilon_n \leq (\epsilon_0 M_C)^{2^n} \tag{6.19}$$

Applying (4.32) to $y_{n+1} - y_n = -R(x, y_n)[\mathcal{F}(x, y_n)]$, and combining with (6.19) and (6.16) one has

$$\|y_{n+1} - y_n\|_{Y_{\sigma_{n+1}}} \leq (\epsilon_0 R_\Omega M_C)^{2^n} \tag{6.20}$$

Note, since $\epsilon_0 R_\Omega M_C \leq 1/3$, we have

$$\sum_{i=0}^{n} (\epsilon_0 R_\Omega M_C)^{2^i} \leq \epsilon_0 R_\Omega M_C \underbrace{\left(\sum_{i=0}^{n} \left(\frac{1}{3}\right)^{2^i - 1} \right)}_{\leq 3/2} \leq \epsilon_0 N$$

and combining this with (6.20) we get

$$\sum_{i=0}^{n} \|y_{i+1} - y_i\|_{Y_{\sigma_{i+1}}} \leq \epsilon_0 N \tag{6.21}$$

Using a telescoping series, (6.21) gives us

$$\|y_{n+1} - y_0\|_{Y_{\sigma_{n+1}}} \leq N \|\mathcal{F}(x, y)\|_{Z_\tau} \tag{6.22}$$

Also, since $\epsilon_0 \leq \delta \text{dist}(y, V^c_{\tau/2})$, from (6.21) we get

$$\|y_{n+1} - y_0\|_{Y_{\sigma_{n+1}}} \leq \frac{1}{2} \min(1, \text{dist}(y, V^c_{\tau/2})) \tag{6.23}$$

so $y_{n+1} \in V_{\sigma_{n+1}}$ and therefore y_n can be defined inductively for all n.

To establish the convergence of $\{y_n\}$ in $Y_{(\tau/2)}$, using the inclusion of $Y_{\sigma_{n+1}} \to Y_{(\tau/2)}$, inequality (6.21) gives

$$\sum_{i=0}^{\infty} \|y_{i+1} - y_i\|_{Y_{(\tau/2)}} \leq \epsilon_0 N$$

so the sequence y_n is Cauchy in $Y_{(\tau/2)}$ and hence converges to some y_∞. Using the inclusion and taking $n \to \infty$ in (6.22) and (6.23) we get

$$\|y_\infty - y\|_{Y_{(\tau/2)}} \leq N \|\mathcal{F}(x, y)\|_{Z_\tau}$$

so y_∞ satisfies inequality (6.2), and

$$\|y_\infty - y\|_{Y_{(\tau/2)}} \leq \frac{1}{2}\min(1, \operatorname{dist}(y, V^c_{\tau/2}))$$

so $y_\infty \in V_{\tau/2}$. Finally, by Hypothesis (F.A0), $\mathcal{F} : U_{\tau/2} \times V_{\tau/2} \to Z_0$ is continuous and hence

$$\|\mathcal{F}(x, y_\infty)\|_{Z_0} = \lim_{n\to\infty} \|\mathcal{F}(x, y_n)\|_{Z_0} \leq \lim_{n\to\infty} \epsilon_n = 0 \tag{6.24}$$

and thus $\mathcal{F}(x, y_\infty) = 0$, which is the other half of (6.2).

To establish extension (A), writing

$$\Omega_Q(s) \equiv C_Q \Psi_Q(s)$$

$$\Omega_R(s) \equiv C_R \Psi_R(s)$$

and

$$\Omega_A(s) \equiv C_A \Psi_A(s)$$

the functions Ψ_Q, Ψ_R and Ψ_A will satisfy the Brjuno-Rüssmann condition so on some $\{\delta_n\}_{n=0}^\infty$ the sequence

$$C^*(n) \equiv \left(\Psi_Q(\delta_{n+1})\,(\Psi_R(\delta_n))^2 + \Psi_A(\delta_n + \delta_{n+1})\right) \tag{6.25}$$

will satisfy property (C1), i.e. there is a constant $M_{C*} \geq 1$ such that

$$\sum_{i=0}^{\infty} 2^{-(i+1)} \log(C^*(i)) \leq \log(M_{C*}) < \infty \tag{6.26}$$

Note that $C(n) \leq \max(C_Q C_R^2, C_A) C^*(n)$ so the constant

$$M_C = \max(C_Q C_R^2, C_A) M_{C*} \tag{6.27}$$

will satisfy (6.15).

We also note that since Ψ_R satisfies the Brjuno-Rüssmann condition on $\{\delta_n\}_{n=0}^{\infty}$, there is a constant $R_\Psi \geq 1$ with

$$\Psi_R(\delta_n) \leq (R_\Psi)^{2^n} \tag{6.28}$$

and, since $C_R \geq 1$, we have $\Omega_R(\delta_n) \leq (C_R R_\Psi)^{2^n}$ so the constant

$$R_\Omega = C_R R_\Psi \tag{6.29}$$

will satisfy (6.16). Substituting (6.27), (6.29) into (6.17), (6.18) and taking $M_\delta \leq \frac{1}{3R_\Psi M_{C^*}}$ and $M_\delta = \frac{3R_\Psi M_{C^*}}{2}$, we get (6.3), (6.4). This proves extension (A).

To establish extension (B), we will use the power growth estimates of Ω_Q, Ω_R, Ω_A to improve estimates (6.19), (6.20) and (6.21). Note that with $\Omega_Q(s) \leq C_Q s^{-\alpha}$, $\Omega_R(s) \leq C_R s^{-\beta}$ and $\Omega_A(s) \leq C_A s^{-\gamma}$, taking $\delta_n \equiv 2^{-n}(\tau/4)$ we can estimate $C(n)$, as defined in (6.14), by

$$\begin{aligned} C(n) &\leq 2^\alpha C_Q C_R^2 (\tau/4)^{-\alpha-2\beta} 2^{n(\alpha+2\beta)} + C_A(\tau/4)^{-\gamma} 2^{n\gamma} \\ &\leq \Bigg(\underbrace{A^* \frac{\max(C_Q C_R^2, C_A)}{\tau^{\max(\alpha+2\beta,\gamma)}}}_{=A}\Bigg) \Big(\underbrace{2^{\max(\alpha+2\beta,\gamma)}}_{=B}\Big)^n \end{aligned} \tag{6.30}$$

for some constant $A^* \geq 1$ which does not depend on C_Q, C_R, C_A or τ but only on α, β, γ. Using A and B, let δ and N satisfying

$$\delta \leq \frac{1}{3AB} = \left(\frac{1}{3A^*B}\right) \frac{\tau^{\max(\alpha+2\beta,\gamma)}}{\max(C_Q C_R^2, C_A)} \tag{6.31}$$

$$N = \left(\frac{3B}{2}\right) C_R \tau^{-\beta} \tag{6.32}$$

be given. Note that taking $M_\delta \leq \frac{1}{3A^*B}$ and $M_N = \frac{3A^*B}{2}$ then δ and N defined in (6.5) and (6.6) will satisfy (6.31) and (6.32). Choosing M_δ sufficiently small so that $M_\delta M_N \leq \eta$, since

$$\frac{\tau^{\max(\alpha+2\beta,\gamma)}}{\max(C_Q C_R^2, C_A)} C_R \tau^{-\beta} \leq \tau^{\alpha+\beta}$$

one has $\delta N \leq \eta\tau^{\alpha+\beta}$.

Using (6.30) and applying Corollary 4.2.12 as in (4.21) from Example 4.2.16 one has

$$\|\mathcal{F}(x, y_n)\|_{Z_{\tau_n}} = \epsilon_n \leq (\epsilon_0 D(n))^{2^n} = \frac{(\epsilon_0 AB)^{2^n}}{AB^{(n+1)}} \tag{6.33}$$

This improves (6.19). Combining the estimate (which is significantly better than (6.16))

$$\Omega_R(\delta_n) \leq C_R \left(\frac{2^{-n}\tau}{4}\right)^{-\beta}$$

(which is significantly better than (6.16)) with (6.33) we can improve estimate (6.20) and get

$$\|y_{n+1} - y_n\|_{Y_{\sigma_{n+1}}} \leq \underbrace{\left(2^\beta \frac{C_R \tau^{-\beta}}{A}\right)}_{\leq 1} \underbrace{\left(\frac{2^\beta}{B}\right)}_{\leq 1}^{n+1} (\epsilon_0 AB)^{2^n} \tag{6.34}$$

Similarly, we can improve (6.21) to

$$\sum_{i=0}^{n} \|y_{i+1} - y_i\|_{Y_{\sigma_{i+1}}} \leq \epsilon_0 B C_R \tau^{-\beta} \underbrace{\left(\sum_{i=0}^{n} (\epsilon_0 AB)^{2^i - 1}\right)}_{\leq 3/2} \tag{6.35}$$

and using $\epsilon_0 \leq \delta \mathrm{dist}(y, V_{\tau/2}^c)$, with δ as in (6.31), we again get (6.23) so $y_{n+1} \in V_{\sigma_{n+1}}$ and the inductive definition of the sequence $\{y_n\}$ can be carried out for all n. Using the inclusion of $Y_{\sigma_{i+1}} \to Y_{\tau/2}$ in (6.35) we again have that y_i is Cauchy in $Y_{\tau/2}$ and its limit y_∞ satisfies

$$\|y - y_\infty\|_{Y_{(\tau/2)}} \leq \underbrace{\left(\frac{3B}{2}\right) C_R \tau^{-\beta}}_{=N} \|\mathcal{F}(x,y)\|_{Z_\tau} \tag{6.36}$$

From (6.23) we again have so $y_\infty \in V_{\tau/2}$ and using the continuity of $\mathcal{F}(x,\cdot)$ we have (6.24) so $\mathcal{F}(x, y_\infty) = 0$. This establishes extension (B) and completes the proof of Theorem 6.1.1. □

Remark 6.1.9. *In Section 7.1 these results (Theorem 6.1.1) are extended to the smooth case (i.e. $X_0^q \times Y_0^q$) using the analytic smoothing discuss in Section 4.1 (see Theorem 7.1.1).*

6.2 Whitney Regularity in analytic spaces

We now establish implicit solutions with Whitney regularity.

Theorem 6.2.1. *Let $\mathcal{F}$ satisfying Hypothesis (F.W1), (F.W2) and (F.W3). For any $(\bar{x}, \bar{y}) \in U_\tau \times V_\tau$ with $\mathcal{F}(\bar{x}, \bar{y}) = 0$, there exists*

$$\bar{\epsilon} > 0 \quad \textit{and} \quad g \in C^\gamma_{Wht}(\mathcal{C}_\tau \cap B_\tau(\bar{x}, \bar{\epsilon}), Y_{\tau/2})$$

with $g : \mathcal{C}_\tau \cap B_\tau(\bar{x}, \bar{\epsilon}) \to V_{\tau/2}$ such that $\mathcal{F}(x, g(x)) = 0$.

Proof. As in Corollary 6.1.2, the proof is simply an application of Theorem 6.1.1. However, we twist around the role played by $x \in X_\sigma$. Specifically, using the notation $C^\gamma_{Wht}(A, Y)$ described in Definition 3.1.1, for fixed $\gamma > 1$,

$k < \gamma \leq k+1$, given any $(\bar{x}, \bar{y}) \in \mathcal{C}_\tau \times V_\tau$ satisfying $\mathcal{F}(\bar{x}, \bar{y}) = 0$, fix $\bar{\epsilon} > 0$ and define the one parameter families of Banach spaces

$$\mathbb{X}_\sigma = \{0\}, \quad \mathbb{Y}_\sigma = C^\gamma_{Wht}(\mathcal{C} \cap B_{X_\tau}(\bar{x}, \bar{\epsilon}), Y_\sigma), \quad \mathbb{Z}_\sigma = C^\gamma_{Wht}(\mathcal{C} \cap B_{X_\tau}(\bar{x}, \bar{\epsilon}), Z_\sigma)$$

for $0 \leq \sigma < \tau \leq 1$. Note that in our definitions of $\mathbb{X}_\sigma$, $\mathbb{Y}_\sigma$ and $\mathbb{Z}_\sigma$, we have made $\mathbb{X}_\sigma$ trivial and placed the $x \in \mathcal{C}_\sigma \cap X_\sigma$ dependence as part of $\mathbb{Y}_\sigma$ and $\mathbb{Z}_\sigma$.

Let $\bar{\mathbf{x}} = 0 \in \mathbb{X}_\sigma$, $\bar{\mathbf{y}} \in \mathbb{Y}_\sigma$ with $\bar{\mathbf{y}}(x) = g^{\leq k}(\bar{x}, \bar{y}; x - \bar{x})$ and define the subsets

$$\mathbb{U}_0 = \{0\} \subseteq \mathbb{X}_0, \quad \mathbb{V}_0 = B_{\mathbb{Y}_0}(\bar{\mathbf{y}}, \bar{\epsilon}) \subseteq \mathbb{Y}_0$$

For $0 \leq \sigma' < \sigma < \tau$, define $\mathbf{z}(x) = F(x, \mathbf{y}(x))$ and note that by Theorem 3.1.8 for $\mathbf{y} \in \mathbb{V}_\sigma$, using (F.W1) we get $\mathbf{z} \in \mathbb{Z}_{\sigma'}$. Furthermore, (3.8) gives us

$$\mathbb{F} : \mathbb{U}_\sigma \times \mathbb{V}_\sigma \to \mathbb{Z}_{\sigma'}$$

is continuous for every $0 \leq \sigma' < \sigma < \tau$, i.e. $\mathbb{F}$ satisfies Hypothesis (F.A0).

Furthermore, note that by (3.9) $\mathbb{F}$ is differentiable in $\mathbb{Y}$ with

$$D_2\mathbb{F}(\mathbf{x}, \mathbf{y})[\mathbf{v}](x) = D_2\mathcal{F}(x, \mathbf{y}(x))[\mathbf{v}(x)]$$

and with

$$\|\mathbb{Q}(\mathbf{x}; \mathbf{y}_1, \mathbf{y}_2)\|_{\mathbb{Z}_{\sigma'}} \leq \Omega_{\mathbb{Q}}(\sigma - \sigma')\|\mathbf{y}_1 - \mathbf{y}_2\|^2_{\mathbb{Y}_\sigma}$$

for $\mathbb{Q}(\mathbf{x}; \mathbf{y}_1, \mathbf{y}_2) = \mathbb{F}(\mathbf{x}, \mathbf{y}_1) - \mathbb{F}(\mathbf{x}, \mathbf{y}_2) - D_2\mathbb{F}(\mathbf{x}, \mathbf{y}_2)[\mathbf{y}_1 - \mathbf{y}_2]$ and $\Omega_{\mathbb{Q}} = M_{\bar{\epsilon}}\Omega_F$. Thus, $\mathbb{F}$ satisfies (F.A1).

In a similar manner, for $0 \leq \sigma' < \sigma < \tau$, given $\mathbf{y} \in \mathbb{V}_\sigma$ and $\mathbf{v} \in \mathbb{Z}_{\sigma'}$, using (F.W2) we can define $\mathbb{R}(\mathbf{x}; \mathbf{y})[\mathbf{z}](x) = R(x, \mathbf{y}(x))[\mathbf{z}(x)]$ and by Theorem 3.1.8 we get

$$\mathbb{R} : \mathbb{U}_\sigma \times \mathbb{V}_\sigma \to L(\mathbb{Z}_\sigma, \mathbb{Y}_{\sigma'})$$

Note that using (4.35) we get

$$\|\mathbb{R}(\mathbf{x}, \mathbf{y})[\mathbf{v}]\|_{\mathbb{Y}_{\sigma'}} \leq \Omega_{\mathbb{R}}(\sigma - \sigma')\|\mathbf{v}\|_{\mathbb{Z}_\sigma}$$

with $\Omega_{\mathbb{R}} = M_{\bar{\epsilon}}\Omega_R$. Furthermore, from (4.36) we get

$$\| \ [\mathrm{Id} - D_2\mathbb{F}(\mathbf{x}, \mathbf{y})\mathbb{R}(\mathbf{x}, \mathbf{y})][\mathbf{v}] \ \|_{\mathbb{Z}_{\sigma'}} \leq \Omega_{\mathbb{A}}(\sigma - \sigma')\|\mathbb{F}(\mathbf{x}, \mathbf{y})\|_{\mathbb{Z}_\sigma}\|\mathbf{v}\|_{\mathbb{Z}_\sigma}$$

with $\Omega_{\mathbb{A}} = M_{\bar{\epsilon}}\Omega_A$ and thus $\mathbb{R}$ satisfies (F.A2).

In order to apply Theorem 6.1.1, all that remains is to show

$$\|\mathbb{F}(0, \bar{\mathbf{y}})\|_{\mathbb{Z}_\tau} \leq \delta\mathrm{dist}(\mathbf{y}, \mathbb{V}^c_{\tau/2})$$

Take $\bar{\delta} = \delta\mathrm{dist}(\mathbf{y}, \mathbb{V}^c_{\tau/2})$ for a fixed $\bar{\epsilon}$ for and note as $\bar{\epsilon}$ decreases so does $\delta\mathrm{dist}(\mathbf{y}, \mathbb{V}^c_{\tau/2})$, hence it is sufficient if we can establish

$$\|\mathbb{F}(0, \bar{\mathbf{y}})\|_{\mathbb{Z}_\tau} \leq \bar{\delta} \tag{6.37}$$

for $\bar{\epsilon}$ sufficiently small.

Note that, for $0 \leq i \leq k$, $\mathcal{W}\mathcal{D}^i_x\mathbb{F}(0, \bar{\mathbf{y}})(\bar{x}) = 0$ and thus, with $\bar{\epsilon}$ sufficiently small we can ensure $|\mathcal{W}\mathcal{D}^i_x\mathbb{F}(0, \bar{\mathbf{y}})(x)| \leq \bar{\delta}$ for all $x \in \mathcal{C}_\tau \cap B_{X_\tau}(\bar{x}, \bar{\epsilon})$ and thus (6.37) holds.

Applying of Theorem 6.1.1, we obtain $\mathbf{y}_\infty$ with $\mathbb{F}(\bar{\mathbf{x}}, \mathbf{y}_\infty) = 0$. Unrolling this, we have $\mathbb{F}(\bar{\mathbf{x}}, \mathbf{y}_\infty)(x) = \mathcal{F}(x, \mathbf{y}_\infty(x)) = 0$ for all $x \in \mathcal{C}_\tau \cap B_{X_\tau}(\bar{x}, \bar{\epsilon})$ and hence $g = \mathbf{y}_\infty$ is our desired C^γ_{Wht} implicit function.

□

Remark 6.2.2. *Note that if one replaces $\bar{y}$ with $g^{\leq k}(\bar{x}, \bar{y}; x - \bar{x})$ in Corollary 6.1.2 and restricts ϵ to the (possibly smaller) $\bar{\epsilon}$, the iterations of Theorem 6.1.1 to obtain in Corollary 6.1.2 actually coincide with the iterations $\mathbf{y}_n$ of Theorem 6.1.1 used to obtain Theorem 6.2.1. Hence, the function $\mathbf{y}_\infty$ obtained in Theorem 6.2.1 and g from Corollary 6.1.2 coincide.*

Given that the zeros of $\mathcal{F}$ are isolated (that is $\mathcal{F}$ has some local uniqueness in y for solutions $\mathcal{F}(x, y) = 0$) we have the following alternative approach to establish the Whitney regularity of any function g which solves $\mathcal{F}(x, g(x)) = 0$ by explicitly verify the estimates for the Whitney Regularity of g:

Theorem 6.2.3. *Let $\mathcal{F}$ be given as in (F0) satisfying the additional Hypotheses (F.A0),(F.A1) and (F.A2).*

Assume $\mathcal{F}$ has local uniqueness in y for solutions $\mathcal{F}(x, y) = 0$. If $\mathcal{F}$ also satisfies (F.W1) and (F.W2) for some $\gamma > 1$, $k < \gamma \leq k + 1$, and either:

(a) $\mathcal{C}$ has the γ density property described in Definition 3.2.5, or

(b) R satisfies (F.W4)

then given any function

$$g : \mathcal{C}_{\tau'} \cap B_{\tau'}(\bar{x}, \epsilon) \to V_{\tau/2}$$

with $\mathcal{F}(x, g(x)) = 0$ *is* C^{γ}_{Wht} *with the k-jet of* g *having the same form as the coefficients* $g_i(x_0, g(x_0))$ *of the polynomial approximate solutions defined in* (5.2) *of Theorem 5.0.12.*

Proof. Fix x and using $(\bar{x}, \bar{y}) = (x, g(x))$ apply Theorem 5.0.12 to construct $g^{\leq k}(x, g(x); \delta)$. For Δ sufficiently small, taking $\sigma < \tau/2$ by Theorem 5.0.12

$$\left\| \mathcal{F}(x + \Delta, g^{\leq k}(x)(x, g(x); \Delta)) \right\|_{Z_\sigma} \leq M \|\Delta\|^{\gamma}_{X_{\tau/2}} \tag{6.38}$$

Provided $x + \Delta \in \mathcal{C}$, (6.38) allows us to apply Theorem 6.1.1 and obtain y_∞ with $\mathcal{F}(x + \Delta, y_\infty) = 0$ and

$$\left\| y_\infty - g_0^{\leq k}(x, g(x); \Delta) \right\|_{Z_\sigma} \leq \eta N M \|\Delta\|^{\gamma}_{X_{\sigma/2}} \tag{6.39}$$

With local uniqueness (for example Corollary 6.3.2) since

$$\mathcal{F}(x + \Delta, g(x + \Delta)) = 0$$

we have $y_\infty = g(x + \Delta)$. Substituting this into (6.39) we have

$$\left\| g(x + \Delta) - g_0^{\leq k}(x, g(x); \Delta) \right\|_{Z_{\sigma/2}} \leq \eta N M \|\Delta\|^{\gamma}_{X_\tau} \tag{6.40}$$

If we are in case (a) of the theorem, combining (6.40) with the Whitney Verification Lemma II (Lemma 3.2.6) gives us that $g \in C^{\gamma}_{Wht}$. On the other hand, given case (b) we can combine (6.40) with the Whitney Verification Lemma I (Lemma 3.1.10) and again obtain that $g \in C^{\gamma}_{Wht}$. □

6.3 Uniqueness in analytic spaces

Now we consider the question of uniqueness for solutions in the analytic spaces.

Theorem 6.3.1. *Let $\mathcal{F}$ as in (F0) satisfying all Hypothesis (F.A0), (F.A1) and (F.AU). There exists constant $\epsilon > 0$ (depending only on τ, Ω_Q and Ω_L) such that for any $y_1, y_2 \in Y_\tau$ with $\mathcal{F}(x, y_i) = 0$, if $\|y_1 - y_2\|_{Y_\tau} < \epsilon$ then $y_1 = y_2$. Moreover, paralleling (A) and (B) in Theorem 6.1.1 we have:*

(A) Writing $\Omega_Q(s) \equiv C_Q\Psi_Q(s)$ and $\Omega_L(s) \equiv C_L\Psi_L(s)$ where C_Q, C_L are constants and $\Psi_Q, \Psi_L : (0, 1] \to [1, \infty)$ are functions which "carry the shape" of Ω_Q, Ω_L, the constant ϵ can be chosen as follows

$$\epsilon = M_\epsilon \frac{1}{C_Q C_L} \tag{6.41}$$

where M_ϵ depends only on Ψ_Q and Ψ_L and τ.

(B) If $\Omega_Q(s) \le C_Q s^{-\alpha}$ and $\Omega_L(s) \le C_L s^{-\beta^*}$ then the constant ϵ can be chosen as follows

$$\epsilon = M_\epsilon \frac{\tau^{\alpha+\beta^*}}{C_Q C_L} \tag{6.42}$$

where M_ϵ depends only on α and β^*.

Proof of Theorem 6.3.1

By Propositions 4.2.2 and 4.2.4 we can assume the Ω_Q and Ω_L satisfy the Brjuno-Rüssmann condition on $\{\delta_n\}$ with $\sum_{i=0}^{\infty} \delta_i < \tau/3$. Set

$$\sigma_n \equiv \tau - 3\left(\sum_{i=1}^{n} \delta_{i-1}\right)$$

and note by Remark 4.2.13 the sequence

$$C(n) \equiv \Omega_L(\delta_n)\Omega_Q(\delta_n)$$

has property (C1) so that one has

$$\sum_{i=1}^{\infty} 2^{-(i+1)} \log(C(i)) \leq \log(M_C) < \infty$$

for some $M_C > 1$. Set

$$\epsilon = \epsilon_0 < \frac{1}{M_C}$$

For any $y_1, y_2 \in Y_\tau$ with $\mathcal{F}(x, y_i) = 0$ and $\|y_1 - y_2\|_{Y_\tau} < \epsilon$. Using the left inverse L, note

$$\begin{aligned}
\|y_1 - y_2\|_{Y_{\sigma_{n+1}}} &\leq \|L(x, y_2)[\mathcal{F}(x, y_1) - \mathcal{F}(x, y_2) - D_2\mathcal{F}(x, y_2)[y_1 - y_2]]\|_{Y_{\sigma_n - 2\delta_2}} \\
&\leq \Omega_L(\delta_n)\|\mathcal{F}(x, y_1) - \mathcal{F}(x, y_2) - D_2\mathcal{F}(x, y_2)[y_1 - y_2]\|_{Z_{\sigma_n - \delta_n}} \\
&\leq \Omega_L(\delta_n)\Omega_Q(\delta_n)\|y_1 - y_2\|^2_{Y_{\sigma_n}}
\end{aligned}$$

Letting $\epsilon_n \equiv \|y_1 - y_2\|_{Y_{\sigma_n}}$ note that, applying Lemma 4.2.11, one gets $\epsilon_n \leq (\epsilon_0 M_C)^{2^n} \to 0$. Thus $\|y_1 - y_2\|_{Y_0} = 0$, i.e. $y_1 = y_2$.

The proofs of the (A) and (B) are straight forward and left to the reader. □

The following corollary establishes uniqueness for the modified Newton method used in the proof of Theorem 6.1.1.

Corollary 6.3.2. *Let $\mathcal{F}$ satisfying the Hypothesis of Theorem 6.1.1 and 6.3.1 be given with N and δ the constants which arise in Theorem 6.1.1 at the τ scale* *and ϵ the constant which arises in Theorem 6.3.1 at the $\tau/2$ scale. Provided δ is taken small enough that*

$$\delta N < \epsilon/3 \tag{6.43}$$

given any

$$(x, y_1), (x, y_2) \in \mathcal{C}_{\tau'} \times V_{\tau'}$$

satisfying (6.1) *with*

$$\|y_1 - y_2\|_{Y_{\tau'}} < \epsilon/3$$

then, the solutions $y_s^1 = y_s(x, y_1)$, $y_s^2 = y_s(x, y_2)$ are equal, i.e. the function $y_s = (x, y)$ is locally constant in y.

Proof of Corollary 6.3.2

Let $y_s^1 = y_s(x, y_1)$ and $y_s^2 = y_s(x, y_2)$ be the solutions which arise by applying Theorem 6.1.1. Note

$$\left\|y_s^1 - y_s^2\right\|_{Y_{\tau/2}} \leq \left\|y_s^1 - y_1\right\|_{Y_{\tau/2}} + \left\|y_s^2 - y_2\right\|_{Y_{\tau/2}} + \|y_1 - y_2\|_{Y_{\tau/2}} < \epsilon$$

so applying Theorem 6.3.1 we get $y_\infty^1 = y_\infty^2$. □

Remark 6.3.3. *Note that writing Ω_Q, Ω_R, Ω_A and Ω_L as in extension (B) of Theorem 6.1.1 and Theorem 6.3.1, if $\beta \geq \beta^*$ and M_δ from Theorem 6.1.1 is chosen so that*

$$M_\delta \leq \frac{M_\epsilon \max(C_Q C_R^2, C_A)}{3 M_N C_R C_Q C_L} \tag{6.44}$$

then condition (6.43) *is satisfied at all scales* $0 \leq \tau \leq 1$.

Chapter 7

Solutions in Smooth Spaces

We now demonstrate how the quantitative estimates in extension (B) of Theorem 6.1.1 can be combined with the analytic smoothing discussed in Section 4.1 to form an iteration scheme which establishes the existence of solutions in the spaces $X_0^q \times Y_0^q$. Such smoothing was used in [Mos66b, Mos66a] and [Zeh75] to establish the existence of smooth solutions. The main difference in our approach is that, rather than developing an implicit function solution around *an analytic solution*, we develop our implicit function solution around any *smooth solution*, i.e. rather than $(\bar{x}, \bar{y}) \in U_\sigma \times V_\sigma$ with $\mathcal{F}(\bar{x}, \bar{y}) = 0$ we only need $(\bar{x}, \bar{y}) \in U_0^q \times V_0^q$ with $\mathcal{F}(\bar{x}, \bar{y}) = 0$.

7.1 Existence in smooth spaces via analytic smoothing

Combining analytic smoothing with the quantitative estimates in extension (B) of Theorem 6.1.1, we can apply Theorem 6.1.1 in an iteration scheme and establish the following:

Theorem 7.1.1. *Given $\mathcal{F}$ and X_σ, Y_σ, Z_σ satisfying (F0), (XYZ.S1) and (XYZ.S2), assume in addition that $\mathcal{F}$ also satisfies Hypotheses (F.S0)-(F.S4).*

Then, for any $q > \max(\alpha + 2\beta, \gamma) + q*$ and $(\bar{x}, \bar{y}) \in \mathcal{C}_0^q \times V_0^q$ there exists *positive constants r, δ and T_0, depending on q, $q*$, α, β, γ and $(\bar{x}, \bar{y})$, such that for any $(x, y) \in (\mathcal{C}_0^q \cap B_0^q(\bar{x}, r)) \times B_0^q(\bar{y}, r)$, with*

$$\|\mathcal{F}(x, y)\|_{Z_0^q} \leq \delta \tag{7.1}$$

there exists a family $y_\infty = y_\infty(x, y, T) \in V_0^{q-\beta}$ for $T_0 \leq T \leq T_\infty$ with $\mathcal{F}(x, y_\infty) = 0$ (here T represents the smoothing taken before applying Theorem 6.1.1 and generating a sequence $y_n \to y_\infty$ in $Y_0^{q-\beta}$). If $\mathcal{F}(x, y) \neq 0$ we have

$$T_\infty = T_0 \left(\frac{\delta}{\|\mathcal{F}(x, y)\|_{Z_0^q}} \right)^{1/\max(\alpha+2\beta,\gamma)} \tag{7.2}$$

while if $\mathcal{F}(x, y) = 0$ we have $T_\infty = \infty$, i.e. the family $y_\infty(x, y, T) \in V_0^{q-\beta}$ exists for all $T_0 \leq T < \infty$.

If $q \geq \max(\alpha + 2\beta, \gamma) + \beta + q*$ there exists a positive constant N, depending on q, $q*$, α, β, γ and $(\bar{x}, \bar{y})$, such that, by optimizing the choice of T, we have $y_\infty = y_\infty(x, y)$ with

$$\|y - y_\infty\|_{Y_0^{q-\beta}} \leq N \|\mathcal{F}(x, y)\|_{Y_0^q} \tag{7.3}$$

Remark 7.1.2. *In Section 7.3 we address the question of uniqueness of y_∞ (especially uniqueness for different y and T).*

As with Theorem 6.1.1 and Corollary 6.1.2, we can use Theorem 7.1.1 to obtain the following useful:

Corollary 7.1.3. *Given $\mathcal{F}$ as in Theorem 7.1.1, for any $q > \max(\alpha + 2\beta, \gamma)$ and $(\bar{x}, \bar{y}) \in \mathcal{C}_0^q \times V_0^q$ with $\mathcal{F}(\bar{x}, \bar{y}) = 0$, there exists a positive constant ϵ and a function $g : \mathcal{C}_0^q \cap B_0^q(\bar{x}, \epsilon) \to V_0^{q-(q*+\beta)}$ with $\mathcal{F}(x, g(x)) = 0$.*

Proof. Identical to the proof of Corollary 6.1.2, using Theorem 7.1.1 in place of Theorem 6.1.1. □

Proof of Theorem 7.1.1

First, we construct suitable choices for r, δ and T_0. Informally, the idea is to choose r and δ sufficiently small so that for a given (x, y), one can apply the smoothing S_T to (x, y) with the smoothing parameter T satisfying $1 \le T_0 \le T \le T_\infty$ and use $(x_0, y_0) = (S_T[x], S_T[y])$ as the starting point of a sequence (x_n, y_n) which will converge to (x, y_∞). We need T, which represents the initial smoothing, to be sufficiently large so that the U_0 and V_0 invariance gives us $(x_0, y_0) \in U_0 \times V_0$. In fact, we take T large enough to ensure that $y_0 = S_T[y]$ is a bounded distance away from the boundary of V_0, see (7.6). We also want T to be large enough so that, as in Remark 4.3.8 we can ensure that $M_4(q)T^{-q+q*}$ is sufficiently small, see (7.9) and (7.10). Finally, T_∞, as defined in (7.2), is an upper bound on T which ensures that one can estimate $(x_0, y_0) \in U_\sigma \times V_\sigma$ for an analytic σ bounded away from 0, in particular $\sigma \ge T_\infty^{-1}$.

Choosing r, δ and T_0: To begin, choose r sufficiently small and $T_0 \ge 1$ sufficiently large to obtain $\mathcal{C}_0$ invariance around $\bar{x}$ as described in Definition 4.1.17. Without loss of generality, assume

$$r < d^*/k(q)$$

where $k(q)$ is the constant which arises in the analytic smoothing (see Definition 4.1.9) and d^* is any positive constant with

$$d^* < d = \min(1, \mathrm{dist}(\bar{x}, U_0^c), \mathrm{dist}(\bar{y}, V_0^c))$$

Similarly, without loss of generality, assume T_0 is sufficiently large so that for $t \geq T_0$

$$\|S_t[\bar{y}] - \bar{y}\|_{Y_0} \leq \left(\frac{d - d^*}{2}\right) \tag{7.4}$$

Note that for any y with $\|y - \bar{y}\|_{Y_0} < r$ one has

$$\|S_t[y - \bar{y}]\|_{Y_{t^{-1}}} \leq k(q)\|y - \bar{y}\|_{Y_0^q} \leq d^*$$

and thus

$$\|S_t[y] - \bar{y}\|_{Y_0} \leq \|S_t[y - \bar{y}]\|_{Y_{t^{-1}}} + \|S_t[\bar{y}] - \bar{y}\|_{Y_0} < \left(\frac{d + d^*}{2}\right) \tag{7.5}$$

so we have

$$\left(\frac{d - d^*}{2}\right) \leq \mathrm{dist}(S_t[y], V_0^c) \leq \mathrm{dist}(S_t[y], V_{(2t)^{-1}}^c) \tag{7.6}$$

Hence, with this choice of r and T_0, given any x with $\|x - \bar{x}\|_{X_0} < r$ and y with $\|y - \bar{y}\|_{Y_0} < r$, for any $t > T_0$, both $S_t[x]$ and $S_t[y]$ remain in $\mathcal{C}_0$ and V_0 and $S_t[y]$ will remain at least a distance of $(d - d^*)/2$ from the boundary.

To choose δ let M_δ and M_N denote the constants in the bounds on δ and N in (6.5) and (6.6) of extension (B) in Theorem 6.1.1 and for $\sigma = t^{-1}$ let $\delta_T(t)$ and $N_T(t)$ denote the RHS of (6.5) and (6.6), i.e.

$$\delta_T(t) \equiv M_\delta \frac{t^{-\max(\alpha+2\beta,\gamma)}}{\max(C_Q C_R^2, C_A)} \quad \text{and} \quad N_T(t) \equiv M_N C_R t^\beta \tag{7.7}$$

Without loss of generality, assume M_δ is sufficiently small so that the η which arises in extension (B) satisfies

$$\eta \le \frac{d-d^*}{4}(1-2^{-(\alpha+\beta)})$$

and hence

$$\sum_{i=1}^{n} \eta(2^{-(i-1)})^{\alpha+\beta} \le \left(\frac{d-d^*}{4}\right) \tag{7.8}$$

Again without loss of generality, assume $T_0 \ge 1$ is sufficiently large that in addition the $\mathcal{C}_0$ invariance and (7.4) above, for $t \ge T_0$ one has

$$M_5 t^{-q+q*} \le \delta_T(t)\left(\frac{d-d^*}{4}\right) \tag{7.9}$$

where

$$M_5 = \max\left(M_3 k(q)(\|\bar{x}\|_{X_0^q} + \mathrm{dist}(\bar{x}, (U_0^q)^c)), M_4(q)\right)$$

Note that combining the smoothing estimate in Hypothesis (F.S4) with (7.9), for $t \ge T_0$ we have

$$\|\mathcal{F}(S_t x, S_t y) - S_t\mathcal{F}(x,y)\|_{Z_{t^{-1}}} \le \delta_T(t)\left(\frac{d-d^*}{4}\right) \tag{7.10}$$

With this T_0 we choose

$$\delta = \delta_T(T_0)\left(\frac{d-d^*}{4k(q)}\right) \tag{7.11}$$

Construction of y_∞: With r, δ and T_0 chosen as above, we are ready to begin. Given

$$(x,y) \in (\mathcal{C}_0^q \cap B_0^q(\bar{x}, r)) \times B_0^q(\bar{y}, r)$$

satisfying (7.1) choose T with $T_0 \leq T \leq T_\infty$ and note from (7.2) we have

$$\|\mathcal{F}(x, y)\|_{Z_0^q} \leq \delta \left(\frac{T_0}{T}\right)^{\max(\alpha+2\beta,\gamma)} \tag{7.12}$$

Define the sequences:

$$\sigma_n = (2^n T)^{-1} \quad \text{and} \quad x_n = S_{2^n T}[x]$$

Note, for $T \geq T_0$ the invariance property described in Definition 4.1.17 ensures $x_n \in \mathcal{C}_{\sigma_n}$. Furthermore, by Definition 4.1.4 we have $x_n \to x$ in X_0. We will use x_n to inductively define a sequence $y_n \in V_{\sigma_n}$ with $y_n \to y_\infty$ in Y_0 and $y_\infty \in Y_0^{q-\beta}$.

We begin the inductive definition of $y_n \in V_{\sigma_n}$ with $y_0 = S_T[y] \in V_{\sigma_0}$. Note that (7.11), (7.12) and the smoothing estimate (4.4) give us

$$\|S_T \mathcal{F}(x, y)\|_{Z_{\sigma_0}} \leq k(q)\|\mathcal{F}(x, y)\|_{Z_0^q} \leq \delta_T(T) \left(\frac{d - d^*}{4}\right)$$

Hence, combining with (7.10) and (7.6), we get

$$\begin{aligned} \|\mathcal{F}(S_T x, S_T y)\|_{Z_{\sigma_0}} &\leq \|\mathcal{F}(S_T x, S_T y) - S_T \mathcal{F}(x, y)\|_{Z_{\sigma_0}} + \|S_T \mathcal{F}(x, y)\|_{Z_{\sigma_0}} \\ &\leq \delta_T(T) \left(\frac{d - d^*}{2}\right) \\ &\leq \delta_T(T) \min(1, \operatorname{dist}(S_T[y], V_{\sigma_1}^c)) \end{aligned} \tag{7.13}$$

Applying extension (B) of Theorem 6.1.1 to $(x_0, y_0) = (S_T x, S_T y) \in \mathcal{C}_{\sigma_0} \times V_{\sigma_0}$, we obtain $y_1 = y_\infty(x_0, y_0) \in V_{\sigma_1}$ with $\mathcal{F}(x_0, y_1) = 0$ and

$$\|y_1 - y_0\|_{Y_{\sigma_1}} \leq N_T(T)\|\mathcal{F}(x_0, y_0)\|_{Z_{\sigma_0}} \leq \eta(T^{-1})^{\alpha+\beta} \tag{7.14}$$

Combining this with (7.8) and (7.5) we have:

$$\begin{aligned}\|y_1 - \bar{y}\|_{Y_0} &\leq \|y_0 - \bar{y}\|_{Y_0} + \|y_1 - y_0\|_{Y_{\sigma_1}} \\ &\leq \left(\frac{d+d^*}{2}\right) + \left(\frac{d-d^*}{4}\right) = d - \left(\frac{d-d^*}{4}\right)\end{aligned} \tag{7.15}$$

and hence

$$\frac{d-d^*}{4} \leq \operatorname{dist}(y_1, V_0^c) \leq \operatorname{dist}(y_1, V_{\sigma_2}^c) \tag{7.16}$$

The inductive step: Inductively, assume that $y_n \in Y_{\sigma_n}$ has been defined for $n \leq m$ with $\mathcal{F}(x_{n-1}, y_n) = 0$ and, as in (7.14),

$$\begin{aligned}\|y_n - y_{n-1}\|_{Y_{\sigma_n}} &\leq N_T(2^n T)\|\mathcal{F}(x_{n-1}, y_{n-1})\|_{Z_{\sigma_{(n-1)}}} \\ &\leq \eta(2^{(n-1)}T)^{-(\alpha+\beta)}\end{aligned} \tag{7.17}$$

As in (7.15) above, we can take a telescoping sequence and combine (7.17) and (7.8), and since $T \geq 1$, we obtain

$$\begin{aligned}\|y_m - \bar{y}\|_{Y_0} &\leq \|y_0 - \bar{y}\|_{Y_0} + \sum_{i=1}^{m} \|y_i - y_{i-1}\|_{Y_{\sigma_i}} \\ &\leq \left(\frac{d+d^*}{2}\right) + \sum_{i=1}^{\infty} \eta(2^{-(i-1)})^{\alpha+\beta} \\ &\leq \left(\frac{d+d^*}{2}\right) + \left(\frac{d-d^*}{4}\right) = d - \left(\frac{d-d^*}{4}\right)\end{aligned} \tag{7.18}$$

and thus, as in (7.16), we have

$$\frac{d-d^*}{4} \leq \operatorname{dist}(y_m, V_0^c) \leq \operatorname{dist}(y_m, V_{\sigma_{(m+1)}}^c) \tag{7.19}$$

Furthermore, using (4.43), (4.5) and (7.9), we have

$$\begin{aligned}
\|\mathcal{F}(x_m, y_m)\|_{Z_{\sigma_m}} &= \|\mathcal{F}(x_m, y_m) - \mathcal{F}(x_{m-1}, y_m)\|_{Z_{\sigma_m}} \\
&\leq M_3\|x_m - x_{m-1}\|_{X_{\sigma_m}} \\
&\leq M_3 k(q)(2^m T)^{-q}\|x\|_{X_0^q} \\
&\leq \underbrace{\Big(M_3 k(q)(\|\bar{x}\|_{X_0^q} + \mathrm{dist}(\bar{x}, (U_0^q)^c))\Big)}_{\leq M_5}(2^m T)^{-q} \\
&\leq \delta_T(2^m T)\left(\frac{d - d^*}{4}\right) \\
&\leq \delta_T(2^m T)\mathrm{dist}(y_m, V_{\sigma_{m+1}}^c)
\end{aligned} \tag{7.20}$$

Applying Theorem 6.1.1 to $(x_m, y_m) \in \mathcal{C}_{\sigma_m} \times V_{\sigma_m}$ we get $y_{m+1} \in V_{\sigma_{(m+1)}}$ with $\mathcal{F}(x_m, y_{m+1}) = 0$ and

$$\begin{aligned}
\|y_{m+1} - y_m\|_{Y_{\sigma_{(m+1)}}} &\leq M_N(2^{(m+1)}T)^\beta\|\mathcal{F}(x_m, y_m)\|_{Z_{\sigma_m}} \\
&\leq \eta(2^m T)^{-(\alpha+\beta)}
\end{aligned}$$

This completes the verification of the inductive hypothesis, so y_n is defined for all n.

Convergence of y_n: Note that

$$\begin{aligned}
\|y_n - y_{n-1}\|_{Y_0} &\leq \|y_n - y_{n-1}\|_{Y_{\sigma_n}} \\
&\leq N_T(2^n T)\|\mathcal{F}(x_{n-1}, y_{n-1})\|_{Z_{\sigma_{(n-1)}}} \\
&\leq N_T(2^n T)M_3\|x_{n-1} - x_{n-2}\|_{X_{\sigma_{(n-1)}}} \\
&\leq \Big(M_N C_R T^\beta M_3 k(q)\|x\|_{X_0^q}\Big)\, 2^{-(q-\beta)n}
\end{aligned} \tag{7.21}$$

from which one can conclude that $y_n \to y_\infty$ in Y_0. Note that from (7.18), we can conclude that $y_\infty \in V_0$ and using continuity of $\mathcal{F}$ one has that

$$\mathcal{F}(x, y_\infty) = \lim_{n\to\infty} \mathcal{F}(x_n, y_{n+1}) = 0$$

Finally, note that from (7.21) we in fact have $y_\infty \in Y_0^{q-(q*+\beta)}$.

Establishing (7.3)**:** Assume that $q \geq \max(\alpha+2\beta, \gamma)+\beta$. Note that if $\|\mathcal{F}(x, y)\|_{Z_0^q} = 0$, rather than constructing y_∞ as above we can simply choose $y_\infty = y$ and trivially satisfy (7.3). On the other hand, if $\|\mathcal{F}(x, y)\|_{Z_0^q} \neq 0$, we can (optimally) choose $T = T_\infty$. Note from (7.2) we obtain

$$T^{-q} \leq T^{-(q-\beta)} \leq T^{-\max(\alpha+2\beta,\gamma)} = \frac{\|\mathcal{F}(x, y)\|_{Z_0^q}}{\delta T_0^{\max(\alpha+2\beta,\gamma)}} \tag{7.22}$$

Note that using the first half of (7.20) in the first half of (7.17) and simplifying we get

$$\|y_n - y_{n-1}\|_{Y_{\sigma_n}} \leq T^{-(q-\beta)}(M_N C_R M_5 2^q)2^{-(q-\beta)n} \tag{7.23}$$

combining this with (7.22) and substituting into (7.23) we get

$$\begin{aligned}
\|y_\infty - y\|_{Y_0^{q-\beta}} &\leq \|y_0 - y\|_{Y_0^{q-\beta}} + \sum_{i=1}^{\infty} \|y_i - y_{i-1}\|_{Y_{\sigma_i}} \\
&\leq T^{-q}k(q)\|y\|_{Y_0^q} + \left(\sum_{i=1}^{\infty}(M_N C_R M_5 2^q)2^{-(q-\beta)i}\right) T^{-(q-\beta)} \\
&\leq N\|\mathcal{F}(x, y)\|_{Z_0^q}
\end{aligned}$$

with

$$N = \left(\frac{k(q)\|y\|_{Y_0^q} + \sum_{i=1}^{\infty}(M_N C_R M_5 2^q)2^{-(q-\beta)i}}{\delta T_0^{\max(\alpha+2\beta,\gamma)}}\right)$$

which establishes (7.3) and completes the proof of Theorem 7.1.1. □

7.2 Whitney regularity in smooth spaces

As in Section 6.2, we have two approaches to obtain the Whitney regularity of the implicit function.

Theorem 7.2.1. *Given $\mathcal{F}$ and X_σ, Y_σ, Z_σ satisfying (F0), (XYZ.S1) and (XYZ.S2), assume in addition that $\mathcal{F}$ also satisfies Hypotheses (F.S0)-(F.S4).*

If $\mathcal{F}$ also satisfies (F.SW1) and (F.SW2) then, for any $q > \max(\alpha + 2\beta, \gamma) + q*$ and $(\bar{x}, \bar{y}) \in \mathcal{C}_0^q \times V_0^q$ with $\mathcal{F}(\bar{x}, \bar{y}) = 0$, there exists

$$\bar{\epsilon} > 0 \quad \textit{and} \quad g \in C^\gamma_{Wht}(\mathcal{C}_0^q \cap B_0^q(\bar{x}, \bar{\epsilon}), Y_0^{q-\beta})$$

with $g : \mathcal{C}_0^q \cap B_0^q(\bar{x}, \bar{\epsilon}) \to V_0^{q-\beta}$ such that $\mathcal{F}(x, g(x)) = 0$.

Proof. As in Theorem 6.2.1, given any $(\bar{x}, \bar{y}) \in \mathcal{C}_0^q \times V_0^q$ satisfying $\mathcal{F}(\bar{x}, \bar{y}) = 0$, for fixed $\gamma > 1$, $k < \gamma \leq k + 1$, and $\bar{\epsilon} > 0$, we twist around the role played by $x \in X_0^q$. Letting

$$A_\sigma = \mathcal{C} \cap B_{X_0}(\bar{x}, \bar{\epsilon}) \cap X_\sigma$$

we define the one parameter families of Banach spaces

$$\mathbb{X}_\sigma = \{0\}, \quad \mathbb{Y}_\sigma = C^\gamma_{Wht}(A_\sigma, Y_{\sigma/2}), \quad \mathbb{Z}_\sigma = C^\gamma_{Wht}(A_\sigma, Z_{\sigma/2})$$

for $0 \leq \sigma \leq 1$. Using the analytic smoothing S_t in X_σ, Y_σ and Z_σ and the $\mathcal{C}_0$ invariance of smoothing in X_σ, with r and T_0 the corresponding constants for invariance around $\bar{x} \in \mathcal{C}_0$, provided $\bar{\epsilon} < r$ we can define analytic smoothing $\mathbb{Y}_\sigma$ and $\mathbb{Z}_\sigma$ via:

$$(\mathbb{S}_t[\bar{y}])(x) \equiv S_t[\bar{y}(S_{t+T_0}x)] \qquad \text{and} \qquad (\mathbb{S}_t[\bar{z}])(x) \equiv S_t[\bar{z}(S_{t+T_0}x)]$$

Note that since $\mathbb{X} = \{0\}$, it trivially has analytic smoothing. With this smoothing we can apply Theorem 7.1.1 and, as in Theorem 6.2.1, provided $\bar{\epsilon}$ is sufficiently small the result follows. □

As in Theorem 6.2.3, provided the zeros of $\mathcal{F}$ are isolated we have the following alternative approach is to establish the Whitney regularity of any function g which solves $\mathcal{F}(x, g(x)) = 0$ by explicitly verify the estimates for the Whitney Regularity of g:

Theorem 7.2.2. *Given $\mathcal{F}$ and X_σ, Y_σ, Z_σ satisfying (F0), (XYZ.S1) and (XYZ.S2), assume in addition that $\mathcal{F}$ also satisfies Hypotheses (F.S0)-(F.S4).*

Assume $\mathcal{F}$ has local uniqueness in y for solutions $\mathcal{F}(x, y) = 0$ (for example, if $\mathcal{F}$ satisfies (F.SU)). If $\mathcal{F}$ also satisfies (F.SW1) and (F.SW2) for some $\gamma \geq 1$, $k < \gamma \leq k + 1$, and either:

(a) $\mathcal{C}$ has the γ density property described in Definition 3.2.5, or

(b) R satisfies (F.SW4)

then given any function

$$g : \mathcal{C}_0^q \cap B_0^q(\bar{x}, \epsilon) \to V_0^{q-\beta}$$

with $\mathcal{F}(x, g(x)) = 0$ is C^γ_{Wht} with with $\mathcal{F}(x, g(x)) = 0$ is C^γ_{Wht} with the k-jet of g having the same form as the coefficients $g_i(x_0, g(x_0))$ of the polynomial approximate solutions defined in (5.2) *of Theorem 5.0.12.*

Proof. The bounds on Ω_F, Ω_R and Ω_A from (F.SW1) and (F.SW2) allow one to apply Theorem 5.0.12 in $X_0^q \times Y_0^q$. Follow the proof of Theorem 6.2.3 using $X_0^q \times Y_0^q$ in place of $X_\sigma \times Y_\sigma$ the result follows. $\square$

7.3 Uniqueness in smooth spaces

Now we consider the question of uniqueness for solutions in the smooth spaces.

Theorem 7.3.1. *Let $\mathcal{F}$ satisfying all the Hypothesis for Theorem 7.1.1 be given and assume that $\mathcal{F}$ satisfies Hypothesis (F2**) with $\Omega_L(s) \leq C_L s^{-\beta^*}$.*

For any $q > \max(\alpha + 2\beta, \gamma)$, $(\bar{x}, \bar{y}) \in \mathcal{C}_0^q \times V_0^q$ there exists positive constants r (as in Theorem 7.1.1) and ϵ such that given

$$(x, y_1), (x, y_2) \in (\mathcal{C}_0^q \cap B_0^q(\bar{x}, r)) \times B_0^q(\bar{y}, r)$$

with $\mathcal{F}(x, y_i) = 0$, if $\|y_1 - y_2\|_{Y_0^q} < \epsilon$ then $y_1 = y_2$.

To prove Theorem 7.3.1, given $(x, y) \in U_0^q \times V_0^q$ with $\mathcal{F}(x, y) = 0$, we analytically approximate (x, y) by a sequence $(x_n, y_n) \in U_{\sigma_n} \times V_{\sigma_n}$ with $\mathcal{F}(x_n, y_n) = 0$. This is done by utilizing the sequences x_n, y_n generated in Theorem 7.1.1. Note that with these sequences one has $\mathcal{F}(x_n, y_{n+1}) = 0$ with $x_n \to x$ and $y_n \to y_\infty$. Furthermore, using Lemma 7.3.2 the uniqueness of Theorem 6.3.1 establishes that $y_\infty = y$. Thus, by re-index the sequence y_n, we get $(x_n, y_n) \in U_{\sigma_n} \times V_{\sigma_n}$ with $x_n \to x$ in X_0^q, $y_n \to y$ in Y_0^q and $\mathcal{F}(x_n, y_n) = 0$.

It is worthwhile to note that if one were to simply apply analytic smoothing to both x and y, one can easily produce sequences $(x_n, y_n) \in$

$U_{\sigma_n} \times V_{\sigma_n}$ with $x_n \to x$ in X_0^q, $y_n \to y$ in Y_0^q. However, with this approach one no longer necessarily has that $\mathcal{F}(x_n, y_n) = 0$.

Lemma 7.3.2. *Let $\mathcal{F}$ satisfy all the Hypotheses for Theorem 7.1.1 as well as (F.SU).*

For any $q > \max(\alpha + 2\beta, \gamma) + q*$ and $(x, y) \in \mathcal{C}_0^q \times V_0^q$ with $\mathcal{F}(x, y) = 0$ there exists positive constant T^* such that for all $T \geq T^* \geq T_0$, the sequence $y_n \in V_{\sigma_n}$ generated in the inductive argument of the proof of Theorem 7.1.1 converges to y in Y_0, i.e. $y_n \to y_\infty(x, y, T) = y$ in Y_0.

Thus, re-indexing y_n, we can approximate $\mathcal{F}(x, y) = 0$ via $\mathcal{F}(x_n, y_n) = 0$ with $x_n \to x$ in X_0 and $y_n \to y$ in Y_0.

Proof. Given (x, y) with $\mathcal{F}(x, y) = 0$, let r, δ and $T_0 \geq 1$ be as in Theorem 7.1.1 where we take $\bar{x} = x$ and $\bar{y} = y$. Without loss of generality, assume that the constant M_δ in the definition of $\delta_T(t)$ given in (7.7) satisfies (6.44) in Remark 6.3.3.

Let x_n^T and y_n^T denote the sequences constructed in Theorem 7.1.1 starting at $x_0^T = S_T[x]$ and $y_0^T = S_T[y]$ and converging to x and $y_\infty = y_\infty(x, y, T)$ in X_0 and Y_0 respectively. Note that

$$x_n^T = S_{2^n T}[x]$$

and set

$$\bar{y}_n^T \equiv S_{2^n T}[y]$$

Define

$$\epsilon_T(t) \equiv \frac{M_\epsilon t^{-(\alpha+\beta^*)}}{C_Q C_L}$$

as in extension (B) of Theorem 6.3.1 and take $T^* \geq T_0$ such that for all $t \geq T^*$

$$M_N C_R t^\beta \delta_T(t) \left(\frac{d-d^*}{4}\right) \leq \epsilon_T(2t)/6 \tag{7.24}$$

and

$$k(q)t^{-q}\|y\|_{Y_0^q} \leq \epsilon_T(2t)/6 \tag{7.25}$$

Note that for $T \geq T^*$ by (7.25)

$$\left\|\bar{y}_{n+1}^T - \bar{y}_n^T\right\|_{Y_{(2^{n+1}T)^{-1}}} \leq \epsilon_T(2^{n+1}T)/6 \tag{7.26}$$

We will iteratively establish

$$\left\|\bar{y}_n^T - y_n^T\right\|_{Y_0} \leq \left\|\bar{y}_n^T - y_n^T\right\|_{Y_{(2^{n+1}T)^{-1}}} \leq \epsilon_T(2^{n+1}T)/3 \tag{7.27}$$

Note for $n = 0$ we have $y_0^T = \bar{y}_0^T$ and thus (7.27) is trivially true.

Inductively assume that (7.27) holds for all $m \leq n$. Using (4.44) and (7.9) note

$$\left\|\mathcal{F}(x_m^T, \bar{y}_m^T)\right\|_{Z_{(2^m T)^{-1}}} \leq \delta_T(2^m T)\left(\frac{d-d^*}{4}\right)$$

so applying Theorem 6.1.1 we obtain a solution which, in light of Corollary 6.3.2 and (7.27), is unique and hence equal to y_{m+1}. By (6.2) and (7.24) we get

$$\left\|y_{m+1}^T - \bar{y}_m^T\right\|_{Z_{(2^{m+1}T)^{-1}}} \leq \epsilon_T(2^{m+2}T)/6$$

and combining this with (7.26) we have

$$\begin{aligned}\left\|\bar{y}^T_{m+1} - y^T_{m+1}\right\|_{Y_{(2^{m+1}T)^{-1}}} &\leq \left\|\bar{y}^T_{m+1} - \bar{y}^T_{m}\right\|_{Y_{(2^{m+1}T)^{-1}}} + \left\|\bar{y}^T_{m} - y^T_{m+1}\right\|_{Y_{(2^{m+1}T)^{-1}}} \\ &\leq \epsilon_T(2^{m+2}T)/3\end{aligned}$$

which inductively establishes (7.27) for $m+1$ and hence (7.27) holds for all $n \geq 0$. Since $y^T_n \to y_\infty$ in Y_0 and $\bar{y}^T_n \to y$ in Y_0, using (7.27) we get $y_\infty = y$. □

Proof of Theorem 7.3.1

Let r, δ and $T_0 \geq 1$ be as in Theorem 7.1.1 and assume that δ is sufficiently small so that the constant M_δ in Extension (B) of Theorem 6.1.1 appearing in the induction argument in the proof of Theorem 7.1.1 satisfies (6.44) and thus Remark 6.3.3 applies. Let T^* to be the maximum T^* arising in Lemma 7.3.2 for y_1 and y_1 and $\bar{\epsilon}$ be the epsilon which appears in Corollary 6.3.2 for $\tau' = 1/2T^*$. Taking $\epsilon = \bar{\epsilon}/(3k(q))$ note

$$\|S_T y_1 - S_T y_2\|_{Y_{(T)^{-1}}} \leq k(q)\|y_1 - y_2\|_{Y^q_0} < \bar{\epsilon}/3$$

so with $y^1_1 = y_\infty(S_T x, S_T y_1)$ and $y^2_1 = y_\infty(S_T x, S_T y_2)$ denoting the solutions which arise in at the first iteration of the induction in the proof of Theorem 7.1.1 (that is y^1_1 and y^2_1 are the first solutions obtained in the proof of Theorem 7.1.1 by the application of Theorem 6.1.1) applying Corollary 6.3.2 we get $y^1_1 = y^2_1$. Thus, all the following terms in the induction of Theorem 7.1.1 for y_1 and y_2 are equal and since these converge in Y_0 to $y^1_\infty = y_\infty(x, y_1) = y_1$ and $y^2_\infty = y_\infty(x, y_2) = y_2$ we have $y_1 = y_2$ as desired. □

Chapter 8

Torus maps

In this chapter we develop several results about torus maps which will used in Chapter 9 to establish that the functional described in Example 4.3.1 satisfies the hypotheses given in Section 4.3. Thus, we will be able to apply the Nash-Moser implicit function theorems (Theorem 6.2.3 and Theorem 7.2.2) to prove a KAM theory for degenerate families of torus maps (see Theorem 9.0.4) which arise in the study of the wave equation in oscillating domains.

8.1 The basics

Note that the universal cover of $\mathbb{T}^n$ is $\mathbb{R}^n$ with the covering map

$$\pi : \mathbb{R}^n \to \mathbb{T}^n, \qquad \pi(x) = x \mod 1$$

Given any continuous torus map

$$F : \mathbb{T}^n \to \mathbb{T}^n \tag{8.1}$$

we can lift F to the universal cover and thus obtain the following commutative diagram

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{\tilde{F}} & \mathbb{R}^n \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ \mathbb{T}^n & \xrightarrow{F} & \mathbb{T}^n \end{array} \tag{8.2}$$

Proposition 8.1.1. *Given a continuous torus map F as in* (8.1)*, the corresponding lift $\tilde{F}$ as in* (8.2) *has the form*

$$\tilde{F}(x) = A[x] + f(x) \tag{8.3}$$

where $A \in M(n,\mathbb{Z})$ and $f \in \mathcal{P}^n$ (here $M(n,\mathbb{Z})$ is the set $n \times n$ integer valued matrices and $\mathcal{P}^n$ is the set of continuous periodic vector valued functions). Thus the "moduli space" of continuous torus maps has the form

$$M(n,\mathbb{Z}) \times \mathcal{P}^n \tag{8.4}$$

Given a continuous family of torus map F_t the corresponding lift $\tilde{F}_t$ has the form

$$\tilde{F}_t(x) = A[x] + f_t(x)$$

for some fixed $A \in M(n,\mathbb{Z})$ and a continuous family $f_t \in \mathcal{P}^n$.

Finally, if F or F_t have additional regularity the corresponding periodic functions f or f_t have the same regularity.

Proof. Straightforward. □

Remark 8.1.2. *While we will usually work with the lift $\tilde{F}$, for the sake of notation we use F to denote both the torus map F and its lift $\tilde{F}$. Furthermore, we will usually use capital letters, such as F and H, to denote the torus maps with the corresponding lower case letters representing the corresponding elements of $\mathcal{P}^n$, e.g. $F = Id + f$, $H = Id + h$.*

Remark 8.1.3. *We will work with families of torus maps, and hence families of periodic functions, possessing some degree or regularity beyond continuity. In section 8.2 we define several one parameter families of Banach spaces (as described in Section 4.1) whose elements are periodic functions with some prescribed regularity. In particular, we distinguish two cases:*

1. *F is smooth, in which case f is smooth, e.g. $[C^q]^n$ for $q \notin \mathbb{Z}$ or $[\hat{C}^q]^n$ for $q \in \mathbb{Z}$ (see section 8.2 for the definition of C^ℓ and $\hat{C}^p$)*

2. *F is analytic, in which case $f \in [A(r\sigma, C^0)]^n$ (see section 8.2 for the definition of $A(r\sigma, C^m)$).*

8.2 The function spaces C^ℓ, $\hat{C}^p$, $A(r\sigma, C^m)$

As in Section 2 of [Zeh75], we make the following:

Definition 8.2.1. *Let $p \geq 0$ an integer and $\alpha \in (0,1)$ and take $\ell = p + \alpha$. Define the Hölder (Banach) spaces $C^\ell(\mathbb{T}^n)$ (often shortened to C^ℓ) to be all functions $u : \mathbb{T}^n \to \mathbb{R}$ with continuous derivatives up to order p for which the norm*

$$\|u\|_{C^\ell} \equiv \sup_{\substack{x\in\mathbb{T}^n \\ |k|\leq p}} \left\{|D^k u(x)|\right\} + \sup_{\substack{x\neq y \\ |k|=p}} \left\{ \frac{|D^k u(x) - D^k u(y)|}{|x-y|^\alpha} \right\} \tag{8.5}$$

is finite. ■

Definition 8.2.2. *Let $p \geq 1$ an integer. Define the Zygmund (Banach) spaces $\hat{C}^p(\mathbb{T}^n)$ (also denoted by $\Lambda_p(\mathbb{T}^n)$ and often shortened to $\hat{C}^p$ or Λ_p) to be all* *functions $u : \mathbb{T}^n \to \mathbb{R}$ with continuous derivatives up to order $p-1$ for which the norm*

$$\|u\|_{\hat{C}^p} \equiv \sup_{\substack{x\in\mathbb{T}^n \\ |k|\le p-1}} \left\{|D^k u(x)|\right\} + \sup_{\substack{x\neq y \\ |k|=p-1}} \left\{\frac{|D^k u(x) + D^k u(y) - 2D^k u(\frac{x+y}{2})|}{|x-y|}\right\} \quad (8.6)$$

is finite. ■

Definition 8.2.3. *Fix $r > 0$ and for $0 < \sigma \le 1$ let $U_{r\sigma}$ denote the complex strips*

$$U_{r\sigma} = \{x + iy \in \mathbb{C}^n : |y_j| \le r\sigma, 1 \le j \le n\}$$

Let $\ell \in \mathbb{R}^+$. Define the Banach spaces $A(r\sigma, C^m)$ to be all holomorphic functions $u : U_{r\sigma} \to \mathbb{C}$ for which

- $\overline{u(x)} = u(\overline{x})$ *(i.e. u is real valued on $\mathbb{R}$)*
- *u is periodic with period 1 in each variable*
- $\|u\|_{\sigma,C^m} < \infty$

Here

$$\|u\|_{\sigma,C^m} = \begin{cases} \|u\|_{C^m(U_{r\sigma})} & \text{if } m \notin \mathbb{Z} \\ \|u\|_{\hat{C}^m(U_{r\sigma})} & \text{if } m \in \mathbb{Z} \end{cases}$$

where by replacing $\mathbb{T}^n$ with $U_{r\sigma}$ in the norms (8.5), (8.6) we mean that the supremums should be taken over the entire complex strip $U_{r\sigma}$. ■

As noted in Example 4.1.8, taking $X_\sigma = A(r\sigma, C^m)$ we get $X_0^q = C^{q+m}$ for $q+m \notin \mathbb{Z}$ and $X_0^q = \hat{C}^{q+m}$ for $q+m \notin \mathbb{Z}$. For a detailed proof of this see Proposition 2.1 in [Zeh75].

Throughout the remainder of this chapter, unless specifically defined otherwise, we take $X_\sigma \equiv A(\sigma, C^0(\mathbb{T}^n))$ with $X_0^q = C^q(\mathbb{T}^n)$ for $q \notin \mathbb{Z}$ and $X_0^q = \hat{C}^q(\mathbb{T}^n)$ for $q \in \mathbb{Z}$.

8.3 Rotations and other foliation preserving torus maps

Let $X_\sigma \equiv A(\sigma, C^0(\mathbb{T}^n))$ with $X_0^q = C^q(\mathbb{T}^n)$ for $q \notin \mathbb{Z}$ and $X_0^q = \hat{C}^q(\mathbb{T}^n)$ for $q \in \mathbb{Z}$.

Given any $A \in M(n, \mathbb{Z})$ which has ω as an eigen vector, note that any map of the form $F = A + \omega f$ with $f \in X_\sigma$ or $f \in X_0^q$ has the property that it preserves the foliations $\{t\omega + x_0 \mod 1 | t \in \mathbb{R}\}$. We refer to such maps as *ω-foliation preserving torus maps.* If the leaves of this foliations wind densely around $\mathbb{T}^n$, the preservation property in some sense forces any foliation preserving map to be "essentially" one dimensional.

An important and basic class of ω-foliation preserving torus maps, are the *rotations* $T_w : \mathbb{T}^n \to \mathbb{T}^n$ defined by

$$T_w(x) = (x + \omega) \mod 1 \tag{8.7}$$

where $\omega \in \mathbb{R}^n$. These maps are clearly invertible and analytic and their dynamics is easy to understand.

An important subset of rotations are those for which the so called "small divisor" problem can be solved. That is, given a rotation T_ω and $r \in X_\sigma$, find $f \in X_{\sigma'}$ such that $f - f \circ T_\omega = r$. For this problem to have a solution, we clearly need $\int_{\mathbb{T}^n} r \, dx = 0$. Taking the Fourier transform diagonalizes this problem and provided $\omega \cdot k \notin \mathbb{Z}$ we can formally determine f. In Section 8.5 we study this problem in more detail. In particular provided ω satisfies certain Diophantine conditions (see Definition 8.5.1) we can make this formal expression for f rigorous provided we loose some regularity in f, i.e. given ω Diophantine, if $r \in X_\sigma$ with $\int_{\mathbb{T}^n} r \; dx = 0$, we can define $f \in X_{\sigma'}$ for $0 \leq \sigma' < \sigma \leq 1$ with $f - f \circ T_\omega = r$.

Before studying the small divisor problem in detail, we define projection operators which give us a decomposition $X_\sigma = \breve{X}_\sigma \oplus \mathbb{R}$ where $r \in \breve{X}_\sigma$ have $\int_{\mathbb{T}^n} r \; dx = 0$.

8.4 Averaging and other projection operators

As we have seen in Section 8.1, the space of sufficiently smooth torus maps has the form $M(n, \mathbb{Z}) \times \mathcal{P}^n$ where we can take $\mathcal{P}$ to be the one parameter Banach spaces $\mathcal{P} = X_\sigma$ or $\mathcal{P} = X_0^q$. We now describe some subspaces of X_σ, X_0^q, $[X_\sigma]^n$ and $[X_0^q]^n$ which will come into play.

Definition 8.4.1. *On X_σ define the functional* $\text{avg} : X_\sigma \to \mathbb{R}$ *by*

$$\text{avg}[f] \equiv \int_{\mathbb{T}^n} f(x) \quad dx$$

Note $|\mathrm{avg}[f]| \le \|f\|_{X_\sigma}$. *Using* avg, *define the closed Banach subspaces*

$$\breve{X}_\sigma = \{f \in X_\sigma : \mathrm{avg}[f] = 0\}$$

Given $g \in X_\sigma \setminus \breve{X}_\sigma$, *define*

$$\Pi_{g\to 1} : X_\sigma \to \mathbb{R} \qquad \Pi_{g\to 1}[f] \equiv \mathrm{avg}[f]/\mathrm{avg}[g]$$
$$\Pi_{g\to 0} : X_\sigma \to \breve{X}_\sigma \qquad \Pi_{g\to 0}[f] \equiv f - (\mathrm{avg}[f]/\mathrm{avg}[g])g$$

Note

$$|\Pi_{g\to 1}[f]| \le \|f\|_{X_\sigma}/|\mathrm{avg}[g]|$$

and

$$\|\Pi_{g\to 1}\|_{X_\sigma} \le \left(1 + \|g\|_{X_\sigma}/|\mathrm{avg}[g]|\right) \|f\|_{X_\sigma}$$

hence we get a continuous splitting

$$X_\sigma = \breve{X}_\sigma \oplus \mathbb{R} \quad \textit{with} \quad Id = \Pi_{g\to 0} + \Pi_{g\to 1} \tag{8.8}$$

Analogous definitions in X_0^q *follow for* avg, $\breve{X}_0^q$, $\Pi_{g\to 1}$ *and* $\Pi_{g\to 0}$ *by replacing* X_σ *with* X_0^q *in the above.* ■

In a manner similar to X_σ and X_0^q, we define an averaging functional and corresponding projection operators in $[X_\sigma]^n$ and $[X_0^q]^n$:

Definition 8.4.2. *On* $[X_\sigma]^n$ *define the functional* $\mathrm{avg} : [X_\sigma]^n \to \mathbb{R}^n$ *by*

$$\mathrm{avg}[f] \equiv \int_{\mathbb{T}^n} f(x) \;\; dx$$

Note $|\text{avg}[f]| \le \|f\|_{[X_\sigma]^n}$. *Using* avg, *define the closed Banach subspaces*

$$[\breve{X}_\sigma]^n = \{f \in [X_\sigma]^n : \text{avg}[f] = 0\}$$

Given M *a matrix with* $m_{i,j} \in X_\sigma$ *such that the matrix* $\bar{M} \in M(n, \mathbb{R})$ *formed by averaging the coefficients, i.e.* $\bar{m}_{i,j} = \text{avg}[m_{i,j}]$, *is invertible, define*

$$\Pi_{M\to 1} : [X_\sigma]^n \to \mathbb{R}^n \qquad \Pi_{M\to 1}[f] \equiv \bar{M}^{-1}\text{avg}[f]$$

$$\Pi_{M\to 0} : [X_\sigma]^n \to [\breve{X}_\sigma]^n \qquad \Pi_{M\to 0}[f] \equiv f - M\bar{M}^{-1}\text{avg}[f]$$

Note

$$|\Pi_{M\to 1}[f]| \le \left\|\bar{M}^{-1}\right\|_{M(n,\mathbb{R})} \|f\|_{[X_\sigma]^n}$$

and

$$\|\Pi_{M\to 0}\|_{[X_\sigma]^n} \le \left(1 + \left\|\bar{M}^{-1}\right\|_{M(n,\mathbb{R})} \|M\|_{M(n,X_\sigma)}\right) \|f\|_{[X_\sigma]^n}$$

(By $\|M\|_{M(n,X_\sigma)}$ *we mean the norm of the matrix* $\tilde{M}$ *where* $\tilde{m}_{i,j} = \|m_{i,j}\|_{X_\sigma}$. *Note that here we uses the fact that* X_σ *is a Banach algebra.) Hence, we have a continuous splitting*

$$[X_\sigma]^n = [\breve{X}_\sigma]^n \oplus \mathbb{R}^n \quad \textit{with} \quad Id = \Pi_{M\to 0} + \Pi_{M\to 1} \tag{8.9}$$

Analogous definitions in $[X_0^q]^n$ *follow for* avg, $[\breve{X}_0^q]^n$, $\Pi_{M\to 1}$ *and* $\Pi_{M\to 0}$ *by replacing* $[X_\sigma]^n$ *with* $[X_0^q]^n$ *in the above.* ■

8.5 Small divisor problems

An important subset of rotations are those for which the following so called "small divisor" problem can be solved.

(SD) Given $\omega \in \mathbb{R}^n$, $r \in \breve{X}_\sigma$ find $f \in \breve{X}_{\sigma'}$ with $f - f \circ T_\omega = r$.

The subspaces $\breve{X}_\sigma \subseteq X_\sigma$ are described in Definition 8.4.1 of the previous section.

In order to solve this problem, we need ω to satisfy certain Diophantine conditions. We will also need to loose some regularity in f, that is given $r \in \breve{X}_\sigma$ we will be able to define $f \in \breve{X}_{\sigma'}$ for $0 \leq \sigma' < \sigma \leq 1$.

Definition 8.5.1. *Let Υ be a Rüssmann Modulus (see Definition 4.2.6). Define the set of* Υ-Diophantine vectors*, $\mathcal{D}_\Upsilon$, as*

$$\mathcal{D}_\Upsilon \equiv \left\{\omega \in \mathbb{R}^n/\mathbb{Z}^n : \forall k \in \mathbb{Z}^n \backslash \{0\}\ \forall m \in \mathbb{Z},\ |\omega \cdot k - m|^{-1} \leq \Upsilon(|k|)\right\} \quad (8.10)$$

Also define

$$\mathcal{R}_\Upsilon^{k,m} \equiv \left\{\omega \in \mathbb{R}^n/\mathbb{Z}^n : |\omega \cdot k - m| < \frac{1}{\Upsilon(|k|)}\right\} \quad (8.11)$$

and note

$$\mathcal{D}_\Upsilon = \mathbb{R}^n/\mathbb{Z}^n \setminus \left(\bigcup_{\substack{k \neq 0 \\ m \in \mathbb{Z}}} \mathcal{R}_\Upsilon^{k,m}\right)$$

■

Proposition 8.5.2. *Let $\Upsilon(r) \geq cr^\nu$ with $\nu > n$ (note that for $c > 0$, $\nu > 0$, the function $\Upsilon(r) = cr^\nu$ is a Brjuno modulus). There exists a positive constant M, depending only on n and ν, such that for any $v \in \mathbb{R}^n$ one has*

$$|B(v, \epsilon) \cap \mathcal{D}_\Upsilon| \geq |B(v, \epsilon)| - M\epsilon^n/c > 0$$

Proof. Fix k and m and note $\mathcal{R}_\Upsilon^{k,m}$ is a the region between two hyperplanes normal to the vector k. The thickness of this region is $\frac{2}{|k|\Upsilon(|k|)}$ and its intersection with $B(v,\epsilon)$ is no larger than $\frac{2\pi_{n-1}\epsilon^{n-1}}{|k|\Upsilon(|k|)}$ (here π_{n-1} is the area of the unit $n-1$ ball). Note that changing m simply translates these hyperplanes by $1/|k|$, hence $\mathcal{R}_\Upsilon^{k,m}$ will intersect $B(v,\epsilon)$ for at most $2\epsilon|k|$ values of m and so the measure of

$$B(v,\epsilon)\bigcap\left(\bigcup_{m\in\mathbb{Z}}\mathcal{R}_\Upsilon^{k,m}\right)$$

is at most $\frac{4\pi_{n-1}\epsilon^n}{\Upsilon(|k|)}$. Using $\Upsilon(r)\geq cr^\nu$ we see this is no larger than $\frac{4\pi_{n-1}\epsilon^n}{c|k|^\nu}$. Finally, letting k vary, we see that the measure of

$$B(v,\epsilon)\bigcap\left(\bigcup_{\substack{k\neq 0\\ m\in\mathbb{Z}}}\mathcal{R}_\Upsilon^{k,m}\right) \tag{8.12}$$

is at most $\sum\limits_{k\neq 0}\frac{4\pi_{n-1}\epsilon^n}{c|k|^\nu}\leq M\epsilon^n/c$. Since the compliment of (8.12) in $B(v,\epsilon)$ is $B(v,\epsilon)\cap\mathcal{D}_\Upsilon$, the result follows. □

Provided $\omega\in\mathcal{D}_\Upsilon$, the small divisor problem (SD) can be solved as described in the following:

Proposition 8.5.3. (Small Divisors) *Given $\omega\in\mathcal{D}_\Upsilon$ with Υ a Rüssmann Modulus (see Definition 4.2.6 and Definition 8.5.1) and $r\in\breve{X}_\sigma$, for any $0\leq\sigma'<\sigma\leq 1$ there exists unique $f\in\breve{X}_{\sigma'}$ such that*

$$\mathcal{S}_\omega[f]=f-f\circ T_\omega=r \tag{8.13}$$

Denoting f by $\mathcal{S}_\omega^{-1}[r]$, for $0\leq\sigma'<\sigma\leq 1$, we obtain a linear operator

$$\mathcal{S}_\omega^{-1}:\breve{X}_\sigma\to\breve{X}_{\sigma'}$$

with operator norm

$$\left\|\mathcal{S}_\omega^{-1}[r]\right\|_{X_{\sigma'}} \leq \Omega_\Upsilon(\sigma-\sigma')\|r\|_{X_\sigma} \tag{8.14}$$

Here Ω_Υ *is as defined in* (4.11) *of Example 4.2.8. Finally, viewing*

$$\mathcal{S}_\omega^{-1} : \mathcal{D}_\Upsilon \to L(\breve{X}_\sigma, \breve{X}_{\sigma'})$$

we have $\mathcal{S}_\omega^{-1} \in C^\infty_{Wht}(\mathcal{D}_\Upsilon, L(\breve{X}_\sigma, \breve{X}_{\sigma'}))$ *with*

$$\left\|\mathcal{S}_\omega^{-1}\right\|_{C^\gamma_{Wht}} \leq [\Omega_\Upsilon(\sigma-\sigma')]^\gamma \tag{8.15}$$

Proof. Expanding r and the unknown f in a Fourier series, i.e.

$$r(x) = \sum_{k\neq 0} \hat{r}_k e^{2\pi i k\cdot x} \qquad \text{and} \qquad f(x) = \sum_{k\neq 0} \hat{f}_k e^{2\pi i k\cdot x}$$

note that (8.13) becomes

$$\hat{f}_k(1-e^{2\pi i k\cdot\omega}) = \hat{r}_k \tag{8.16}$$

Formally, dividing (8.16) by $(1-e^{2\pi i k\cdot\omega})$ gives the Fourier coefficients of w. To establish convergence, note that for all $k \in \mathbb{Z}^n$ with $k \neq 0$

$$|1-e^{2\pi i k\cdot\omega}| \geq |\sin 2\pi k\cdot\omega| \geq 2/\pi \min\{|k\cdot\omega - m| : m \in \mathbb{Z}\}$$

and since ω is irrational we can invert this equation to get

$$|1-e^{2\pi i k\cdot\omega}|^{-1} \leq \frac{\pi}{2} \max\{|k\cdot\omega - m|^{-1} : m \in \mathbb{Z}\} \tag{8.17}$$

Since $\omega \in \mathcal{D}_\Upsilon$ we have $|k\cdot\omega - m|^{-1} \leq \Upsilon(|k|)$ for all $m \in \mathbb{Z}$ and $k \in \mathbb{Z}^n$, $k \neq 0$. Combining with (8.17) (and absorbing the $\pi/2$ into Υ), for all $k \in \mathbb{Z}^n$, $k \neq 0$, we get

$$|1-e^{2\pi i k\cdot\omega}|^{-1} \leq \Upsilon(|k|) \tag{8.18}$$

Cauchy estimate for the Fourier coefficients of r give us

$$|\hat{r}_k| \leq e^{-2\pi|k|\sigma}\|r\|_{X_\sigma} \tag{8.19}$$

and combining with (8.18) we get

$$|\hat{f}_k| = \left|\frac{\hat{r}_k}{(1-e^{2\pi ik\cdot\omega})}\right| \leq \Upsilon(|k|)e^{-2\pi|k|\sigma}\|r\|_{X_\sigma} \tag{8.20}$$

Restricting $|\mathrm{Im}\,(z)\,| \leq \sigma'$ the Fourier series for $f(z)$ converges and using (8.20) we have

$$|f(z)| \leq \underbrace{\left(\sum_{k\neq 0}\Upsilon(|k|)e^{-2\pi|k|r(\sigma-\sigma')}\right)}_{=\Omega_\Upsilon(\sigma-\sigma')}\|r\|_{X_\sigma}$$

This establishes (8.14).

To establish $\mathcal{S}_\omega^{-1} \in C^\infty_{Wht}(\mathcal{D}_\Upsilon, L(\breve{X}_\sigma, \breve{X}_{\sigma'}))$ with (8.15), apply Proposition 8.5.5 with $s_k(\omega) = \frac{1}{1-e^{2\pi ik\cdot\omega}}$ and $A = \mathcal{D}_\Upsilon$. $\square$

Remark 8.5.4. *Note that given any k for which estimate* (8.18) *is sharp, those k' near k will satisfy much better estimates. Using this observation one can obtain estimates which are sharper than* (8.14) *(see [Rüs75], [Rüs76b]).*

Proposition 8.5.5. *Let γ and A be given with $k < \gamma \leq k+1$ and A an arbitrary subset of a Banach space. For $k \in \mathbb{Z}^n$ let $s_k \in C^\gamma_{Wht}(A, \mathbb{R})$ with and $\|s_k\|_{C^\gamma_{Wht}} \leq \Upsilon(|k|)$ for Υ a Rüssmann Modulus (see Definition 4.2.6). Given $f \in X_\sigma$ with $f(x) = \sum_{k\in\mathbb{Z}^n} f_k e^{2\pi ik\cdot x}$ define*

$$S(\omega)[f](x) = \sum_{k\in\mathbb{Z}^n} s_k(\omega) f_k e^{2\pi ik\cdot x}$$

Then $S \in C^{\gamma}_{Wht}(A, L(X_\sigma, X_{\sigma'}))$ *with* $\|S\|_{C^{\gamma}_{Wht}} \le \Omega_\Upsilon(\sigma-\sigma')$*. Denoting the* k*-jet of* s_k *with* $\{s_{k,j}\}_{j=0}^k$ *the* k*-jet of* S *has the form* $\{S_j\}_{j=0}^k$ *where*

$$S_j(\omega)[f](x) = \sum_{k\in\mathbb{Z}^n} s_{k,j}(\omega) f_k e^{2\pi i k\cdot x}$$

Proof. Straightforward. □

Combining Proposition 8.5.3 with Definition 8.4.1 we get the following:

Definition 8.5.6. *Given* $g \in X_\sigma \setminus \check{X}_\sigma$ *define the operator*

$$\mathcal{S}_{\omega,g} : \check{X}_\sigma \times \mathbb{R} \to X_\sigma \qquad \mathcal{S}_{\omega,g}[f,c] \equiv f - f\circ T_\omega + cg \tag{8.21}$$

Using (8.8) *we can have the following (unbounded) inverse*

$$\mathcal{S}^{-1}_{\omega,g} : X_\sigma \to \check{X}_{\sigma'} \times \mathbb{R}$$

defined by

$$\mathcal{S}^{-1}_{\omega,g}[r] \equiv \left(\mathcal{S}^{-1}_{\omega}[\Pi_{g\to 0}[r]], \Pi_{g\to 1}[r]\right) \tag{8.22}$$

For $0 \le \sigma' < \sigma \le 1$ *we have*

$$\left\|\mathcal{S}^{-1}_{\omega,g}[r]\right\|_{\check{X}_{\sigma'}\times\mathbb{R}} \le C\Omega_\Upsilon(\sigma-\sigma')\|r\|_{X_\sigma} \tag{8.23}$$

with $C = \left(1 + \frac{\max(\|g\|_{X_\sigma},1)}{|\mathrm{avg} g|}\right)$. ■

Similarly, combining Proposition 8.5.3 with Definition 8.4.2 we get the following:

Definition 8.5.7. *Given M a matrix with $m_{i,j} \in X_\sigma$ such that the matrix $\bar{M} \in M(n, \mathbb{R})$ formed by averaging the coefficients, i.e. $\bar{m}_{i,j} = \operatorname{avg}[m_{i,j}]$, is invertible, define the operator*

$$\mathcal{S}_{\omega,M} : [\breve{X}_\sigma]^n \times \mathbb{R}^n \to [X_\sigma]^n \qquad \mathcal{S}_{\omega,M}[f, v] \equiv f - f \circ T_\omega + Mv \tag{8.24}$$

Using (8.9) *we have the following (unbounded) inverse*

$$\mathcal{S}_{\omega,M}^{-1} : [X_\sigma]^n \to [\breve{X}_{\sigma'}]^n \times \mathbb{R}^n$$

defined by

$$\mathcal{S}_{\omega,M}^{-1}[r] \equiv \left(\mathcal{S}_\omega^{-1}[\Pi_{M\to 0}[r]], \Pi_{M\to 1}[r]\right) \tag{8.25}$$

Here $\mathcal{S}_\omega^{-1} : [\breve{X}_\sigma]^n \to [\breve{X}_{\sigma'}]^n$ is simply applying $\mathcal{S}_\omega^{-1}$ from Lemma 8.5.3 on each component. Note, for $0 \le \sigma' < \sigma \le 1$ we have the estimate

$$\left\|\mathcal{S}_{\omega,M}^{-1}[r]\right\|_{[\breve{X}_{\sigma'}]^n \times \mathbb{R}^n} \le C\Omega_\Upsilon(\sigma - \sigma')\|r\|_{[X_\sigma]^n} \tag{8.26}$$

with $C = \left(1 + \max(\|M\|_{M(n,X_\sigma)}, 1)\left\|\bar{M}^{-1}\right\|_{M(n,\mathbb{R})}\right)$. ■

We will use $\mathcal{S}_{\omega,M}^{-1}$ in Chapter 9.

8.6 Analytic smoothing

As noted in Example 4.1.12, there exists analytic smoothing S_t in $X_\sigma = A(r\sigma, C^m(\mathbb{T}^n))$ with respect to the X_0^q where $X_0^q = C^{q+m}(\mathbb{T}^n)$ for $q + m \notin \mathbb{Z}$ and $X_0^q = \hat{C}^{q+m}(\mathbb{T}^n)$ for $q + m \notin \mathbb{Z}$. See Lemma 2.1 of [Zeh75] for a proof.

Throughout the remainder of this section, let $X_\sigma \equiv A(\sigma, C^0(\mathbb{T}^n))$ with $X_0^q = C^q(\mathbb{T}^n)$ for $q \notin \mathbb{Z}$ and $X_0^q = \hat{C}^q(\mathbb{T}^n)$ for $q \in \mathbb{Z}$. In order to establish the smoothing estimates for the composition of torus diffeomorphisms in the next section (Section 8.7), we explicitly construct the smoothing S_t in X_σ with respect to X_0^q as described in Section 2 of [Zeh75].

Choose $\tilde{\rho} : \mathbb{R} \to [0,1]$, C^∞, even, $\tilde{\rho} \equiv 1$ on $[-1/(2\pi), 1/(2\pi)]$, non-increasing on $[0,1]$, with support in $[-(1+\epsilon)/(2\pi), (1+\epsilon)/(2\pi)]$. Let $\rho : \mathbb{R} \to \mathbb{R}$ be the Fourier transform of $\tilde{\rho}$ and note that using the definition of the Fourier transform, ρ has an analytic continuation to an entire holomorphic function on $\mathbb{C}$. Define the functions $\tilde{s} : \mathbb{R}^n \to \mathbb{R}$ and $s : \mathbb{C}^n \to \mathbb{C}$ by

$$\tilde{s}(x_1, \ldots, x_n) = \tilde{\rho}(x_1) \cdots \tilde{\rho}(x_n)$$

and

$$s(z_1, \ldots, z_n) = \rho(z_1) \cdots \rho(z_n)$$

Note that s is the Fourier transform of $\tilde{s}$ and like ρ can be extended to an entire holomorphic function. With the scaling $s_t(z) = t^n s(tz)$, we define the analytic smoothing $S_t \in L(X_0, X_1)$ by

$$S_t[f] = s_t * f$$

We can also write this as

$$S_t[f](z) = t^n \int_{\mathbb{R}} s(t(y-z)) f(y) dy \tag{8.27}$$

or, using the change of variables $\xi = t\mathrm{Re}\,(y-z) = ty - t\mathrm{Re}\,(z)$,

$$S_t[f](z) = \int_{\mathbb{R}} s(\xi - it\mathrm{Im}\,(z)) f(\mathrm{Re}\,(z) + \xi/t) d\xi \tag{8.28}$$

Another useful expression of $S_t[f]$ is

$$S_t[f](z) = \sum_{k \in \mathbb{Z}^n} \tilde{s}(k/t) f_k e^{2\pi i k \cdot z} \tag{8.29}$$

Here f_k are the Fourier coefficients of f, i.e. $f(x) = \sum_{k \in \mathbb{Z}^n} f_k e^{2\pi i k \cdot x}$. From (8.27) or (8.29) it is clear that $S_t[f]$ is an entire function, while from (8.28) or (8.29) it is clear that S_t maps periodic functions to periodic functions.

The proof of estimates (4.3), (4.4) and (4.5), which establish that S_t is indeed an analytic smoothing in X_σ with respect to X_0^q, can be found in Lemma 2.1 of [Zeh75]. In addition to these estimates, we have the following:

Lemma 8.6.1. *Let $S_t : X_0 \to X_1$ be analytic smoothing as defined above.*

- *Given constants $r, C \geq 1$ there exists a constant M such that for all $g \in X_1$, and $t \geq 1$ with $t^{-1}(C + r\log(t)) \leq 1$*

$$\|(1 - S_t)g\|_{X_{Ct^{-1}}} \leq M t^{-r+n} \|g\|_{X_{t^{-1}(C + r\log(t))}} \tag{8.30}$$

- *There exists a constant M such that given $f \in X_0^q$ with $q = p + \alpha$, p a positive integer and $\alpha \in (0,1)$, and $r(1+\epsilon) + n \leq q$, for all $t > 1$*

$$\|S_t f\|_{X_{t^{-1}(C + r\log(t))}} \leq M \|f\|_{X_0^q} \tag{8.31}$$

Corollary 8.6.2. *Given $q \geq q* = (2+\epsilon)n$ there exists positive constants t_0 such that for any $f \in X_0^q$ and $t > t_0$*

$$\|(1 - S_t)S_t[f]\|_{X_{Ct^{-1}}} \leq M t^{-q+q*} \|f\|_{X_0^q} \tag{8.32}$$

Proof. Take $r = (q - n)/(1 + \epsilon)$ and apply (8.30) and (8.31). ∎

Proof of Lemma 8.6.1

First we consider (8.30). To simplify notation, let M be a "generic" constant that does not depend on t or g but can depend on n, r and C. By "generic" we mean that if M is multiplied by another constant we continue to denote the product by M. With $g \in X_1$, by shifting the contour of integration we can estimate the Fourier coefficients of g as follows

$$|g_k| \leq \|g\|_{X_{t^{-1}(C+r\log(t))}} e^{-2\pi|k|t^{-1}(C+r\log(t))}$$

Using (8.29) for $|\mathrm{Im}\,(z)\,| \leq Ct^{-1}$ we have

$$\begin{aligned}
|(1-S_t)[g](z)| &\leq \left(\sum_{t/(2\pi)\leq|k|} e^{-2\pi|k|t^{-1}(C+r\log(t))} e^{2\pi|k|Ct^{-1}} \right) \|g\|_{X_{t^{-1}(C+r\log(t))}} \\
&\leq M \left(\int_{t/(2\pi)}^{\infty} s^{n-1} e^{-(2\pi r\log(t)/t)s}\ ds \right) \|g\|_{X_{t^{-1}(C+r\log(t))}} \\
&\leq Mt^{-r+n}\mathrm{Polynomial}\left(\frac{1}{\log(t)}\right) \|g\|_{X_{t^{-1}(C+r\log(t))}} \\
&\leq Mt^{-r+n}\|g\|_{X_{t^{-1}(C+r\log(t))}}
\end{aligned}$$

This establishes (8.30).

Next, we consider (8.31). Again, to simplify notation, let M be a "generic" constant that does not depend on t or f but can depend on n, q, r and C. For $f \in X_0^q$, with $q = p + \alpha$, p a positive integer and $\alpha \in (0,1)$, using integration by parts we can estimate the Fourier coefficients as follows

$$|f_k| \leq \frac{M}{(1+|k|)^q}\|f\|_{X_0^q}$$

Using (8.29) for $|\mathrm{Im}\,(z)\,| \leq t^{-1}(C + r\log(t))$ and provided $r(1+\epsilon)+n \leq q$ we have

$$
\begin{aligned}
|S_t[g](z)| &\leq M\left(\sum_{|k|\leq t(1+\epsilon)/(2\pi)} \frac{e^{2\pi|k|t^{-1}(C+r\log(t))}}{(1+|k|)^q}\right)\|f\|_{X_0^q} \\
&\leq M\left(\int_0^{t(1+\epsilon)/(2\pi)} s^{n-1}\frac{e^{2\pi st^{-1}(C+r\log(t))}}{(1+s)^q}\ ds\right)\|f\|_{X_0^q} \\
&\leq M\left(M + \int_1^{t(1+\epsilon)/(2\pi)} s^{n-q-1}e^{2\pi t^{-1}r\log(t)s}\ ds\right)\|f\|_{X_0^q} \\
&\leq M\left(M + \left(\frac{t}{(2\pi r\log(t))}\right)^{n-q}\int_1^{(1+\epsilon)r\log(t)} u^{n-q-1}e^{u}\ du\right)\|f\|_{X_0^q} \\
&\leq M\left(M + t^{n+r(1+\epsilon)-q}\right)\|f\|_{X_0^q} \\
&\leq M\|f\|_{X_0^q}
\end{aligned}
$$

Which establishes (8.31) and completes the lemma. □

8.7 Composition of torus diffeomorphisms

Let X_σ and X_0^q be as in the previous section. Let $A \in M(n,\mathbb{Z})$ be given. Note for any $f, g \in [X_0^q]^n$, the composition $f\circ(A+g)(x) = f(Ax+g(x))$ makes sense and is in fact an element of $[X_0^q]^n$. For $f, g \in [X_\sigma]^n$, one needs to ensure the range of $A+g$ remains in the domain of analyticity for f. Note that

$$
\begin{aligned}
&\{|\mathrm{Im}\,(Az+g(z))\,| : |\mathrm{Im}\,(z)\,| \leq \sigma\} \\
&\qquad \leq \sup\{|\frac{\partial}{\partial y}\mathrm{Im}\,(A[x+iy]+g(x+iy))\,| : |y| \leq \sigma\}\sigma \\
&\qquad\qquad \leq (\|A\|_{M(n,\mathbb{Z})} + \|Dg\|_{X_\sigma})\sigma
\end{aligned}
$$

Thus, for any $C \geq 1$, given $f \in [X_{\sigma'}]^n$ and taking $\sigma C \leq \sigma'$, for any $g \in X_\sigma$ with $(\|A\|_{M(n,\mathbb{Z})} + \|Dg\|_{X_\sigma}) \leq C$ we have $f \circ (A+g) \in X_\sigma$.

Proposition 8.7.1. *We have the following basic estimates:*

1. *For any $C \geq 1$, if $f \in [X_{\sigma'}]^n$ and $\sigma C \leq \sigma'$, then, for any $g \in X_\sigma$ with $(\|A\|_{M(n,\mathbb{Z})} + \|Dg\|_{X_\sigma}) \leq C$,*

$$\|f \circ (A+g)\|_{[X_\sigma]^n} \leq \|f\|_{[X_{\sigma'}]^n} \tag{8.33}$$

2. *For any $C \geq 1$, if $f \in [X_{\sigma'}]^n$ and $\sigma C \leq \sigma'$, then, for any $g_1, g_2 \in X_\sigma$ with $(\|A\|_{M(n,\mathbb{Z})} + \|Dg_i\|_{X_\sigma}) \leq C$,*

$$\|f \circ (A+g_1) - f \circ (A+g_2)\|_{[X_\sigma]^n} \leq \|Df\|_{[X_{\sigma'}]^n} \|g_1 - g_2\|_{[X_\sigma]^n} \tag{8.34}$$

3. *If $f \in [X_0^q]^n$ and $g \in [X_0^q]^n$ with $q \geq 1$ then*

$$\|f \circ (A+g)\|_{[X_0^q]^n} \leq C\|f\|_{[X_0^q]^n} (1 + (\|A\|_{M(n,\mathbb{Z})} + \|g\|_{[X_0^q]^n})^q) \tag{8.35}$$

4. *If $f \in [X_0^{q'}]^n$ and $g_1, g_2 \in [X_0^q]^n$ where $q' > q \geq 1$ then there exists positive numbers M, δ and ρ so that, for $\|g_1 - g_2\|_{[X_0^q]^n} < \delta$,*

$$\|f \circ (A+g_1) - f \circ (A+g_2)\|_{[X_0^q]^n} \leq M\|f\|_{[X_0^{q'}]^n} \|g_1 - g_2\|^\rho_{[X_0^q]^n} \tag{8.36}$$

Proof. Statements (1) and (2) are immediate. See Theorem 4.3 in [dlLO99] for (3) and Theorem 6.2 in [dlLO99] for (4). □

The following lemma establishes that composition satisfies Hypothesis (F.S4) defined in Section 4.3.3.

Lemma 8.7.2. *Given $A \in M(n,\mathbb{Z})$ and $C > 1$, for $q > 1$ consider the open sets*

$$U_0^q = \{f \in [X_0^q]^n : \|Df\|_{[X_0]} < C\}$$

and

$$V_0^q = \{g \in [X_0^q]^n : \|Dg\|_{[X_0]} < C^*\}$$

where $C^* = (C - \|A\|_{M(n,\mathbb{Z})})/\max(k(0), M)$ with $k(0)$ the constant in (4.4) and M the constant in (8.31).

With smoothing as defined in Section 8.6, for any $f \in U_0^q$ and $g \in V_0^q$ one has

$$\|(S_t[f])\circ(A + S_t[g]) - S_t[f\circ(A+g)]\|_{X_{t^{-1}}} \leq M_4(q)t^{-q+q*} \tag{8.37}$$

Proof. The proof follows by combining the estimates from Lemma 8.6.1 and Proposition 8.7.1. In order to apply these estimates, we break

$$E \equiv (S_t[f])\circ(A + S_t[g]) - S_t[f\circ(A+g)]$$

into several terms as follows. First, write

$$f = \underbrace{S_t[f]}_{f_a} + \underbrace{(1 - S_t)[f]}_{f_s}$$

and note

$$\begin{aligned} E \equiv & \underbrace{(S_t[f_a])\circ(A + S_t[g]) - S_t[f_a\circ(A+g)]}_{E_a} \\ & + \underbrace{(S_t[f_s])\circ(A + S_t[g]) - S_t[f_s\circ(A+g)]}_{E_s} \end{aligned}$$

We further break down E_a as follows

$$E_a \equiv \underbrace{(S_t[f_a])\circ(A+S_t[g]) - f_a\circ(A+S_t[g])}_{E_{a1}} + \underbrace{f_a\circ(A+S_t[g]) - S_t[f_a\circ(A+S_t[g])]}_{E_{a2}} + \underbrace{S_t[f_a\circ(A+S_t[g])] - S_t[f_a\circ(A+g)]}_{E_{a3}}$$

Note that $\|A\|_{M(n,\mathbb{Z})} + \|S_t[Dg]\|_{[X_{t-1}]^n} < C$ and hence

$$\begin{aligned} \|E_{a1}\|_{[X_{t-1}]^n} &= \|[(1-S_t)f_a]\circ(A+S_t[g])\|_{[X_{t-1}]^n} \\ &\le \|(1-S_t)f_a\|_{[X_{Ct-1}]^n} \\ &\le M(q)t^{-q+q*}\|f\|_{[X_0^q]^n} \end{aligned} \tag{8.38}$$

Also, with r as in Corollary 8.6.2, note $\|A\|_{M(n,\mathbb{Z})} + \|S_t[Dg]\|_{[X_{t^{-1}(1+r\log(t))}]^n} < C$ and hence

$$\begin{aligned} \|E_{a2}\|_{[X_{t-1}]^n} &= \|(1-S_t)[f_a\circ(A+S_t[g])]\|_{[X_{t-1}]^n} \\ &\le Mt^{-r+n}\|f_a\circ(A+S_t[g])\|_{[X_{t^{-1}(1+r\log(t))}]^n} \\ &\le Mt^{-r+n}\|f_a\|_{[X_{Ct^{-1}(1+r\log(t))}]^n} \\ &\le Mt^{-q+q*}(\|f\|_{[X_0^q]^n}) \end{aligned} \tag{8.39}$$

Next, for $q' < q$, we have

$$\begin{aligned} \|E_{a3}\|_{[X_{t-1}]^n} &= \|S_t[f_a\circ(A+S_t[g]) - f_a\circ(A+g)]\|_{[X_{t-1}]^n} \\ &\le M\|f_a\circ(A+S_t[g]) - f_a\circ(A+g)\|_{[X_0^0]^n} \\ &\le M\|f_a\|_{[X_0^{q'}]^n}\|S_t[g]-g\|^\rho_{[X_0^0]^n} \\ &\le Mt^{-q\rho}\|f_a\|_{[X_0^{q'}]^n}\|g\|^\rho_{X_0^q} \end{aligned} \tag{8.40}$$

Finally, note $\|A\|_{M(n,\mathbb{Z})} + \|S_t[Dg]\|_{[X_{t-1}]^n} < C$ and hence

$$\begin{aligned} \|E_s\|_{[X_{t-1}]^n} &\leq \|S_t[f_s]\circ(A+S_t[g])\|_{[X_{t-1}]^n} + \|S_t[f_s\circ(A+g)]\|_{[X_{t-1}]^n} \\ &\leq \|S_t[f_s]\|_{[X_{Ct-1}]^n} + M\|f_s\circ(A+g)\|_{[X_0^0]^n} \\ &\leq M\|f_s\|_{[X_0^0]^n} + M\|f_s\|_{[X_0^0]^n}(1+\|A\|_{M(n,\mathbb{Z})} + \|g\|_{[X_0^0]^n}) \\ &\leq Mt^{-q}\|[f]\|_{[X_0^q]^n}(2+\|A\|_{M(n,\mathbb{Z})} + \|g\|_{[X_0^0]^n}) \end{aligned} \tag{8.41}$$

Combining (8.38), (8.39), (8.40) and (8.41) we get (8.37). □

Theorem 8.7.3. *Given $A \in M(n,\mathbb{Z})$ and $C > 1$, define $Y_\sigma = X_{C^*\sigma}$ where*

$$C^* = (C - \|A\|_{M(n,\mathbb{Z})})/\max(k(0), M)$$

with $k(0)$ the constant in (4.4) and M the constant in (8.31). Also, we define $U_0 = \{f \in [X_0^q]^n : \|Df\|_{[X_0]} < C\}$, $V_0 = \{g \in [X_0^q]^n : \|Dg\|_{[X_0]} < C^*\}$ and the functional

$$\mathcal{F} : U_0 \times V_0 \to Z_0 \quad \mathcal{F}(f,g) \equiv f\circ(A+g) \tag{8.42}$$

For $0 < \sigma \leq 1$ take $U_\sigma = U_0 \cap X_\sigma$ and $V_\sigma = V_0 \cap Y_\sigma$ (note U_σ and V_σ are open). For $0 \leq \sigma \leq 1$, we have

$$\mathcal{F} : U_\sigma \times V_\sigma \to Z_\sigma$$

Similarly, for $0 \leq q < \infty$ take $U_0^q = U_0 \cap X_0^q$ and $V_0^q = V_0 \cap X_0^q$ (note that for $q > 1$ the sets U_0^q and V_0^q are open). For $0 \leq q < \infty$, we have

$$\mathcal{F} : U_0^q \times V_0^q \to Z_0^q$$

Finally, we have:

(C0) The functional $\mathcal{F}$ defined in (8.42) *satisfies Hypotheses (F0), (F.A0) and (F.S0) described in Section 4.3.*

(C1) The functional $\mathcal{F}$ defined in (8.42) *satisfies Hypotheses (F.P1), (F.A1), (F.W1), (F.S1) and (F.SW1) described in Section 4.3.*

(C3) The functional $\mathcal{F}$ defined in (8.42) *satisfies Hypothesis (F.S3) described in Section 4.3.*

(C4) With smoothing as defined in Section 8.6, the functional $\mathcal{F}$ defined in (8.42) *satisfies Hypothesis (F.S4) described in Section 4.3.*

Chapter 9

An Application

The main application of Theorem 6.2.3 and Theorem 7.2.2 we now present is to prove the following KAM theorem for the torus maps discuss in Example 4.3.1.

Theorem 9.0.4. *Fix $\omega_0 \in \mathcal{D}_\Upsilon$ with $\Upsilon(s) \geq cs^\nu$ for $\nu > n$ (see equation 8.10). Let $F_\mu = Id + f_\mu$ with $\mu \in \Omega \equiv B(0, r_0) \subseteq \mathbb{R}^d$ and $f_\mu \in [X_\sigma]^n$.*

Assume:

(i) *the map $\mu \to f_\mu$ is C^γ for $\gamma > 1$ with $k < \gamma \leq k+1$*

(ii) $f_0 = \omega_0$

(iii) $\operatorname{avg}[f_\mu] = \omega_0 + A\mu^{\otimes m} + O(\mu^\eta)$ *with $1 \leq m \leq k$, $\eta > m$ and $A \in \operatorname{Sym}_m(\mathbb{R}^d, \mathbb{R}^n)$, $A \neq 0$.*

Then, there exists a Cantor set $\mathcal{C}_F \subseteq B(0, r_*) \subseteq B(0, r_0)$ such that

(a) *For each $\mu \in \mathcal{C}_F$, there exists $h_\mu \in [X_{\sigma/2}]^n$ and $a_\mu \in \mathcal{D}_\Upsilon$ such that with $H_\mu = Id + h_\mu$ we have*

$$F_\mu \circ H_\mu = H_\mu \circ T_{a_\mu} \tag{9.1}$$

Furthermore, $h_\mu : \mathcal{C}_F \to X_{\sigma/2}$ and $a_\mu : \mathcal{C}_F \to \mathbb{R}$ are C^γ_{Wht}.

(b) Provided $c > 0$ is sufficiently large then there exists positive constants M and r_* such that, for all r with $0 < r \le r_*$,

$$|B_{\mathbb{R}^d}(0,r) \cap \mathcal{C}_F| \ge M r^{d/m} > 0 \tag{9.2}$$

Finally, if $\Upsilon(s) = cs^\nu$ the above holds then we can take $f_\mu \in [X_0^q]^n$ for q sufficiently large and obtain $h_\mu \in [X_0^{q'}]^n$ for some $q' < q$.

Informally, Theorem 9.0.4 states that if $f_0 \in \mathcal{D}_\Upsilon$ and $\mathrm{avg}[f_\mu]$ is not very degenerate then there is a cantor set $\mathcal{C}_F$ of large density such that for $\mu \in \mathcal{C}_F$ there exists a change of variables $H_\mu = Id + h_\mu$ which takes $F_\mu = Id + f_\mu$ to the rotation $T_{a_\mu} = Id + a_\mu$. The proof of Theorem 9.0.4 will be done by rephrasing equation (9.1) as a zero of some functional and apply Theorem 6.2.3 (or, if $f_\mu \in [X_0^q]^n$ and $\Upsilon(s) = cs^\nu$, Theorem 7.2.2). Section 9.1 contains a detailed sketch of the steps involved in obtaining Theorem 9.0.4.

Note that Theorem 9.0.4 applies in the case we take

$$f_\mu \in w_0 X_\sigma \subseteq [X_\sigma]^n \quad \text{or} \quad f_\mu \in w_0 X_0^q \subseteq [X_0^q]^n \tag{9.3}$$

Note that for such f_μ, the corresponding torus maps $F_\mu = \mathrm{Id} + f_\mu$ preserve the foliation whose leaves are given by the lines $\{x_0 + t\omega_0 | t \in \mathbb{R}\}$. Since $\omega_0 \cdot k \neq 0$ for all $k \in \mathbb{Z}^n \backslash \{0\}$, each leaf is dense in $\mathbb{T}^n$ and thus, even when $n > 1$, from the dynamical point of view these maps are essentially one-dimensional. In particular, since f_μ are continuous is suffices to know how F_μ acts on $\{t\omega_0 | t \in \mathbb{R}\}$. At this stage we are not making any mathematical claim on what happened when μ lies in the gaps of the cantor set $\mathcal{C}$, but the paper

[Pet02] contains numerical evidence and conjectures. Also, in Example 2.8 of the Prologue of [Gar83], a family of this form is studied around $f_0 = 0$ as an example of a map possessing a "weak" type of strange attractor exhibiting sensitive dependence to initial conditions.

Torus maps of the form (9.3) arise in the study of resonators with quasi-periodically moving walls. In these maps, some degree of degeneracy is unavoidable. Specifically, one has

$$\frac{d}{d\mu}\mathrm{avg}[f_\mu] = 0$$

In this setting, given $\mu \in \mathcal{C}$, a solution H_μ to (9.1) implies the energy of the electric field in the cavity remains uniformly bounded in time. The periodic case ($n = 1$) was studied in [dlLP99] (see also [DDG98], [DDG96], [CC95]).

In Section 9.1 we define a functional $\mathcal{F}$, which we will be used to establish Theorem 9.0.4. We show that $\mathcal{F}$ is differentiable and we construct an approximate right (and left) inverse R to $D_2\mathcal{F}$ and apply Theorem 6.2.3. In Section 9.2 we use the implicit function we obtained in Section 9.1 to construct $\mathcal{C}_F$ and $h_\mu : \mathcal{C}_F \to X_{\sigma/2}$, $a_\mu : \mathcal{C}_F \to \mathbb{R}$ both C^γ_{Wht}, thus establish Theorem 9.0.4.

9.1 Definition of the Functional

Lifting to the universal cover and re-arranging terms, we can express (9.1) as

$$\bar{\mathcal{F}}(f_\mu; h, T_a) = 0 \tag{9.4}$$

where

$$\bar{\mathcal{F}}(f_\mu; h, a) \equiv h - h \circ T_a + f_\mu \circ (Id + h) - a \tag{9.5}$$

Here for a given $f_\mu \in X_\sigma$ (i.e. the independent variable) we want to find $h \in X_{\sigma'}$ and $a \in \mathbb{R}^n$ (i.e. dependent variables) so that (9.4) holds. In light of the small small divisor problem (SD) discuss in Section 8.5, if turns out to be much easier to consider a as a dependent variable and restrict $\mathcal{D}_\Upsilon$, i.e. $x = (f_\mu, a) \in X_\sigma \times \mathcal{D}_\Upsilon$. Furthermore, since for any given f_μ there can be at most at most one a satisfying (9.4), even when $a \in \mathcal{D}_\Upsilon$, in order to have $\mathcal{F}(x, g(x)) = 0$ for all $x = (f_\mu, w) \in \mathcal{C}_\sigma = X_\sigma \times \mathcal{D}_\Upsilon$, we add an additional dependent parameter $v \in \mathbb{R}^n$. We refer to this process of converting a to a dependent variable and adding v as "borrowing parameters." We now rigorously define a concrete F satisfying the hypotheses in Section 4.3.

Let $C > 1$ and define $C^* = (C-1)/\max(k(0), M)$ with $k(0)$ the constant in (4.4) and M the constant in (8.31). Also, for $q > 1$ define

$$U_0 = \{f \in [X_0^q]^n : \|Df\|_{[X_0]} < C^*\} \quad V_0 = \{g \in [X_0^q]^n : \|Dg\|_{[X_0]} < C^*\}$$

and using C^*, U_0 and V_0, define

$$\mathbb{X}_\sigma = [X_{C^*\sigma}]^n \times \mathbb{R}^n, \quad \mathbb{Y}_\sigma = [X_{C^*\sigma}]^n \times \mathbb{R}^n, \quad \mathbb{Z}_\sigma = [X_\sigma]^n$$

$$\mathbb{U}_0 = U_0 \times \mathbb{R}^n, \quad \mathbb{V}_0 = V_0 \times \mathbb{R}^n$$

$$\mathcal{C}_0 = U_0 \times \mathcal{D}_\Upsilon \subseteq \mathbb{U}_\sigma$$

As in (4.23) of Example 4.3.1, define $\mathcal{F} : \mathbb{U}_0 \times \mathbb{V}_0 \to \mathbb{Z}_0$ by

$$\mathcal{F}(\underbrace{f, a}_{x}; \underbrace{h, v}_{y}) \equiv h - h \circ T_a + v + f \circ (Id + h) - a \tag{9.6}$$

Given $0 < \sigma' < \sigma \leq 1$ define $D_2\mathcal{F} : \mathbb{U}_\sigma \times \mathbb{V}_\sigma \to L(\mathbb{Y}_\sigma, \mathbb{Z}_{\sigma'})$ by

$$D_2\mathcal{F}(\underbrace{f, a}_{x}; \underbrace{h, v}_{y})[\underbrace{\Delta h, \Delta v}_{\Delta y}] \equiv \Delta h - (\Delta h)\circ T_a + (D_\theta f)\circ(Id + h)[\Delta h] + \Delta v \quad (9.7)$$

and, for $0 < \sigma' < \sigma \leq 1$, with $\Delta z \in Z_\sigma$ define $R : \mathcal{C}_\sigma \times \mathbb{V}_\sigma \to L(\mathbb{Z}_\sigma, \mathbb{Y}_{\sigma'})$ by

$$R(\underbrace{f, \omega}_{x}; \underbrace{h, v}_{y})[\Delta z] \equiv ((Id + D_\theta h)[\mathcal{S}_\omega^{-1}[\Pi_{\mathrm{Id}\to 0}[\Delta z]]], \Pi_{\mathrm{Id}\to 1}[\Delta z]) \quad (9.8)$$

Before showing that $\mathcal{F}$ satisfies the Hypotheses of Theorem 6.2.3, we make the following important:

Remark 9.1.1. *Note that $\bar{\mathcal{F}}(f; h, a) = \mathcal{F}(f, a; h, 0)$. In particular, if $v = 0$ then the functional equation $\mathcal{F}(f, a; h, 0) = 0$ implies* (9.1).

Theorem 9.1.2. *Let $\mathbb{X}_\sigma$, $\mathbb{Y}_\sigma$, $\mathbb{Z}_\sigma$, $\mathbb{U}_0$, $\mathbb{V}_0$ and $\mathcal{C}_0$ be defined as above. Then we have:*

(A) The functional $\mathcal{F}$ defined in (9.6) *has $D_2\mathcal{F}$ as in* (9.7) *and satisfies hypothesis (F0), (F.A0), (F.P1), (F.A1), (F.W1) and (F.A2) described in Section 4.3.*

(B) The approximate right inverse R defined in (9.8) *satisfies hypothesis (F.P2), (F.A2), (F.W2) and (F.W4) described in Section 4.3.*

Thus, taking $\bar{x} \equiv (\omega_0, \omega_0)$ and $\bar{y} \equiv (0, 0)$, since $\mathcal{F}(\bar{x}, \bar{y}) = 0$ we can can apply Corollary 6.1.2 and Theorem 6.2.3 (or Corollary 7.1.3 and Theorem 7.2.2) *and obtain* $g : \mathcal{C}_{\tau'} \cap B_{\tau'}(\bar{x}, \epsilon) \to V_{\tau/2}$ *with* $\mathcal{F}(x, g(x)) = 0$, *i.e.*

$$\mathcal{F}(\underbrace{f, a}_{x}; \underbrace{h_{f,a}, v_{f,a}}_{y}) = h_{f,a} - h_{f,a} \circ T_a + f \circ (Id + h_{f,a}) - a + v_{f,a} = 0 \qquad (9.9)$$

with $(f, a) \to h_{f,a}$ *and* $(f, a) \to v_{f,a}$ *both* C^{γ}_{Wht}.

Proof. Note that for $0 < \sigma \leq 1$ the sets $\mathbb{U}_\sigma$, $\mathbb{V}_\sigma$ are open.

The proof of (A) follows from (C0) and (C1) in Theorem 8.7.3 and Lemma 6 in [Mey75].

To prove (B), note that R is composed of bounded linear operators and thus, using (8.14) we have

$$\begin{aligned}
\|R(x; y)[\Delta z]\|_{\mathbb{Y}_{\sigma'}} &\leq \left\|(\mathrm{Id} + D_\theta h)[\mathcal{S}_\omega^{-1}[\Pi_{\mathrm{Id}\to 0}[\Delta z]]]\right\|_{[X_{C\sigma'}]^n} + \|\Pi_{\mathrm{Id}\to 1}[\Delta z]\|_{\mathbb{R}^n} \\
&\leq C\left\|\mathcal{S}_\omega^{-1}[\Pi_{\mathrm{Id}\to 0}[\Delta z]]\right\|_{[X_{C\sigma'}]^n} + \|\Delta z\|_{\mathbb{Z}_{\sigma'}} \\
&\leq C\Omega_\Upsilon(\sigma - \sigma')\|\Pi_{\mathrm{Id}\to 0}[\Delta z]\|_{[X_{C\sigma}]^n} + \|\Delta z\|_{\mathbb{Z}_{\sigma'}} \\
&\leq \Omega_R(\sigma - \sigma')\|\Delta z\|_{\mathbb{Z}_{\sigma'}}
\end{aligned}$$

with $\Omega_R(s) = (C\Omega_\Upsilon(s) + 1)$. Thus R satisfies (4.32).

To establish (4.33), note that by differentiating the functional $\mathcal{F}$ defined in (9.6) with respect to θ we have

$$D_\theta \mathcal{F}(f, \omega; h, v) = D_\theta h - (D_\theta h) \circ T_\omega + (D_\theta f) \circ (\mathrm{Id} + h)[\mathrm{Id} + D_\theta h] \qquad (9.10)$$

Substituting $\Delta h = (\mathrm{Id} + D_\theta h)[W]$ with $W \in X_\sigma$ into (9.7) and using (9.10) we get

$$D_2\mathcal{F}(x; y)[\Delta y] \equiv \underbrace{[W - W \circ T_\omega + \Delta v]}_{\text{(solvable)}} + D_\theta \mathcal{F}(f, \omega; h, v)[W] \qquad (9.11)$$

The above follows section 5 of [Zeh75] which describes how (9.10) can be used to compute an approximate right inverse for functionals, such at $\mathcal{F}$, which posses a "group structure."

Note in (9.8), we have $\Delta h = (Id + D_\theta h)[\mathcal{S}_\omega^{-1}[\Pi_{\text{Id}\to 0}[\Delta z]]]$ and thus, combining (9.7) and (9.8) and using (9.11) we get

$$\Delta z - D_2\mathcal{F}(x;y)R(x;y)[\Delta z] = D_\theta\mathcal{F}(f,\omega;h,v)[\mathcal{S}_\omega^{-1}[\Pi_{\text{Id}\to 0}[\Delta z]]] \qquad (9.12)$$

Again using (8.14), note

$$\begin{aligned}\|\Delta z - D_2\mathcal{F}(x;y)[R(x;y)[\Delta z]]\|_{\mathbb{Z}_{\sigma'}} &\leq \left\|D_\theta\mathcal{F}(f,\omega;h,v)[\mathcal{S}_\omega^{-1}[\Pi_{\text{Id}\to 0}[\Delta z]]]\right\|_{\mathbb{Z}_{\sigma'}} \\ &\leq \|D_\theta\mathcal{F}(f,\omega;h,v)\|_{\mathbb{Z}_{\sigma'}}\left\|\mathcal{S}_\omega^{-1}[\Pi_{\text{Id}\to 0}[\Delta z]]\right\|_{\mathbb{Z}_{\sigma'}} \\ &\leq \Omega_A(\sigma-\sigma')\|\mathcal{F}(f,\omega;h,v)\|_{\mathbb{Z}_\sigma}\|\Delta z\|_{\mathbb{Z}_\sigma}\end{aligned}$$

with $\Omega_A(s) = Cs\Omega_\Upsilon(s)$. Thus R also satisfies (4.33).

Finally, the fact that R satisfies (F.W2) and (F.W4) follows from Proposition 8.5.3.

□

9.2 Obtaining KAM from IFT

In this section we present the proof of Theorem 9.0.4 using Theorem 9.1.2. Informally, the basic idea in going from Theorem 9.1.2 to Theorem 9.0.4 is to find $\mu \in \Omega$ so that for the corresponding $f_\mu \in U_\sigma$ there exists $a \in \mathcal{D}_\Upsilon$ with $v_{f,a} = 0$. Then, as in Remark 9.1.1, (9.9) reduces to (9.1) and hence the change of variables $\text{Id} + h_{f,a}$ transforms $F_\mu = \text{Id} + f$ into the rotation T_a.

Proof of Theorem 9.0.4

Applying Theorem 9.1.2, note that the Whitney derivative of the resulting $v_{f,a}$ has the form

$$Dv_{f,a}[\Delta\mu, \Delta a] = \text{avg}[D_1\mathcal{F}(f_\mu, a; h, v)]$$

with

$$D_1\mathcal{F}(f, a; h, v)[\Delta f, \Delta a] = (\Delta f)\circ(\text{Id} + h) - (\text{Id} + D_\theta h \circ T_a)[\Delta a]$$

Thus, we have

$$D_a v_{f,a}[\Delta a] = -\Delta a$$

and

$$D_f v_{f,a}[\Delta f] = \text{avg}[(\Delta f)\circ(\text{Id} + h)]$$

Taking $v_{\mu,a}$ for $(\mu, a) \in \Omega \times \mathcal{D}_\Upsilon$ and using the Whitney extension theorem (see Theorem 3.3.1) to extend $v_{\mu,a} = v_{f_\mu,a}$ to all of $\Omega \times \mathbb{R}^n$ we get

$$D_a \left[v_{\mu,a}\right]_{\mu=0,a=\omega_0} = -\text{Id} \tag{9.13}$$

and

$$D_\mu \left[v_{\mu,a}\right]_{\mu=0,a=\omega_0} = D_\mu \text{avg}[f_\mu] \tag{9.14}$$

By (9.13) we can apply the classical implicit function theorem to $v_{\mu,a}$ around $\mu = 0$, $a = \omega_0$ and obtain $a(\mu)$ with $v_{\mu,a(\mu)} = 0$. Define

$$\mathcal{C}_F \equiv v^{-1}(\mathcal{D}_\Upsilon)$$

and

$$h_\mu \equiv h_{\mu,\alpha(\mu)}$$

and observe that for any $\mu \in \mathcal{C}_F$ h_μ and a_μ satisfy assertion (a) of Theorem 9.0.4.

Using (9.14) and condition (iii) from Theorem 9.0.4 note that we have $D_\mu^m[v_{\mu,\alpha}]_{\mu=0,a=\omega_0} = A \neq 0$ and thus $D_\mu^m[a(\mu)]_{\mu=0} = A \neq 0$ and hence $a(\mu) = w_0 + A\mu^{\otimes m} + O(\mu^\nu)$. Using Proposition 8.5.2, note that

$$|B(\omega_0, \epsilon) \cap \mathcal{D}_\Upsilon| \geq |B(\omega_0, \epsilon)| - M\epsilon^n/c > 0$$

and hence, applying Proposition B.7, assertion (b), i.e. (9.2), follows. $\square$

Appendix

Appendix A

Open Questions

1. Compare the definition of a scale of Banach spaces X_σ a la [Zeh75] in Section 4.1 with what you would get by completing the various semi-norms of a "tame Frechet space" a la [Ham82].

2. Is there an abstract version of the Arzela-Ascoli in X_0^ℓ, i.e. is the embedding of $X_0^{\ell+m}$ into X_0^ℓ is compact?

3. Given analytic-smoothing in the family X_σ with respect to X_0^ℓ, viewing S_t as acting in one-parameter family X_0^ℓ, one obtains C^∞-smoothing.

4. What reasonable conditions can be to placed on Υ given in Example 4.2.8 to ensure that the function Ψ_{Ω_Υ} described in Definition 4.2.5 has $\Psi_{\Omega_\Upsilon}(s) \leq Cs^{-\alpha}$ for some α (other than taking $\Upsilon(t) = Ct^\alpha$ which simply leads to $\Omega_\Upsilon(s) = As^{-\alpha}$ as in Example 4.2.9).

Appendix B

Results about The Density of Pullbacks

Proposition B.1. *Given a Lipschitz function* $f : \mathbb{R} \to \mathbb{R}$ *and a measurable set* $D \subseteq \mathbb{R}$ *one has*

$$|f(D)| \leq \|f\|_{lip} |D| \tag{B.1}$$

Proof. Since f is Lipschitz, it is also of bounded variation and so $f'(x)$ exists almost everywhere. Using this we have

$$|f(D)| = \int_{f(D)} 1du = \int_D 1f'(x)dx \leq \int_D \|f\|_{\text{lip}} dx = \|f\|_{\text{lip}} |D|$$

which establishes (B.1). □

Definition B.2. *Given a set* $C \subseteq \mathbb{R}$, $\kappa \in \mathbb{R}$ *and* $0 \leq \gamma \leq 1$ *define*

$$d_\kappa(C) \equiv \limsup_{\epsilon \to 0^+} \epsilon^{-\kappa} |C \cap (0, \epsilon)| \tag{B.2}$$

and

$$d_{\kappa,\gamma}(C) \equiv \limsup_{\epsilon \to 0^+} \epsilon^{-\kappa} |C \cap (\gamma\epsilon, \epsilon)| \tag{B.3}$$

■

Remark B.3. *The function* $d_\kappa(C)$ *measures a one sided "upper density" at* 0*. That is, if* $d_\kappa(C) = c < \infty$ *then given any* $\delta > 0$ *there is an* $\epsilon_* > 0$ *so that* *for any* $0 < \epsilon < \epsilon_*$ *one has*

$$|C \cap (0, \epsilon)| \leq (c + \delta)\epsilon^\kappa$$

Similarly, one could use liminf to define a "lower density" measure to provide estimates of the form

$$|C \cap (0, \epsilon)| \geq (c - \delta)\epsilon^\kappa$$

Other variations for "densities" include using intervals

$$(-\epsilon, 0), (-\epsilon, -\gamma\epsilon), (-\epsilon, \epsilon), \text{ or } (-\epsilon, -\gamma\epsilon) \cup (\gamma\epsilon, \epsilon)$$

in place of $(0, \epsilon)$ *and* $(\gamma\epsilon, \epsilon)$ *or using other functions in place of* $c\epsilon^\kappa$ *to measure* $|C \cap (0, \epsilon)|$.

Proposition B.4. (Properties of $d_\kappa(C)$ and $d_{\kappa,\gamma}(C)$)

1. $d_{\kappa,\gamma}(C) = d_\kappa(C) = 0$ *for any* $\kappa < 1$

2. $d_\kappa(C)$ *and* $d_{\kappa,\gamma}(C)$ *are increasing as functions of* κ

3. *For any* $\epsilon > 0$

$$d_\kappa(C) > 0 \implies d_{\kappa+\epsilon}(C) = \infty$$
$$d_\kappa(C) < \infty \implies d_{\kappa-\epsilon}(C) = 0$$

   and thus for every $C \subseteq \mathbb{R}$ *there is a unique "critical" value of* κ *such that for every* $\epsilon > 0$, $d_{\kappa-\epsilon}(C) = 0$ *and* $d_{\kappa+\epsilon}(C) = \infty$

4. *$d_\kappa(C)$ and $d_{\kappa,\gamma}(C)$ are related via*

$$0 \le (1-\gamma^\kappa) d_\kappa(C) \le d_{\kappa,\gamma}(C) \le d_\kappa(C) \le \infty$$

5. *For any invertible orientation preserving Lipschitz function*

$$f : (-T, T) \to \mathbb{R}$$

with $f(0) = 0$ and Lipschitz inverse f^{-1} such that

$$\left\| f|_{(0,\epsilon)} \right\|_{lip} \to 1 \text{ as } \epsilon \to 0 \tag{B.4}$$

and

$$\left\| f^{-1}|_{(0,\epsilon)} \right\|_{lip} \to 1 \text{ as } \epsilon \to 0 \tag{B.5}$$

one has

$$d_\kappa(f(C)) = d_\kappa(C) \tag{B.6}$$

Proof. Properties 1, 2 and 3 are clear from the definitions.

The only non-trivial inequality in Property 4 is $(1-\gamma^\kappa)d_\kappa(C) \le d_{\kappa,\gamma}(C)$ and the only situation in which this asserts a non-vacuous statement is when $d_{\kappa,\gamma}(C) = c < \infty$ and $0 < \gamma < 1$. To establish the inequality in this case let $\delta > 0$ be given and choose $R > 0$ so that for all $0 < r \le R$

$$r^{-\kappa} \left| C \cap (\gamma r, r) \right| \le c + \delta$$

Note that for $0 < t \le R$ we have

$$t^{-\kappa} \left| C \cap (0,t) \right| = \sum_{n=0}^{\infty} (\gamma^\kappa)^n \left((\gamma^n t)^{-\kappa} \left| C \cap (\gamma(\gamma^n t), (\gamma^n t)) \right| \right) \le \frac{c+\delta}{1-\gamma^\kappa}$$

Taking the limsup as $t \to 0$ implies $(1-\gamma^\kappa)d_\kappa(C) \leq d_{\kappa,\gamma}(C) + \delta$ and since δ was arbitrary one has $(1-\gamma^\kappa)d_\kappa(C) \leq d_{\kappa,\gamma}(C)$.

To prove property 5, note that since f^{-1} is Lipschitz given any x with $|x| \leq r$ one has

$$|f^{-1}(x)| = |f^{-1}(x) - f^{-1}(0)| \leq \big\|f^{-1}|_{(0,r)}\big\|_{\mathrm{lip}} |x-0| = \big\|f^{-1}|_{(0,r)}\big\|_{\mathrm{lip}} |x|$$

so for $r' \geq r\big\|f^{-1}|_{(0,r)}\big\|_{\mathrm{lip}}$ one has

$$f(C) \cap (0,r) \subseteq f\left(C \cap (0,r')\right)$$

Using this inclusion along with Proposition B.1, one has

$$|f(C) \cap (0,r)| \leq |f\left(C \cap (0,r')\right)| \leq \big\|f|_{(0,r')}\big\|_{\mathrm{lip}} |C \cap (0,r')| \tag{B.7}$$

Let $\delta > 0$ and pick $R > 0$ such that for all $0 < r' \leq R$

$$(r')^\kappa |C \cap (0,r')| \leq d_\kappa(C) + \delta \tag{B.8}$$

Multiplying (B.7) by r^κ and using (B.8), for $r \leq r'/\big\|f^{-1}|_{(0,r)}\big\|_{\mathrm{lip}}$ with $r' \leq R$ one has

$$\begin{aligned} r^\kappa |f(C) \cap (0,r)| &\leq \big\|f|_{(0,r')}\big\|_{\mathrm{lip}} r^\kappa |C \cap (0,r')| \\ &\leq \big\|f|_{(0,r')}\big\|_{\mathrm{lip}} (\big\|f^{-1}|_{(0,r)}\big\|_{\mathrm{lip}})^{-\kappa} (r')^\kappa |C \cap (0,r')| \\ &\leq \big\|f|_{(0,r')}\big\|_{\mathrm{lip}} (\big\|f^{-1}|_{(0,r)}\big\|_{\mathrm{lip}})^{-\kappa} (d_\kappa(C) + \delta) \end{aligned} \tag{B.9}$$

Taking the limsup as $r \to 0$ on both sides of (B.9) one obtains

$$d_\kappa(f(C)) \leq \big\|f|_{(0,r')}\big\|_{\mathrm{lip}} (d_\kappa(C) + \delta)$$

and letting $r' \to 0$ and $\delta \to 0$ gives

$$d_\kappa(f(C)) \le d_\kappa(C) \tag{B.10}$$

Replacing f and C with f^{-1} and $f(C)$, (B.10) also gives

$$d_\kappa(C) = d_\kappa(f^{-1}(f(C))) \le d_\kappa(f(C)) \tag{B.11}$$

Together, (B.10) and (B.11) establish equality (B.6).

□

Proposition B.5. *For any $a, d \in \mathbb{R}$ with $a > 0, d \ge 1$ define*

$$f : \mathbb{R} \to \mathbb{R} \qquad f(x) = ax^d$$

Let $D \subseteq \mathbb{R}$ be given and define

$$C \equiv f^{-1}(D)$$

Then for any $0 < \gamma < 1$

$$\left(\frac{a^{(\kappa-1)/d}}{d}\right) d_{\kappa',\gamma^d}(D) \le d_{\kappa,\gamma}(C) \le \gamma^{-(d-1)} \left(\frac{a^{(\kappa-1)/d}}{d}\right) d_{\kappa',\gamma^d}(D) \tag{B.12}$$

where $\kappa' = 1 + (\kappa - 1)/d \le \kappa$. In particular

$$\left(\frac{a^{(\kappa-1)/d}}{d}\right) d_{\kappa'}(D) \le d_\kappa(C) \le c \left(\frac{a^{(\kappa-1)/d}}{d}\right) d_{\kappa'}(D) \tag{B.13}$$

with

$$c = \begin{cases} \frac{(\kappa+d-1)^{(\kappa+d-1)/\kappa}}{\kappa(d-1)^{(d-1)/\kappa}} & \text{if } d \ne 1 \\ 1 & \text{if } d = 1 \end{cases}$$

Proof. Note that

$$\left\| f|_{(\gamma\epsilon,\epsilon)} \right\|_{\mathrm{lip}} = ad\epsilon^{d-1}$$

so

$$\begin{aligned}\left|D \cap (a\gamma^d\epsilon^d, a\epsilon^d)\right| &= \left|f\left(C \cap (\gamma\epsilon, \epsilon)\right)\right| \\ &\leq \left\| f|_{(\gamma\epsilon,\epsilon)} \right\|_{\mathrm{lip}} \left|C \cap (\gamma\epsilon, \epsilon)\right| \\ &\leq ad\epsilon^{d-1} \left|C \cap (\gamma\epsilon, \epsilon)\right|\end{aligned}$$

in particular

$$\left(\frac{a^{(\kappa-1)/d}}{d}\right)(a\epsilon^d)^{-(1+(\kappa-1)/d)} \left|D \cap (\gamma^d a\epsilon^d, a\epsilon^d)\right| \leq \epsilon^{-\kappa} \left|C \cap (\gamma\epsilon, \epsilon)\right| \tag{B.14}$$

Similarly

$$\left\| f^{-1}|_{(a\gamma^d\epsilon^d, a\epsilon^d)} \right\|_{\mathrm{lip}} = \frac{1}{a^{1/d}d}(a\gamma^d\epsilon^d)^{(1-d)/d} = \frac{\gamma^{(1-d)}}{a^{1/d}d}(a\epsilon^d)^{-(1-1/d)}$$

hence

$$\begin{aligned}\left|C \cap (\gamma\epsilon, \epsilon)\right| &= \left|f^{-1}\left(D \cap (a\gamma^d\epsilon^d, a\epsilon^d)\right)\right| \\ &\leq \left\| f|_{(\gamma\epsilon,\epsilon)} \right\|_{\mathrm{lip}} \left|D \cap (a\gamma^d\epsilon^d, a\epsilon^d)\right| \\ &\leq \frac{\gamma^{(1-d)}}{a^{1/d}d}(a\epsilon^d)^{-(1-1/d)} \left|D \cap (a\gamma^d\epsilon^d, a\epsilon^d)\right|\end{aligned}$$

so that

$$\epsilon^{-\kappa} \left|C \cap (\gamma\epsilon, \epsilon)\right| \leq \gamma^{-(d-1)} \left(\frac{a^{(\kappa-1)/d}}{d}\right)(a\epsilon^d)^{-(1+(\kappa-1)/d)} \left|D \cap (a\gamma^d\epsilon^d, a\epsilon^d)\right| \tag{B.15}$$

Combining (B.14) and (B.15) one has

$$\left(\frac{a^{(\kappa-1)/d}}{d}\right)(a\epsilon^d)^{-(1+(\kappa-1)/d)}\left|D\cap(\gamma^d a\epsilon^d, a\epsilon^d)\right|$$
$$\leq \epsilon^{-\kappa}\left|C\cap(\gamma\epsilon,\epsilon)\right|$$
$$\leq \gamma^{-(d-1)}\left(\frac{a^{(\kappa-1)/d}}{d}\right)(a\epsilon^d)^{-(1+(\kappa-1)/d)}\left|D\cap(\gamma^d a\epsilon^d, a\epsilon^d)\right| \quad \text{(B.16)}$$

Taking the limsup as $\epsilon \to 0$ in (B.16) we obtain (B.12). Using Property 4 of Proposition B.4, (B.12) can be written as in terms of $d_\kappa(C)$ and $d_\kappa(D)$. Specifically,

$$(1-\gamma^{d\kappa})\left(\frac{a^{(\kappa-1)/d}}{d}\right)d_{\kappa'}(D) \leq d_\kappa(C)$$

so letting $\gamma \to 0$ one gets

$$\left(\frac{a^{(\kappa-1)/d}}{d}\right)d_{\kappa'}(D) \leq d_\kappa(C)$$

which establishes the left hand inequality in (B.13). Similarly note

$$(1-\gamma^\kappa)d_{\kappa,\gamma}(C) \leq \gamma^{-(d-1)}\left(\frac{a^{(\kappa-1)/d}}{d}\right)d_{\kappa'}(D)$$

so if $d = 1$ the we can again take $\gamma \to 0$ and obtain the right hand inequality of (B.13) with $c = 1$. If $d \neq 1$, choosing $\gamma = ((d-1)/(d+\kappa-1))^{1/\kappa}$ so as to minimizes the expression $\gamma^{-(d-1)}/(1-\gamma^\kappa)$, one obtains the right hand inequality of (B.13) with $c = \gamma^{-(d-1)}/(1-\gamma^\kappa)$. □

Corollary B.6. *Let*

$$f : \mathbb{R} \to \mathbb{R}$$

be a differentiable function with

$$f(x) = ax^n + o(|x|^n),$$

for some $a > 0$ *and* $n \geq 1$ *. Let* $D \subseteq \mathbb{R}$ *be given and define*

$$C \equiv f^{-1}(D)$$

Then for any $0 < \gamma < 1$ *equations* (B.12) *and* (B.13) *still hold.*

Proof. Define

$$\gamma(x) \equiv \frac{x}{|x|}\left(\frac{|x|}{a}\right)^{1/d},\ x \neq 0$$
$$\gamma(0) \equiv 0$$

Note that $\gamma \circ f$ is differentiable for $x \neq 0$ and

$$\lim_{|x| \to 0} \frac{\gamma \circ f(x) - \gamma \circ f(0)}{x} = \lim_{|x| \to 0} \frac{\gamma \circ f(x)}{x} = 1$$

so $\gamma \circ f$ is differentiable on $\mathbb{R}$ and $(\gamma \circ f)'(0) = 1$. Applying the (finite dimensional) inverse function theorem to $\gamma \circ f$ for some $T > 0$ one obtains a differentiable function $g : (-T, T) \to \mathbb{R}$ such that for x sufficiently small $f \circ \gamma \circ g(x) = x$ and $g'(0) = 1$. Since g and g^{-1} are orientation preserving Lipschitz functions with $g(0) = 0$ satisfying (B.4) and (B.5), by Property 5 of Proposition B.4 one has

$$d_\kappa(D) = d_\kappa(g(D))$$

Note that

$$C \equiv f^{-1}(D) = \gamma \circ g(D)$$

and for $x \geq 0$,

$$\gamma^{-1}(x) = a x^n$$

so applying Proposition B.5 with $D' \equiv g(D)$ and $f' = \gamma^{-1}$ one obtains the desired control over the set $C' \equiv (f')^{-1}(D') = \gamma(g(D)) = C$. □

Proposition B.7. *Given a function* $f : \mathbb{R}^d \to \mathbb{R}^n$ *with*

1. $f(0) = 0$

2. $f(x) = A[x]^{\otimes m} + o(|x|^m)$*, where* $w = A[v]^{\otimes m} \neq 0$ *for some* $v \in \mathbb{R}^d$

and a set $D \subseteq \mathbb{R}^n$ *with the property that for some* $0 < \kappa < 1$

$$d_\kappa(\{t : tw \notin D\}) = d < \infty$$

then there is positive constants M *and* r_* *such that the set*

$$C \equiv f^{-1}(D)$$

has

$$|C \cap B(0, r)| \geq M r^{d/m}$$

for all $r < r_*$.

Proof. By Corollary B.6, we have

$$|\{t : tv \notin (C \cap B(0, r))\}| \leq c r^{1+m(\kappa-1)}$$

The same estimate hold, with a smaller c, uniformly on the cone

$$\{v' : |A[v']^{\otimes n} - w| \geq |w|/2\}$$

Estimating the volume of $C \cap B(0, r)$ on this cone the result follows. □

Bibliography

[AR67] Ralph Abraham and Joel Robbin. *Transversal mappings and flows.* W. A. Benjamin, Inc., New York-Amsterdam, 1967.

[BHS96a] H. W. Broer, G. B. Huitema, and M. B. Sevryuk. Families of quasi-periodic motions in dynamical systems depending on parameters. In *Nonlinear dynamical systems and chaos (Groningen, 1995)*, pages 171–211. Birkhäuser, Basel, 1996.

[BHS96b] H. W. Broer, G. B. Huitema, and M. B. Sevryuk. *Quasi-Periodic Motions in Families of Dynamical Systems.* Order Amidst Chaos. Springer-Verlag, Berlin, 1996.

[BHS96c] Hendrik W. Broer, George B. Huitema, and Mikhail B. Sevryuk. *Quasi-periodic motions in families of dynamical systems.* Springer-Verlag, Berlin, 1996. Order amidst chaos.

[Brj71] A. D. Brjuno. Analytic form of differential equations. I. *Trudy Moskov. Mat. Obšč.*, 25:119–262, 1971. English translation: *Trans. Moscow Math. Soc.*, 25(1971):131–288, 1973.

[Brj72] A. D. Brjuno. Analytic form of differential equations. II. *Trudy Moskov. Mat. Obšč.*, 26:199–239, 1972. English translation: *Trans. Moscow Math. Soc.*, 26(1972):199–239, 1974.

[CC95] C. K. Cole and W. C.Schieve. Radiation modes of a cavity with a moving boundary. *Phys. Rev. A (3)*, 52:4405–4415, 1995.

[CS94] Chong Qing Cheng and Yi Sui Sun. Existence of KAM tori in degenerate Hamiltonian systems. *J. Differential Equations*, 114(1):288–335, 1994.

[DDG96] J. Dittrich, P. Duclos, and N. Gonzalez. d'Alembert field driven by a periodically moving wall. In *Problems of quantum field theory (Alushta, 1996)*, pages 166–169. Joint Inst. Nuclear Res., Dubna, 1996.

[DDG98] J. Dittrich, P. Duclos, and N. Gonzalez. Stability and instability of the wave equation solutions in a pulsating domain. *Rev. Math. Phys.*, 10(7):925–962, 1998.

[DeL97] David DeLatte. Diophantine conditions for the linearization of commuting holomorphic functions. *Discrete Contin. Dynam. Systems*, 3(3):317–332, 1997.

[Die69] J. Dieudonné. *Foundations of modern analysis.* Academic Press, New York, 1969. Enlarged and corrected printing, Pure and Applied Mathematics, Vol. 10-I.

[dlLO99] R. de la Llave and R. Obaya. Regularity of the composition operator in spaces of Hölder functions. *Discrete Contin. Dynam. Systems*, 5(1):157–184, 1999.

[dlLP99] Rafael de la Llave and Nikola P. Petrov. Theory of circle maps and the problem of one-dimensional optical resonator with a periodically moving wall. *Phys. Rev. E (3)*, 59(6):6637–6651, 1999.

[Gar83] Luis Garrido, editor. *Dynamical systems and chaos*, Berlin, 1983. Springer-Verlag.

[GJ71] Robert E. Greene and Howard Jacobowitz. Analytic isometric embeddings. *Ann. of Math. (2)*, 93:189–204, 1971.

[Gla58] Georges Glaeser. Étude de quelques algèbres tayloriennes. *J. Analyse Math.*, 6:1–124; erratum, insert to 6 (1958), no. 2, 1958.

[Had98] J. Hadamard. Sur le module maximum d'une fonction et de ses derives. *Bull. Soc. Math. France*, 42:68–72, 1898.

[Ham82] Richard S. Hamilton. The inverse function theorem of Nash and Moser. *Bull. Amer. Math. Soc. (N.S.)*, 7(1):65–222, 1982.

[Hör76] Lars Hörmander. The boundary problems of physical geodesy. *Arch. Rational Mech. Anal.*, 62(1):1–52, 1976.

[Hör77] L. Hörmander. Correction to: "The boundary problems of physical geodesy" (Arch. Rational Mech. Anal. **62** (1976), no. 1, 1–52). *Arch. Ration. Mech. Anal.*, 65(4):395, 1977.

[Hör85] Lars Hörmander. On the Nash Moser implicit function theorem. *Ann. Acad. Sci. Fenn. Ser. A I Math.*, 10:255–259, 1985.

[Hör90] Lars Hörmander. The Nash-Moser theorem and paradifferential operators. In *Analysis, et cetera*, pages 429–449. Academic Press, Boston, MA, 1990.

[Jac72] Howard Jacobowitz. Implicit function theorems and isometric embeddings. *Ann. of Math. (2)*, 95:191–225, 1972.

[KM97] Andreas Kriegl and Peter W. Michor. *The convenient setting of global analysis*. American Mathematical Society, Providence, RI, 1997.

[Kol49] A. Kolmogoroff. On inequalities between the upper bounds of the successive derivatives of an arbitrary function on an infinite interval. *Amer. Math. Soc. Translation*, 1949(4):19, 1949.

[Las93] Jacques Laskar. Frequency analysis for multi-dimensional systems. Global dynamics and diffusion. *Phys. D*, 67(1-3):257–281, 1993.

[LN92] P. Lochak and A. I. Neĭshtadt. Estimates of stability time for nearly integrable systems with a quasiconvex Hamiltonian. *Chaos*, 2(4):495–499, 1992.

[Mey75] K. R. Meyer. *The implicit function theorem and analytic differential equations*. Springer, Berlin, 1975.

[Mos61] Jürgen Moser. A new technique for the construction of solutions of nonlinear differential equations. *Proc. Nat. Acad. Sci. U.S.A.*, 47:1824–1831, 1961.

[Mos66a] Jürgen Moser. A rapidly convergent iteration method and non-linear differential equations. II. *Ann. Scuola Norm. Sup. Pisa (3)*, 20:499–535, 1966.

[Mos66b] Jürgen Moser. A rapidly convergent iteration method and non-linear partial differential equations. I. *Ann. Scuola Norm. Sup. Pisa (3)*, 20:265–315, 1966.

[Nas56] John Nash. The imbedding problem for Riemannian manifolds. *Ann. of Math. (2)*, 63:20–63, 1956.

[Neĭ81] A. I. Neĭshtadt. Estimates in the Kolmogorov theorem on conservation of conditionally periodic motions. *Prikl. Mat. Mekh.*, 45(6):1016–1025, 1981.

[Nel69] Edward Nelson. *Topics in dynamics. I: Flows.* Princeton University Press, Princeton, N.J., 1969. Mathematical Notes.

[Nir74] L. Nirenberg. *Topics in nonlinear functional analysis.* Courant Institute of Mathematical Sciences New York University, New York, 1974. With a chapter by E. Zehnder, Notes by R. A. Artino, Lecture Notes, 1973–1974.

[Pet02] Nikola Petrov. Methods of dynamical systems, harmonic analysis and wavelets applied to several physical systems (thesis). *UMI*, 2002.

[Pös82] J. Pöschel. Integrability of Hamiltonian systems on Cantor sets. *Comm. Pure Appl. Math.*, 35(5):653–696, 1982.

[Pös01] Jürgen Pöschel. A lecture on the classical KAM theorem. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, pages 707–732. Amer. Math. Soc., Providence, RI, 2001.

[Rüs75] Helmut Rüssmann. *On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus*, pages 598–624.Lecture Notes in Phys., Vol. 38. Springer, Berlin, 1975.

[Rüs76a] Helmut Rüssmann. Note on sums containing small divisors. *Comm. Pure Appl. Math.*, 29(6):755–758, 1976.

[Rüs76b] Helmut Rüssmann. On optimal estimates for the solutions of linear difference equations on the circle. In *Proceedings of the Fifth Conference on Mathematical Methods in Celestial Mechanics (Oberwolfach, 1975), Part I*, volume 14, pages 33–37, 1976.

[Rüs80] Helmut Rüssmann. On the one-dimensional Schrödinger equation with a quasiperiodic potential. In *Nonlinear dynamics (Internat. Conf., New York, 1979)*, pages 90–107. New York Acad. Sci., New York, 1980.

[Rüs87] Helmut Rüssmann. Über die Stabilität periodischer Lösungen bei zeitabhängigen Hamiltonschen Differentialgleichungen von einem

Freiheitsgrad. *Z. Angew. Math. Phys.*, 38(2):279–291, 1987.

[Rüs90] H. Rüssmann. Nondegeneracy in the perturbation theory of integrable dynamical systems. In *Stochastics, Algebra and Analysis in Classical and Quantum Dynamics (Marseille, 1988)*, pages 211–223. Kluwer Acad. Publ., Dordrecht, 1990.

[Rüs94] Helmut Rüssmann. On the frequencies of quasi-periodic solutions of analytic nearly integrable Hamiltonian systems. In *Seminar on Dynamical Systems (St. Petersburg, 1991)*, pages 160–183. Birkhäuser, Basel, 1994.

[Rüs01] H. Rüssmann. Invariant tori in non-degenerate nearly integrable Hamiltonian systems. *Regul. Chaotic Dyn.*, 6(2):119–204, 2001.

[Sch60] J. Schwartz. On Nash's implicit functional theorem. *Comm. Pure Appl. Math.*, 13:509–530, 1960.

[Sch69] J. T. Schwartz. *Nonlinear Functional Analysis.* Gordon and Breach Science Publishers, New York, 1969.

[Ser72] F. Sergeraert. Un théorème de fonctions implicites sur certains espaces de Fréchet et quelques applications. *Ann. Sci. École Norm. Sup. (4)*, 5:599–660, 1972.

[Ser73] F. Sergeraert. Un théorème de fonctions implicites. Applications. *Ann. Inst. Fourier (Grenoble)*, 23(2):151–157, 1973. Colloque

International sur l'Analyse et la Topologie Différentielle (Colloq. Internat. CNRS, No. 210, Strasbourg, 1972).

[Sev96] M. B. Sevryuk. Invariant tori of Hamiltonian systems that are nondegenerate in the sense of Rüssmann. *Dokl. Akad. Nauk*, 346(5):590–593, 1996.

[Sev99] M. B. Sevryuk. The lack-of-parameters problem in the KAM theory revisited. In *Hamiltonian Systems with Three or More Degrees of Freedom (S'Agaró, 1995)*, pages 568–572. Kluwer Acad. Publ., Dordrecht, 1999.

[Sim98] C. Simó. Effective computations in celestial mechanics and astrodynamics. In *Modern Methods of Analytical Mechanics and Their Applications (Udine, 1997)*, pages 55–102. Springer, Vienna, 1998.

[Ste70] Elias M. Stein. *Singular integrals and differentiability properties of functions*. Princeton University Press, Princeton, N.J., 1970.

[TLRF02] A. M. Tarquis, J.C. Losada, Benito R.M., and Borondo F. Multifractal analysis of tori destruction in a molecular hamiltonian system. *Phys. Rev. E (3)*, 65(1):016213–1 – 016213–9, 2002.

[Whi34] Hassler Whitney. Analytic extensions of differentiable functions defined in closed sets. *Trans. Amer. Math. Soc.*, 36(1):63–89, 1934.

[Whi36] Hassler Whitney. Differentiable functions defined in arbitrary subsets of Euclidean space. *Trans. Amer. Math. Soc.*, 40(2):309–317, 1936.

[Zeh75] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. I. *Comm. Pure Appl. Math.*, 28:91–140, 1975.

[Zeh76a] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. II. *Comm. Pure Appl. Math.*, 29(1):49–111, 1976.

[Zeh76b] E. Zehnder. Moser's implicit function theorem in the framework of analytic smoothing. *Math. Annalen*, 219:105–121, 1976.

Index

Vita

John Andrew Vano was born to parents Andrew and Sally on April 17, 1974. Oldest of three children, John has always had a inquisitive nature and recalls, as a child, asking his father about what types of "advanced mathematics" existed. During high school, his mathematical career was foreshadowed by his computer experimentation with diffraction patterns and Mandelbrot sets.

John graduated from Jefferson Senior High School in Alexandria, MN in the spring of 1992. He went on to attend Gustavus Adolphus College in St. Peter, MN where he enrolled in almost every class offered by the mathematics and physics departments. While at Gustavus, he met his future wife, Jennie. John wrote two honors theses and received his B.A. from Gustavus in the spring of 1996 with a double major in honors mathematics and honors physics.

In the summer of 1996, John traveled to Austin, Texas to enter the Ph.D. program in the Department of Mathematics at The University of Texas at Austin. That fall, Jennie joined him in Austin to pursue a M.S. in Library and Information Science. On July 4, 1998, John and Jennie were married in Jennie's home town of Portland, Oregon.

John's hobbies include collecting and solving Rubik's cube type puzzles, juggling, flying, playing with his cats, and taking road trips with Jennie. In the fall of 2002, John and Jennie will hit the road and head north to Madison Wisconsin with much vim and (a three year) VIGRE post doc!

Permanent address: 6950 E Lk. Carlos Dr. NE / Carlos, MN 56319

This dissertation was typeset with LaTeX[†] by the author.

[†]LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.