A NUMERICALLY ACCESSIBLE CRITERION FOR THE BREAKDOWN OF QUASI-PERIODIC SOLUTIONS AND ITS RIGOROUS JUSTIFICATION

RENATO CALLEJA AND RAFAEL DE LA LLAVE

ABSTRACT. We formulate and justify rigorously a numerically efficient criterion for the computation of the analyticity breakdown of quasi-periodic solutions in Symplectic maps (any dimension) and 1-D Statistical Mechanics models. Depending on the physical interpretation of the model, the analyticity breakdown may correspond to the onset of mobility of dislocations, or of spin waves (in the 1-D models) and to the onset of global transport in symplectic twist maps in 2-D.

The criterion proposed here is based on the blow-up of Sobolev norms of the hull functions. We prove theorems that justify the criterion. These theorems are based on an abstract implicit function theorem, which unifies several results in the literature. The proofs also lead to fast algorithms, which have been implemented and used elsewhere. The method can be adapted to other contexts.

CONTENTS

2000 *Mathematics Subject Classification.* 37A60, 70K43, 37M20, 47J08,

Key words and phrases. Analyticity breakdown, Nash-Moser method, quasi-periodic solutions, breakdown of tori.

1. Introduction

The celebrated KAM (Kolmogorov-Arnold-Moser) theory establishes the persistence under perturbation of analytic quasi-periodic solutions in a variety of mathematical contexts which model many different physical systems (e.g. celestial mechanics, solid state physics, etc.)

Depending on the physical system, the presence or not of these quasi-periodic analytic solutions has deep physical consequences. In celestial mechanics, it means an abundance of stable orbits, in solid state physics it could mean the presence of extended states and therefore, possibility of transport.

As it is well known, the existence or not of these solution depends on the values of parameters. Locating the range of parameters where these quasi-periodic solutions exist is a problem (very similar to describing a phase diagram) to which considerable attention has been devoted. See references later and Appendix B.

The goal of this paper is to present a numerically practical method to compute the values of parameters where the KAM theory breaks down which admits a rigorous mathematical justification. A more precise description of the method is in Section 2.

We will consider families of problems indexed by a parameter λ and seek to identify the set of parameters for which there is an analytic quasi-periodic solution of a fixed Diophantine frequency ω.

Roughly speaking the criterion we present (See Section 2 and Theorem 2.2 for a general overview and to Theorems 5.3, 6.1 for a justification for twist mappings (in any dimension) and models in $1-D$ statistical mechanics respectively) says that (provided some non-degeneracy conditions that can be readily checked) a parameter value λ_0 is on the boundary of the parameters with KAM tori of a fixed frequency if and only if, when we consider a Sobolev norm of high enough order of the KAM tori for nearby parameters λ, it blows up as λ approaches λ_0.

We anticipate that the justification is based on several KAM theorems: 1), an *"a-posteriori"* KAM theorem for Sobolev regularity with some special features 2) a bootstrap of regularity theorem that shows that Sobolev solutions of a high enough index are analytic. The choice of Sobolev spaces is not dictated by the theory (C^α spaces could work just as well from the theory point of view) but from practical considerations depending on the numerical implementation. See the observations after Definition 4.1.

Of course, for each different context, the proof of the KAM theorems 1), 2) above that justify the method is different and requires different arguments.

In this paper we present proofs of the KAM theorems that justify rigorously the criterion in two different contexts: Symplectic Twist maps (See Section 5) and Equilibrium models in Statistical Mechanics of 1-D systems. (See Section 6.). What we mean by twist mappings is that the derivative of the frequency map is non-degenerate in some neighborhood of the torus. Some times, "twist" is used to mean that the derivative of the frequency map is positive definite. We use a KAM iteration scheme and, therefore, non-degeneracy is enough.

We point out that these two different contexts, with very different physical interpretation, have a significant overlap (e.g. Frenkel-Kontorova models, which are equivalent to twist mappings). Nevertheless, there are models which are in one class but not in the other (See the discussion in [dlL08]). The same time of theorems have been developed in yet another context (dissipative systems) in [CCdlL09]. Of course, different contexts require very different proofs.

It should be noted that the proofs of the KAM theorems presented here, also lead to efficient algorithms (see Algorithms 6.4 and 5.5) to compute the quasi-periodic solutions if one chooses appropriate discretizations and efficient algorithms for the sub-steps.

These algorithms have been implemented in the contexts described here in [CdlL09b, CdlL09a]. In the context of dissipative systems there are implementations in [CC09] and a full justification in [CCdlL09]. For whiskered tori in Hamiltonian systems, we refer to [FdlLS09, HdlLS09]. We expect that the criterion for breakdown presented here (and the development of efficient algorithms) can be adapted to other contexts.

We also note that, if all the error of the numerics (truncation, round off) could be estimated and shown to be small enough with respect to the non-degeneracy assumptions, the KAM theorems establish rigorously that the tori exist. This is the basis of many *computer-assisted* proofs.

To make it clear that the method is very general, in Appendix A, we present an abstract Nash-Moser theorem which can be used to prove both theorems 1), 2). We hope that this will help show the deep unity and the generality of the methods.

In Appendix B, we provide a comparison between all the methods that we know of to compute invariant circles for twist mappings. Even if this survey is, necessarily incomplete, from the publication of [CdlL09b] we received many requests for such a comparison.

2. Criterion for the breakdown

The criterion for the breakdown of quasi-periodic solutions we propose is summarized in the following algorithm.

Given a family of problems indexed by some parameters, we try to locate the boundary of the set of parameters for which there is an analytic quasi-periodic solution.

We note that the following algorithm is a continuation algorithm which relies on an iterative step (e.g. a Newton method) that finds the solution once one starts close to a solution. We stop when the solutions have large Sobolev norm.

Algorithm 2.1.
Choose a path in the parameter space starting in the integrable case.
Initialize
The parameters and the solution at the integrable case
Repeat

Increase the parameters along the path
Run the iterative step
***If** (Iterations of the Newton step do not converge)*
Decrease the increment in parameters
***Else** (Iteration success)*
Record the values of the parameters
and the Sobolev norm of the solution.
***If** Non-degeneracy conditions fail*
Return "inconclusive"
***Until** Sobolev norm exceeds a threshold*

The Meta-Theorem that guarantees the correctness of the method is:

Meta-Theorem 2.2. *Let f_λ be a family of analytic problems, satisfying appropriate hypothesis.*

Assume that f_{λ_0} has a Sobolev regular quasi-periodic solution u_{λ_0} that satisfies some non-degeneracy assumptions.

Then if $|\lambda - \lambda_0|$ is small enough depending on the size of the Sobolev norm of u_{λ_0}, then f_λ has an analytic solution, which is locally unique.

Of course, depending on the exact context of the problem, one has to prove actual theorems that implement the Meta-Theorem 2.2. In this paper, we will present Theorems that implement it for twist maps and for some models of Statistical mechanics.

Given the Meta-Theorem 2.2, it is clear that, the algorithm 2.1 will continue progressing till it gets as close to breakdown as allowed by the resources of the computer.

2.1. **Some comments on the Algorithm 2.1.** • It is important that one identifies the exact conditions for the validity of algorithms. As any numerical algorithm, Algorithm 2.1, returns inconclusive results if the non-degeneracy conditions fail. In the case of twist mappings, a KAM torus may stop satisfying the twist condition as we continue over parameters.

• On the other hand, if the algorithm has progressed and produced a torus with a small error, reasonably small norm and which satisfies the non-degeneracy conditions, the a-posteriori Theorem 3.1 will prove that such an invariant torus exists, so that if one implements the algorithm bounding the error in the computation, one obtains rigorous existence of tori.

• Of course, the rigorous theorem 2.2 only concludes the fact that the KAM torus cannot be continued from the fact that a norm tends to infinity. This is not verified numerically. One could solipsistically argue that we are not excluding that the norm stops growing when the parameter is closer to the limiting value than the values computed. Of course, similar arguments can be done for almost any computer algorithm. As we will see, the best that the present machinery can do is to provide a series of rigorous lower bounds that are guaranteed to converge to the correct value. In the case of twist mappings with one degree of freedom, in Appendix B, we will discuss algorithms that provide rigorous upper bounds for the breakdown which also converge to the right value.

• In practice, one can make the results more convincing by observing that the norms blow up according to a power law Renormalization group predicts that there is a power law blow up for each Sobolev norm and that there is a simple relation between the scaling exponents corresponding to Sobolev norm, [dlL92, CdlL09b]. If the scaling relations are found, and the relation between the power exponents is as predicted, then one can be rather confident that the value of breakdown is obtained by extrapolating, and that indeed there is a renormalization description.

• Algorithm 2.1 is a continuation algorithm and shares all the shortcomings common to continuation algorithms. One can compute the connected component of the set of parameters for which there is an analytic quasi-periodic solution.

Indeed, [CdlL09a] presents numerical evidence that the twist maps

$$T_\varepsilon(p,q) = (p+\varepsilon[\alpha\cos(2\pi x)+\beta\cos(4\pi x)], q+p+\varepsilon[\alpha\cos(2\pi x)+\beta\cos(4\pi x)]) \tag{1}$$

the set of parameters with KAM tori has several components when $\alpha = 1, \beta \gtrapprox \frac{\sqrt{2}-1}{4}$.

Of course, this non-connected regions were computed using appropriate paths in the two parameter family (1). We think that it is reasonable to conjecture that the set of analytic twist maps with an analytic KAM torus of frequency ω, with ω Diophantine, is connected.

2.2. **Verifying Theorem 2.2.**

The verification of Theorem 2.2 can be reduced to two theorems.

(1) Showing that given a Sobolev solution, then all nearby maps have a locally unique Sobolev solution.
(2) All maps with a Sobolev solution have an analytic solution.

Note that, in contrast with many formulations of KAM theory, we are not allowed to consider only maps which are close to integrable.

As it turns out, the key step to both results is a theorem that shows that given an approximate solution (either in analytic sense or in a Sobolev sense) that satisfies appropriate non-degeneracy conditions, then there is a locally unique solution in the same spaces. These theorems called *"a-posteriori"* will be discussed in the next Section 3.

The bootstrap result is obtained by observing that that, if the Sobolev regularity of a solution is high enough, then, a truncation will be an approximate solution in the analytic sense and, hence, applying the a-posteriori theorem, we can conclude that there is an analytic solution which is close to the truncated one. Again, if the Sobolev regularity is high enough, the local uniqueness result in Sobolev spaces shows that the original solution and the analytic one agree.

For more details on this argument, see Theorem 5.8 and see [GEdlL08] for complete implementation of these ideas in twist mappings and [dlL08] for implementations for models in Statistical Mechanics. In the references above, one can find the details for C^r regularity. Because of the Sobolev embedding theorem, the bootstrap of C^r regularity to analytic implies the bootstrap of Sobolev to analytic. We point out, however, that the Sobolev argument is easier.

3. A posteriori KAM estimates

The rigorous justification of the criterion is based on an *a posteriori* formulation of a KAM theorem and continuity properties of a functional equation.

The theorem with its proof can be interpreted as an implicit function theorem, [Zeh75, Zeh76a, Bos86]. It is possible to find a finite set of explicit conditions [Kol54, Nas56] (non-degeneracy conditions) that guarantee that a *Newton method* started on a sufficiently approximate solution will converge to a true solution.

In numerical analysis theorems of this type are often called *a posteriori estimates*. The prototype of such a Meta-Theorem [dlLR91] is

Meta-Theorem 3.1. *Let $\mathcal{X}_0 \subset \mathcal{X}_1$ be Banach spaces and $\mathcal{U} \subset \mathcal{X}_0$ and open set. For certain maps*

$$\mathcal{F} : \mathcal{U} \subset \mathcal{X}_0 \to \mathcal{X}_0,$$

there exists an explicit function $\varepsilon^* : \mathbb{R}^+ \times (\mathbb{R}^+)^n \to \mathbb{R}^+$ such that

$$\lim_{t\to\infty} \varepsilon^*(t, \cdot) = 0,$$

explicit functionals $f_1, ..., f_n : \mathcal{X}_0 \to \mathbb{R}^+$, and $M_0, M_1, ..., M_n \in \mathbb{R}^+$ satisfying the following property. Suppose $x_0 \in \mathcal{X}_0$ with $\|x_0\|_{\mathcal{X}_0} \le M_0$, *and that*

$$f_1(x_0) \le M_1, ..., f_n(x_0) \le M_n,$$
$$\|\mathcal{F}(x_0)\|_{\mathcal{X}_0} < \varepsilon^*(M_0, ..., M_n).$$

Then there exists an $x^* \in \mathcal{X}_1$ *such that* $\mathcal{F}(x^*) = 0$ *and*

$$\|x_0 - x^*\|_{\mathcal{X}_1} \le C_{M_0,...,M_n} \|\mathcal{F}(x_0)\|_{\mathcal{X}_0}$$

The Meta-Theorem 3.1 states that we can conclude the existence of a true solution if we find an $x_0 \in \mathcal{X}_0$ that solves the functional equation very approximately as long as a finite number of explicit non degeneracy conditions are satisfied. In the applications considered here, we work in subsets of $\mathcal{X}_0$ where the functionals $f_1, ..., f_n$ are bounded. Therefore, the true solution can only break down when the norm of the approximate solution blows up.

Of course, in the actual applications, we have to prove actual theorems that implement Meta-Theorem 3.1 and, in particular, give the explicit expressions of the non-degeneracy conditions.

Remark 3.2. For our applications it is important that the Newton method starts on an arbitrary approximate solution and not necessarily form an integrable system. Some versions of Nash–Moser implicit function theorems, notably those of [Zeh75, Sch60], have this feature, but they require a special structure in the equations or invertibility of the derivative of the functional in a neighborhood about the initial guess.

Remark 3.3. Other examples of *a posteriori* KAM theorems have appeared in the literature. For example [FdlLS09] considers whiskered tori, [CC08] considers dissipative systems in analytic regularity.

In the subsequent sections we present two complete implementations of the above ideas. We present two different contexts both of which lead to criteria for breakdown.

In section 5, we present a KAM theorem for symplectic mappings based in *automatic reducibility.* A version for the analytic spaces was presented in [dlLGJV05], but now we present a result for Sobolev spaces, which is well suited for numerical implementations. Of course, the numerical implementations involve another layer of complications that we will not discuss here. An implementation needs to describe the discretizations and the algorithms to accomplish several tasks. The efficiency and accuracy depend on these choices that we will not consider here.

In section 6, we present another KAM theorem based on automatic reducibility for models in Statistical Mechanics. We present a version based on Sobolev spaces which is well suited for our criterion. The analytic version was presented in [dlL08].

Both theorems will be proved by applying an abstract Nash-Moser implicit function theorem which we present in Appendix A so that it can be read independently. Theorem A.6 combines features of the abstract theorems in [Zeh75, Sch69]. Of course, there is a large literature of implicit function theorems, which are useful in many contexts (for example, [Ser73, Ham82, Hör85]) but for us the feature that the iteration could start on any approximate solutions was very important.

For simplicity, in this paper we will present results for maps only, but there are analogous results for flows. [Dou82, dlLGJV05] show how to deduce results for flows from results from maps.

4. Some standard definitions and results

In this section, we include some of the definitions that we will require for stating our breakdown criterion.

In the abstract discussion of an *a posteriori* theorem in Section 3 and in the rigorous version in Appendix A, we have not specified which norm to use. This allows extra flexibility. There are several norms which one can use in applications. In mathematical applications, one often encounters norms in analytic spaces or C^r norms. For numerical applications we have found convenient to use Sobolev norms rather than C^r norms, see the remarks after Definition 4.1.

4.1. **Sobolev Spaces.**

We denote the Fourier expansion of a periodic mapping $u : \mathbb{T}^n \to \mathbb{R}^d$ by

$$u(\theta) = \sum_{k \in \mathbb{Z}^n} \hat{u}_k \, \exp(2\pi i \, k \cdot \theta),$$

where $\cdot$ is the Euclidean scalar product in $\mathbb{R}^n$, and the Fourier coefficients $\hat{u}_k$ can be computed

$$\hat{u}_k := \int_{\mathbb{T}^n} u(\theta) \, \exp(-2\pi i \, k \cdot \theta) \, d\theta \, .$$

The average of u is the 0-Fourier coefficient, we denote it by

$$\operatorname{avg} \{u\}_\theta := \int_{\mathbb{T}^n} u(\theta) \, d\theta = \hat{u}_0 \, .$$

Definition 4.1. *For $m \in \mathbb{R}^+$, the Sobolev Space H^m is the Banach space of functions from $\mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n$ taking values in $\mathbb{R}^d$ defined by*

$$H^m = \{ u \, | \, \|u\|_m^2 = \sum_{k \in \mathbb{Z}^n} (1 + |k|^2)^m |\hat{u}_k|^2 < \infty \}$$

Here $|\cdot|$ *represents the maximum norm on the spaces* $\mathbb{R}^d$ *and* $\mathbb{C}^d$*, i.e. if* $x = (x_1, \ldots, x_d) \in \mathbb{C}^d$*, then*

$$|x| := \max_{j=1,\ldots,m} |x_j| \,.$$

Similar notation for the norm will be also used for real or complex matrices of arbitrary dimension, and it will refer to the matrix norm induced by the vectorial one.

We note that for $m \in \mathbb{Z}^+$ *we have*

$$C_1(\|D^m u\|^2 + |avg\,\{u\}_\theta\,|^2) \le \|u\|_m^2 \le C_2(\|D^m u\|2 + |avg\,\{u\}_\theta\,|^2)$$

so that the norms are equivalent.

There are several advantages of using Sobolev norms for our applications

- In our computations, which involve handling of Fourier series, the computation of Sobolev norms is extremely fast and reliable. (the computation of C^r norms seems more involved and is more unreliable because it is affected by errors at one point. Sobolev norms are more immune to local errors since they are weighted sums of averages) [dlLP02].
- Analytic norms involve the choice of a domain. It is known that near the breakdown the analyticity domain shrinks so that one gets different values of the breakdown depending on the domain chosen to study (indeed the analyticity breakdown is defined as the value for which there is no domain of analyticity)
- Sobolev norms transform well under rescaling transformations. It is easy to show that

  $$\|D^r u \circ \lambda\| = \|D^r u\|_{L^2} \lambda^{r-1/2}$$

  These scaling properties of Sobolev norms are very useful to study breakdown [CdlL09b] since it is known that for the *universality class* the analyticity breakdown satisfies some scaling relations [Mac83]. The scaling relations of Sobolev norms allow us to identify more accurately the breakdown point for families in the same universality class. Of course verifying the scaling relations for Sobolev norms gives indications on the validity of the renormalization group.
- It has been shown in [GEdlL08, dlL08] that solutions of the invariance equations for twist mappings and for models in Statistical Mechanics which are Sobolev of a sufficiently high order are analytic. The proof is an easy consequence of the *a-posteriori* format and local uniqueness. See Theorems 5.8 and 6.8 for a

precise formulation. An overview of the argument is given after Theorem 5.8.

4.1.1. *Some properties of Sobolev spaces.* In the rest of the section we collect some standard material to set the notation and is just meant to be used as reference.

The Sobolev spaces we have introduced are a Banach Algebra for large enough m. The following result is proven in [Ada75] as a straightforward application of the Sobolev Imbedding Theorem.

Lemma 4.2. *Let $m > \frac{n}{2}$. There is a constant K_1 depending on m and n, such that for $u, v \in H^m$ the product $u \cdot v$, belongs to H^m and satisfies*

$$\|u \cdot v\|_m \leq K_1 \|u\|_m \|v\|_m \tag{2}$$

In particular, H^m is a commutative Banach algebra with respect to the point-wise multiplication and the norm

$$\|u\|_m^* = K_1 \|u\|_m$$

An elementary consequence of H^m being a Banach algebra under multiplication when $m > n/2$ is that if M is a matrix valued function $M, M^{-1} \in H^m$ and $\|M - \tilde{M}\|_m$ is sufficiently small, then $\tilde{M}^{-1} \in H^m$. The proof is just using the Neumann series

$$\tilde{M}^{-1} = \left[\sum_{n=0}^{\infty} M^{-1}(M - \tilde{M})^n \right] M^{-1} \tag{3}$$

and the Banach algebra properties.

It is also useful to have the following estimates on compositions (see for example [Tay97, Section 13.3]).

Lemma 4.3. *Let $f \in C^m$ and assume $f(0) = 0$. Then, for $u \in H^m \cap L^\infty$*

$$\|f(u)\|_m \leq K_2(\|u\|_{L^\infty})(1 + \|u\|_m) \tag{4}$$

where

$$K_2(\lambda) = \sup_{|x| \leq \lambda, \mu \leq m} |D^\mu f(x)|$$

In the case that $m > n/2$, if $f \in C^{m+2}$, we have that

$$\begin{aligned} ||f \circ (u+v) - f \circ (u) - & Df \circ (u) v||_m \\ & \leq C_{n,m}(\|u\|_{L^\infty})(1 + \|u\|_m) ||f||_{C^{m+2}} ||v||_m^2 \end{aligned} \tag{5}$$

Notice that by the Sobolev Imbedding Theorem we also have that

$$H^m \subset C^0 \cap L^\infty, \quad \text{for } m > \frac{n}{2} \tag{6}$$

so if $f \in \mathcal{C}^m$ and $u \in H^m$ with $m > \frac{n}{2}$ hypotheses that $f \in H^m \cup L^\infty$ of Theorem 4.3 reduce to $f \in H^m$.

4.2. **Number Theory.** A standard definition in KAM theory is the Diophantine condition.

Definition 4.4. *We say that $\omega \in \mathbb{R}^n$ is Diophantine if for some $\nu > 0$ and $\tau > n$ we have that*

$$|p \cdot \omega - q| \geq \nu |q|_1^{-\tau} \qquad q \in \mathbb{Z}^n - \{0\}, \quad \forall p \in \mathbb{Z} \tag{7}$$

where $|l|_1 = |l_1| + |l_n| + ... + |l_n|$. We define $D(\nu, \tau)$ as the set of all frequency vectors satisfying (7)*.*

4.3. **Cohomology equations.** It is also standard in KAM theory to solve for φ given ξ in the equation

$$\varphi(\theta \pm \omega) - \varphi(\theta) = \xi(\theta) \tag{8}$$

where $\omega \in D(\nu, \tau)$ and ξ a function of zero average.

Estimates for (8) in Sobolev spaces are given by the following lemma, which is straight forward compared to the version in analytic or C^m spaces.

Lemma 4.5. *Let $\omega \in D(\nu, \tau)$. Given any function $\xi \in H^{m+\tau}$ satisfying $avg\{\xi\}_\theta = 0$ there is one and only one function $\varphi \in H^m$ satisfying*

$$\varphi(\theta \pm \omega) - \varphi(\theta) = \xi, \quad avg\{\varphi\}_\theta = 0 \tag{9}$$

Moreover,

$$\|\varphi\|_m \leq C\nu^{-1} \|\xi\|_{m+\tau} \tag{10}$$

Proof. ¿From (9) the Fourier coefficients of φ and ξ satisfy

$$\hat{\varphi}_k = \frac{\hat{\xi}_k}{e^{\pm 2\pi i k \cdot \omega} - 1}, \quad k \neq 0$$

So

$$\begin{aligned} \|\varphi\|_m^2 &= \sum_{k \in \mathbb{Z}} (1 + |k|^2)^m |\hat{\varphi}_k|^2 = \sum_{k \in \mathbb{Z}} (1 + |k|^2)^m \frac{|\hat{\xi}_k|^2}{|e^{\pm 2\pi i k \cdot \omega} - 1|^2} \\ &\leq C \sum_{k \in \mathbb{Z}} (1 + |k|^2)^m \frac{|\hat{\xi}_k|^2}{(2\pi k \cdot \omega)^2} \leq C\nu^{-2} \sum_{k \in \mathbb{Z}} (1 + |k|^2)^{m+\tau} |\hat{\xi}_k|^2 \\ &= C\nu^{-2} \|\xi\|_{m+\tau}^2 \end{aligned} \tag{11}$$

5. The criterion for the breakdown in Symplectic Maps

In this section we introduce a justification of the criterion of breakdown of analyticity for Symplectic maps with a Diophantine rotation vector. Similar result hold for vector fields, see [dlLGJV05] for an argument that shows how to obtain results for flows from results for maps. Hence, we have the straight forward adaptation of results and proofs to the reader.

The criterion we present works also in the context of Variational Problems in Statistical Mechanics discussed in Section 6 for which no analogue of periodic orbits and stable manifolds seems to be available.

The proof is based in the constructive proof for the analytic case presented in [dlLGJV05]. The guiding principle of the proof is the observation that the geometry of the problem implies that KAM tori are *reducible* and approximate invariant tori are *approximately reducible*. This leads to a solution of the linearized equations without transformation theory. Here we will summarize the main ideas to construct and find the estimates for the quasi-Newton method.

5.1. **Formulation of the invariance equation.** The results for an exact symplectic map f of a $2n$-dimensional manifold $\mathbf{U}$ are based on the study of the equation

$$(f \circ K)(\theta) = K(\theta + \omega) \tag{12}$$

where $K : \mathbb{T}^n = (\mathbb{R}/\mathbb{Z})^n \to \mathbf{U}$ is the function to be determined and $\omega \in \mathbb{R}^n$ satisfies a Diophantine condition.

We will assume that $\mathbf{U}$ is either $\mathbb{T}^n \times U$ with $U \subset \mathbb{R}^n$ or $B \subset \mathbb{R}^{2n}$, so that we can use a system of coordinates. In the case that $\mathbf{U} = \mathbb{T}^n \times U$, we note that the embedding K could be non-trivial.

Let $\Omega = d\alpha$ be an exact symplectic structure on U and let $a : U \to \mathbb{R}^{2n}$ be defined by

$$\alpha_z = a(z)dz \quad \forall z \in U \tag{13}$$

For each $z \in U$ let $J(z) : T_zU \to T_zU$ be a linear isomorphism satisfying

$$\Omega_z(\xi, \eta) = <\xi, J(z)\eta> \tag{14}$$

where $< \cdot, \cdot >$ is the Euclidean scalar product on $\mathbb{R}^{2n}$. Since Ω is antisymmetric, J satisfies $J(z)^T = -J(z)$.

Notice that (12) implies that the range of K is invariant under f. The map K gives a parameterization of the invariant torus which makes the dynamics of f restricted to the torus into a rigid rotation.

We will consider the set of functions $K : \mathbb{T}^n = (\mathbb{R}/\mathbb{Z})^n \to \mathbf{U}$ satisfying

$$K(\theta + k) = K(\theta) + (k, 0) \qquad k \in \mathbb{Z}^n. \tag{15}$$

Notice that it is equivalent to say that K satisfies (15) than to say that $K(\theta) - (\theta, 0)$ is periodic. Hence, we can consider $\tilde{H}^r$ as an affine space modeled on H^r.

Definition 5.1. *Given a symplectic map f and $\omega \in D(\nu, \tau)$. A mapping $K \in H^m$ is said to be non-degenerate if it satisfies the following conditions*

N1: *There exists an $n \times n$ matrix-valued function $N(\theta)$, such that*

$$N(\theta)\ \left(DK(\theta)^\top DK(\theta)\right) = I_n, \tag{16}$$

where I_n is the n-dimensional identity matrix.

N2: *The average of the matrix-valued function*

$$S(\theta) := P(\theta+\omega)^\top \left[\, Df(K(\theta)) J(K(\theta))^{-1} P(\theta)\ - J(K(\theta+\omega))^{-1} P(\theta+\omega) \,\right] \tag{17}$$

with

$$P(\theta) := DK(\theta)\, N(\theta)\,,$$

is non-singular.

N3: *The $2n \times 2n$ matrix $M(\theta)$ obtained by juxtaposing the two $2n \times n$ matrices $Dk(\theta)$ and $J(K(\theta))^{-1} DK(\theta) N(\theta)$ as follows*

$$M(\theta) := \begin{pmatrix} DK(\theta) & J(K(\theta))^{-1} DK(\theta) N(\theta) \end{pmatrix}, \tag{18}$$

is invertible in H^{m-1} and that

$$\|M^{-1}\|_{m-1} < \infty$$

for some $m > n/2 + 1$.

We will denote the set of functions in $\mathcal{P}_m$ satisfying conditions **N1-N3** *by $\mathcal{ND}(m)$.*

Remark 5.2. By the Rank Theorem, Condition **N1** guarantees that $\dim K(\mathbb{T}^n) = n$. For the KAM theorem, the main non-degeneracy condition is **N2**, which is a twist condition. Its role will become clear in 5.3. Note that **N1** depends only on K whereas **N2** depends on K and f.

Note also that by the observation on series (3), the Condition **N3** is an open condition in H^m. As we will see condition **N3** will be implied for functions K which satisfy the invariance equation with good accuracy.

5.2. **Statement of an *a posteriori* theorem for symplectic maps.**

Theorem 5.3. *Let* $m > \frac{n}{2} + 2\tau + 1$ *and* $f \in C^{m+34\tau-17}$, $f : \mathbf{U} \to \mathbf{U}$ *be an exact symplectic map, and* $\omega \in D(\nu, \tau)$, *for some* $\nu > 0$ *and* $\tau > n$. *Assume that the following hypotheses hold*

H1. $K_0 \in H^{m+34\tau-17}$ *(i.e.* K_0 *satisfies 5.1).*

H2. *The map* $f \in C^{m+34\tau-17}$ *in* $\mathcal{B}_r$, *a neighborhood of radius* r *of the image under* K_0 *of* $\mathbb{T}^n$ *for some* $r > 0$.

H3. *If* a *and* J *are as in* (13) *and* (14), *respectively,*

$$\|a\|_{C^{r+3}(\mathcal{B}_r)} < \infty\,, \qquad \|J\|_{C^{r+3}(\mathcal{B}_r)} < \infty\,, \qquad \|J^{-1}\|_{C^{r+3}(\mathcal{B}_r)} < \infty\,.$$

Define the error function e_0 *by*

$$e_0 := f \circ K_0 - K_0 \circ T_\omega.$$

where $T_\omega(\theta) = \theta + \omega$.

There exists a constant $c > 0$ *depending on* σ, n, r, ρ_0, $|f|_{C^2,\mathcal{B}_r}$, $\|a\|_{C^{r+3}}$, $\|J\|_{C^{r+3}}$, $\|J^{-1}\|_{C^{r+3}}$, $\|DK_0\|_{m-2\tau+1}$, $\|N_0\|_{m-2\tau+1}$, $\left|(avg\{S_0\}_\theta)^{-1}\right|$ *(where* N_0 *and* S_0 *are as in 5.1, replacing* K *by* K_0*) such that, if* $\|e_0\|_{m-2\tau}$ *verifies the following inequalities*

$$c\,\nu^{-4}\|e_0\|_{m-2\tau+1} < 1, \tag{19}$$

and

$$c\,\nu^{-2}\|e_0\|_{m-2\tau+1} < r\,,$$

then there exists $K^* \in H^m$ *such that*

$$f \circ K^* - K^* \circ T_\omega = 0\,. \tag{20}$$

Moreover,

$$\|K^* - K_0\|_m \le c\,\nu^{-2}\|e_0\|_{m-2\tau+1}.$$

The function K^* *is the only function – up to translation – in a neighborhood of* K_0 *of size* $c\nu^{-2}$, *in* $H^{m+4\tau-2}$ *satisfying 20.*

Remark 5.4. In [GEdlL08], it is shown that if f is analytic, $K \in H^m$, for $m \ge m_0$ and satisfies the invariance equation and Definition 5.1 then K is analytic. Here, m_0 depends on τ and n.

5.3. **Quasi-Newton method for symplectic maps.** Here we describe the procedure to improve approximate solutions in the case of symplectic maps. We use the methods developed in [dlLGJV05]. This leads both to a practical algorithm (see Algorithm 5.5) and a theorem stating the convergence of the procedure when started in a sufficiently approximate solution (see Theorem 5.3).

Since equation (12) can be formulated as finding zeros of

$$F(K) := f \circ K - K \circ T_\omega\,, \tag{21}$$

then the improvement $K+\Delta$ of an approximate solution K is obtained by solving the linearized equation for Δ

$$DF(K)\Delta(\theta) = Df(K(\theta))\Delta(\theta) - \Delta(\theta+\omega) = -e(\theta) \tag{22}$$

with error function

$$e(\theta) = F(K)(\theta) \quad \forall\theta \in \mathbb{R}^n \tag{23}$$

In [dlLGJV05], the authors introduce a change of variables given by the $2n \times 2n$ matrix $M(\theta)$ constructed by juxtaposing the $2n \times n$ matrices $DK(\theta)$ and $J(K(\theta))^{-1}DK(\theta)N(\theta)$ as in (18). that transforms the derivative Df into

$$M(\theta+\omega)^{-1}Df(K(\theta))M(\theta) = C(\theta) + B(\theta)$$

with

$$C(\theta) := \begin{pmatrix} I_n & S(\theta) \\ 0 & I_n \end{pmatrix}, \tag{24}$$

and

$$\|B\|_{m-2\tau-1} \leq C\nu^{-2}\|M^{-1}\|_{m-1}\|e\|_m.$$

Since B is linear in the error e, then we can solve the modified equation

$$\begin{pmatrix} I_n & S(\theta) \\ 0 & I_n \end{pmatrix} M^{-1}(\theta)\Delta(\theta) - M^{-1}(\theta+\omega)\Delta(\theta+\omega) = -M^{-1}(\theta+\omega)e(\theta). \tag{25}$$

Notice that equation (25) differs from the Newton step equation by a term $BM^{-1}\Delta$ which is quadratic in e (we will show latter the Δ is also estimates by $\|e\|$) so the modified Newton Method (25) gives rise to a quadratically convergent scheme.

Moreover, equation (25) can be solved in two steps. We are therefore led to the following algorithm.

Algorithm 5.5. The iterative step is constructed as follows:

1) Compute

$$e(\theta) = f \circ K - K \circ T_\omega,$$
$$N(\theta) = (DK(\theta)^T DK(\theta))^{-1},$$
$$M(\theta) = (DK(\theta) \quad J^{-1}(K(\theta))DK(\theta)N(\theta)),$$
$$M^{-1}(\theta),$$

and

$$E(\theta) = M^{-1}(\theta)e(\theta).$$

2) Find a normalized function W_2 (i.e. $\operatorname{avg}\{W_2\}_\theta = 0$) solving the equation
$$W_2(\theta) - W_2(\theta+\omega) = E_2(\theta)$$
We can choose $\mathcal{T} \in \mathbb{R}^n$ such that
$$\operatorname{avg}\{E_1\}_\theta = \operatorname{avg}\{S(\cdot)(E_2(\cdot)+\mathcal{T})\}_\theta$$
3) We solve for W_2 from
$$W_1(\theta) - W_1(\theta+\omega) = E_1(\theta) - S(\theta)(E_1(\theta)+\mathcal{T})$$
and set $\operatorname{avg}\{W_1\}_\theta = 0$.
4) Set $\Delta(\theta) = M(\theta)W(\theta)$ and
$$\tilde{K}(\theta) = K(\theta) + \Delta(\theta)$$
$\tilde{K}$ is the improved solution.

We call attention that all the steps are diagonal either in Fourier space of in Real space. The fact that the steps are diagonal and that one can switch from Real space to Fourier using efficient Fast Fourier Transforms allows to have fast numerical implementations. Notice also the that the Newton step does not need to store full matrices.

5.4. **Estimates for the Quasi-Newton Method.** In this section we provide estimates for the iterative step described in Algorithm 5.5. The form of the estimates we will prove is typical of the Nash-Moser strategy. We will show that the new error will be bounded (in a less smooth norm) by the square of the original error.

Actually, we will follow the formulation of the abstract implicit function theorem, Theorem A.6, and we will describe the Algorithm 5.5 by a linear operator η that produces the correction Δ out of the true error e. That is

$$\Delta = \eta[K]e \tag{26}$$

According to the strategy of Theorem A.6 we will check that

1) The operator η can be defined for all K in a ball.
2) We will provide estimates for η.
3) We will show that η is an approximate left inverse for the derivative of the functional.

As is discussed in [dlLGJV05], an approximate solution K of (12) with error e defined in (23). Define S, M, and C by (17), (18), and (24), respectively, and let us define $\mathcal{E}$ as

$$\mathcal{E}(\theta) := Df(K(\theta))M(\theta) - M(\theta+\omega)\,C(\theta). \tag{27}$$

and we notice that

$$B = M^{-1} \circ T_\omega \mathcal{E}$$

so

$$\|B\|_{m-2\tau-1} \leq C\nu^{-2}\|M^{-1}\|_{m-1}\|e\|_m$$

Lemma 5.6. *Let $m > \frac{n}{2} + 2\tau + 1$, $F[K] \in H^m$ and $\eta : H^m \to H^{m-2\tau}$ the operator constructed in Algorithm* (5.5)*.*

Then

$$\|\eta[K]F[K]\|_{m-2\tau-1} \leq C\nu^{-2}\|M\|_{m-1}\|M^{-1}\|_{m-2\tau-1}\|F[K]\|_{m-1}$$

We will also need estimates on $DF[K]\eta[K]$. This estimates establish that $\eta[K]$ is an approximate left inverse of $DF[K]$ as we show in the following lemma.

Lemma 5.7. *Let $m > \frac{n}{2}+2\tau+1$, $F[K] \in H^m$, and $F[K]$, $\eta[K]$ defined in 26.*

Then we have the estimates

(28)
$$\begin{aligned}&\|(DF[K]\eta[K] - \mathrm{Id})F[K]\|_{m-2\tau-1}\\ &\qquad\leq C\nu^{-2}\|M\|^2_{m-1}\|M^{-1}\|_{m-2\tau-1}\|F[K]\|_{m-1}\|F[K]\|_m\end{aligned}$$

Proof. We have

(29)
$$\begin{aligned}&\|(DF[K]\eta[K] - \mathrm{Id})F[K]\|_{m-2\tau-1}\\
&\leq \|DF\eta[k]F[K] + M \circ T_\omega B M^{-1}\eta[K]F[K]\\
&- M \circ T_\omega B M^{-1}\eta[K]F[K] - F[K]\|_{m-2\tau-1}\\
&\leq \|M \circ T_\omega[(C+B)M^{-1}\eta[k]F[K] - (M^{-1}\eta[K]F[K]) \circ T_\omega]\\
&- M \circ T_\omega B M^{-1}\eta[K]F[K] - F[K]\|_{m-2\tau-1}\\
&\leq \|M \circ T_\omega B M^{-1}(\eta[K]F[K]) \circ T_\omega\|_{m-2\tau-1}\\
&\leq C\nu^{-3}\|M\|^2_{m-1}\|M^{-1}\|_{m-2\tau-1}\|F[K]\|_{m-1}\|F[K]\|_m\end{aligned}$$

which completes the estimates for the approximate inverse.

The final result follows from an application of Theorem A.6. In the context of Theorem A.6 we consider the previous estimates with $\alpha = 2\tau - 1$ and the result from of Theorem 4.

5.5. **Bootstrap of regularity.** In this section we state the bootstrap of regularity theorem. We will provide a sketch of the proof. A full proof for the case of C^m spaces is given in [GEdlL08].

Theorem 5.8. *Assume that the hypotheses of Theorem 5.3 hold and that the map f is a real analytic map. Consider the map $K \in H^m$ with $m > \frac{n}{2} + 2\tau + 1$ solving*

$$F(K) = f \circ K - K \circ T_\omega = 0.$$

Then $K \in \mathcal{A}_{\frac{\rho_N}{4}}$ where $\rho_N > 0$ is a explicit function of $N \in \mathbb{N}$.

Proof. (Sketch) We start with $K \in H^m$, $m > \frac{n}{2} + 2\tau + 1$, a solution of the invariance equation obtained in Theorem 5.3, i.e.,$F(K) = 0$.

Then, we consider an approximation to the solution K obtained by truncating the Fourier series at the N-th Fourier mode.

$$K^{\leq N}(\theta) = \sum_{n=-N}^{N} K_n e^{2\pi i k\theta}.$$

a) We have the following estimates in the C^0 and C^1 norm for every N sufficiently large

$$\|K - K^{\leq}\|_{C^0} \leq CN^{-(m-\frac{1}{2}-\frac{n}{4})} \quad \text{and } \|K - K^{\leq}\|_{C^1} \leq CN^{-(m-1-\frac{n}{4})}. \tag{30}$$

Then for N large enough we can guarantee that the function $K^{\leq N}$ satisfies the non degeneracy conditions (16), (17), (18) in $\mathcal{A}_{\rho_N}$ as long as K satisfies these non degeneracy conditions in H^m.

b) to get estimate in the $\mathcal{A}_{\rho_N}$ norm, for some $\rho_N > 0$, we will consider an $\alpha > 1$ so that if $\rho_N = \frac{1}{N}\frac{\log \alpha}{2\pi}$

$$\|K^{\leq N}\|_{\mathcal{A}_{\rho_N}} \leq \alpha \|K\|_r \tag{31}$$

c) Indeed, using the estimates in *a*) and *b*), we obtain estimates for the invariance equation in the space of analytic functions $\mathcal{A}_{\frac{\rho_N}{2}}$ as follows

$$\|F(K^{\leq N})\|_{C^0} \leq CN^{-(m-\frac{1}{2}-\frac{n}{4})} \quad \text{and } \|F(K^{\leq N})\|_{\mathcal{A}_{\rho_N}} \leq C,$$

then using the log-convexity of the supremum of the analytic functions (Hadamard three circle theorem), we can interpolate the previous inequalities and obtain estimates in $\mathcal{A}_{\frac{\rho_N}{2}}$,

$$\|F(K^{\leq N})\|_{\mathcal{A}_{\frac{\rho_N}{2}}} \leq CN^{-(\frac{m}{2}-\frac{1}{4}-\frac{n}{8})}. \tag{32}$$

c) We apply the Theorem 1 in [dlLGJV05] in the analytic case. To obtain the non-degeneracy conditions (16), (17), and (18) for $K^{\leq N}$, we use estimates in (30) and the same non-degeneracy conditions for K. We also have that if N is large enough then

$$C\nu^{-4}\rho_N^{-4\tau}\|K^{\leq N}\|_{\frac{\rho_N}{2}} < 1 \quad \text{and } C\nu^{-2}\rho_N^{-2\tau}\|K^{\leq N}\|_{\frac{\rho_N}{2}} < r$$

where $r > 0$ satisfies $\|f\|_{C^2,B_r} < \infty$, with $B_r = \left\{ z \in \mathbb{C}^{2n} : \sup_{|\mathrm{Im}\theta| < \frac{\rho_N}{2}} |z - K^{\leq N}(\theta)| < r \right\}$. Then there exists a $K^* \in \mathcal{A}_{\frac{\rho_N}{4}}$ such that $F(K^*) = 0$. Moreover

$$\|K^* - K^{\leq N}\|_m \leq C\nu^{-2} N^{-\frac{m}{2}+1/2+2\tau+\frac{n}{4}} \tag{33}$$

For $K^{\leq N} - K$ we chose $\delta, t \in \mathbb{R}^+$ as follows

$$\frac{\delta}{2} = \frac{m}{2} - \frac{n}{4} - \tau - \frac{1}{2}, \quad t = (\nu^{-2}\|K\|_r^{-1})^{\frac{1}{\delta}}$$

and we use the smoothing operators of Sobolev spaces (57).

$$\begin{aligned} \|K^{\leq} - K\|_m &\leq \|S_t(K^{\leq N} - K)\|_m + \|(Id - S_t)(K^{\leq} - K)\|_m \\ &\leq Ct^{\delta}\|K^{\leq N} - K\|_{m-\delta} \leq C\nu^{-2} N^{-\frac{m}{2}+1/2+2\tau+\frac{n}{4}} \end{aligned} \tag{34}$$

d) Finally, combining (33) and (34) we obtain

$$\|K^* - K\|_{H^r} \leq C\nu^{-2} N^{-\frac{r}{2}+1/2+2\tau}.$$

So for N large enough and by the uniqueness in H^m (see Theorem (A.9)), we get that $K^* = K$. In particular, we have that $K \in \mathcal{A}_{\frac{\rho_N}{4}}$. $\square$

6. The criterion of breakdown for models in statistical mechanics

In this section we present a formulation and a full mathematical justification of a criterion for the breakdown of analyticity in 1-D models coming from statistical mechanics. This includes as particular cases the breakdown of KAM tori for twist mappings. In section 6.1, we introduce the models considered. In section 6.5, we state the theorem that justifies the criterion for these models.

As we anticipated, the proof of this statement is based on an abstract Nash-Moser implicit function theorem (see Theorem A.6). The statement of Theorem 6.1 asserts that if the Sobolev norms of an approximate solution are small enough and it satisfies the functional equation very approximately then there is a true solution. Thus, the existence of a true solution is validated. The algorithm which is the basis of the Newton step (and which is a practical algorithm for numerical computation) is detailed in section 6.6. The estimates used for the step are in 6.7 and the convergence is established using Theorem A.6 from the Appendix A.

6.1. **Models considered.** We will consider one dimensional systems. At each integer, there is one site, whose state is described by one real variable. Hence, the configuration of the system is characterized by sequence of real values (equivalently a function $x : \mathbb{Z} \to \mathbb{R}$). Following [Rue99], the physical properties of a model are determined by an energy which is a formal sum of the energy of every group of particles (we allow multi-body interactions).

In this paper, we will be concerned with the existence of equilibrium states (see Definition 6.2) with density $1/\omega$.

We will assume that the interaction is invariant under translations. Hence, we will consider models whose formal energy is of the form:

$$\mathcal{S}(\{x_n\}) = \sum_{L\in\mathbb{N}} \sum_{k\in\mathbb{Z}} H_L(x_k, ..., x_{k+L}) \tag{35}$$

This sum is purely formal, but there are well defined ways of making sense of several quantities of interest. We will furthermore make the following assumptions in our models.

i) The following periodicity condition holds.

$$H_L(x_k, ..., x_{k+L}) = H_L(x_k + 1, \dots, x_{k+L} + 1) \tag{36}$$

The property (36) is a rather weak periodicity condition. It is implied by the stronger property

$$H_L(x_k, ..., x_{k+L}) = H_L(x_k + \ell_0, \dots, x_{k+L} + \ell_L) \tag{37}$$

for all $\ell_i \in \mathbb{Z}$. The latter property (37) is natural when the variables x_i are angles. For example, spin variables. The weaker property (36) has appeared in many situations. It is natural when considering twist maps of the annulus [MF94a] or monotone recurrences [Ang90].

ii) We will also require a decay condition for the criterion to hold. In section 6.7, we present the detailed description of the decay condition and in 2 we state its relevance for the persistence of quasi-periodic solutions.

6.2. **Equilibrium equations.** Equilibrium configurations, by definition are solutions of the Euler-Lagrange equations indicated formally as

$$\partial_{x_i} \mathcal{S}(\{x_n\}) = 0 \tag{38}$$

The physical meaning of the equilibrium equations is that the total force on each of the sites exerted from the other ones vanish.

The equilibrium states have a direct physical relevance (they are sometime called meta-stable states, instantons). In dynamical systems, when $\mathcal{S}$ has the physical interpretation of an action, equilibrium states correspond to orbits of a dynamical system.

For models of the form (35) the Euler-Lagrange equations are:

$$\sum_{L\in\mathbb{N}}\sum_{j=0}^{L}\partial_j H_L(x_{k-j},...,x_{k-j+L}) = 0 \quad \forall k\in\mathbb{Z} \tag{39}$$

We call attention that, in contrast with the sums defining $\mathcal{S}$ which are merely formal, the sums involved in equilibrium equations (39) are meant to converge.

A practical case of equilibria that has attracted a great deal of attention is ground states [Mat82, Ban89, MF94b] (also known as Class A minimizers) we note that under convexity assumptions, using Hilbert integrals all critical points given by a continuous hull functions are ground states [CdlL98]. The Frenkel-Kontorova and twist mappings satisfy these assumptions.

6.3. **Plane-like configurations and hull functions.** We are interested in equilibrium configurations $\{x_n\}$ that can be written as

$$x_n = h(n\omega) \tag{40}$$

for $\omega\in\mathbb{R}\setminus\mathbb{Q}$ and where h satisfies the periodicity condition

$$h(x+e) = h(x)+e \quad \forall e\in\mathbb{Z} \tag{41}$$

and is analytic.

The function h is often called the *hull function*. The periodicity condition (41) means that h can be considered as a map of the circle to itself. We will use the notation

$$h(\theta) = \theta + u(\theta).$$

Were u is a periodic function.

6.4. **Equilibrium equations in terms of hull functions.** For configurations of the form (40), the equilibrium equations become:

$$E[u](\theta) \equiv \sum_L \sum_{j=0}^{L} \partial_j H_L(\theta - j\omega + u(\theta - j\omega), \dots,$$
$$\theta + u(\theta), \dots, \theta + (L-j)\omega + u(\theta + (L-j)\omega))$$
$$= 0 \tag{42}$$

6.5. **Statement of an *a posteriori* theorem for statistical mechanics models.**

Theorem 6.1. *Let $m > \frac{1}{2} + 2\tau$ and $H_L \in C^{m+34\tau}$ be translation invariant interactions as in* (35) *satisfying the periodicity condition* (36)*. Let $\omega \in \mathbb{R}$. Let $h = \mathrm{Id} + u$, with $u \in H^{m+32\tau}$, avg$\{u\}_\theta = 0$ be a diffeomorphism of $\mathbb{T}$. Assume:*

H1) *ω is Diophantine, i.e., for some $\nu > 0$, $\tau > 0$*

$$|q\omega - p| \geq \nu |q|^{-\tau} \;\; \forall p, q \in \mathbb{Z} \tag{43}$$

H2) *The interactions $H_L \in C^{m+34\tau}$.*
Denote

$$M_L = K_m \|H_L\|_{C^{m+3}} \|(\mathrm{Id} + u')\|_m^2$$
$$a = \sum_{L \geq 2} M_L L^4$$

H3) *Assume that the inverses indicated below exist and that:*

$$\|(\partial_0 \partial_1 H_1)^{-1}(u(\theta), u(\theta + \omega))\|_m \leq T.$$
$$\left(avg\left\{(\partial_0 \partial_1 H_1)^{-1}(u(\theta), u(\theta + \omega))\right\}_\theta\right)^{-1} \leq U.$$

The following bounds measure the non-degeneracy of the problem.

a1) $\|(\mathrm{Id} + u')\|_m \leq N^+$.
a2) $\|(\mathrm{Id} + u')^{-1}\|_m \leq N^-$.
b) $\|E[u]\|_{m-2\tau} \leq \varepsilon$.

Assume furthermore that the above upper bounds satisfy the following relations:

i) *Let $T(1 + a) < A$, $UT(1 + a) < B$*
ii) *$\varepsilon \leq \varepsilon^*(N^-, N^+, \nu, \tau, a, T, A, B)$ where ε^* is a function which we will make explicit along the proof. The function ε^* makes quantitative the relation between the smallness conditions and the non-degeneracy conditions.*

Then, there exists a periodic function $u^* \in H^m$ *such that*

$$E[u^*] = 0 \tag{44}$$

Moreover

$$\|u - u^*\|_m \leq C\nu^{-2}(N^+)^2\varepsilon$$

The function u^* *is the only function in a neighborhood of* u *in* $H^{m+4\tau}$ *satisfying* (44) *and* $avg\{u^*\}_\theta = 0$.

Remark 6.2. In [dlL08], it is shown that if H_L are analytic and satisfy analogs of H2) and H3) and i) with analytic norms in place of C^{m+3} norms then, if $m > m_0$ with m_0 depending only on τ, then any solution of the equilibrium equations in H^m is, in fact, analytic.

Consider the case when H1), H2), and H3) are satisfied. The statement of the theorem asserts that if there is a numerical solution and its Sobolev norm H^r is not too large (and that of $(\mathrm{Id} + u')^{-1}$), there is a true solution nearby. Furthermore, there is an open set of parameters with invariant solutions. Hence, if the solutions cease to exist, the Sobolev norms of the numerically computed solutions have to blow up.

Remark 6.3. In the special case of twist mappings with Diophantine rotation numbers, the non-degeneracy conditions, H1), H2), and H3), are trivially satisfied. Therefore, the only thing that has to be checked is that for ε small enough, the Sobolev norms of the approximate solution, u, are finite, and that $\mathrm{Id} + u'$ is bounded away from zero.

6.6. **Quasi-Newton method for statistical mechanics models.** In this section we describe an iterative procedure (a quasi-Newton method) to improve approximate solutions.

This method is the basis of very practical algorithms and it is the key to the proof of Theorem 6.1 which we use to justify the criterion. The improvement $u+v$ of an approximate solution u is given by solving for v from the following equation

$$h'(\theta)(DE[u]v)(\theta) - v(\theta)(DE[u]h')(\theta) = -h'(\theta)E[u](\theta). \tag{45}$$

Note that equation (45) differs from the Newton step equation by the term $v(\theta)(DE[u]h')(\theta)$. Using the identity

$$\frac{d}{d\theta}E[u](\theta) = DE[u]h'(\theta). \tag{46}$$

We see that this neglected term is quadratic in $E[u]$ so that adding a term of this form to a standard Newton method will give rise to a quadratically convergent iterative scheme given that we can solve for v form equation (45). The advantage of (45) comes from the fact that the left hand side can be factored into a sequence of invertible operators. For a detailed exposition of this factorization we refer the reader to [dlL08]. Here we give a brief summary.

Introducing the operator

$$[\mathcal{L}_l f](\theta) \quad = \quad f(\theta + l\omega) - f(\theta) \ .$$

and the new variable w related to v by $v(\theta) = h'(\theta)w(\theta)$, the equation (45) transforms into:

$$\mathcal{L}_1[(\mathcal{C}_{0,1,1} + \mathcal{G})\mathcal{L}_{-1}w] = -h'E[u]. \tag{47}$$

where

$$\mathcal{C}_{i,j,L} = \partial_i\partial_j H_L \circ \gamma_L(\theta - j\omega)h'(\theta)h'(\theta - (i-j)\omega) \tag{48}$$

with

$$\gamma_L(\theta) = (h(\theta), h(\theta + \omega), ..., h(\theta + L\omega))$$

and

$$\mathcal{G} = \sum_{L\le 2}\sum_{i>j} \mathcal{L}_1^{-1}\mathcal{L}_{i-j}\mathcal{C}_{i,j,L}\mathcal{L}_{j-i}\mathcal{L}_{-1}^{-1}$$

We note that the operators $\mathcal{L}_{\pm 1}$ are invertible on functions with average 0. That is, given a function ξ with average 0, we can solve for φ satisfying

$$\varphi(\theta \pm \omega) - \varphi(\theta) = \xi(\theta) \tag{49}$$

Thus, equation (47) can be solved following the next algorithm:

Algorithm 6.4. a) Check that $\operatorname{avg}\{h'E[u]\}_\theta = 0$.

b) Find a normalized function φ (i.e. $\operatorname{avg}\{\varphi\}_\theta = 0$) solving the equation

$$\mathcal{L}_1\varphi = -h'E[u] \tag{50}$$

Therefore, if φ is a solution for (50) then for any $\mathcal{T} \in \mathbb{R}$ the equation $\mathcal{L}_1(\varphi + \mathcal{T}) = h'E[u]$ holds. In particular, we choose $\mathcal{T}$ such that

$$\operatorname{avg}\left\{(\mathcal{C}_{0,1,1} + \mathcal{G})^{-1}(\varphi + \mathcal{T})\right\}_\theta = 0.$$

c) We solve for w from

$$\mathcal{L}_{-1}w = (\mathcal{C}_{0,1,1} + \mathcal{G})^{-1}(\varphi + \mathcal{T}) \tag{51}$$

d) Finally we obtain the improved solution

$$\tilde{u}(\theta) = u(\theta) + h'(\theta)w(\theta)$$

We call attention that all the steps are diagonal either in Fourier space of in Real space. The fact that the steps are diagonal allows to have fast numerical implementations which were done in [CdlL09b].

6.7. **Estimates for the Quasi-Newton Method.** The goal of this section is to provide precise estimates for the iterative step described in Section 6.6. Throughout this section we will assume that $\omega \in \mathbb{R}$ satisfies the Diophantine condition given in Definition 4.4.

The following lemma is proven in [dlL08]

Lemma 6.5. *For every $m > 0$ we have:*

$$\begin{array}{rcl} \|\mathcal{L}_\ell \mathcal{L}_{\pm 1}^{-1}\|_m & \leq & |\ell| \\ \|\mathcal{L}_{\pm 1}^{-1}\mathcal{L}_\ell\|_m & \leq & |\ell| \end{array} \tag{52}$$

¿From these estimates we get the following

$$\begin{aligned} \|\mathcal{G}\|_m &\leq \sum_{L\geq 2}\sum_{j<i}^{L} \|\mathcal{L}_1^{-1}\mathcal{L}_{j-i}\mathcal{C}_{i,j,L}\mathcal{L}_{i-j}\mathcal{L}_{-1}^{-1}\|_m \\ &\leq \mathcal{C}\sum_{L\geq 2} L^4 M_L = a \end{aligned}$$

where

$$M_L = K_1 \|H_L\|_{C^{m+3}} \|\mathrm{Id} + u'\|_m^2$$

and K_1 is the constant of (2) depending only on m.

Then if $T(1+a) < 1$ we get

$$\|(\mathcal{C}+\mathcal{G})^{-1}\|_m < \frac{T}{1-Ta} < 1$$

Similarly we get estimates for $\mathcal{T}$ since

$$\left(\mathrm{avg}\left\{C^{-1}\right\}_\theta\right)\mathcal{T} + \left(\mathrm{avg}\left\{(\mathcal{C}+\mathcal{G})^{-1} - \mathcal{C}^{-1}\right\}_\theta\right)\mathcal{T} = \mathrm{avg}\left\{(\mathcal{C}+\mathcal{G})^{-1}\varphi\right\}_\theta$$

The second term in the left hand side can be treated as a perturbation of the first term. Therefore

$$|\mathcal{T}| \leq U/(1-2UTa)\|\varphi\|_m \leq 2U\|\varphi\|_m$$

The operator $\eta[u]$ is the operator obtained by applying the procedure 6.4.

To apply the abstract implicit function theorem we will need the following estimates on the approximate inverse η. The estimates obtained from the construction of the operator η are given in the following lemma.

Consider $r \in \mathbb{N}$.

Lemma 6.6. *Let $m > \frac{n}{2} + 2\tau$, $E[u] \in H^m$, and $\eta : H^m \to H^{m-2\tau}$ the operator constructed in Algorithm* (6.4).

Then we have the following estimates on η

$$\|\eta[u]E[u]\|_{m-2\tau} \le C\nu^{-2}(N^+)^2\|E[u]\|_m. \tag{53}$$

We will also need estimates on $DE[u]\eta[u]$.

Lemma 6.7. *Let $m > \frac{n}{2} + 2\tau$, $E[u] \in H^m$, and $E[u]$, and $\eta[u]$ defined above.*

Then we have the estimates

(54)

$$\|(DE[u]\eta[u] - \mathrm{Id})E[u]\|_{m-2\tau} \le C\nu^{-2}(N^+)^2N^-\|E[u]\|_{m-2\tau-1}\|E[u]\|_m$$

Proof. Let $\psi = DE[u]\eta[u]E[u] - E[u]$, then we have that

$$\begin{aligned} \psi &= DE[u]\eta[u]E[u] + (-h')^{-1} \cdot \eta[u]E[u] \cdot DE[u]h' \\ &\quad - (-h')^{-1} \cdot \eta[u]E[u] \cdot DE[u]h' - E[u] \\ &= (-h')^{-1} \cdot (h'DE[u]\eta[u]E[u] - \eta[u]E[u] \cdot DE[u]h') \\ &\quad + (h')^{-1} \cdot \eta[u]E[u] \cdot DE[u]h' - E[u] \\ &= (h')^{-1} \cdot \eta[u]E[u] \cdot DE[u]h' \end{aligned} \tag{55}$$

So we have that

$$\begin{aligned} &\|(DE[u]\eta[u] - \mathrm{Id})E[u]\|_{m-2\tau} \\ &\qquad \le \|\,(h')^{-1} \cdot \eta[u]E[u] \cdot DE[u]h'\|_{m-2\tau} \\ &\qquad \le C\nu^{-2}(N^+)^2N^-\|E[u]\|_{m-2\tau-1}\|E[u]\|_m, \end{aligned} \tag{56}$$

which completes the estimates for the approximate inverse.

The final result follows from an application of Theorem A.6. In the context of Theorem A.6 we consider the previous estimates with $\alpha = 2\tau$ and the estimates of Theorem 4.

6.8. **Bootstrap of regularity.** In the present section we state the theorem of bootstrap of regularity for the case of $1 - D$ statistical mechanics models. For a sketch of the proof the reader can follow the steps presented the proof of Theorem 5.8. A complete proof is presented in [dlL08].

Theorem 6.8. *Assume that the hypotheses of Theorem 6.1 hold and that the H_L are real analytic. Consider the function $u \in H^m$ with $m > \frac{1}{2} + 2\tau$ solving*

$$E[u] = 0.$$

Then $u \in \mathcal{A}_{\frac{\rho}{4}}$ where $\rho_N > 0$ is an explicit function of $N \in \mathbb{N}$.

Appendix A. An abstract Nash-Moser implicit function theorem

In this appendix we prove Theorem A.6, an abstract Nash-Moser implicit function theorem that is very well suited for the proof of Theorem 6.1. We hope that this theorem can have other applications.

In contrast with the elementary implicit function theorems, which assume that the derivative of the functional considered is invertible, the Nash-Moser implicit function theorems can cope with derivatives which do not have a bounded inverse from one space to itself. In our applications this unboundedness of the inverse of the derivative arises because the linearized equation involves solving equations with small divisors. It has become standard to think of the problem as a functional equation acting on a scale of Banach spaces, so that the linearization is boundedly invertible from one space to another (with some appropriate quantitative bounds).

The main technique is to combine the Newton step – which losses derivatives – with some smoothing that restores them. It is remarkable that, when the inverses of the linearization have a bounded order, the whole procedure converges. This has become a basic tool of non-linear analysis. The main hypothesis is that the initial guess satisfies the equation very approximately (as well as some other explicit non-degeneracy conditions. We call these theorems *a posteriori* following the notation that one uses in numerical analysis.

The theorem A.6 is very close to the main theorem in [Sch60] (see also the exposition in [Sch69]) but we allow an extra term in the reminder as in [Zeh75]. We also remark that it suffices to estimate the approximate inverse in the range of the operator. As pointed out in [Zeh75] allowing the extra term, permits us to deal very comfortably with equations with a *group structure*. In particular with conjugacy equations. Even if our problem is not a conjugacy problem (except in the case of twist maps), it is close enough to it so that it fits in the scheme. One feature of Theorem A.6 which is very important for our purposes is that Theorem A.6 does not assume that the initial system is close to integrable.

We will assume that there is a family of Banach spaces endowed with smoothing operators. The non-linear operator will satisfy some assumptions.

In our applications the scale of spaces will be the Sobolev spaces, and we will denote the scale of spaces by H^m. Nevertheless Theorem A.6 works for general scales of spaces and indeed, the scheme of the proof can also produce results with analytic regularity.

We will consider scales of Banach spaces $\mathcal{X}^r$ such that $\mathcal{X}^r \subset \mathcal{X}^{r'}$ whenever $r' \le r$ and the inclusions are continuous.

Remark A.1. There are many other spaces that can be used in place of H^m. Notably, [Zeh75] emphasizes spaces $\widehat{C}^m$ (which are just the familiar Hölder spaces when $m \notin \mathbb{N}$. The spaces $\widehat{C}^m$ have the extra property, that they can be characterized by the speed of approximation by some concrete smoothing operators.

This allows to apply the technique of "double smoothing" (already mentioned in [Mos66b, Mos66a]). The technique of "double smoothing" consists in approximating the problem by a sequence of analytic problems. Of course, each of these problems are solved by a Newton method and smoothing.

As it is argued in [Mos66b, Mos66a, Zeh75] the method of double smoothing leads to good estimates on regularity.

In this paper we have decided not to follow the double smoothing technique. The main reason is that we wanted a simple theorem that follows closely the numerical algorithm and for which it is straightforward to obtain numerical values. (For numerical estimates for double smoothing see [Zeh76b].)

We also note that the main applications we have in mind involve geometric properties, so that the smoothing of the problem has to be done in the classes of problems which preserve the geometric structure. See [GEdlL08] for a treatment of symplectic and volume preserving cases. Other geometric structures would require different treatments.

A.1. **Smoothing operators.**

Definition A.2. *Given a scale of spaces, we say that $\{S_t\}_{t\in\mathbb{R}^+}$ is a family of smoothing operators when*

i)
$$\lim_{t\to\infty} \|(S_t - \mathrm{Id})u\|_0 = 0$$
ii)
$$\|S_t u\|_m \le C t^{m-n} \|u\|_n$$
iii)
$$\|(\mathrm{Id} - S_t)u\|_m \le C t^{-n} \|u\|_{n+m}$$

In the concrete case of Sobolev spaces, some very convenient smoothing operators S_t are defined for $t > 1$ as follows

$$\widehat{(S_t u)}_k = e^{-|k|/t} \hat{u}_k \tag{57}$$

Note that $S_t u$ is analytic. In our concrete applications we will use (57).

Lemma A.3. *The operators S_t defined in* (57) *are smoothing operators in the sense of Definition A.2.*

Proof. Notice that for $0 \leq m, n < \infty$

$$
\begin{aligned}
\|S_t u\|_m^2 &= \sum e^{-2|k|/t}(1+|k|^2)^m |u_k|^2 \\
&= \sum e^{-2|k|/t}(1+|k|^2)^{m-n}(1+|k|^2)^n |u_k|^2 \\
&\leq \sum e^{-2|k|/t} t^{2m-2n}(1+\tfrac{|k|^2}{t^2})^{m-n}(1+|k|^2)^n |u_k|^2 \\
&\leq C t^{2(m-n)} \|u\|_n^2
\end{aligned}
\tag{58}
$$

We also have that

$$
\begin{aligned}
\|(S_t - \mathrm{Id})u\|_m^2 &= \sum (e^{-|k|/t}-1)^2 (1+|k|^2)^m |u_k|^2 \\
&\leq \sum \frac{(e^{-|k|/t}-1)^2}{\left(1+\frac{|k|^2}{t^2}\right)^n} t^{-2n}(1+|k|^2)^{m+n}|u_k|^2 \\
&\leq C_n t^{-2n}\|u\|_{m+n}^2
\end{aligned}
\tag{59}
$$

with

$$C_n = \sup_x \frac{(e^{-x}-1)^2}{(1+x^2)^n}.$$

One important consequence of the existence of smoothing operators are interpolation inequalities [Zeh75].

Lemma A.4. *Let $\mathcal{X}^r$ be a scale of Banach spaces with smoothing operators. For any $0 \leq n \leq m$, $0 \leq \theta \leq 1$, denoting*

$$l = (1-\theta)n + \theta m$$

we have that for any $u \in \mathcal{X}^m$:

$$\|u\|_l \leq C_{n,m}\|u\|_n^{1-\theta}\|u\|_m^{\theta} \tag{60}$$

Proof. For all $t > 0$, $u = S_t u + (\mathrm{Id} - S_t)u$ and therefore, if $u \in \mathcal{X}^m$,

$$\|u\|_l \leq \|S_t u\|_l + \|(\mathrm{Id} - S_t)u\|_l \leq t^{l-n}C_{n,l}\|u\|_n + t^{-(m-l)}C_{l,m}\|u\|_m.$$

Computation of the minimum in t of the RHS above leads to the result.

□

A.2. **Formulation of Theorem A.6.** In this section, we formulate and prove the abstract implicit function theorem TheoremA.6. Following standard practice in KAM theory, we use the letter C to denote arbitrary constants that depend only on the uniform assumptions of the theorem. In particular, the meaning of C can change from line to line.

Remark A.5. The proof we present follows very closely [Sch60]. In particular, we have followed the choices of [Sch60] in the loss of regularity. Clearly, these choices are far from optimal. In particular, we have assumed that the functional and the approximate inverse loose α derivatives. This is natural for PDE applications, but in our case the functional itself does not loose any derivatives (even if the approximate inverse does).

Theorem A.6. *Let $m > \alpha$ and $\mathcal{X}^r$ for $m \le r \le m + 17\alpha$ be a scale of Banach spaces with smoothing operators. Let $\mathcal{B}_r$ be the unit ball in $\mathcal{X}^r$, $\tilde{\mathcal{B}}_r = u_0 + \mathcal{B}_r$ the unit ball translated by $u_0 \in \mathcal{X}^r$, and $\mathcal{B}(\mathcal{X}^r, \mathcal{X}^{r-\alpha})$ is the space of bounded linear operators from $\mathcal{X}^r$ to $\mathcal{X}^{r-\alpha}$. Consider a map*

$$\mathcal{F} : \tilde{\mathcal{B}}_r \to \mathcal{X}^{r-\alpha}$$

and

$$\eta : \tilde{\mathcal{B}}_r \to \mathcal{B}(\mathcal{X}^r, \mathcal{X}^{r-\alpha})$$

satisfying:

i) $\mathcal{F}(\tilde{\mathcal{B}}_r \cap \mathcal{X}^r) \subset \mathcal{X}^{r-\alpha}$ *for* $m \le r \le m + 17\alpha$.

ii) $\mathcal{F}|_{\tilde{\mathcal{B}}_m \cap \mathcal{X}^r} : \tilde{\mathcal{B}}_r \cap \mathcal{X}^r \to \mathcal{X}^{r-\alpha}$ *has two continuous Fréchet derivatives, both bounded by* M, *for* $m \le r \le m + 17\alpha$.

iii) $\|\eta[u]z\|_{r-\alpha} \le C\|z\|_r$, $u \in \tilde{\mathcal{B}}_r, z \in \mathcal{X}$, *for* $r = m - \alpha, m + 16\alpha$.

iv) $\|(D\mathcal{F}[u]\eta[u] - \mathrm{Id})z\|_{r-\alpha} \le C\|\mathcal{F}[u]\|_r \|z\|_r$, $u \in \tilde{\mathcal{B}}_r, z \in \mathcal{X}^r$, *for* $r = m$

v) $\|\mathcal{F}[u]\|_{m+16\alpha} \le C(1 + \|u\|_{m+17\alpha})$, $u \in \tilde{\mathcal{B}}_m$

Then if $\|\mathcal{F}[u_0]\|_{m-\alpha}$ is sufficiently small, there exists $u^* \in \mathcal{X}^m$ such that $\mathcal{F}[u^*] = 0$. Moreover,

$$\|u - u^*\|_m < C\|\mathcal{F}[u_0]\|_{m-\alpha}$$

Remark A.7. Note that in *iii*) we are only requiring estimates for two values of r. When *iii*) acts in the whole space, and the spaces $\mathcal{X}^r$ are interpolation spaces [Tay97], this implies the bounds for the intermediate spaces.

Remark A.8. Note that conditions *iii*) and *iv*) can be replaced by the weaker conditions

iii)′ $\|\eta[u]\mathcal{F}[u]\|_{r-\alpha} \le C\|\mathcal{F}[u]\|_r$, $u \in \tilde{\mathcal{B}}_r$, for $r = m - \alpha, m + 16\alpha$.

iv)′ $\|(D\mathcal{F}[u]\eta[u] - \mathrm{Id})\mathcal{F}[u]\|_{r-\alpha} \le C\|\mathcal{F}[u]\|_r^2$, $u \in \tilde{\mathcal{B}}_r$, for $r = m$. Indeed in both of our applications the definition of the approximate inverse requires that some average of z is small which is true for functions in the range of $\mathcal{F}$ but not in general.

A.3. **Proof of Theorem A.6.** The proof is based on an iterative procedure combining the ideas of [Sch60, Zeh75]. Given a function $u \in \mathcal{X}^{m+13\alpha}$ so that $\|\mathcal{F}[u]\|_{m-\alpha}$ is sufficiently small compared with the other properties of the function, the iterative procedure constructs another function u^* such that $\mathcal{F}[u^*] = 0$.

Let $\kappa > 1$, $\beta, \mu, \delta > 0$, $0 < \nu < 1$ be real numbers to be specified later. We will need that they satisfy a finite set of inequalities among them and with the quantities appearing in the assumptions of the problem. We will indicate them during the proof when they have been specified and we will check that they can be satisfied simultaneously. We anticipate that there will be a finite number of inequalities of κ, μ, and δ for β large enough and other finite set of inequalities for the remaining quantities.

We construct a sequence $\{u_n\}_{n\geq 0}$ by taking

$$u_{n+1} = u_n - S_{t_n}\eta[u_n]\mathcal{F}[u_n] \tag{61}$$

where $t_n = e^{\beta\kappa^n}$. We will prove that this sequence satisfies

(p1;n)

$$(u_n - u_0) \in \mathcal{B}_m$$

(p2;n)

$$\|\mathcal{F}[u_n]\|_{m-\alpha} \leq \nu e^{-\mu\alpha\beta\kappa^n}$$

(p3;n)

$$1 + \|u_n\|_{m+16\alpha} \leq \nu e^{\delta\alpha\beta\kappa^n}$$

Suppose that conditions $(p1; j)$, $(p2; j)$, and $(p3; j)$ are true for $j < n$. We start by establishing $(p1; n)$.

Notice that $(p2; n-1)$ implies that

$$\begin{aligned}\|u_{n-1} - u_n\|_m &= \|S_{t_{n-1}}\eta[u_{n-1}]\mathcal{F}[u_{n-1}]\|_m \leq Ce^{2\alpha\beta\kappa^{n-1}}\|\eta[u_{n-1}]\mathcal{F}[u_{n-1}]\|_{m-2\alpha}\\ &\leq C\nu e^{\alpha\beta\kappa^{n-1}(2-\mu)}\end{aligned}$$

Then if

$$\mu > 2, \tag{62}$$

$\{u_n\} \subset \mathcal{X}^m$ converges to some $u \in \mathcal{X}^m$.

Now, to prove $(p1; n)$, we notice that, using $\kappa^j \leq j(\kappa - 1)$ we have that

$$
\begin{aligned}
\|u_n - u_0\|_m &\le C\nu \sum_{j=1}^{\infty} e^{\alpha\beta\kappa^j(2-\mu)} \\
&\le C\nu \sum_{j=1}^{\infty} e^{\alpha\beta j(\kappa-1)(2-\mu)} \\
&\le C\nu \frac{e^{\alpha\beta(\kappa-1)(2-\mu)}}{1-e^{\alpha\beta(\kappa-1)(2-\mu)}}
\end{aligned} \tag{63}
$$

Therefore, $\|u_n - u_0\|_m \le C\nu$ for $\mu > 2$ and β large enough.

To prove $(p2; n)$ start by adding and subtracting the terms $\mathcal{F}[u_{n-1}]$, $D\mathcal{F}[u_{n-1}]\eta[u_{n-1}]\mathcal{F}[u_{n-1}]$, and $D\mathcal{F}[u_{n-1}]S_{t_{n-1}}\eta[u_{n-1}]\mathcal{F}[u_{n-1}]$ to $\mathcal{F}[u_n]$. Then we group the terms in three groups obtaining the following inequality

$$
\begin{aligned}
\|\mathcal{F}[u_n]\|_{m-\alpha} \le\ & \|\mathcal{F}[u_n] - \mathcal{F}[u_{n-1}] + D\mathcal{F}[u_{n-1}]S_{t_{n-1}}\eta[u_{n-1}]\mathcal{F}[u_{n-1}]\|_{m-\alpha} \\
&+\|(\mathrm{Id} - D\mathcal{F}[u_{n-1}]\eta[u_{n-1}])\mathcal{F}[u_{n-1}]\|_{m-\alpha} \\
&+\|D\mathcal{F}[u_{n-1}](\mathrm{Id} - S_{t_{n-1}})\eta[u_{n-1}]\mathcal{F}[u_{n-1}]\|_{m-\alpha}.
\end{aligned} \tag{64}
$$

We will estimate the terms on the left hand side of 64 separately. We estimate the first term of (64) using assumption $iii)$ and the quadratic remainder of Taylor's theorem.

$$
\begin{aligned}
\|\mathcal{F}[u_n] - &\mathcal{F}[u_{n-1}] + D\mathcal{F}[u_{n-1}]S_{t_{n-1}}\eta[u_{n-1}]\mathcal{F}[u_{n-1}]\|_{m-\alpha} \\
&\le C\|S_{t_{n-1}}\eta[u_{n-1}]\mathcal{F}[u_{n-1}]\|_m^2 \\
&\le Ce^{2\alpha\beta\kappa^{n-1}}\|\eta[u_{n-1}]\mathcal{F}[u_{n-1}]\|_{m-2\alpha}^2 \\
&\le C\nu^2 e^{2\alpha\beta\kappa^{n-1}(2-\mu)}
\end{aligned} \tag{65}
$$

For the second term of (64) by assumption $iv)$ we get

$$
\|(D\mathcal{F}[u_{n-1}]\eta[u_{n-1}] - \mathrm{Id})\mathcal{F}[u_{n-1}]\|_{m-\alpha} \le C\|\mathcal{F}[u_{n-1}]\|_m^2. \tag{66}
$$

We can estimate, $\|\mathcal{F}[u_{n-1}]\|_m^2$, the using interpolation inequality (60) and induction hypotheses $(p2; n-1)$ and $(p3; n-1)$.

$$
\begin{aligned}
\|\mathcal{F}[u_{n-1}]\|_m^2 &\le C\|\mathcal{F}[u_{n-1}]\|_{m-\alpha}^{36/17}\|\mathcal{F}[u_{m-1}]\|_{m+16\alpha}^{2/17} \\
&\le C\|\mathcal{F}[u_{n-1}]\|_{m-\alpha}^{36/17}(1+\|u_{n-1}\|_{m+17\alpha})^{2/17} \\
&\le C\nu^2 e^{\alpha\beta\kappa^{n-1}(-\frac{36\mu}{17}+\frac{2\delta}{17})}
\end{aligned} \tag{67}
$$

For the third term of (64), we use the properties of the smoothing operators and the fact that the Fréchet derivative, $D\mathcal{F}[u_{n-1}]$, is bounded.

$$
\begin{aligned}
&\|D\mathcal{F}[u_{n-1}](\mathrm{Id}-S_{t_{n-1}})\eta[u_{n-1}]\mathcal{F}[u_{n-1}]\|_{m-\alpha}\\
&\quad\le C\|(\mathrm{Id}-S_{t_{n-1}})\eta[u_{n-1}]\mathcal{F}[u_{n-1}]\|_{m}\\
&\quad\le Ct^{-15\alpha}\|\eta[u_{n-1}]\mathcal{F}[u_{n-1}]\|_{m+15\alpha}\\
&\quad\le Ct^{-15\alpha}\|\mathcal{F}[u_{n-1}]\|_{m+16\alpha}\\
&\quad\le C\nu e^{-15\alpha\beta\kappa^{n-1}}(1+\|u_{n-1}\|_{m+17\alpha})\\
&\quad\le C\nu e^{\alpha\beta\kappa^{n-1}(\delta-15)}
\end{aligned}
\tag{68}
$$

The desired inequality $(p2;n)$ is satisfied if

$$
C(\nu^2 e^{2\alpha\beta\kappa^{n-1}(2-\mu)}+\nu^2 e^{\alpha\beta\kappa^{n-1}(\frac{2\delta}{17}-\frac{36\mu}{17})}+\nu e^{\alpha\beta\kappa^{n-1}(\delta-15)})\le \nu e^{-\mu\alpha\beta\kappa^{n}}
$$

or equivalently

$$
C(\nu e^{-\alpha\beta\kappa^{n-1}(2(\mu-2)-\mu\kappa)}+\nu e^{-\alpha\beta\kappa^{n-1}(-\frac{2\delta}{17}+\frac{36\mu}{17}-\mu\kappa)}+e^{-\alpha\beta\kappa^{n-1}(15-\delta-\mu\kappa)})\le 1.
\tag{69}
$$

Condition (69) is true whenever ν is small enough and

$$
\begin{aligned}
\mu(2-\kappa) &> 4,\\
\mu(36-17\kappa) &> 2\delta,\\
15-\mu\kappa &> \delta,
\end{aligned}
\tag{70}
$$

and β is sufficiently large. This establishes $(p2;n)$.

Finally we note that

$$
\begin{aligned}
1+\|u_n\|_{m+17\alpha} &\le 1+\sum_{j=0}^{n-1}\|S_{t_j}\eta[u_j]\mathcal{F}[u_j]\|_{m+17\alpha}\\
&\le 1+C\sum_{j=0}^{n-1}e^{2\alpha\beta\kappa^j}\|\eta[u_j]\mathcal{F}[u_j]\|_{m+15\alpha}\\
&\le 1+C\sum_{j=0}^{n-1}e^{2\alpha\beta\kappa^j}\|\mathcal{F}[u_j]\|_{m+16\alpha}\\
&\le 1+C\sum_{j=0}^{n-1}e^{2\alpha\beta\kappa^j}(1+\|u_j\|_{m+17\alpha})\\
&\le 1+C\sum_{j=0}^{n-1}e^{\alpha\beta(2+\delta)\kappa^j}.
\end{aligned}
\tag{71}
$$

Thus

$$
(1+\|u_n\|_{m+13\alpha})e^{-\delta\alpha\beta\kappa^n}\le e^{-\delta\alpha\beta\kappa^n}+C\sum_{j=0}^{n-1}e^{\alpha\beta\kappa^j(2+\delta-\kappa\delta)}
\tag{72}
$$

To have $(p3; n)$ it suffices that the RHS of (72) is less than 1. If $\delta > \frac{2}{\kappa-1}$ the right side of (72) will be less than 1 for sufficiently large β.

If we consider $\kappa = 4/3$, $\delta = 6$, and $\mu = 61/10$ then (70) and

$$\delta > \frac{2}{\kappa - 1} \tag{73}$$

are satisfied at the same time. To complete the induction, we fix β large enough so that (72) and (69) are satisfied.

Finally we consider with our choices of β and μ, and fix ν to be

$$\nu = \|\mathcal{F}[u_0]\|_{m-\alpha} e^{\alpha\beta\mu}. \tag{74}$$

¿From this choice of ν, together with (63) we have that

$$\|u^* - u_0\|_m \leq C\nu \frac{e^{\alpha\beta(\kappa-1)(2-\mu)}}{1 - e^{\alpha\beta(\kappa-1)(2-\mu)}} \leq C_{\mu,\alpha,\beta,\kappa} \|\mathcal{F}[u_0]\|_{m-\alpha},$$

which completes the proof. □

A.4. **Uniqueness and dependence with respect to parameters.**

Theorem A.9. *With the notation of Theorem A.6, assume that, when $u \in \mathcal{X}_r$ is such that $F(u) = 0$ we can find $\gamma(u) : \mathcal{X}_r \to \mathcal{X}_{r-\alpha}$ for $r \in [r_0 - 2\alpha, r_0 + 2\alpha]$. such that*

$$\|\gamma(u) DF(u) z - z\|_{r-\alpha} \leq C \|z\|_r^2 \tag{75}$$

Then, there is $\rho > 0$ such that if

$$\|u - v\|_{r_0 + 2\alpha} \leq \rho \ , \qquad F(v) = 0 \ .$$

Then, $u = v$.

As we will see, the rather elementary proof we present here gives that one can take $\rho = \left(C(1 + \frac{1}{2}\|D^2 F\|_{\mathcal{X}_r \to \mathcal{X}_{r-\alpha}})\right)^{-1}$.

Remark A.10. Both in Theorem A.6 and A.9 one could assume that F, DF, η, γ have different losses of differentiability and obtain similar results. For example, in the applications in Section 5 and Section 6 F, DF do not have any loss of differentiability. Of course, one can take α to be the maximum of all these losses.

Remark A.11. Both in Theorem A.6 and A.9 one can improve the hypothesis on the approximate inverses γ and η. In place of (75) one could just assume

$$\|\gamma(u) DF(u) z - z\|_{r-\alpha} \leq C \|z\|_r^{1+\nu}$$

with $\nu > 0$.

Remark A.12. In the applications to KAM theory, rather than assuming (75) we have the stronger assumptions

$$\gamma(u)DF(u)z = z$$

of $z \in \mathcal{X}_r$. That is, we can assume that the derivative has a left inverse In Theorem A.9 we just assume that there is an approximate inverse. Note that we only assume the existence of γ on the set of solutions.

Proof. Define R by

$$0 = F(v) - F(u) - DF(u)(v-u) + R\ .$$

By Taylor's theorem, we have:

$$\|R\|_{r-\alpha} \le \frac{1}{2}\|D^2F\|_{\mathcal{X}_r\to\mathcal{X}_{r-\alpha}}\|u-v\|_r^2 \tag{76}$$

If $F(u) = F(v) = 0$, we have

$$\gamma(u)DF(u)(v-u) + \gamma(u)R = 0$$

Using (75) and ii) of Theorem A.6 we have

$$\begin{aligned} \|v-u\|_{r-2\alpha} &\le C\|v-u\|_r^2 + \|\gamma(u)R\|_{r-2\alpha} \\ &\le C\|v-u\|_r^2 + C\|R\|_{r-\alpha} \\ &\le \tilde{C}\|v-u\|_r^2 \end{aligned} \tag{77}$$

with $\tilde{C} = C(1+\frac{1}{2}\|D^2F\|_{\mathcal{X}_r\to\mathcal{X}_{r-\alpha}})$.

Using the interpolation inequalities (60), we have

$$\|v-u\|_r \le \|v-u\|_{r-2\alpha}^{1/2}\|v-u\|_{r+2\alpha}^{1/2}$$

Hence, substituting in (77), we obtain:

$$\|v-u\|_{r-2\alpha} \le \tilde{C}\|v-u\|_{r-2\alpha}\|v-u\|_{r+2\alpha}$$

Hence, if $\tilde{C}\|v-u\|_{r+2\alpha} < 1$, we conclude that $\|v-u\|_{r-2\alpha} = 0$. □

Remark A.13. The proof presented here is very similar to the proof of uniqueness in [Ham82] The statements in [Ham82] assume the existence of an inverse defined on a neighborhood, but the argument for uniqueness only uses the existence of a left inverse defined on the solutions. We also point out that [Ham82] uses an improved version of (76) and assumes different bounds on γ than those assumed here.

A.4.1. *Lipschitz dependence on parameters.* As a Corollary of Theorem A.6 and A.9 is that if we assume Lipschitz dependence of $\mathcal{F}$ with respect to a parameter f in a Banach space $\mathcal{Y}$, we obtain Lipschitz dependence of the solution with respect to the parameter f.

Let $\mathcal{Y}$ be a Banach space of parameters. in Theorem A.6 that $\tilde{\mathcal{B}}_{\Lambda,r} = \{(f,u) \in \mathcal{Y} \times (u_o + \mathcal{B}_r) | \|f - f_0\|_{\mathcal{Y}} < \Lambda\}$, for some fixed $\Lambda > 0$.

Corollary A.14. *Consider the map*

$$\mathcal{F} : \mathcal{B}'_{\Lambda,r} \to \mathcal{X}^{r-\alpha}$$

and

$$\eta : \mathcal{B}'_{\Lambda,r} \to \mathcal{B}(\mathcal{X}^r, \mathcal{X}^{r-\alpha}).$$

Assume that for each value of f, the map $\mathcal{F}(f,\cdot)$ satisfies the hypothesis of Theorem A.6.

Assume also Lipschitz dependence with respect to the parameter f.

$$\|\mathcal{F}[f,u] - \mathcal{F}[f',u]\|_{m-\alpha} \leq L\|f - f'\|_{\mathcal{Y}}$$

for $(f,u),(f',u) \in \tilde{\mathcal{B}}_{\Lambda,m-\alpha}$.

Then, there exists a Λ such that the solution $(f,u(f))$ $\mathcal{F}[f,u(f)] = 0$ produced applying Theorem A.6 is Lipschitz with respect to the parameter f.

Notice that, because of Theorem A.9, these solutions are unique in the neighborhood we are considering.

Proof. Notice that for $(f',u) \in \tilde{\mathcal{B}}_{\Lambda,m-\alpha}$ then

$$\|\mathcal{F}[f',u]\|_{m-\alpha} = \|\mathcal{F}[f,u(f)] - \mathcal{F}[f',u]\|_{m-\alpha} \leq L\|f - f'\|_{\mathcal{Y}}$$

We choose Λ small enough so that by Theorem A.6 we have that there exists $u(f') \in \mathcal{X}^m$ so that $\mathcal{F}[f',u(f')] = 0$ and we have the Lipschitz dependence with respect to the parameter $f \in \mathcal{Y}$, i.e., for $f,f' \in \tilde{\mathcal{B}}_{\Lambda,r}$

$$\|u(f) - u(f')\|_m < CL\|f - f'\|_{\mathcal{Y}}.$$

□

Remark A.15. The argument presented here admits several generalizations, which can be found in [Van02], which under appropriate hypothesis establish differentiability in the sense of Whitney. If we do not assume uniqueness, but assume Lipschitz dependence of the approximate inverse, it is possible to conclude the existence of a Lipschitz family of solutions (even if there may be others).

Appendix B. Some remarks on the literature on computation of breakdown of analyticity

Since the problem of the breakdown of analyticity has importance both in Mathematics and in Physics, there is an extensive literature on its computation.

In the following, we will describe the main methods that we know of and sketch a comparison with the present method. We emphasize that we cannot claim to give details of all the methods. In particular, we do not mention the all important issue of how can one assess the range of applicability of the methods nor the numerical precision. We certainly hope that more qualified author will make a more systematic survey.

B.0.2. *Scalings and renormalization group.* We do not discuss the (very important) phenomena that occur at breakdown. Notably, we do not cover asymptotic scalings and their explanations by a renormalization group. We just refer to [Mac82, Koc04], to the survey [CJ02], and the references there.

When scalings and renormalization group are present at breakdown, all the algorithms discussed here can be improved in two ways. First, using the scaling relations it is possible to compute better the objects near breakdown since the scaling provides good initial points for the iterative methods. Second, using the scalings, one can post-process the results and fit powers law to the data. Furthermore, there are conjectures on the behavior of the renormalization group that show that some of these criterion are sharp, [dlLO06].

These improvements based on renormalization are very important for some of the methods and, it is quite customary to use them. For example, see [Mac82] for a discussion on the relation of Greene's method and renormalization. Of course, when methods rely on scaling relations, they are powerless to assess whether indeed there are scalings. Fortunately, the method presented in the paper can work quite comfortably in situations when there is no scaling. Since there is a rigorous justification allows us to make sure that the computed solution corresponds to a true solution. See [CdlL09b, CdlL09a] for reports on the numerical findings.

B.1. **Several methods used in the literature.** In this section we describe succinctly the methods that have been proposed in the literature. In Section B.2 we will present some comparisons among the different methods. We note that the descriptions here are rather terse and the interested reader is urged to consult the original references.

B.1.1. *Greene Method (GM).* This method was introduced in [Gre79] for twist mappings. Partial justifications appear in [FdlL92b, Mac92]. extensions (with justification) to complex values in [FdlL92b], to non-twist mappings in [DdlL00] and to higher dimensions in [Tom96].

The method (GM) consists in searching for periodic orbits and computing the *residue* (i.e $\frac{1}{4}(TrDf^n(x_n)-2)$ The breakdown happens when the residue grows as n approaches to infinity. In [Tom96] it is shown that the correct analogue in higher dimension is to study the spectrum of $Df^n(x_n)$ and to check whether all the eigenvalues remain close to 1 (equivalently, that the coefficients of the characteristic polynomial remain closer to those of $(\lambda-1)^{2d}$.

B.1.2. *Obstruction method (OM).* This method was introduced in [OS87]. The method searches for homoclinic connections between periodic orbits of nearby period. It is shown in [OS87] that if there is such a connection, there are no invariant circles with a rotation in the interval bounded by the rotation numbers of the periodic orbits exhibiting homoclinic connections.

In [dlLO06] it is shown that, under some hypothesis on the renormalization group appropriate for the rotation number ρ^*, the criterion is sharp in the universality class. Maps in the universality class either have an invariant circle of the a homoclinic connection between periodic orbits whose rotation numbers straddle ρ^*.

B.1.3. *Well ordered orbits (WOO).* In [BH87] it was shown that, for twist mappings, there are invariant circles if and only if the periodic orbits of all the convergents are *well ordered.* An orbit $\{x_k\}_{k\in\mathbb{Z}}$ is *well ordered* if when we fix $l, m \in \mathbb{Z}$ then for all $k \in \mathbb{Z}$ either $x_{k+m}+l \geq x_k$ or $x_{k+m}+l \leq x_k$. Computing all periodic orbits is, of course impossible, but finding a badly ordered periodic orbit, excludes the existence of invariant circles of several rotation numbers.

A preliminary assessment of the numerical efficiency of (WOO) is done in [LM86].

B.1.4. *Peierls-Nabarro barrier (PNB).* In [Mat86] it is shown that for twist mappings an invariant circle exists if and only if, the so-called Peierls-Nabarro barrier vanishes. This barrier is the difference in action between minimizers and other critical orbits of the same rotation number.

A different proof which extends to PDE's and to models with long range interactions in [dlLV07b, dlLV07c, dlLV07a].

Note that for the case of rational rotations the lower bounds of the PN barrier can be computed just by considering the minimizing periodic orbit and another one. For twist mappings there are a-priori bounds on the modulus of continuity of the PN barrier [Mat87], so that, for twist mappings, one can, in principle, use this method to show that there are no invariant circles with a finite computation. We are not aware of any numerical implementation of these ideas.

B.1.5. *A priori bounds for minimizers (APB).* When the mapping is a twist mapping, it was shown in [Mat84] that invariant circles have a-priori bounds on the slope. Geometrically, this means that the circles passing through a point are contained in a cone with a-priori bounds in the slopes. The key step in the argument is to show that the invariant circles have to be graphs over the angle coordinate. Then, one can observe that if we compose with an shear close enough to the identity, the map is still a twist map, so that the invariant circle has to be a graph over a tilted coordinate. For further improvements of this line of argument see [Her83, LC91, Her89, Arn08].

It was observed in [MP85] that, by studying carefully the maps one can show that if the a-priori bounds are satisfied in a region, then its image violates them (because the map rotates). By repeating this argument enough times (using computer assisted calculation) one can show that the circle does not exist [MP85].

In [Mac89] it was observed, following [Mos86], the theory of the calculus of variations implies that continuous families of critical points (an invariant circle) has to be a minimizer. This implies that the second derivative of the action along an orbit has to be positive definite. This gives a different proof to the cone criterion. Nevertheless, it is rather simple to compute orbits and decide whether the orbits are minimizers of the action.

This criterion for non-existence was shown to be sharp in [MMS89] in the sense that, for twist maps, if there is no invariant circle, the cone method will prove this result with a finite computation.

B.1.6. *Variational Shadowing (VS).* A method based on the variational theory of shadowing was introduced in [Jun91]. We believe that the theory could be profitably recast in terms of viscosity sub-solutions and super-solutions of the Hamilton-Jacobi theory [Fat07].

This method can incorporate discussions of round off error and lead to rigorous proofs of nonexistence of invariant circles. Computer assisted results of are obtained in [Jun91]. They produce results which seem to be extremely accurate and extremely fast. We are not aware of extensions of the method to statistical mechanics models.

B.1.7. *The Padé method (PM).* In the case that the system is analytic and depends analytically on parameters, often one can compute some Lindstedt series expansion of the invariant torus. These series can be computed traditionally term by term, but there are also a quadratically convergent algorithms. It suffices to apply the quadratically convergent algorithms presented in this paper to power series. A Newton step will double the number of exact terms. Hence, it is possible to obtain polynomial approximations with high degree of accuracy.

These Lindstedt series were shown in [Mos67] to have positive radii of analyticity when the rotation is Diophantine. The proof in [Mos67] uses KAM methods. A later proof of convergence by exhibiting explicitly cancellations in the series was obtained [Eli96]. The method of exhibiting explicitly cancellations in the series has lead to a very large literature, even if it is limited only to analytic systems. Among the first papers of these literature, we point out [CF94, Gal94] which deal with perturbations that do not depend on the action.

Strictly speaking, the domain of convergence of a power series expansion is a disk. Nevertheless, for analytic functions, whose domain of definition is not a disk, the power series determines the function and one can wonder whether the domain of definition of the function can be assessed from the Taylor expansion at one point.

From the numerical point of view, given a numerically computed power series (of course, one only computes a polynomial) $F(\epsilon)$, a well known method to estimate the domain of convergence of the polynomial is to obtain a rational function $P(\epsilon)/Q(\epsilon)$ with P, Q polynomials of coefficients of degree N, M such that

$$F(\epsilon) = P(\epsilon)/Q(\epsilon) + O(\epsilon^{N+M+1}) \tag{78}$$

It can be readily seen that, there are unique polynomials of the indicated degrees that satisfy (78) and the normalization condition $Q(0) = 1$. The roots of Q give an approximation of the singularities of F.

There is a very large literature on the convergence of Padé methods for analytic functions. See, [BGM96]. The application for the analyticity domains of KAM theory was started in [BC90, BCCF92] and they found that the domain of analyticity of the standard map is very close to a disk. The Padé method was compared with Greene's method for complex values in [FdlL92a], not only for the standard map, but also for other maps with other harmonics. Further comparisons of this method with other methods were undertaken in [dlLT94b, dlLT94a, dlLT95]. In [dlLT95] one can find extensions of the method such as the multi-point Padé method.

It should be remarked that the Padé method is numerically very unstable. One can easily see that if one considers functions $F(\epsilon) = \sum_j a_j/(b_j - \epsilon)$, for which the Padé approximation indeed converges, we have that $F_n = \sum_j a_j(b_j)^{-n-1}$ so that, the poles farther away contribute much less than the closest poles to the value of the coefficients. Equivalently, to get information on the poles farther away, one needs a very accurate computation of the coefficients F_n. It is, therefore, rather fortunate that the standard map family has a domain of analyticity which is a circle.

Another problem with the Padé method is that it assumes that the singularities are poles. An easy numerical experiment – performed in [dlLT94b, dlLT95] – is to study the Padé approximants of $F(\epsilon) = \sum_j a_j/(b_j - \epsilon) + d_j\sqrt{e_j - \epsilon}$ and check whether the Padé method succeeds in finding the singularities b_j, e_n. It was empirically found that the square roots generate lines of zeros and poles in the Padé approximant accumulating at the branch point, a behavior that had already been conjectured in the mathematical literature. Note that the accumulation of zeros and poles makes the numerical method more unstable. It is interesting to note that this accumulation of poles and zeros had been found in Lindstedt series in [BM94].

Furthermore, [dlLT94b, dlLT95] presented heuristic arguments showing that one should expect that the boundary of analyticity of KAM tori is better described by an accumulation of branch points than by an accumulation of poles.

For example, if one considers periodic orbits, they satisfy an analytic equation depending on parameters, the generic bifurcation – easy to verify in many cases – is a branch point. Similarly, if one adds a small imaginary component to the frequency, one obtains that the invariance operator is compact and again, the bifurcation one expects is branch points (and indeed this is what one finds numerically). Since the analyticity domain is known to be well approximated by that of periodic orbits, see Section B.1.1, it follows that it is better to approximate by sums of branch points.

One variation of the Padé method that deals well with branch points is the logarithmic Padé method introduced in [dlLT94b]. The main observation is that if $f(\varepsilon)$ has a branch point singularity $f'(\varepsilon)/f(\varepsilon)$ has a pole at ε_0. Note also that given a power series expansion, the power series expansion of $g(\varepsilon) = f'(\varepsilon)/f(\varepsilon)$ can be computed matching powers in $g(\varepsilon)f(\varepsilon) = f'(\varepsilon)$. Hence the LPM proceeds by

Algorithm B.1. *LPM*

(1) *Compute* $g(\varepsilon)$ *solving* $f(\varepsilon)g(\varepsilon) = f'(\varepsilon)$.

(2) *Compute Padé approximant of* $g(\varepsilon) = N(\varepsilon)/P(\varepsilon)$.
(3) *Compute this zeros of* $P(\varepsilon)$.

Implementations and comparisons of the Logarithmic Padé Method can be found in [dlLT94b].

B.2. **Comparison among the methods.** Of course, in practice, all these different methods have different ranges of applicability, differ in their possibilities of extending them to other situations etc.

In the following remarks, we try to summarize some of the differences. Of course, we cannot hope to be completely systematic and, in particular, we omit many issues related to the implementation and numerical efficiency.

We concentrate on just a few issues. From the practical point of view, one question that interested us is whether the method depends on computing periodic orbits, (a task that seems to lead to be numerically difficult for some maps). In [Gre79], following [DV58], it was shown that the calculation of periodic orbits is much simpler in the case that the map is *reversible*. Nevertheless, even for standard mappings with two frequencies it was shown [LC06] that, for some parameter values, the periodic orbits appear in complicate orders and there are many periodic orbits so that continuation methods have difficulty following them. Similar calculations were done in [FdlL92a] for complex values. At the moment, it is not known if the smooth solutions for the models in statistical mechanics or in PDE's are approximated by periodic orbits.

From a more theoretical point of view it is interesting to know whether the method leads to conclusions after a finite computation (provided, of course, that one controls the round-off and truncation errors). In other words, whether one can turn the method into a computer assisted proof. Some of the methods discussed in this section exclude the existence of invariant tori under a finite computation. The methods presented here, allow to conclude existence after a finite computation.

As for the conditions of applicability, we discuss whether the method depends on the system to be positive definite (as is the usual case in variational methods) or whether it is enough that the derivative of the frequency with respect of the action is invertible. Similarly, some of the methods extend to higher dimensional symplectic maps while others do not. Some methods extend to long range interactions and PDE's and others do not.

From the point of view of theoretical Physics, it is interesting to know whether the methods behave well under renormalization. Of course, the fact that there is no renormalization theory so far does not mean that there could not be one in the future.

A comparison between the methods discussed above is included in following table. We refer to the headings of the sections for the meaning of the abbreviations of the names of the method.

	GM	OM	WOO	PNB	APB	VS	Present
Conclusions after finite computation	N	$\nexists$	$\nexists$	$\nexists$	$\nexists$	$\nexists$	$\exists$
Rigorous justification	$\Rightarrow$	$\Rightarrow$	$\Leftrightarrow$	$\Leftrightarrow$	$\Leftrightarrow$	$\Rightarrow$	$\Leftrightarrow$
Requires periodic orbits	Y	Y	Y	Y	N	N	N
Interpretation of the Renor. Group	Y	Y	N	?	?	Y	?
Requires Positive Def. twist	N	N	Y	Y	Y	Y	N
Long Range Models/ PDE	N	?	Y	Y	?	Y	Y
Higher dimensional Symplectic Maps and PDE	Y	?	N	N	N	?	Y

Acknowledgments

The work of both authors has been supported by NSF grants. The last version of the paper was prepared at Centre de Recerca Matemàtica (Bellaterra). RL enjoyed a FRA from UT Austin.

We thank Prof. Y. Sire, Dr. G. Hughet and Mrs. T. Blass, J. James for many discussions.

References

[Ada75] Robert A. Adams. *Sobolev spaces*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.

[Ang90] Sigurd B. Angenent. Monotone recurrence relations, their Birkhoff orbits and topological entropy. *Ergodic Theory Dynam. Systems*, 10(1):15–41, 1990.

[Arn08] Marie-Claude Arnaud. Fibrés de Green et régularité des graphes C^0-lagrangiens invariants par un flot de Tonelli. *Ann. Henri Poincaré*, 9(5):881–926, 2008.

[Ban89] V. Bangert. On minimal laminations of the torus. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 6(2):95–138, 1989.

[BC90] Alberto Berretti and Luigi Chierchia. On the complex analytic structure of the golden invariant curve for the standard map. *Nonlinearity*, 3(1):39–44, 1990.

[BCCF92] A. Berretti, A. Celletti, L. Chierchia, and C. Falcolini. Natural boundaries for area-preserving twist maps. *J. Statist. Phys.*, 66(5-6):1613–1630, 1992.

[BGM96] George A. Baker, Jr. and Peter Graves-Morris. *Padé approximants*, volume 59 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 1996.

[BH87] Philip L. Boyland and Glen Richard Hall. Invariant circles and the order structure of periodic orbits in monotone twist maps. *Topology*, 26(1):21–35, 1987.

[BM94] Alberto Berretti and Stefano Marmi. Scaling near resonances and complex rotation numbers for the standard map. *Nonlinearity*, 7(2):603–321, 1994.

[Bos86] Jean-Benoît Bost. Tores invariants des systèmes dynamiques hamiltoniens (d'après Kolmogorov, Arnol′d, Moser, Rüssmann, Zehnder, Herman, Pöschel, ...). *Astérisque*, (133-134):113–157, 1986. Seminar Bourbaki, Vol. 1984/85.

[CC08] A. Celletti and L. Chierchia. Quasi-periodic attractors and spin-orbit resonances. In *SPT 2007—Symmetry and perturbation theory*, pages 56–63. World Sci. Publ., Hackensack, NJ, 2008.

[CC09] R. Calleja and A. Celletti. Breakdown of invariant attractors for the dissipative standard map. *Preprint. Avaliable at mp_arc. http://www.ma.utexas.edu/mp_arc-bin/mpa?yn=09-124*, 2009.

[CCdlL09] R. Calleja, A. Celletti, and R. de la Llave. KAM theory for dissipative systems. *Manuscript*, 2009.

[CdlL98] A. Candel and R. de la Llave. On the Aubry-Mather theory in statistical mechanics. *Comm. Math. Phys.*, 192(3):649–669, 1998.

[CdlL09a] R. Calleja and R. de la Llave. Computation of the breakdown of analyticity in statistical mechanics models. *Preprint. Avaliable at mp_arc. http://www.ma.utexas.edu/mp_arc-bin/mpa?yn=09-56*, 2009.

[CdlL09b] Renato Calleja and Rafael de la Llave. Fast numerical computation of quasi-periodic equilibrium states in 1d statistical mechanics, including twist maps. *Nonlinearity*, 22(6):1311–1336, 2009.

[CF94] L. Chierchia and C. Falcolini. A direct proof of a theorem by Kolmogorov in Hamiltonian systems. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 21(4):541–593, 1994.

[CJ02] C. Chandre and H. R. Jauslin. Renormalization-group analysis for the transition to chaos in Hamiltonian systems. *Phys. Rep.*, 365(1):1–64, 2002.

[DdlL00] Amadeu Delshams and Rafael de la Llave. KAM theory and a partial justification of Greene's criterion for nontwist maps. *SIAM J. Math. Anal.*, 31(6):1235–1269 (electronic), 2000.

[dlL92] R. de la Llave. A renormalization group explanation of numerical observations of anal yticity domains. *J. Statist. Phys.*, 66(5-6):1631–1634, 1992.

[dlL08] Rafael de la Llave. KAM theory for equilibrium states in 1-D statistical mechanics models. *Annales Henri Poincare*, 9(5):835–880, 2008.

[dlLGJV05] R. de la Llave, A. González, À. Jorba, and J. Villanueva. KAM theory without action-angle variables. *Nonlinearity*, 18(2):855–895, 2005.

[dlLO06] Rafael de la Llave and Arturo Olvera. The obstruction criterion for non-existence of invariant circles and renormalization. *Nonlinearity*, 19(8):1907–1937, 2006.

[dlLP02] Rafael de la Llave and Nikola P. Petrov. Regularity of conjugacies between critical circle maps: an experimental study. *Experiment. Math.*, 11(2):219–241, 2002.

[dlLR91] R. de la Llave and D. Rana. Accurate strategies for K.A.M. bounds and their implementation. In *Computer aided proofs in analysis (Cincinnati, OH, 1989)*, volume 28 of *IMA Vol. Math. Appl.*, pages 127–146. Springer, New York, 1991.

[dlLT94a] Rafael de la Llave and Stathis Tompaidis. Computation of domains of analyticity for some perturbative expansions of mechanics. *Phys. D*, 71(1-2):55–81, 1994.

[dlLT94b] Rafael de la Llave and Stathis Tompaidis. Nature of singularities for analyticity domains of invariant curves. *Phys. Rev. Lett.*, 73(11):1459–1463, 1994.

[dlLT95] Rafael de la Llave and Stathis Tompaidis. On the singularity structure of invariant curves of symplectic mappings. *Chaos*, 5(1):227–237, 1995.

[dlLV07a] Rafael de la Llave and Enrico Valdinoci. Critical points inside the gaps of ground state laminations for some models in statistical mechanics. *J. Stat. Phys.*, 129(1):81–119, 2007.

[dlLV07b] Rafael de la Llave and Enrico Valdinoci. Ground states and critical points for generalized Frenkel-Kontorova models in $\mathbb{Z}^d$. *Nonlinearity*, 20(10):2409–2424, 2007.

[dlLV07c] Rafael de la Llave and Enrico Valdinoci. Multiplicity results for interfaces of Ginzburg-Landau-Allen-Cahn equations in periodic media. *Adv. Math.*, 215(1):379–426, 2007.

[Dou82] R. Douady. Une démonstration directe de l'équivalence des théorèmes de tores invariants pour difféomorphismes et champs de vecteurs. *C. R. Acad. Sci. Paris Sér. I Math.*, 295(2):201–204, 1982.

[DV58] René De Vogelaere. On the structure of symmetric periodic solutions of conservative systems, with applications. In *Contributions to the theory of nonlinear oscillations, Vol. IV*, Annals of Mathematics Studies, no 41, pages 53–84. Princeton University Press, Princeton, N.J., 1958.

[Eli96] L. H. Eliasson. Absolutely convergent series expansions for quasi periodic motions. *Math. Phys. Electron. J.*, 2:Paper 4, 33 pp. (electronic), 1996.

[Fat07] Albert Fathi. *Weak Kam Theorem In Lagrangian Dynamics*. Cambridge Studies In Advanced Mathematics. Cambridge University Press, Cambridge, 2007.

[FdlL92a] Corrado Falcolini and Rafael de la Llave. Numerical calculation of domains of analyticity for perturbation theories in the presence of small divisors. *J. Statist. Phys.*, 67(3-4):645–666, 1992.

[FdlL92b] Corrado Falcolini and Rafael de la Llave. A rigorous partial justification of Greene's criterion. *J. Statist. Phys.*, 67(3-4):609–643, 1992.

[FdlLS09] E. Fontich, R. de la Llave, and Y. Sire. Construction of invariant whiskered tori by a parameterization method. part I: maps and flows in finite dimensions. *Jour. Diff. Eq.*, 246(8), 2009.

[Gal94] Giovanni Gallavotti. Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable Hamiltonian systems. A review. *Rev. Math. Phys.*, 6(3):343–411, 1994.

[GEdlL08] A. González-Enríquez and R. de la Llave. Analytic smoothing of geometric maps with applications to KAM theory. *J. Differential Equations*, 245(5):1243–1298, 2008.

[Gre79] John M. Greene. A method for determining a stochastic transition. *Journal of Mathematical Physics*, 20(6):1183–1201, 1979.

[Ham82] Richard S. Hamilton. The inverse function theorem of Nash and Moser. *Bull. Amer. Math. Soc. (N.S.)*, 7(1):65–222, 1982.

[HdlLS09] Gemma Huguet, Rafael de la Llave, and Yannick Sire. Fast numerical algorithms for the computation of invariant tori in hamiltonian systems. *MP_ARC # 09-02*, 2009.

[Her83] Michael-R. Herman. *Sur les courbes invariantes par les difféomorphismes de l'anneau. Vol. 1*, volume 103 of *Astérisque*. Société Mathématique de France, Paris, 1983. With an appendix by Albert Fathi, With an English summary.

[Her89] Michael-R. Herman. Inégalités "a priori" pour des tores lagrangiens invariants par des difféomorphismes symplectiques. *Inst. Hautes Études Sci. Publ. Math.*, (70):47–101 (1990), 1989.

[Hör85] Lars Hörmander. On the Nash-Moser implicit function theorem. *Ann. Acad. Sci. Fenn. Ser. A I Math.*, 10:255–259, 1985.

[Jun91] Irwin Jungreis. A method for proving that monotone twist maps have no invariant circles. *Ergodic Theory Dynam. Systems*, 11(1):79–84, 1991.

[Koc04] Hans Koch. A renormalization group fixed point associated with the breakup of golden invariant tori. *Discrete Contin. Dyn. Syst.*, 11(4):881–909, 2004.

[Kol54] A. N. Kolmogorov. On conservation of conditionally periodic motions for a small change in Hamilton's function. *Dokl. Akad. Nauk SSSR (N.S.)*, 98:527–530, 1954.

[LC91] Patrice Le Calvez. Propriétés dynamiques des difféomorphismes de l'anneau et du tore. *Astérisque*, (204):131, 1991.

[LC06] Héctor E. Lomelí and Renato Calleja. Heteroclinic bifurcations and chaotic transport in the two-harmonic standard map. *Chaos*, 16(2):023117, 8, 2006.

[LM86] I. Leage and R. S. MacKay. Badly ordered periodic orbits of the standard map. *Phys. Lett. A*, 118(6):274–278, 1986.

[Mac82] R.S. MacKay. *Renormalisation in area preserving maps*. PhD thesis, Princeton Univ., 1982.

[Mac83] R. S. MacKay. A renormalisation approach to invariant circles in area-preserving maps. *Phys. D*, 7(1-3):283–300, 1983. Order in chaos (Los Alamos, N.M., 1982).

[Mac89] R. S. MacKay. A criterion for nonexistence of invariant tori for Hamiltonian systems. *Phys. D*, 36(1-2):64–82, 1989.

[Mac92] R. S. MacKay. Greene's residue criterion. *Nonlinearity*, 5(1):161–187, 1992.

[Mat82] John N. Mather. Existence of quasiperiodic orbits for twist homeomorphisms of the annulus. *Topology*, 21(4):457–467, 1982.

[Mat84] John N. Mather. Nonexistence of invariant circles. *Ergodic Theory Dynam. Systems*, 4(2):301–309, 1984.

[Mat86] John Mather. A criterion for the nonexistence of invariant circles. *Inst. Hautes Études Sci. Publ. Math.*, (63):153–204, 1986.

[Mat87] John N. Mather. Modulus of continuity for Peierls's barrier. In *Periodic solutions of Hamiltonian systems and related topics (Il Ciocco, 1986)*, volume 209 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pages 177–202. Reidel, Dordrecht, 1987.

[MF94a] John N. Mather and Giovanni Forni. Action minimizing orbits in Hamiltonian systems. In *Transition to chaos in classical and quantum mechanics (Montecatini Terme, 1991)*, volume 1589 of *Lecture Notes in Math.*, pages 92–186. Springer, Berlin, 1994.

[MF94b] John N. Mather and Giovanni Forni. Action minimizing orbits in Hamiltonian systems. In *Transition to chaos in classical and quantum mechanics (Montecatini Terme, 1991)*, volume 1589 of *Lecture Notes in Math.*, pages 92–186. Springer, Berlin, 1994.

[MMS89] R. S. MacKay, J. D. Meiss, and J. Stark. Converse KAM theory for symplectic twist maps. *Nonlinearity*, 2(4):555–570, 1989.

[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.

[Mos67] Jürgen Moser. Convergent series expansions for quasi-periodic motions. *Math. Ann.*, 169:136–176, 1967.

[Mos86] Jürgen Moser. Monotone twist mappings and the calculus of variations. *Ergodic Theory Dynam. Systems*, 6(3):401–413, 1986.

[MP85] R. S. MacKay and I. C. Percival. Converse KAM: theory and practice. *Comm. Math. Phys.*, 98(4):469–512, 1985.

[Nas56] John Nash. The imbedding problem for Riemannian manifolds. *Ann. of Math. (2)*, 63:20–63, 1956.

[OS87] Arturo Olvera and Carles Simó. An obstruction method for the destruction of invariant curves. *Phys. D*, 26(1-3):181–192, 1987.

[Rue99] David Ruelle. *Statistical mechanics*. World Scientific Publishing Co. Inc., River Edge, NJ, 1999. Rigorous results, Reprint of the 1989 edition.

[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. Notes by H. Fattorini, R. Nirenberg and H. Porta, with an additional chapter by Hermann Karcher, Notes on Mathematics and its Applications.

[Ser73] Francis 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).

[Tay97] Michael E. Taylor. *Partial differential equations. III*, volume 117 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1997. Nonlinear equations, Corrected reprint of the 1996 original.

[Tom96] Stathis Tompaidis. Approximation of invariant surfaces by periodic orbits in high-dimensional maps: some rigorous results. *Experiment. Math.*, 5(3):197–209, 1996.

[Van02] John A. Vano. *A Nash-Moser Implicit Function Theorem with Whitney Regularity and Applications*. PhD thesis, Univ. of Texas at Austin, 2002. MP_ARC # 02-276.

[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] Eduard Zehnder. Moser's implicit function theorem in the framework of analytic smoothing. *Math. Ann.*, 219(2):105–121, 1976.

Dept. of Mathematics, 1 University Station C1200, Austin TX 78712-0257

E-mail address: `rcalleja@math.utexas.edu`

Dept. of Mathematics, 1 University Station C1200, Austin TX 78712-0257

E-mail address: `llave@math.utexas.edu`