Quantum ergodicity in mixed and KAM Hamiltonian systems

Seán P Gomes

A thesis submitted for the degree of
Doctor of Philosophy of
The Australian National University

May, 2017
Declaration

This thesis is an account of research undertaken between August 2012 and May 2017 at the Mathematical Sciences Institute, The Australian National University, Canberra, Australia.

Except where acknowledged in the customary manner, the material presented in this thesis is, to the best of my knowledge, original and has not been submitted in whole or part for a degree in any university.

The bulk of the original work in this thesis is contained in Chapters 3 and 6. Chapters 1 and 2 are a summary of known background material and Chapters 4 and 5 closely follow the work in [36] and [37].

Seán P Gomes
May, 2016
Acknowledgements

First and foremost I would like to express my immense appreciation and gratitude to my thesis adviser, Professor Andrew Hassell, you have been a fantastic mentor and role model in these early stages of academic life. Your advice and support on matters both related to mathematics and career progress has been of immeasurable value to me.

I would also like to thank my committee members, Professors Ben Andrews and Xu-Jia Wang and the the HDR convenor Associate Professor Scott Morrison for making my thesis defence an enjoyable experience, and for providing useful feedback.

Thanks also go to Professor Georgi Popov for several enlightening email exchanges elaborating on aspects of his work in the KAM setting, and to Assistant Professor Semyon Dyatlov for a fruitful discussion that motivated a weakening of the “slow torus” condition in the the results of Chapter 6.

Thanks to the Australian federal government, whose funding via the Australian post-graduate award and research training program stipend have made this research possible.

Thanks to all of my friends and colleagues, for always encouraging me to strive for my goals.

A special thanks to Frank, Ronette, Karen and Tanya. The unconditional love, support, and encouragement that a family like ours provides is something that cannot be overstated.

Last but certainly not least, I must acknowledge my partner, Adeline. You have been a boundless source of personal support at the times when it was needed most. Being a thesis-widow is no easy burden, yet you have done everything in your power to support me on this journey and have rode the highs and lows alongside me, making even the most challenging obstacles seem surmountable.
Abstract

In this thesis, we investigate quantum ergodicity for two classes of Hamiltonian systems satisfying intermediate dynamical hypotheses between the well understood extremes of ergodic flow and quantum completely integrable flow. These two classes are mixed Hamiltonian systems and KAM Hamiltonian systems.

Hamiltonian systems with mixed phase space decompose into finitely many invariant subsets, only some of which are of ergodic character. It has been conjectured by Percival that the eigenfunctions of the quantisation of this system decompose into associated families of analogous character. The first project in this thesis proves a weak form of this conjecture for a class of dynamical billiards, namely the mushroom billiards of Bunimovich for a full measure subset of a shape parameter $t \in (0, 2]$.

KAM Hamiltonian systems arise as perturbations of completely integrable Hamiltonian systems. The dynamics of these systems are well understood and have near-integrable character. The classical-quantum correspondence suggests that the quantisation of KAM systems will not have quantum ergodic character. The second project in this thesis proves an initial negative quantum ergodicity result for a class of positive Gevrey perturbations of a Gevrey Hamiltonian that satisfy a mild slow torus condition.
Contents

6 Eigenvalue Localisation Results 105
   6.1 Introduction .................................................. 105
   6.2 Eigenvalue and quasieigenvalue variation .................. 109
   6.3 Spectral non-concentration ................................... 113

A Estimates for analytic functions 119

B Properties of anisotropic Gevrey classes 123

C Whitney extension theorem 125

D Miscellaneous 133

Bibliography 135
List of Figures

1.1 ................................................................. 4
1.2 ................................................................. 6
1.3 ................................................................. 9
3.1 ................................................................. 23
Chapter 1

Introduction to Quantum Ergodicity

The central objective in quantum chaos is to understand how chaotic dynamical assumptions about a classical mechanical system manifest themselves in the behaviour of its quantum mechanical analogue.

A natural setting for studying this correspondence is that of Hamiltonian flow on a compact Riemannian manifold $M$, and this is the setting of the original work in this thesis. In this setting, our dynamical assumption is based on the measure-theoretic concept of ergodicity.

We shall begin in Section 1.1 by summarising the aspects of the Hamiltonian formalism relevant to our work. A more comprehensive treatment can be found in the book [3]. In particular, we shall highlight the opposing concepts of ergodicity and complete integrability.

In Section 1.2 we introduce Schrödinger’s equation, the quantum mechanical counterpart to Hamilton’s equations. We shall then discuss the semiclassical formalism and its relevance to studying the classical-quantum correspondence.

In Section 1.3 we define the quantum mechanical analogue to ergodicity of Hamiltonian flow and survey the major results and conjectures in this field.

In Section 1.4 we discuss the quantisations of Hamiltonian systems that are either completely integrable, or are small perturbations of completely integrable systems. As the Hamiltonian flow in these settings is far from ergodic, intuition suggests that the eigenfunctions for such a system will be far from equidistributed.

1.1 Hamiltonian flow

Suppose that we have a smooth $n$-dimensional compact Riemannian manifold $(M, g)$ (possibly with boundary). Given a smooth Hamiltonian function $H : T^*M \rightarrow \mathbb{R}$ which we interpret as an energy, we obtain the Hamiltonian flow $\Phi^t : T^*M \times [0, \infty) \rightarrow T^*M$ generated by Hamilton’s equations

$$
\dot{x} = -\nabla_x H(x, \xi) \quad \dot{\xi} = \nabla_\xi H(x, \xi)
$$

(1.1.1)

with coordinates $(x, \xi)$ corresponding to the cotangent vector $\sum_j \xi_j \, dx_j$. In this work we shall assume that our Hamiltonians are such that the flow $\Phi^t$ does not blow up in finite time. We denote the Hamiltonian vector field given by (1.1.1) as $X_H$.

The primary Hamiltonians of interest in this thesis will be Schrödinger Hamiltonians...
of the form
\[ H(x, \xi) = ||\xi||_g^2 + V(x, \xi), \quad (1.1.2) \]
however the results of Chapter 6 could be generalised to symbols of more general classes of self-adjoint pseudodifferential operators. In Chapter 5, the function \( V \) is a compactly supported symbol in the Gevrey class \( S_\ell \) from Definition 2.2.5 with self-adjoint quantisation. In the special case of \( V = 0 \) as in Chapter 3, this Hamiltonian system is referred to as billiards on \( M \). The trajectories of billiard flow can be identified with the geodesics of \( M \) under the canonical isomorphism between the tangent and co-tangent bundles of Riemannian manifolds.

A major advantage of the Hamiltonian formulation of mechanics over the Newtonian and Lagrangian formulations is the duality of the variables \( x \) and \( \xi \), best highlighted through the lens of symplectic geometry.

**Definition 1.1.1** (Symplectic Form). A symplectic form \( \omega \) on a smooth manifold \( M \) is a closed non-degenerate differential 2-form.

**Definition 1.1.2** (Symplectomorphism). A symplectomorphism \( \chi : N_1 \to N_2 \) between two symplectic manifolds \( (N_1, \omega_1) \) and \( (N_2, \omega_2) \) is a diffeomorphism such that \( \chi^* \omega_2 = \omega_1 \).

**Definition 1.1.3.** An exact symplectic form \( \omega \) on a smooth manifold \( M \) is an exact non-degenerate differential 2-form.

**Definition 1.1.4** (Exact Symplectomorphism). A symplectomorphism \( \chi : N_1 \to N_2 \) between two exact symplectic manifolds \( (N_1, d\alpha_1) \) and \( (N_2, d\alpha_2) \) is a diffeomorphism such that \( \chi^* \alpha_2 - \alpha_1 \) is an exact 1-form.

Given a smooth function \( H \) on a smooth manifold \( N \) equipped with a symplectic form \( \omega \), we can obtain a Hamiltonian vector field \( X_H \) on \( N \) defined implicitly by
\[ dH(Y) = \omega(X_H, Y). \quad (1.1.3) \]

In the case \( N = T^*M \) there is a canonical choice of symplectic form
\[ \omega = d\xi \wedge dx = \sum_j d\xi_j \wedge dx_j \quad (1.1.4) \]
and the vector field \( X_H \) generates the flow given by Hamilton’s equations (1.1.1).

In fact for a general symplectic manifold \( (N, \omega) \), Darboux’s theorem asserts that local coordinates \( (x, \xi) \) can be chosen such that (1.1.4) holds. Such coordinates are said to be canonical coordinates.

Writing \( z = (x, \xi), w = (y, \eta) \) in a canonical coordinate system on \( (N, \omega) \) allows us to write
\[ \omega(z, w) = \langle Jz, w \rangle \quad (1.1.5) \]
where
\[ J := \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \in \mathbb{R}^{2n \times 2n}. \quad (1.1.6) \]

Indeed we say that a matrix \( A \) is symplectic if
\[ JA^TJ = A^{-1}, \quad (1.1.7) \]
and a diffeomorphism on a symplectic manifolds is a symplectomorphism if and only if its Jacobian with respect to canonical coordinates is a symplectic matrix.

Symplectomorphisms are the natural class of coordinate transformations of a Hamiltonian system to work with as they preserve Hamilton’s equations.

**Proposition 1.1.5.** If \( \chi : N_1 \to N_2 \) is a symplectomorphism, and \( H : N_2 \to \mathbb{R} \) is a smooth Hamiltonian then the transformed Hamiltonian \( \tilde{H} := H \circ \chi : N_1 \to \mathbb{R} \) generates a Hamiltonian flow in the coordinates \((y, \eta) = \chi^{-1}(x, \xi)\) given by

\[
\dot{y} = -\nabla_y \tilde{H}(y, \eta) \quad \dot{\eta} = \nabla_\eta \tilde{H}(y, \eta)
\]  

(1.1.8)

The Hamiltonian vector field \( X_{\tilde{H}} \) on \( N_1 \) is the pullback of the Hamiltonian vector field \( X_H \) on \( N_2 \).

Hamiltonian flows give rise to symplectomorphisms in a natural way.

**Proposition 1.1.6.** If \( H \) is a Hamiltonian on the symplectic manifold \( N \), then the Hamiltonian flow \( \Phi_t \) on \( N \) is a one-parameter family of symplectomorphisms.

Another useful method of constructing symplectomorphisms is through the use of a generating function.

**Proposition 1.1.7.** If \( \Phi(x, \xi) \in \mathcal{C}^\infty(N) \) is such that the Hessian \( \partial^2_{x, \xi} \Phi \) is non-singular, then a solution to the implicit equation

\[
\chi(\partial_\xi \Phi(x, \xi), \xi) = (x, \partial_x \Phi(x, \xi))
\]  

(1.1.9)

is symplectic on its domain.

For the particularly simple symplectic manifold \( T^n \times D \), with \( D \subset \mathbb{R}^n \), we can make this construction global, provided that \( \Phi(x, \xi) - \langle x, \xi \rangle \) is \( 2\pi \)-periodic in \( x \). The resulting symplectomorphism \( \chi \) is then an exact symplectomorphism.

Provided that all \( E \in [a, b] \) are regular values for the Hamiltonian \( H \), the canonical symplectic form \( d\xi \wedge dx \) on \( T^*M \) determines a family of measures \( \mu_E \) on each of the energy hypersurfaces

\[
\Sigma_E = \{ z = (x, \xi) \in T^*M : H(z) = E \}
\]  

(1.1.10)

defined implicitly by

\[
\int_a^b \int_{\Sigma_E} f \, d\mu_E \, dE = \int_{\{(x, \xi) : \eta^{-1} \in [a, b]\}} f \, d\xi \, dx
\]  

(1.1.11)

for \( f \in \mathcal{C}^c_c(T^*M) \). In the special case of of \( H = ||\xi||_g^2 \) and \( E = 1 \), upon normalisation we obtain the Liouville measure \( \mu_L \) on \( S^*M \).

The measures \( \mu_E \) allow us to study the ergodic properties of the Hamiltonian flow \( \phi_t \).

**Definition 1.1.8.** If \( \phi^t : X \to X \) denotes a measure preserving flow on a finite measure space \((X, A, \mu)\), we say that \( \phi^t \) is ergodic if

\[
\lim_{T \to \infty} \frac{m(\{t \in [0, T] : \phi^t(x) \in A\})}{T} = \frac{\mu(A)}{\mu(X)} \quad \text{for each } A \in A
\]  

(1.1.12)

for \( \mu \)-almost all \( x \in X \).
That is to say, a flow is ergodic if and only if almost all trajectories equidistribute in the measure space. An equivalent characterisation can be made in terms of flow invariant subsets.

**Proposition 1.1.9.** A measure preserving flow $\phi_t$ on a finite measure space $(X, \mathcal{A}, \mu)$ is ergodic if and only if the only $\phi_t$-invariant measurable sets are of full measure or null measure.

In particular, we say that the Hamiltonian flow $\Phi^t$ generated by $H : T^*M \to \mathbb{R}$ is ergodic on the energy surface $\Sigma_E = H^{-1}(E)$ if $\Phi^t$ satisfies 1.1.8 on $\Sigma_E$ with respect to the measure $\mu_E$.

Two particularly famous examples of ergodic Hamiltonian systems are the Bunimovich stadium billiard [6] and the Sinai billiard [42].

![Figure 1.1: An ergodic trajectory on the Bunimovich stadium billiard. Image from [45].](image)

From Definition 1.1.8, we can see that for ergodic flows, the time average of a smooth classical observable $a : T^*M \to \mathbb{R}$ tends to its space average. That is, we have

$$
\lim_{T \to \infty} \frac{1}{T} \int_0^T (a \circ \phi_t)(x_0, \xi_0) \, dt \to \int_{\Sigma_E} a \, d\mu_E
$$

for $\mu_E$-almost all $(x_0, \xi_0) \in \Sigma_E$.

A strictly stronger property of a flow is the mixing property which asserts that for smooth classical observables $a, b : T^*M \to \mathbb{R}$ we have

$$
\int_{\Sigma_E} (a \circ \phi_t) \cdot b \, d\mu_E \to \left( \int_{\Sigma_E} a \, d\mu_E \right) \cdot \left( \int_{\Sigma_E} b \, d\mu_E \right)
$$

as $t \to \infty$. The strong assumption of Anosov flow leads to (1.1.14) with an exponential rate of convergence. Thus, manifolds with non-positive sectional curvature also give rise to ergodic billiards. In this thesis we shall not discuss the stronger property of (1.1.14), and restrict ourself to the study of ergodicity.

In order to study the Hamiltonian evolution of functions on our phase space $T^*M$, we define the Poisson bracket.

**Definition 1.1.10 (Poisson Bracket).** If $f, g \in C^\infty(T^*M)$, we define

$$
\{f, g\} := \omega(X_f, X_g) = \sum_{j=1}^n \frac{\partial f}{\partial x_j} \frac{\partial g}{\partial \xi_j} - \frac{\partial f}{\partial \xi_j} \frac{\partial g}{\partial x_j}
$$

(1.1.15)
An immediate consequence of the chain rule is that if \((\xi(t), x(t))\) is a trajectory of the Hamiltonian flow, then for a smooth function \(f : T^*M \to \mathbb{C}\), we have

\[
\frac{d}{dt} (f(\xi(t), x(t))) = \{f, H\}(\xi(t), x(t)).
\] (1.1.16)

Motivated by this calculation we can define invariants of our flow.

**Definition 1.1.11.** An invariant or first integral of the Hamiltonian flow \(X_H\) is a smooth function \(f : T^*M \to \mathbb{C}\) such that \(\{f, H\} = 0\).

The Hamiltonian \(H\) itself is of course an invariant of the flow that it generates. Hence Hamiltonian flow is constrained to energy shells \(\Sigma_E = H^{-1}(\{E\})\). Often we can find additional invariants that are mathematical manifestations of conservation laws from physics, such as that of angular momentum. Symmetries in a system lead to an abundance of such flow invariants, as follows from Noether’s theorem (See Chapter 4, Section 20 of \[3\]).

We say that a collection of flow invariants \(\{f_i\}\) are in involution if their pairwise Poisson brackets vanish and we say that they are independent if their differentials \(df_i\) are linearly independent.

**Definition 1.1.12.** If an invariant subset of Hamiltonian flow admits \(n\) independent invariants that are in involution, we say that the corresponding Hamiltonian system is completely integrable.

In a completely integrable Hamiltonian system, trajectories are constrained to \(n\)-dimensional submanifolds and are thus far from equidistributed on the \((2n-1)\)-dimensional energy surfaces. The Liouville–Arnold theorem from classical mechanics asserts that for completely integrable systems, we can find a neighbourhood \(U\) of an arbitrary invariant manifold \(\Lambda \subset T^*M\) and a symplectomorphism \(\chi : \Lambda \to \mathbb{T}^n \times D\) for some \(D \subset \mathbb{R}^n\) such that the transformed Hamiltonian \(\tilde{H}(\theta, I) = H(\chi(\theta, I))\) is independent of \(\theta\). Thus the invariant manifolds are diffeomorphic to \(n\)-dimensional tori. Moreover, the Hamiltonian flow is quasiperiodic, with trajectories given by

\[
I(t) = I(t_0); \quad \theta(t) = \theta(t_0) + t\nabla I \tilde{H}
\] (1.1.17)

in the \((\theta, I)\)-coordinates, referred to as action-angle variables. A construction of these coordinates can be found in Section 50 of \[3\].

We note that the invariant tori of a completely integrable Hamiltonian system are Lagrangian, that is the restriction of the symplectic form to any of the invariant tori \(\Lambda_c = \{(x, \xi) : f(x, \xi) = c \in \mathbb{R}^n\}\) vanishes.

An example of a completely integrable Hamiltonian system is geodesic billiards on an ellipsoid, pictured in Figure 1.2.

We can extend the above discussion to manifolds with boundary by extending Hamiltonian flow by reflection at non-tangential boundary collisions. We shall postpone this somewhat technical discussion until we require it in Chapter 3.

### 1.2 Quantum dynamics

The quantum mechanical analogue of the system (3.1.1) is the evolution of a wave-function \(\psi \in L^2(M)\) governed by Schrödinger’s equation.
Introduction to Quantum Ergodicity

Figure 1.2: Geodesic flow on an ellipsoid. Image generated using GeographicLib.

\[
(i \hbar \frac{\partial}{\partial t} - \hbar^2 \Delta) \psi = 0 \tag{1.2.1}
\]

where \( \hbar \approx 6.626 \times 10^{-34} \text{ m}^2 \text{ kg s}^{-1} \) (1.2.2) is Planck’s constant and

\[
\Delta = -\text{div grad} \tag{1.2.3}
\]

is the Laplace–Beltrami operator with the positive sign convention.

The Bohr correspondence principle asserts that a classical Hamiltonian system is in a vague sense the macroscopic or high-energy limit of the corresponding quantum dynamical system. By scaling the units in (1.2.1), we may instead consider \( \hbar \) to be a small parameter in our problem. This is known as the semiclassical formalism. In the semiclassical formalism we replace differential operators with semiclassical differential operators \( \hbar \frac{\partial}{\partial x_j} \) to account for the scaling factor. By the Bohr correspondence principle, we then expect chaotic behaviour of the classical system to be manifest in the quantum system in the limit \( \hbar \to 0 \).

In solving (1.2.1), we can expand \( \psi \) in terms of the basis of \( L^2(M) \) comprised by eigenfunctions of \( \hbar^2 \Delta \), and so localisation of quantum dynamics can be understood by the study of the localisation of these eigenfunctions.

For Hamiltonian systems that are more general than billiards, such as the Schrödinger type Hamiltonians in (1.1.2), we consider eigenfunction of the semiclassical Schrödinger operator

\[
P_h = H(x, \hbar D) \tag{1.2.4}
\]

obtained by formally applying the Hamiltonian function to the operators \( x \) and \( \hbar D = \frac{\hbar \partial}{\partial x} \).

One can then ask how the localisation properties of the Hamiltonian flow on \( M \) are reflected by the spectral theory of the associated Schrödinger operator \( P_h \) in the limit \( \hbar \to 0 \).

Typically it is not possible to find exact, or even approximate expressions for chaotic eigenfunctions. Nevertheless, the machinery of microlocal analysis allows us to rigorously prove state and prove phase space equidistribution properties. The key tools here are pseudodifferential operators and Fourier integral operators, which correspond to quantisations of classical observables and symplectomorphisms respectively.
1.3 Quantum ergodicity

Under the assumption of ergodic Hamiltonian flow, one can make a remarkable statement of phase space equidistribution of the steady-state solutions to the corresponding Schrödinger’s equation (3.1.5).

A primitive version of this theorem asserts that the sequence of probability measures $|u_j|^2$ on $M$ must have a full density subsequence which tends to the uniform measure.

The stronger statement of phase space equidistribution requires some additional machinery to state, such as the pseudodifferential calculus which we introduce in Chapter 2.

**Theorem 1.3.1** (Quantum Ergodicity). If the Hamiltonian $H(x, hD) = h^2 \Delta + V(x)$ on the smooth compact boundaryless Riemannian manifold $(M, g)$ generates ergodic flow on the regular energy band $H^{-1}([a, b])$ for a smooth real potential $V(x)$ that is bounded below, then there exists a family of subsets $\Lambda(h) \subset [a, b]$ of eigenvalues of $P_h$ such that

$$\lim_{h \to 0} \#\Lambda(h) = 1 \quad (1.3.1)$$

and

$$\langle A(x, hD)u_j(h), u_j(h) \rangle \to \int_{H^{-1}([a, b])} A(x, \xi) d\xi dx \quad (1.3.2)$$

uniformly for $j \in S(h)$ for any zero-th order semiclassical pseudodifferential operator $A(x, hD)$ with the property that

$$\frac{1}{\mu_E} \int_{\Sigma_E} A(x, \xi) d\mu_E \quad (1.3.3)$$

is independent of $E$.

A semiclassical proof of this theorem can be found in [50], whilst the initial result goes back to [43],[48],[9]. For billiards on manifolds with boundary, there are additional technical considerations even from the purely dynamical perspective. Nevertheless, the quantum ergodicity theorem generalises to this setting [49],[17].

We can also define a notion of quantum ergodicity localised to an individual energy surface that is motivated by the results of [23]. We will make use of this definition in the proof of the negative quantum ergodicity result Theorem 6.1.3, in Chapter 6.

Suppose the semiclassical pseudodifferential operator $P_h = h^2 \Delta + V(x, hD)$ on the smooth manifold $M$ has principal symbol $p(x, \xi) \in C^\infty(T^*M)$ and has purely point spectrum, with eigenpairs $(u_j(h), E_j(h))$ in increasing order.

If $E \in \mathbb{R}$ is a regular value of $p$ with nonempty preimage, then we can define quantum ergodicity localised to the energy surface $\Sigma_E$ as follows.

**Definition 1.3.2.** We say that $P_h$ is quantum ergodic at energy $E$ if for each sufficiently small $h > 0$, there exists a family $S(h) \subset \{j \in \mathbb{N} : E_j(h) \in [E - h, E + h]\}$ such that

$$\frac{\#S(h)}{\#\{j \in \mathbb{N} : E_j(h) \in [E - h, E + h]\}} \to 1 \quad (1.3.4)$$

and

$$\langle A(x, hD)u_j(h), u_j(h) \rangle \to \int_{\Sigma_E} A(x, \xi) d\mu_E \quad (1.3.5)$$
uniformly for \( j \in S(h) \) for any zero-th order semiclassical pseudodifferential operator \( A(x, hD) \).

**Remark 1.3.3.** We could replace the \( O(h) \) width energy windows in Definition 1.3.2 with \( O(h^\beta) \) for \( 0 < \beta < 1 \). The \( O(h) \) size energy windows in Definition 1.3.2 are the smallest windows in which quantum ergodicity results are possible because the Weyl asymptotics break down in smaller energy windows.

An alternate formulation of quantum ergodicity can be made in the language of **semiclassical measures**. We shall state this version of quantum ergodicity in the special case \( H = \|\xi\|^2_g \) as we only make use of it in this setting of billiards in Chapter 3.

To each subsequence of \((u_j)\), we can associate at least one non-negative Radon measure \( \mu \) on \( S^*M \) which provides a notion of phase space concentration in the semiclassical limit. We say that the eigenfunction subsequence \((u_{jk})\) has unique semiclassical measure \( \mu \) if

\[
\lim_{k \to \infty} \langle a(x, E_j^{-1/2}D)u_{jk}, u_{jk} \rangle = \int_{S^*M} a(x, \xi) \, d\mu
\]

for each semiclassical pseudodifferential operator with principal symbol \( a \) compactly supported supported away from the boundary of \( S^*M \). In Chapter 5 of [50], the existence and basic properties of semiclassical measures are established using the calculus of semiclassical pseudodifferential operators (see also [17]).

A billiard \( M \) can then be said to be **quantum ergodic** if there is a full density subsequence of eigenfunctions \((u_{nk})\) such that the the Liouville measure on \( S^*M \) is the unique semiclassical measure associated to the sequence \((u_{nk})\). This statement can be interpreted as saying that the sequence of eigenfunctions equidistributes in phase space with the possible exception of a sparse subsequence.

Under the stronger assumption of Anosov flow, it is conjectured that the full sequence of eigenfunctions equidistributes in the sense of Theorem 1.3.1. This is known as the quantum unique ergodicity conjecture.

The prize-winning work of Lindenstrauss [29] verified this conjecture in certain arithmetic cases where we work with the Hecke joint eigenfunctions. The study of quantum ergodicity where we have this additional arithmetic structure is known as arithmetic QE. Sarnak [40] has written a survey on the recent developments in this field.

On the other hand, it is known that quantum ergodicity is strictly weaker than quantum unique ergodicity. Indeed, Hassell [22] showed that on the Bunimovich stadium there exist semiclassical measures that have positive mass on the union of the bouncing ball trajectories in phase space.

### 1.4 Negative results

In the extreme case of quantum unique ergodicity, there is a unique semiclassical measure, which is the Liouville measure. It is natural to ask what we can say about the semiclassical measures associated with sequences of eigenfunctions of Hamiltonian systems that are not quantum uniquely ergodic.

For the Bunimovich stadium, the quantum ergodicity theorem implies that any non-uniform limit can only arise from a density-zero subsequence.

Whilst Burq–Zworski [8] showed that concentration in a strict subrectangle is not possible, numerical evidence suggests that there could well be a sparse sequence of eigenfunctions with semiclassical limit supported in the rectangle itself. Rigorous proof of this
phenomenon remains an open problem, with the most notable progress being Hassell’s proof that a semiclassical measure exists with positive mass on the union of bouncing ball trajectories [22].

![Figure 1.3: An apparent “bouncing ball” eigenfunction in the quarter stadium corresponding to the eigenvalue $\approx 2859.47$. Image courtesy of Dr. Barnett.](image)

On the other hand, if a Hamiltonian system is assumed to be completely integrable, any trajectory is constrained to a single invariant torus corresponding to the intersection of the level sets of the $n$ conserved quantities. The intuition stemming from the classical-quantum correspondence suggests that this extreme concentration of trajectories should manifest itself in a statement about the concentration of eigenfunctions onto the Lagrangian tori. Such a result is proven in [46] for systems satisfying a stronger notion of quantum integrability and tori satisfying a certain nonresonance condition, however rigorous results in the general setting of complete integrability seem to be elusive.

At this point we introduce the notion of approximate eigenfunctions, or quasimodes.

**Definition 1.4.1.** Given a semiclassical pseudodifferential operator $P(h)$, a $O(h^\beta)$ quasimode is a family of functions $u_h \in L^2(M)$ such that

$$\|(P(h) - E)u_h\|_{L^2(M)} = O(h^\beta)$$  \hspace{1cm} (1.4.1)

for some $\beta > 0$ and some real $E$, referred to as the quasi-eigenvalue.

**Remark 1.4.2.** As a consequence of the semiclassical rescaling, it should be noted that $O(h^2)$ for the semiclassical Laplacian $h^2\Delta$ correspond to $O(1)$ quasimodes of the Laplacian $\Delta$.

We can of course replace the $O(h^\beta)$ with a stronger bound in this definition. The uses of quasimodes are plentiful. Most results about eigenfunctions apply just as well to quasimodes, and it is easier to construct quasimodes than exact eigenfunctions. In fact we can often construct quasimodes with desirable localisation properties, as they are better behaved with respect to taking cutoffs than exact eigenfunctions which are generally destroyed. The

In [10], Colin de Verdière established that for completely integrable Hamiltonian systems, there exist quasimodes with exponentially small error term that localise onto the certain individual invariant Langrangian tori. This result relies on the construction of a quantum Birkhoff normal form.

### 1.5 Mixed and KAM systems

Between the extremes of completely integrable Hamiltonian dynamics and ergodic Hamiltonian dynamics, results are rather sparse. The original work in this thesis explores two
intermediate classes of Hamiltonian dynamical systems for which questions of quantum ergodicity are tractable.

The two main original results in this thesis, Theorem 3.5.4 and Theorem 6.1.3, both make use of known quasimode constructions and perturbation arguments in order to prove eigenfunction localisation statements in the cases of mixed billiards and KAM system respectively. We now outline these two classes of Hamiltonian systems.

**Mixed billiards**

If a dynamical billiard can be separated into multiple invariant subsets, only some of which are ergodic, it is said that the billiard is *mixed*. In this case, it is conjectured that we can divide the sequence of eigenfunctions into corresponding families, with the eigenfunctions corresponding to an ergodic invariant subset satisfying a suitable equidistribution property. For the sake of simplicity, we shall state the conjecture in the case of billiards with exactly two invariant subsets, one ergodic and one completely integrable.

**Conjecture 1.5.1** (Percival’s Conjecture). For every compact Riemannian manifold $M$ such that $S^*M$ is the disjoint union of two invariant subsets $U, S^*M \setminus U$, with $U$ ergodic and $S^*M \setminus U$ completely integrable, we can find two subsets $A, B \subset \mathbb{N}$ such that

1. $A \cup B$ has density 1;
2. $(u_k)_{k \in A}$ equidistributes in the ergodic region $U$;
3. Each semiclassical measure associated to the subset $B$ is supported in the completely integrable region $D \setminus U$;
4. The density of $A$ is equal to $\mu_L(U)$.

Numerical evidence [4] strongly supports this conjecture, but no rigorous proofs have been discovered, even for concrete examples.

A weaker version of Percival’s conjecture is formulated by slightly relaxing the density requirements of the subsets $A$ and $B$.

**Conjecture 1.5.2** (Weak Percival’s Conjecture). For every compact Riemannian manifold $M$ such that $S^*M$ is the disjoint union of two invariant subsets $U, S^*M \setminus U$, with $U$ ergodic and $S^*M \setminus U$ completely integrable, we can find two subsets $A, B \subset \mathbb{N}$ such that

1. $A \cup B$ has upper density 1
2. $(u_k)_{k \in A}$ equidistributes in the ergodic region $U$
3. Each semiclassical measure associated to the subset $B$ is supported in the completely integrable region $S^*M \setminus U$
4. The upper densities of $A$ and $B$ are equal to $\mu_L(U)$ and $1 - \mu_L(U)$ respectively.

In Chapter 3, I provide the first verification of the weak Percival’s conjecture for a family of “mushroom” billiards, defined in Section 3.1. The main result is

**Theorem 1.5.3.** Conjecture 1.5.2 holds for the mushroom billiard $M_t$ for any fixed inner and outer radii, and almost all “stalk lengths” $t \in (0,2]$. 
KAM Hamiltonian systems

A particularly interesting class of Hamiltonian systems arise if we apply a small perturbation to a completely integrable real analytic Hamiltonian in action-angle coordinates.

\[ H(\theta, I) := H^0(I) + \epsilon H^1(\theta, I). \]  

(1.5.1)

Motivated by the geometry of completely integrable systems, where our phase space \( \mathbb{T}^n \times D \) is foliated by the Lagrangian tori

\[ \Lambda_{\omega} := \{(\theta, I) \in \mathbb{T}^n \times D : \theta \in T^n\} \]  

(1.5.2)

where \( \nabla_I H^0 = \omega \), it is natural to ask whether there are any such invariant Lagrangian tori that survive the perturbation. This problem is one of real physical significance, as one application is the study of celestial stability by viewing the dynamics of the solar system as a small perturbation of the completely integrable system that results from neglecting forces between pairs of planets. This perturbation is of course “small” because of the considerably greater mass of the sun compared to the planets.

The initial significant breakthrough in this problem was due to Kolmogorov [28], with the conclusion that although a dense set of tori is indeed generally destroyed by such a perturbation, a large measure collection of the invariant tori survive, precisely those whose frequency \( \omega = \nabla_I H^0 \) of quasiperiodic flow (1.1.17) satisfy the Diophantine condition

\[ |\langle \omega, k \rangle| \geq \frac{\kappa}{|k|^\tau} \]  

(1.5.3)

for all nonzero \( k \in \mathbb{Z}^n \) and fixed \( \kappa > 0 \) and \( \tau > n - 1 \). The tori satisfying this Diophantine condition are said to be nonresonant.

The field of KAM theory developed from this problem as a broad class of techniques applicable to perturbation problems in classical mechanics, founded by Kolmogorov, Arnold and Moser.

More recent work by Popov [36] proved a version of the KAM theorem for Hamiltonian systems in the Gevrey regularity class, with the purpose of constructing a Birkhoff normal form. This led to a quantum Birkhoff normal form construction, and a proof of the existence of quasimodes with exponentially small error localising onto the nonresonant tori in [37].

The details of Popov’s construction are summarised in Chapter 4 and Chapter 5 for families of Hamiltonians of the form

\[ H(x, \xi; t) = \|\xi\|^2 + V(x, \xi) + tQ(x, \xi) \]  

(1.5.4)

for smooth real-valued symbols \( V, Q \) in a suitable Gevrey class \( S_\ell(T^* M) \) in the notation of (2.2.5).

In Chapter 6, we prove the following main result. The formal statement is Theorem 6.1.3.

**Theorem 1.5.4.** Suppose \( M \) is a compact boundaryless Gevrey smooth Riemannian manifold, and the perturbation \( Q \) is such that \( Q(x, hD) \) is a positive operator and there exists a slow torus in the energy band \([a, b]\).

Then there exists \( \delta > 0 \) such that for almost all \( t \in [0, \delta] \) the quantisation \( P_h(t) \) of \( H(x, \xi; t) \) is non-quantum ergodic over the energy surface \( \Sigma_E \) for a positive Lebesgue
measure subset of energies $E \in [a, b]$.

A slow torus, defined formally in 6.1.1, is an invariant nonresonant Lagrangian torus $\Lambda$ in the energy surface $\mathcal{E} := \{(x, \xi) \in T^*M : \|\xi\|^2_\theta + V(x, \xi) = E\}$ such that the average of $Q$ over $\Lambda$ is strictly smaller than the average of $Q$ over $\mathcal{E}$. The assumption of the existence of such a torus is a mild one, and will typically be satisfied by perturbations whose symbols are nonconstant on energy surfaces.
Semiclassical Analysis

In this chapter, we briefly collect some of the necessary machinery of the semiclassical pseudodifferential calculus necessary for the results in the remainder of this thesis. The Fourier transform on \( \mathbb{R}^n \) from classical harmonic analysis allows us to pass from the \( n \)-dimensional position space to the \( n \)-dimensional frequency space. Pseudodifferential operators allow us to formulate and prove statements in the full \( 2n \)-dimensional phase space, and generalise the procedure of using a Fourier multiplier as a frequency cutoff. Some standard references for the classical pseudodifferential calculus include [18] [41] [25]. Our presentation shall be in the semiclassical formalism, for which an extensive account can be found in [50] and [12].

A key application of the pseudodifferential calculus to spectral theory is Weyl’s law, which provides an asymptotic for eigenvalues of a Schrödinger operator.

### 2.1 Semiclassical pseudodifferential operators

We begin by presenting the semiclassical pseudodifferential calculus on \( \mathbb{R}^{2n} \).

The semiclassical pseudodifferential calculus provides a correspondence between *classical observables* (smooth functions on the phase space \( \mathbb{R}^{2n} \)) and *quantum observables* (integral operators on position space \( \mathbb{R}^n \)).

The classical observables in this correspondence are traditionally referred to as symbols, and estimates on their derivatives are required to obtain desirable mapping properties for their quantisations.

For such operators, we can define semiclassical pseudodifferential operators on \( \mathbb{R}^n \).

\[
 a(x, hD)u(x) := (2\pi h)^{-n} \int \int e^{i(x-y) \cdot \xi / h} a(x, \xi; h) u(y) \, dy \, d\xi
\]  

(2.1.1)

The quantisation (2.1.1) is referred to as the standard quantisation. It is sometimes more convenient to work with a formally self-adjoint operator however. This motivates the definition of the Weyl quantisation.

\[
 a^w(x, hD)u(x) = (2\pi h)^{-n} \int \int e^{i(x-y) \cdot \xi / h} a\left(\frac{x+y}{2}, \xi; h\right) u(y) \, dy \, d\xi
\]  

(2.1.2)

which is a formally self-adjoint operator if \( a \) is real.

These formally defined integral operators are clearly convergent if \( a \) and \( u \) are of Schwartz class, but is otherwise understood in the sense of oscillatory integrals ([50] Theorem 3.8). In this fashion, the Kohn–Nirenberg symbol class leads to a class of semiclassical pseudodifferential operators bounded on semiclassical Sobolev spaces.
The standard class of $m$-th order Kohn–Nirenberg symbols on $\mathbb{R}^{2n}$ is given by:

$$S^m := \{ a(x, \xi) \in C^\infty(\mathbb{R}^{2n}) : |\partial^\alpha_x \partial^\beta_\xi a| \leq C_{\alpha, \beta} (1 + |\xi|^2)^{m-|\beta|/2} \text{ for all } \alpha, \beta \}.$$ \hfill (2.1.3)

**Remark 2.1.1.** It is also sometimes useful to consider classes $S^m_\delta$, where the right-hand side of (2.1.3) is multiplied by $\hbar^{-\delta}$ with each differentiation, but we shall not require these symbol classes.

This is a sufficiently broad symbol class for most applications, and includes semiclassical differential operators

$$\sum_{|\alpha| \leq n} a_\alpha(x)(\hbar D)^\alpha$$ \hfill (2.1.4)

as a special case by quantising polynomials in $\xi$.

The index $m$ in Definition 2.1.3 corresponds to the mapping properties of the associated pseudodifferential operator.

**Proposition 2.1.2.** If $a \in S^m$, then

$$a(x, \hbar D) : H^s \to H^{s-m}$$ \hfill (2.1.5)

is a bounded operator for any $s \in \mathbb{R}$, where $H^s$ denotes the Sobolev space of order $s$. In particular, zero-th order semiclassical pseudodifferential operators are bounded on $L^2$, and negative order semiclassical pseudodifferential operators are compact on $L^2$.

In practice, symbols of semiclassical pseudodifferential operators are often constructed using formal power series in $\hbar$. Indeed, if $a \in S^m(\mathbb{R}^{2n})$ and $a_j \in S^m(\mathbb{R}^{2n})$ for each $j$, we introduce the notation

$$a(x, \xi) \sim \sum_{j=0}^\infty a_j(x, \xi) \hbar^j,$$ \hfill (2.1.6)

to mean that

$$\left| \partial^\alpha_x \partial^\beta_\xi \left( a - \sum_{j=0}^N a_j \hbar^j \right) \right| \leq C_{\alpha, \beta, N} \hbar^{N+1} (1 + |\xi|^2)^{m-|\beta|/2}$$ \hfill (2.1.7)

for each $n$ uniformly in some interval $\hbar \in (0, h_{N, \alpha, \beta}]$.

The key point is that for an arbitrary formal series $\sum_j a_j \hbar^j$ with symbols $a_j \in S^m$, we can find a $\sim$-equivalent symbol $a \in S^m$.

**Proposition 2.1.3.** Given an arbitrary sequence of symbols $a_j \in S^m(\mathbb{R}^{2n})$, there exists a symbol $a \in S^m(\mathbb{R}^{2n})$ satisfying (2.1.6). We call $a$ the Borel resummation of the formal series $\sum_j a_j \hbar^j$.

A proof of Proposition 2.1.3 can be found in Theorem 4.15 of [50].

We refer the the leading term $a_0$ in (2.1.6) as the principal symbol of $a$, and write $a_0 = \sigma(a)$. This is of course only well-defined modulo $O(\hbar)$.

From repeated integration by parts in (2.1.1), we see that if the dist(spt$(u)$, spt$(a)$) = $c > 0$, then the function $(a(x, \hbar D)u)(x)$ is of size $O(\hbar^n)$ for any $n \in \mathbb{N}$. We denote such a size estimate by $O(\hbar^\infty)$, and note that these terms can be regarded as negligible in the semiclassical limit $\hbar \to 0$.

At this point we introduce the notion of a semiclassical wavefront set for $L^2$ functions.
**Definition 2.1.4.** Suppose \( u(x; h) \) is a collection of smooth functions on \( \mathbb{R}^n \) for \( h \in (0, h_0] \). Then the semiclassical wavefront set \( WF_h(u) \subset \mathbb{R}^{2n} \) is defined as follows. \( (x_0, \xi_0) \in (WF_h(U))^c \) if there exists a \( v \in C_0^\infty(\mathbb{R}^{2n}) \) with \( |v(x_0, \xi_0)| > 0 \) such that we have

\[
v(x, hD)u = O_{L^2}(h^k).
\]

(2.1.8)

for any \( k \in \mathbb{N} \).

Such a definition is possible in considerably more general classes of distributions (See Section 8.4 of [50]), but we shall not require it in this generality.

**Remark 2.1.5.** In fact, it suffices to prove that \( v(x, h_j D)u = O_{L^2}(h_j^k) \) for any \( k \in \mathbb{N} \) and a single sequence \( h_j \to 0 \).

A crucial formula in the pseudodifferential calculus is the composition formula, which asserts that if \( a \in S_m^1 \) and \( b \in S_m^2 \), then the composition

\[
a(x, hD) \circ b(x, hD)
\]

is a semiclassical pseudodifferential operator of order \( m_1 + m_2 \), and its symbol is given by

\[
a \circ b := \exp\left(\frac{h}{i} \langle \partial_\xi, \partial_y \rangle \right) (a(x, \xi)b(y, \eta))_{(y, \eta) = (x, \xi)}.
\]

(2.1.10)

as is shown in Theorem 9.5 of [50].

Expanding the symbols \( a, b, \) and \( a \circ b \) in (2.1.10) as semiclassical series yields the following

**Proposition 2.1.6.** Given two symbols \( a \in S_m^1 \) and \( b \in S_m^2 \), their composition \( p \circ q \in S_{m_1+m_2} \) as the Borel resummation of

\[
a \circ b \sim \sum_{j=0}^\infty c_j h^j
\]

(2.1.11)

where

\[
\sum_{r+s+|\gamma|=j} \frac{1}{\gamma!} \frac{\partial_\gamma^r a_r(x, \xi) \partial_\gamma^s b_s(x, \xi)}.
\]

(2.1.12)

A key feature of the pseudodifferential calculus that immediately lends itself to PDE applications is that of the invertibility of elliptic operators.

**Proposition 2.1.7.** If we have a symbol \( a \in S_m^1 \) with

\[
a(x, \xi) \geq c(1 + |\xi|^2)^{m/2}
\]

(2.1.13)

for some \( c > 0 \), then there exists a symbol \( b \in S_{-m} \) with

\[
a(x, hD) \circ b(x, hD) = O(h^\infty).
\]

(2.1.14)

The proof of Proposition 2.1.7 is an application of (2.1.10) and can be found in Proposition 2.6.10 of [30].

Importantly, for an arbitrary diffeomorphism from \( \gamma : U \to V \) with \( U, V \subset \mathbb{R} \) open, the symbols classes \( S^m \) and invariant under the pullback of the lift of \( \gamma \) to a symplectomorphism \( \tilde{\gamma} : U \times \mathbb{R}^n \to V \times \mathbb{R}^n \). This invariance allows for the construction of semiclassical pseudodifferential operators on compact manifolds, as is done in ([50] Chapter 14).
Definition 2.1.8. We write $S^m(T^*M)$ to denote the class of $m$-th order Kohn-Nirenberg symbols on a compact manifold $M$ and we write $S^{m,k}(T^*M) := h^{-k}S^{m,k}(T^*M) := h^{-k}S^m(T^*M)$ to denote the class of Kohn-Nirenberg symbols of differential order $m$ and semiclassical order $k$.

Definition 2.1.9. We write $\Psi^m(M)$ to denote the class of $m$-th order semiclassical pseudodifferential operators on $M$ in the sense of ([50] Chapter 14). We write $\Psi^{m,k}(M) := h^{-k}\Psi^m(M)$ to denote the class of semiclassical pseudodifferential operators of differential order $m$ and semiclassical order $k$.

One significant difference between the calculus on compact manifolds and on Euclidean space however, is that the symbol of a semiclassical pseudodifferential operator $A_h \in \Psi^{m,k}(M)$ is only invariantly defined modulo $S^{m-1,k-1}$. One can also define semiclassical pseudodifferential operators on the space of half-densities on a compact manifold $M$.

Definition 2.1.10. A half-density $\rho$ on an $n$ dimensional vector space $V$ is a map $\rho : V^n \to \mathbb{R}$ such that
\[
\rho(Av_1, Av_2, \ldots, Av_n) = |\det(A)|^{1/2} \rho(v_1, v_2, \ldots, v_n),
\]
for any $A$ is a linear transformation on $V$. We denote the space of half-densities on $V$ by $\Omega^{1/2}(TM)$.

Definition 2.1.11. The space $C^\infty(M, \Omega^{1/2})$ of smooth half-densities on a compact Riemannian manifold is given by the collection of maps $u : M \to \cup_{x \in M} \Omega^{1/2}(TM)$ such that $u(x) \in \Omega^{1/2}(T_xM)$ for each $x \in M$, and $u(X_1, X_2, \ldots, X_n) \in C^\infty(M)$ for any $n$ smooth vector fields $X_1, X_2, \ldots, X_n$.

Since half-densities are given in local coordinates on a Riemannian manifold by
\[
C^\infty(M; \Omega^{1/2}) := \{f|dx|^{1/2} : f \in C^\infty(M)\}
\]
where $dx$ is the Riemannian volume form, we can identify half-densities with functions in this setting, however note that their pullbacks as half-densities will involve a Jacobian factor.

Thus we can locally define semiclassical pseudodifferential operators on half-densities by setting
\[
P(x,hD)(u|dx|^{1/2}) = (P(x,hD)u)|dx|^{1/2},
\]
and they can be defined globally in a similar fashion to semiclassical pseudodifferential operators acting on functions. (See Section 14.2.5 of [50]).

An advantage of working with half-densities is that principal symbols of semiclassical pseudodifferential operators on half-densities are invariantly in $S^{m,k}/S^{m-2,k-2}$, and subprincipal symbols of operators are thus invariantly defined. (See Section 1.3 of [21] for a further discussion of this invariance).

In Section 5.2, we work with semiclassical Fourier integral operators, which are generalisations of semiclassical pseudodifferential operators obtained locally by replacing the phase function $i(x-y) \cdot \xi/h$ in the oscillatory integral expression (2.1.1) with more general phase functions $i\psi(x,y,\xi)/h$. 
The kernels of such operators are then special cases of Fourier integrals
\[ u(x; h) = (2\pi h)^{-m} \int_{\mathbb{R}^m} e^{i\phi(x, \xi)/h} a(x, \xi) \, d\xi \]  
(2.1.18)
with \( m = 2n \). For such a phase function, we can associate a Lagrangian submanifold of \( \mathbb{R}^{2m} \) given by
\[ \Lambda_\phi = \{(x, \partial_x \phi(x, \xi)) : \partial_\xi(x, \xi) = 0\}. \]  
(2.1.19)
Indeed, stationary phase asymptotics show that \( WF_h(u) \subset \Lambda_\phi \) as in [24].

Defining a canonical relation \( \chi : T^*M_1 \to T^*M_2 \) to be a relation with flipped graph
\[ \Gamma_\chi := \{(x, \xi, y, -\eta) : (x, \xi, y, \eta) \in \chi\} \]  
(2.1.20)
a Lagrangian submanifold of \( T^*(M_1 \times M_2) \), we can then define a Fourier integral operator associated to a given relation \( \mathcal{C}^\infty(M_1) \to \mathcal{C}^\infty(M_2) \) to be a finite sum of Fourier integrals associated to \( \Gamma_\chi \).

The global theory of Fourier integrals is complicated by the fact that different phase functions \( \phi \) can parametrise the same Lagrangian manifold \( \Lambda \) locally, yet for different \( \phi \) a different symbol \( a \) will be required in (2.1.18) in order to represent the same Fourier integral \( u(x; h) \). In order to invariantly define the notion of a principal symbol for a Fourier integral operator, it must be defined as an object on a certain line bundle over the Lagrangian submanifold \( \Lambda \), known as the Maslov bundle.

A thorough account of Fourier integral operators can be found in the seminal paper [24] in the classical setting, and in [21] in the semiclassical setting. We shall summarise the relevant details in our exposition of Popov’s construction of the quantum Birkhoff normal form for KAM Hamiltonians [35][37] in Section 5.2.

### 2.2 Gevrey class symbols

Our application of the semiclassical pseudodifferential calculus in Chapter 5 involves working with Gevrey class symbols. We outline the relevant differences from the theory in Section 2.1 here.

We suppose \( D \) is a bounded domain in \( \mathbb{R}^n \), and take \( X = \mathbb{T}^n \) or a bounded domain in \( \mathbb{R}^m \). We fix the parameters \( \sigma, \mu > 1 \) and \( \varrho \geq \sigma + \mu - 1 \), and denote the triple \((\sigma, \mu, \varrho)\) by \( \ell \).

**Definition 2.2.1.** A formal Gevrey symbol on \( X \times D \) is a formal sum
\[ \sum_{j=0}^{\infty} p_j(\theta, I)h^j \]  
(2.2.1)
where the \( p_j \in \mathcal{C}^\infty_0(X \times D) \) are all supported in a fixed compact set and there exists a \( C > 0 \) such that
\[ \sup_{X \times D} |\partial_\theta^\alpha \partial_I^\beta p_j(\theta, I)| \leq C^{j+|\alpha|+|\beta|+1} \beta! \alpha! j! \varrho. \]  
(2.2.2)

**Definition 2.2.2.** A realisation of the formal symbol (2.2.1) is a function \( p(\theta, I; h) \in \)
\[ C_0^\infty(X \times D) \text{ for } 0 < h \leq h_0 \text{ with} \]
\[
\sup_{X \times D \times (0, h_0)} \left| \partial_\theta^\alpha \partial_I^\beta \left( p(\theta, I; h) - \sum_{j=0}^N p_j(\theta, I) h^j \right) \right| \leq h^{N+1} C_1^{N+|\alpha|+|\beta|+2} \beta! \alpha! \mu (N + 1)! \nu. \]

(2.2.3)

**Lemma 2.2.3.** Given a formal symbol (2.2.1), one choice of realisation is
\[
p(\theta, I; h) := \sum_{j \leq \epsilon h^{-1/\nu}} p_j(\theta, I) h^j
\]
(2.2.4)

where \( \epsilon \) depends only on \( n \) and \( C_1 \).

**Definition 2.2.4.** We define the residual class of symbols \( S_{-\nu}^\infty \) as the collection of realisations of the zero formal symbol.

**Definition 2.2.5.** We write \( f \sim g \) if \( f - g \in S_{-\nu}^\infty \). It then follows that any two realisations of the same formal symbol are \( \sim \)-equivalent. We denote the set of equivalence classes by \( S_{\ell}(X \times D) \).

An important feature of the Gevrey symbol calculus is that the symbol class \( S\ell(X \times D) \) is closed under composition.

We can now discuss the pseudodifferential operators corresponding to these symbols.

**Definition 2.2.6.** To each symbol \( p \in S\ell(X \times D) \), we associate a semiclassical pseudodifferential operator defined by
\[
(2\pi h)^{-n} \int_{X \times \mathbb{R}^n} e^{i(x-y) \cdot \xi/h} p(x, \xi; h) u(y) \, d\xi \, dy.
\]
(2.2.5)

for \( u \in C_0^\infty(X) \).

The above construction is well defined modulo \( \exp(-c h^{-1/\nu}) \), as for any \( p \in S_{-\infty}(X \times D) \) we have
\[
\| P_h u \| = O_{L^2}(\exp(-c h^{-1/\nu}))
\]
(2.2.6)

for some constant \( c > 0 \).

**Remark 2.2.7.** The exponential decay of residual symbols is a key strengthening that comes from working in a Gevrey symbol class.

The operations of symbol composition and conjugation then correspond to composing operators and taking adjoints respectively. Moreover, if \( p \in S_{(\sigma, \sigma, 2\sigma - 1)} \), then \( G^\sigma \)-smooth changes of variable preserve the symbol class of \( p \). This coordinate invariance allows us to extend the Gevrey pseudodifferential calculus to compact Gevrey manifolds.

At this point we introduce the notion of a microsupport in the Gevrey sense.

**Definition 2.2.8.** Suppose \( u(x; h) \) is a collection of smooth functions on the \( G^\sigma \)-manifold \( M \) for \( h \in (0, h_0] \). Then the \( G^\sigma \) microsupport \( MS^\sigma(u) \subset T^*M \) is defined as follows.
(x_0, \xi_0) \in (MS\Theta)^c if there exists a product of compact sets U \times V \subset X \times \mathbb{R}^n with (x_0, \xi_0) \in U \times V inside a single coordinate chart and there exists a c > 0 such that for any v \in G^2 \cap C_0^\infty(M) we have
\[
\int_{\mathbb{R}^n} e^{ix \cdot \xi/h} v(x) u(x; h) \, dx = O(e^{-ch^{-1/\rho}}) \tag{2.2.7}
\]
uniformly in V.

It follows from stationary phase that if a symbol p is $S_{\ell}^{-\infty}$ in a neighbourhood of a point $(x_0, \xi_0)$, then the point $(x_0, \xi_0, x_0, -\xi_0)$ lies outside the $G^2$ microsupport of the distribution kernel of $P_h$.

### 2.3 Weyl law

An application of the semiclassical pseudodifferential calculus that is particularly important to us is the semiclassical Weyl law, which provides asymptotics for the counting functions of eigenvalues for suitable semiclassical pseudodifferential operators $P(h)$ in fixed energy bands $[a, b]$ or shrinking energy bands $[a, a + \epsilon(h)]$ with $\epsilon(h) \to 0$ as $h \to 0$.

We consider semiclassical pseudodifferential operators of the form
\[
P(h) := h^2 \Delta_g + Q(x, hD) \in \Psi^2(M) \tag{2.3.1}
\]
on a compact Riemannian manifold $M$, where $Q(x, \xi) \in S^0(M)$ is real valued.

For each fixed $h > 0$, the operator $P(h)$ is a self-adjoint operator
\[
P(h) : H^2(M) \subset L^2(M) \to L^2(M) \tag{2.3.2}
\]
with compact inverse, where $H^2(M)$ is the Sobolev space of order 2.

Basic spectral theory then tells us that the spectrum of $P(h)$ is real and discrete, consisting of a countable orthonormal basis of eigenpairs $(u_j(h), E_j(h))$, with $E_j \to \infty$ as $j \to \infty$.

Weyl’s law is then the statement that
\[
(2\pi h)^{-n} \# \{ j \in \mathbb{N} : E_j(h) \in [a, b] \} \to \mu(\{(x, \xi) \in T^*M : P(x, \xi) \in [a, b]\}). \tag{2.3.3}
\]
where $\mu$ denotes the symplectic measure $d\xi \, dx$ on $T^*M$.

A standard proof relies on a trace formula for a Schwartz class functional calculus for semiclassical pseudodifferential operators. If $f \in \mathcal{S}(\mathbb{R})$, then we can define
\[
f(P(h))(u) := \sum_{j=1}^\infty f(E_j) \langle u, u_j \rangle u_j. \tag{2.3.4}
\]
The rapid decay of $f$ in fact implies that $f(P(h))$ is a semiclassical pseudodifferential operator in the class
\[
\Psi^{-\infty}(M) = \bigcap_{m \in \mathbb{Z}} \Psi^m(M). \tag{2.3.5}
\]

In fact it can be shown that $f(P(h))$ is a trace-class operator on $L^2(M)$, with principal
Semiclassical Analysis

symbol

\[ \sigma(f(P(h))) = f(P(x, \xi)) \]  \hspace{1cm} (2.3.6)

and trace

\[ \text{tr}(f(P(h))) = (2\pi h)^{-n} \int_{\mathcal{T}^* M} f(P(x, \xi)) \, d\xi \, dx. \]  \hspace{1cm} (2.3.7)

The equation (2.3.3) then follows from (2.3.7) and regularisation of the indicator functions. Full details can be found in Chapter 14 of [50].

Remark 2.3.1. In the special case of the Laplace–Beltrami operator \( p(x, hD) = h^2 \Delta_g \), rescaling yields the classical Weyl law which gives counting asymptotics for the Laplacian eigenvalues.

The Weyl law can also be localised in phase space by a semiclassical pseudodifferential operator. That is, for any \( B \in \Psi^0(M) \), we have

\[ (2\pi h)^n \sum_{E_j(h) \in [a, b]} \langle B(x, hD)u_j(h), u_j(h) \rangle \to \int_{P^{-1}([a, b])} B(x, \xi) \, d\xi \, dx. \]  \hspace{1cm} (2.3.8)

as \( h \to 0 \). The proof of this generalisation again makes use of (2.3.7), and can be found in Section 15.3 of [50].

A version of the semiclassical Weyl law was proven by Petkov and Robert [32] for boundaryless manifolds that is localised to \( O(h) \) sized energy bands. That is, for regular values \( E \) of the Hamiltonian \( P \), we have

\[ \#\{ E_j(h) \in [E - ch, E - c\theta] \} \sim \int_{P^{-1}([E - ch, E + c\theta])} P(x, \xi) \, d\xi \, dx. \]  \hspace{1cm} (2.3.9)

Remark 2.3.2. This result requires the dynamical assumption that the set of trapped trajectories is of measure zero. Without this assumption, we only obtain a uniform upper bound for \( \#\{ E_j(h) \in [E - ch, E - c\theta] \} \).
Chapter 3

Quantum Ergodicity in Mixed Systems

3.1 Introduction

In this chapter, we turn our attention to mixed billiards. We begin by recalling the relevant definitions.

If \((M, g)\) is a compact boundaryless Riemannian manifold, we define dynamical billiards on \(M\) to be the Hamiltonian flow \(\phi_t\) on the cotangent bundle \(T^*M\) of the manifold given by Hamilton’s equations

\[
\dot{x}_j = \frac{\partial H}{\partial \xi_j}, \quad \dot{\xi}_j = \frac{\partial H}{\partial x_j}
\]

for the Hamiltonian \(H(x, \xi) := |(x, \xi)|^2_{g^{-1}}\) where \(g^{-1}\) is the dual metric tensor.

Since the Hamiltonian is an invariant of motion for the flow \(\phi_t\), it is natural to restrict the domain of this flow to the cosphere bundle

\[
S^*M := \{z = (x, \xi) \in T^*M : |z|_{g^{-1}} = 1\}.
\]

More generally, one can define billiards on compact Riemannian manifolds with piecewise smooth boundary in the sense of Chapter 6 of [11], see also [49].

To be precise, we assume that we can smoothly embed \(M\) in a boundaryless manifold \(\tilde{M}\) of the same dimension and that there exist finitely many smooth functions \(f_j \in C^\infty(\tilde{M})\) such that the following conditions are satisfied.

1. \(df_i|_{f_j^{-1}(0)} \neq 0\),
2. \(df_i, df_j\) are linearly independent on \(f_i^{-1}(0) \cap f_j^{-1}(0)\),
3. \(M = \{x \in \tilde{M} : f_j(x) \geq 0\}\) for all \(j\).

We can then write

\[
\partial M = \cup_j \partial M_j := \cup_j (f_j^{-1}(0) \cap M)
\]

and denote by \(\mathcal{S} \subset \partial M\) the set of points that lie in \(\partial M_j\) for multiple \(j\).

We define the broken Hamiltonian flow \(\phi_t\) on \(S^*M\) locally by extending the boundaryless Hamiltonian flow by reflection at non-tangential and non-singular boundary collisions.

That is, if \(\phi_{t_0}(z) = (x, \xi_+)\) with \(x \in \partial M \setminus \mathcal{S}\) and \(\langle \xi_+, N_x \rangle > 0\) where \(N_x\) is the outgoing unit normal covector, we extend \(\phi_t\) to sufficiently small \(t > t_0\) by defining
\[ \phi_t(z) = \phi_{t-t_0}(x, \xi_-), \] where \( \xi_- \in S^*_+M \) is the unique covector such that \( \xi_+ + \xi_- \in T^*\partial M \) and \( \pi(\xi_+) = \pi(\xi_-) \) where \( \pi : T^*_\partial M \to T^*\partial M \) is the canonical projection. We terminate all trajectories that meet \( \partial M \) in any other manner.

There are four subsets \( \{B_j\}_{j=1}^4 \) of phase space for this class of manifolds which present an obstruction to obtaining a globally defined broken Hamiltonian flow or to the application of tools from microlocal analysis. We enumerate these sets below.

1. \( B_1 = \{z \in S^*M : \phi_t(z) \in S\} \)
2. \( B_2 = \{z \in S^*M : \phi_t(z) \in \partial M \text{ for infinitely many } t \text{ in a bounded interval}\} \)
3. \( B_3 = \{z \in S^*M : \phi_t(z) \notin \partial M \text{ for any } t > 0 \text{ or } \phi_t(z) \notin \partial M \text{ for any } t < 0\} \)
4. \( B_4 = \{z \in S^*M : \phi_t(z) \text{ meets } \partial M \text{ tangentially for some } t \in \mathbb{R}\} \)

Removing these sets from our flow domain, we then obtain a globally defined billiard flow on \( D = S^*M \setminus (\bigcup_{j=1}^4 B_j) \). For manifolds without boundary, we simply take \( D = S^*M \).

The canonical symplectic form \( d\xi \wedge dx \) on \( T^*M \) determines a family of measures \( \mu_c \) on each of the energy hypersurfaces

\[ \Sigma_c = \{z = (x, \xi) \in T^*M : |z|_{g^{-1}} = c\} \quad (3.1.3) \]

defined implicitly by

\[ \int_a^b \int_{\Sigma_c} f \, d\mu_c \, dc = \int_{\{(x, \xi) \in [a,b] : |(x, \xi)|_{g^{-1}} \in [a,b]\}} f \, d\xi \wedge dx \quad (3.1.4) \]

for \( f \in C^\infty_c(T^*M) \).

Upon normalisation of \( \mu_1 \) we then obtain the Liouville measure \( \mu_L \) on \( S^*M \), which allows us to define ergodicity for the billiard flow \( \phi_t \) as in Definition 1.1.8.

**Remark 3.1.1.** It is shown in Section 6.2 of [11] that the sets \( B_1, B_2 \) are of Liouville measure zero, and it is shown in [49] that the set \( B_4 \) is of Liouville measure zero for the class of manifolds considered. That the remaining set \( B_3 \) is Liouville null is usually taken as an assumption. In particular, it is clear that this assumption is satisfied by bounded domains in \( \mathbb{R}^n \).

The quantum mechanical analogue of the system (3.1.1) is the evolution of a wave function \( \psi \in L^2(M) \) according to the rescaled Schrodinger’s equation

\[ -\Delta_g \psi = i \frac{\partial \psi}{\partial t} \quad (3.1.5) \]

with boundary conditions to ensure self-adjointness of the Laplacian. We shall work with the most studied and technically easiest choice of Dirichlet boundary conditions.

Since the boundary of \( M \) is Lipschitz, it follows that the Laplacian \(-\Delta_g\) is self adjoint on \( L^2 \) when given the standard domain \( H^2(M) \cap H^1_0(M) \). Standard spectral theory then shows that \(-\Delta_g\) has purely discrete spectrum (counting multiplicity) \( \{0 < E_1 \leq E_2 \leq \ldots\} \subset \mathbb{R}^+ \).
\textbf{Figure 3.1:} The half-mushroom billiard, with a high energy eigenfunction that extends by odd symmetry to the mushroom billiard. This particular eigenfunction appears to live in the ergodic region of phase space. Image courtesy of Dr Barnett.

The phase space localisation of the high energy eigenfunctions of $-\Delta_g$ can then be described using the calculus of semiclassical pseudodifferential operators, as defined in Chapters 4 and 14 of [50].

For the reader’s sake, we recap the definition of semiclassical measures from Section 1.3. To each subsequence of $(u_j)$, we can associate at least one non-negative Radon measure $\mu$ on $S^*M$ which provides a notion of phase space concentration in the semiclassical limit.

We say that the eigenfunction subsequence $(u_{jk})$ has unique \textit{semiclassical measure} $\mu$ if

$$\lim_{k \to \infty} \langle a(x, E_{jk}^{-1/2}D)u_{jk}, u_{jk} \rangle = \int_{S^*M} a(x, \xi) \, d\mu \quad (3.1.6)$$

for each semiclassical pseudodifferential operator with principal symbol $a$ compactly supported away from the boundary of $S^*M$. In Chapter 5 of [50], the existence and basic properties of semiclassical measures are established using the calculus of semiclassical pseudodifferential operators (see also [17]).

A billiard $M$ is then said to be \textit{quantum ergodic} if there is a full density subsequence of eigenfunctions $(u_{nk})$ such that the Liouville measure on $S^*M$ is the unique semiclassical measure associated to the sequence $u_{nk}$. This statement can be interpreted as saying that the sequence of eigenfunctions equidistributes in phase space with the possible exception of a sparse subsequence.

In this chapter we consider the family of mushroom billiards $M_t = R_t \cup S \subset \mathbb{R}^2$ where $R_t = [-r_1, r_1] \times [-t, 0]$ and $S$ is the closed upper semidisk of radius $r_2 > r_1$ centred at the origin. We denote the area of $M_t$ by $A(t)$.

This billiard, proposed by Bunimovich [7] is neither classically ergodic nor completely integrable for $t > 0$ and is rather one of the simplest billiards that satisfies the following mixed dynamical assumptions.
• \( M \) is a smooth Riemannian manifold with piecewise smooth boundary

• The flow domain \( \mathcal{D} \) is the union of two invariant subsets, each of positive Liouville measure and one of which, \( U \), has ergodic geodesic flow

• The billiard flow is completely integrable on \( \mathcal{D} \setminus U \).

In the mushroom billiard, \( U_t \) consists of \( \mu_L \)-almost all trajectories that enter \( R_t \cup \overline{B}(0,r_1) \) before their first boundary collision. The trajectories that do not enter \( R_t \cup \overline{B}(0,r_1) \) before their first boundary collision lie entirely within the upper semi-annulus \( S \setminus B(0,r_1) \) and are just reflected trajectories of the disk billiard. The integrability of the geodesic flow on \( \mathcal{D} \setminus U_t \) then follows from the integrability of the disk billiard.

In the case of such mixed systems, we do not yet have a satisfactory analogue to the quantum ergodicity theorem. It is a long-standing conjecture of Percival [31] that a full density subset of a complete system of eigenfunctions of the Laplace–Beltrami operator can be divided into two disjoint subsets, one corresponding to the ergodic region of phase space and the other corresponding to the completely integrable region. Moreover, the natural density of these subsets is conjectured to be in proportion to the Liouville measures of the corresponding flow-invariant subsets of \( \mathcal{D} \).

**Conjecture 3.1.2** (Percival’s Conjecture). For every compact Riemannian manifold \( M \) such that \( \mathcal{D} \) is the disjoint union of two invariant subsets \( U, \mathcal{D} \setminus U \), with \( U \) ergodic and \( \mathcal{D} \setminus U \) completely integrable, we can find two subsets \( A, B \subset \mathbb{N} \) such that

1. \( A \cup B \) has density 1
2. \( (u_k)_{k \in A} \) equidistributes in the ergodic region \( U \)
3. Each semiclassical measure associated to the subset \( B \) is supported in the completely integrable region \( \mathcal{D} \setminus U \)
4. The density of \( A \) is equal to \( \mu_L(U) \).

Numerical evidence due to Barnett-Betcke [4] has strongly supported this conjecture for the mushroom billiard, yet rigorous proof has remained elusive.

A weaker version of the conjecture can be formulated by slightly relaxing the density requirements of the subsets \( A \) and \( B \).

**Conjecture 3.1.3** (Weak Percival’s Conjecture). For every compact Riemannian manifold \( M \) such that \( \mathcal{D} \) is the disjoint union of two invariant subsets \( U, \mathcal{D} \setminus U \), with \( U \) ergodic and \( \mathcal{D} \setminus U \) completely integrable, we can find two subsets \( A, B \subset \mathbb{N} \) such that

1. \( A \cup B \) has upper density 1
2. \( (u_k)_{k \in A} \) equidistributes in the ergodic region \( U \)
3. Each semiclassical measure associated to the subset \( B \) is supported in the completely integrable region \( \mathcal{D} \setminus U \)
4. The upper densities of \( A \) and \( B \) are equal to \( \mu_L(U) \) and \( 1 - \mu_L(U) \) respectively.

In this chapter, we prove Conjecture 3.1.3 is indeed true for the mushroom billiard, at least for almost all \( t \in (0, 2] \). Essential in our work is the following result due to Galkowski [15].
Theorem 3.1.4. For any compact Riemannian manifold with boundary satisfying (3.1), there exists a full density subsequence of \((u_j)\), such that every associated semiclassical measure \(\mu\) satisfies
\[
\mu|_U = a\mu_L|_U
\]
for some constant \(a\).

Our strategy for this proof is motivated by that used by Hassell in constructing the first known example of a non-QUE ergodic billiard [22].

We begin in Section 2 by using the Dirichlet eigenfunctions on the semicircle to construct a family \((v_n, \alpha_n^2)\) of \(O(n^{-\infty})\) quasimodes that are almost orthogonal and are microlocally supported in the completely integrable region \(S^*M_t \setminus U_t\).

Using the well-known asymptotics of the Bessel function zeroes, we obtain a lower bound (3.2.11) for the counting function of this quasimode family.

In Section 3, the main result is Proposition 3.3.1, an abstract spectral theoretic result that allows us to approximate certain eigenfunctions by linear combinations of quasimodes of similar energy given that the numbers of each are comparable. This is the essential ingredient for passing from localisation properties about our explicit family of quasimodes to localisation properties of a family of eigenfunctions with asymptotically equivalent counting function.

In Section 4 we commence our study of the variation of eigenvalues as the stalk length \(t\) varies in \((0, 2]\). In order to simplify the nomenclature, we often interpret \(t\) as a time parameter.

The Hadamard variational formula asserts that
\[
\dot{E}(t) = -\int_{\partial M_t} \rho_t(s)(d_nu(t)(s))^2 ds
\]
where \(\rho_t(s)\) is the unit normal variation of the domain at a boundary point \(s\). For normally expanding domains such as ours, (3.1.8) directly implies that individual eigenvalues are non-increasing in \(t\).

However, using an interior formulation of the Hadamard variational formula from Proposition 3.4.1, we can also quantify the variation of the eigenvalue \(E_j(t)\) by
\[
E_j^{-1}(t)\dot{E_j}(t) = \langle Qu_j(t), u_j(t) \rangle
\]
for an appropriate pseudodifferential operator \(Q\) supported in the stalk \(R_t \subset M_t\).

Proposition 3.4.2 then establishes that for a full density subset of the eigenvalues, the quantity \(\langle Qu_j(t), u_j(t) \rangle\) can be approximated up to an error of \(o(E_j)\) by cutting off \(Q\) sufficiently close to the boundary \(\partial M_t\). This result is shown by using analysis of the wave kernel to establish the key spectral projector estimates (3.4.9) and (3.4.10).

We can then use the equidistribution result of Galkowski’s Theorem 3.1.4 to asymptotically control \(\langle Qu_j(t), u_j(t) \rangle\) and hence provide us with an upper bound (3.4.17) on the speed of eigenvalue variation for almost all eigenvalues.

Section 5 completes the argument in two parts.

In the first of these parts, we define a set \(\mathcal{G} \subset (0, 2]\) such that for \(t \in \mathcal{G}\), we have a
certain spectral non-concentration property on $M_t$. This property implies that that

$$\text{the number of eigenvalues lying in the union } \bigcup_{j=1}^{n} [\alpha_j^2 - c, \alpha_j^2 + c]$$

can exceed $n$ by at most a small proportion, for infinitely many $n$. (3.1.10)

Proposition 3.3.1 then implies that for $t \in G$ and $n$ such that (3.1.10) is satisfied, a large proportion of the corresponding eigenfunctions are asymptotically well-approximated by linear combinations of the previously constructed family of quasimodes $(v_n)$, which are microlocally supported in the completely integrable region $S^* M_t \setminus U_t$ of phase space.

In fact, the explicit computation (3.2.11) of the counting function of these quasimodes leads to a proof that the corresponding family of eigenfunctions must have the maximal upper density allowed by the contraint of Weyl’s law. We show this in Theorem 3.5.4.

Consequently, we show in Proposition 3.5.5 that a subset of the complementary family of eigenfunctions with full upper density must have all semiclassical mass in the ergodic region $U_t$. From Theorem 3.1.4, this family must then equidistribute in $U_t$ as required.

The final part of the chapter establishes via contradiction that $(0, 2] \setminus G$ is Lebesgue-null. As in [22] we can choose the eigenvalue branches $E_j(t)$ to be in increasing order and piecewise smooth in $t$. The crucial ingredient here is then the asymptotic bound (3.4.17) on the speed of eigenvalue variation.

If $G$ is not of full measure, we can construct a small interval $I = [t_1, t_2]$ in which the average number of eigenvalues $E_j(t)$ lingering near quasi-eigenvalues $\alpha_i^2$ exceeds $d = d(t_1) = 1 - \mu_L(U_{t_1})$ by using the negation of (3.1.10).

We can use (3.1.11) together with the fact that the small windows about quasi-eigenvalues are comparatively sparse in the interval $[E_j(t_1), E_j(t_2)]$ to show that the upper bound (3.4.17) on eigenvalue speed provides a lower bound of $(1 - d)$ on the time they must spend travelling outside of quasi-eigenvalue windows.

This implies that the average proportion of time spent by large eigenvalues lingering near quasi-eigenvalues for $t \in I$ cannot exceed $d$, and consequently that the proportion of lingering eigenvalues cannot exceed $d$. This contradiction concludes the proof.

### 3.2 Quasimodes

In polar coordinates, the Dirichlet eigenfunctions for the semidisk are given by

$$u_{n,k} := \sin(n \theta) J_n(\alpha_n k r / r_2)$$

where $\alpha_{n,k}$ is the $k$-th positive zero of the $n$-th order Bessel function $J_n$. 
Proposition 3.2.1. If we define
\[ v_{n,k} := \frac{\chi(r)u_{n,k}}{\|\chi u_{n,k}\|_{L^2}}, \]  
(3.2.2)
where
\[ \chi(r) = \begin{cases} 0 & \text{for } r \leq r_1 \\ 1 & \text{for } r \geq (r_1 + \epsilon)\sqrt{1 - \epsilon^2} > r_1. \end{cases} \]  
(3.2.3)
then the family
\[ \{(v_{n,k}, \alpha_{n,k}^2/r_2^2) : \alpha_{n,k} < nr_2/r_1 + \epsilon\} \]  
(3.2.4)
forms an \(O(n^{-\infty})\) family of quasimodes, with all semiclassical mass contained in the completely integrable region \(S^*M_t \setminus U_t\) of the billiard.

Moreover, these quasimodes are almost orthogonal, in the sense that
\[ |\langle v_{n,k}, v_{m,l} \rangle| = O(\min(n, m)^{-\infty}) = O(\min(\alpha_{n,k}, \alpha_{m,l})^{-\infty}). \]  
(3.2.5)

Proof. The restriction on \(k\) in our family implies that the error incurred in cutting off only depends on the values of the Bessel function \(J_n(x)\) for \(x \in [0, n\sqrt{1 - \epsilon^2}]\).

Then from [1], we have the estimates
\[ |J_n(nx)| \leq \frac{x^n e^{\sqrt{1-x^2}}}{(1 + \sqrt{1-x^2})^n} \quad \text{for } x \leq 1 \]  
(3.2.6)
and
\[ |J_n'(nx)| \leq \frac{(1 + x^2)^{1/4} x^n e^{\sqrt{1-x^2}}}{x \sqrt{2\pi n}(1 + \sqrt{1-x^2})^n} \quad \text{for } x \leq 1 \]  
(3.2.7)
for bounding the Bessel function near 0.

Together these estimates imply that the error incurred by cutting off is \(O(n^{-\infty})\). Furthermore, as the \(u_{n,k}\) are pairwise orthogonal, these bounds also show that the \(v_{n,k}\) are almost orthogonal in the sense claimed.

Now for any smooth compactly supported symbol \(a\) spatially supported in \(R_t \cup B(0, r_1)\), the disjointness of supports from our family of quasimodes implies that
\[ \langle a(x, (r_2/\alpha_{n,k})D)v_{n,k}, v_{n,k} \rangle = O(n^{-\infty}). \]  
(3.2.8)

In particular, we have that any semiclassical measure \(\mu\) associated to these quasimodes cannot have mass in the region \(\{(x, \xi) \in S^*M_t : x \in R_t \cup B(0, r_1)\}\).

Moreover, by the flow invariance of semiclassical measures (See Theorem 5.4 in [50]), this implies that any corresponding semiclassical measure cannot have mass in the ergodic region \(U_t\) because the pre-images under geodesic flow of \(\{(x, \xi) \in D_t : x \in R_t \cup B(0, r_1)\}\) cover \(U_t\).

\[ \square \]

Proposition 3.2.2. We can index these quasimodes as \((v_n, \alpha_n^2)\) so that the quasi-eigenvalues are in increasing order, whilst having
\[ (\Delta + \alpha_n^2)v_n = O(n^{-\infty}) = O(\alpha_n^{-\infty}) \]  
(3.2.9)
Quantum Ergodicity in Mixed Systems

|⟨vn,vk⟩| = O(min(n,k)−∞) = O(min(αn,αk)−∞). \hspace{1cm} (3.2.10)

Moreover, as \(\epsilon \to 0\), the counting function of these quasimodes has the following asymptotic bound.

**Proposition 3.2.3.**

\[
\lim_{\epsilon \to 0} \liminf_{\lambda \to \infty} \frac{\#\{(n,k) : \alpha_{n,k}/r_2 < \lambda, \alpha_{n,k} < \frac{nr_2}{r_1 + \epsilon}\}}{\lambda^2} \geq \left(1 - \frac{\mu_L(U_t)}{\mu_L(D_t)}\right) \cdot \frac{A(t)}{4\pi}. \hspace{1cm} (3.2.11)
\]

where \(A(t)\) is the area of the mushroom \(M_t\).

**Proof.** To simplify our calculations, we scale \(\mu_L\) so that \(\mu_L(M_t) = 2\pi A(t)\).

From an arbitrary point \((r,\theta)\) in the annulus, the trajectories that never enter the stalk have measure \((2\pi - 4 \sin^{-1}(r_1/r))\) out of the full measure \(2\pi\) of the unit cosphere at that point.

Hence

\[
\mu_L(D_t) - \mu_L(U_t) = \int_0^\pi \int_{r_1}^{r_2} r(2\pi - 4 \sin^{-1}(r_1/r)) \, dr \, d\theta
= \pi^2(r_2^2 - r_1^2) - 4\pi \int_{r_1}^{r_2} r \sin^{-1}(r_1/r) \, dr
= \pi^2r_2^2 - 2\pi r_1^2 \sqrt{C^2 - 1} - 2\pi r_2^2 \sin^{-1}(C^{-1})
\]

where \(C = r_2/r_1\).

This implies that

\[
\left(1 - \frac{\mu_L(U_t)}{\mu_L(D_t)}\right) \cdot \frac{A(t)\lambda^2}{4\pi} = \frac{r_2^2}{8} \left(1 - \frac{2}{\pi} \sqrt{C^2 - 1} - \frac{2}{\pi} \sin^{-1}(C^{-1})\right) \lambda^2. \hspace{1cm} (3.2.12)
\]

To estimate the left hand side of (3.2.11), we use the leading order uniform asymptotics for Bessel function zeros found in [1].

As \(n \to \infty\), we have

\[
\alpha_{n,k} = nz(n^{-2/3}a_k) + o(n) \hspace{1cm} (3.2.13)
\]

uniformly in \(k\), where \(z : (-\infty,0] \to [1,\infty)\) is defined implicitly by

\[
\frac{2}{3}(-\zeta)^{3/2} = \sqrt{z(\zeta)^2 - 1} - \sec^{-1}(z(\zeta)) \hspace{1cm} (3.2.14)
\]

and the \(a_k\) are the negative zeros of the Airy function, which have asymptotic

\[
a_k = -\left(\frac{3\pi k}{2}\right)^{2/3} + O(k^{-1/3}). \hspace{1cm} (3.2.15)
\]

We now write \(C_\epsilon = r_2/(r_1 + \epsilon)\).

We count the left hand side of (3.2.11) by separating into two regimes based on the size of \(n/\lambda\). In each of these two regimes, a single one of the inequalities defining our
§3.2 Quasimodes

family (3.2.4) implies the other. More precisely, we have

$$|\{(n,k) : \alpha_{n,k}/r_2 \leq \lambda, \alpha_{n,k} \leq C_r n\}|$$
$$= |\{(n,k) : n \leq r_2 \lambda/C_r, \alpha_{n,k} \leq C_r n\}|$$
$$+ |\{(n,k) : r_2 \lambda/C_r < n \leq r_2 \lambda, \alpha_{n,k}/r_2 \leq \lambda\}|$$
$$= N_A(\lambda; \epsilon) + N_B(\lambda; \epsilon) \quad (3.2.16)$$

where

$$N_A(\lambda; \epsilon) := |\{(n,k) : n \leq r_2 \lambda/C_r, \alpha_{n,k} \leq C_r n\}| \quad (3.2.17)$$

and

$$N_B(\lambda; \epsilon) := |\{(n,k) : r_2 \lambda/C_r < n \leq r_2 \lambda, \alpha_{n,k}/r_2 \leq \lambda\}| \quad (3.2.18)$$

respectively.

For \(n, k\) sufficiently large, a sufficient condition for being in regime \(A\) of (3.2.16) is to have

$$z(n^{-2/3} a_k) \leq C_r - \epsilon = \hat{C}_r$$

and

$$n \leq r_2 \lambda/C_r. \quad (3.2.19)$$

Also, from the Airy function asymptotics we have

$$\frac{2}{3} (-n^{-2/3} a_k)^{3/2} = \frac{2}{3n} \left( \frac{3\pi k}{2} \right)^{2/3} + O(k^{-1/3}) \right)^{3/2} = \frac{\pi k}{n} + O(n^{-1}) \quad (3.2.21)$$

where the error is uniform in \(k\).

Hence from the monotonicity of \(z\), for all \(n, k\) sufficiently large with \(n \leq r_2 \lambda/C_r\), a sufficient condition for being in regime \(A\) of (3.2.16) is

$$k \leq \frac{\sqrt{\hat{C}_r^2 - 1 - \sec^{-1}(\hat{C}_r) - \epsilon}}{\pi} n. \quad (3.2.22)$$

Noting that the contribution from small \(n\) and \(k\) is finite, we can conclude that

$$\liminf_{\lambda} \frac{N_A(\lambda^2)}{\lambda^2} \geq \liminf_{\lambda} \left( \frac{1}{\lambda^2} \sum_{n \leq r_2 \lambda/C_r} n \right) \cdot \frac{\sqrt{C_r^2 - 1 - \sec^{-1}(\hat{C}_r) - \epsilon}}{\pi}$$
$$= \frac{r_2^2(\sqrt{\hat{C}_r^2 - 1 - \sec^{-1}(\hat{C}_r) - \epsilon})}{2C_r^2 \pi}. \quad (3.2.23)$$

Similarly, for sufficiently large \(n, k\), a sufficient condition for being in regime \(B\) of (3.2.16) is to have

$$z(n^{-2/3} a_k) \leq \lambda r_2 / n - \epsilon = D_r(\lambda, n). \quad (3.2.23)$$

and

$$r_2 \lambda/C_r < n \leq r_2 \lambda. \quad (3.2.24)$$

Hence, for all \(n, k\) sufficiently large with \(r_2 \lambda/C_r < n \leq r_2 \lambda/(1+\epsilon)\), a sufficient condition
for being in regime $B$ of (3.2.16) is

$$k \leq \frac{\sqrt{D_\epsilon(\lambda, n)^2 - 1}}{\pi} - \sec^{-1}(D_\epsilon(\lambda, n)) - \epsilon n. \quad (3.2.25)$$

Again throwing away a finite number of small pairs, we obtain

$$\liminf_{\lambda} \frac{N_B(\lambda^2)}{\lambda^2} \geq \liminf_{\lambda} \frac{1}{\pi \lambda^2} \sum_{\frac{r_2}{2\lambda} < n \leq \frac{r_2}{\lambda}} \left( n \sqrt{D_\epsilon(\lambda, n)^2 - 1} - n \sec^{-1}(D_\epsilon(\lambda, n)) - \epsilon n \right)$$

$$= \liminf_{\lambda} \frac{1}{\pi \lambda^2} \int_{\frac{r_2}{2\lambda}}^{\frac{r_2}{\lambda}} \left( \sqrt{D_\epsilon(\lambda, t)^2 - 1} - t \sec^{-1}(D_\epsilon(\lambda, t)) - \epsilon t \right) dt.$$

Each of the three summands in the integrand has elementary primitive, so we can explicitly compute this quantity.

Noting that $C_\epsilon, \hat{C}_\epsilon \to C$, we compute

$$\lim_{\epsilon \to 0} \liminf_{\lambda} \frac{N_A(\lambda^2) + N_B(\lambda^2)}{\lambda^2}$$

$$\geq \left( \frac{r_2^2 \sqrt{C^2 - 1}}{2\pi} - \frac{r_2^2 \pi}{2\pi} \frac{1}{2} \sin^{-1}(C^{-1}) \right)$$

$$+ \frac{1}{\pi \lambda^2} \int_{\frac{r_1}{\lambda}}^{\frac{r_2}{\lambda}} \sqrt{\frac{\lambda^2 r^2 - t^2}{2\pi}} \ dt - \frac{1}{\pi \lambda^2} \int_{\frac{r_1}{\lambda}}^{\frac{r_2}{\lambda}} \frac{t \sec^{-1}(\frac{\lambda r^2}{t}) dt}{2\pi^2}$$

$$= \frac{r_2}{8} \left( 1 - \frac{2}{\pi C^2} \sqrt{C^2 - 1} - \frac{2}{\pi} \sin^{-1}(C^{-1}) \right).$$

as required.

3.3 Spectral theory

We next establish the following key spectral theoretic result.

**Proposition 3.3.1.** Let $\mathcal{H}$ be a Hilbert space. Suppose $T \in \mathcal{L}(\mathcal{H})$ has a complete orthonormal system of eigenvectors $(u_i, E_i)_{i \in \mathbb{N}}$ with the sequence $(E_i)$ non-negative and increasing without bound. Suppose further that we have a family of normalised quasimodes $(v_i, E'_i)_{i=1}^{n}$ with

$$\| (T - E'_i) v_i \| < \epsilon_1 \quad (3.3.1)$$

and

$$| \langle v_i, v_j \rangle | < \epsilon_2 \quad \text{for } i \neq j \quad (3.3.2)$$

for some positive $\epsilon_1, \epsilon_2 > 0$.

We write

$$V = \text{Span}\{v_i\}_{i=1}^{n} \quad (3.3.3)$$
and
\[ U = \operatorname{Span}\{u_j : E_j \in \bigcup_{i=1}^{n} [E'_i - c, E'_i + c]\}. \tag{3.3.4} \]

We denote the orthogonal projection onto a subspace \( S \subseteq \mathcal{H} \) by \( \pi_S \).

If for some \( c > 0 \) and some \( 0 < \epsilon, \delta < 1/2 \) we have
\[ m = \#\{j \in \mathbb{N} : E_j \in \bigcup_{i=1}^{n} [E'_i - c, E'_i + c]\} < n(1 + \epsilon) \tag{3.3.5} \]
and
\[ \frac{c^2}{\epsilon^2} + \epsilon < \frac{\delta}{n} \tag{3.3.6} \]
then at least \( n(1 - \sqrt{\epsilon}) \) of the corresponding eigenvectors \( u_i \) satisfy
\[ \|u_i - \pi_U(u_i)\| < \epsilon^1/4 + 2\delta^{3/2}. \tag{3.3.7} \]

**Proof.** The idea behind the proof of the estimate (3.3.7) consists of several successive approximations.

We first show that the projections \( \pi_U(v_i) \) are almost orthogonal and can be transformed into an orthonormal basis \( (w_i)_{i=1}^{n} \) of their span by a matrix \( A \) that is approximately the identity.

We then show that excluding some exceptional eigenvectors, the remaining eigenvectors are necessarily rather close to the space \( W \). This implies that the non-exceptional eigenvectors can be well approximated by their projections, which leads us to conclude \( u \approx \pi_W(u) = Bw = BA\pi_U(v) \approx BA v \) for some matrix \( B \).

To begin, we reindex the eigenpairs \((u_i, E_i)\) so that \( E_j \in \bigcup_{i=1}^{n} [E'_i - c, E'_i + c] \) precisely for \( j = 1, 2, \ldots, m \).

The assumptions (3.3.1) and (3.3.2) then imply
\[
\|(T - E'_i) \sum_{j \in \mathbb{N}} \langle v_i, u_j \rangle u_j \|^2 < \epsilon_1^2 \]
\[ \Rightarrow \sum_{j=m+1}^{\infty} |E_j - E'_i|^2 |\langle v_i, u_j \rangle|^2 < \epsilon_1^2 \]
\[ \Rightarrow \sum_{j=m+1}^{\infty} |\langle v_i, u_j \rangle|^2 < \frac{\epsilon_1^2}{c^2} \]
\[ \Rightarrow \|\pi_U(v_i)\|^2 > 1 - \frac{\epsilon_1^2}{c^2}. \]

and
\[
|\langle \pi_U(v_i), \pi_U(v_j) \rangle| \leq |\langle v_i, v_j \rangle| + |\langle \pi_U(v_i), \pi_U(v_j) \rangle| \]
\[ < \epsilon_2 + \sqrt{(1 - \|\pi_U(v_i)\|^2)(1 - \|\pi_U(v_j)\|^2)} \]
\[ < \epsilon_2 + \frac{\epsilon_1^2}{c^2} \]
for \( i \neq j \).
Together with (3.3.6) we obtain
\[ \| \pi_U(v_i) \|^2 > 1 - \frac{\delta}{n} \] (3.3.8)
and
\[ |\langle \pi_U(v_i), \pi_U(v_j) \rangle| < \frac{\delta}{n} \] (3.3.9)
for \( i \neq j \).

The Gram matrix \( M \) with entries \( M_{ij} = \langle \pi_U(v_i), \pi_U(v_j) \rangle \) satisfies
\[ \| M - I \|_{HS} : = \| E \|_{HS} < \delta < 1/2. \] (3.3.10)

Note that if the collection \( \{ \pi_U(v_i) \} \) were linearly dependent, then the matrix \( M \) would be singular. The estimate (3.3.10) precludes this possibility, because we can invert \( M = I - (I - M) \) as a Neumann series. In particular, this implies that \( m \geq n \).

We now write \( W = \text{Span}\{ \pi_U(v_i) \}_{i=1}^n \) and suppose that \( (w_i)_{i=1}^n \) is an orthonormal basis for \( W \) which can be given by the transformation \( w = A\pi_U(v) \), where \( A \) is an \( n \times n \) real matrix that acts on the Hilbert space \( \mathcal{H}^n \) via matrix multiplication.

Expanding out the matrix equation \( \langle w_i, w_j \rangle = \delta_{ij} \) we obtain
\[ A M A^* = \sum_{k=1}^n \sum_{l=1}^n a_{ik} a_{lj} \langle \pi_U(v_k), \pi_U(v_l) \rangle = I \] (3.3.11)
which has a solution
\[ A = M^{-1/2} = (I + E)^{-1/2} = \sum_{k=0}^\infty \left( -\frac{1}{2} \right)^k E^k = \sum_{k=0}^\infty (-1)^k \binom{2k}{k} A^{-k} E^k. \] (3.3.12)

From (3.3.10), we then deduce
\[ \| A - I \|_{HS} \leq \sum_{k=1}^\infty \| E \|^k = \| E \| (1 - \| E \|)^{-1} < 2\delta. \] (3.3.13)

In the case \( m = n \), that is when \( W = U \), we can find a unitary matrix \( B \) with \( Bw = u \). We now assume \( m > n \), recalling that the assumptions of the proposition imply that this excess is small as a proportion of \( n \).

We have
\[ \sum_{i=1}^m \| \pi_W(u_i) \|^2 = m - \sum_{i=1}^m \| \pi_W(u_i) \|^2 = m - \sum_{i=1}^m \sum_{j=1}^n \| u_i, w_j \|^2 = m - n < n\epsilon \]
which implies that
\[ \# \{ i : \| \pi_W(u_i) \|^2 \geq \sqrt{\epsilon} \} < n\sqrt{\epsilon} \] (3.3.14)
and consequently
\[ \# \{ i : \| \pi_W(u_i) \|^2 > 1 - \sqrt{\epsilon} \} \geq m - n\sqrt{\epsilon} > n(1 - \sqrt{\epsilon}). \] (3.3.15)

We again re-index the eigenpairs for convenience, so that the first \( n' = \lceil n(1 - \sqrt{\epsilon}) \rceil \)
3.4 Results on eigenvalue flow

eigenvectors \( u_i \) satisfy the estimate in (3.3.15).

In this case, we define the \( n' \times n \) matrix \( B \) to have entries

\[
B_{ij} = \langle \pi W(u_i), w_j \rangle \quad (3.3.16)
\]

and the vector \( u \in H^{n'} \) by

\[
(u_i)_{i=1}^{n'} \quad (3.3.17)
\]

We then have

\[
\pi W(u) = Bw \quad (3.3.18)
\]

and the \( i \)-th row \( B_i \) of \( B \) has \( \ell^2 \) norm trivially bounded by 1.

This leaves us with

\[
u = (u - \pi W(u)) + BA\pi U(v) = (u - \pi W(u)) + BA\pi (v) \quad (3.3.19)
\]

which implies

\[
\| (u - BAv)_i \|_H \leq \| u_i - \pi W(u_i) \|_H + \| B_i \|_2 \| A \|_HS \| \pi U(v) - v \|_{H^n} < \epsilon^{1/4} + (2\delta) \cdot \sqrt{\delta}
\]

\[
< \epsilon^{1/4} + 2\delta^{3/2}.
\]

This estimate shows that each \( u_i \) has distance less than \( \epsilon^{1/4} + 2\delta^{3/2} \) to some element of \( V \).

Consequently

\[
\| u_i - \pi V(u_i) \| < \epsilon^{1/4} + 2\delta^{3/2} \quad (3.3.20)
\]

as required.

Our strategy to prove the main theorem is to control the number of eigenvalues in most clusters formed by finite unions of overlapping intervals of the form \( [\alpha_i^2 - c, \alpha_i^2 + c] \), and then repeatedly employ Proposition 3.3.1 to establish the existence of a subsequence of these eigenfunctions with upper density \( d \) that localises in the semidisk.

### 3.4 Results on eigenvalue flow

Central to the argument is the analysis of how eigenvalues flow as we vary \( t \).

Weyl’s law provides us with the asymptotic

\[
N_t(\lambda^2) \sim \frac{\lambda^2 |M_t|}{4\pi} \quad (3.4.1)
\]

where \( N_t \) is the counting function of the Dirichlet eigenvalues on \( M_t \).

To obtain a more precise statement about the change of individual eigenvalues, we employ an interior version of the Hadamard variational formula and Theorem 3.1.4.

In order to make use of Theorem 3.1.4, we choose \( \phi(y) \in C_c^\infty(\mathbb{R}) \) non-negative, supported near \( y = -1/2 \), and with integral 1, and we define the family of metrics

\[
g_t = dx^2 + (1 + (t - 1)\phi)^2 dy^2 \quad (3.4.2)
\]

on \( M_1 \).

This metric induces a natural isometry \( I_t : (M_1, g_t) \to (M_t, g_1) \).
If we define \( R_t = (1 + (t - 1)\phi)^{-1/2} \), then we have the following result from Proposition 7 of the appendix of [22].

**Proposition 3.4.1.** Let \( u(t) \) be an \( L^2 \)-normalised real Dirichlet eigenfunction of \( \Delta \) on \( M_t \) with corresponding eigenvalue \( E(t) \). We then have

\[
\dot{E}(t) = -\frac{1}{2}\langle Qu(t), u(t) \rangle
\]

(3.4.3)

where the operator \( Q \) is given by

\[
Q = -4\partial_y\partial_t + [\partial_y, [\partial_y, \phi_t]] = \phi''_t - 4(\phi'_t\partial_y + \phi_t\partial^2_y)
\]

(3.4.4)

on \( M_t \).

Here, \( \phi_t : M_t \to \mathbb{R} \) is given by:

\[
\phi_t = (\phi R^2_t) \circ I^{-1}_t.
\]

(3.4.5)

We now cut \( Q \) off away from the vertical sides of the stalk so that we can use the interior equidistribution result Theorem 3.1.4 to control the quantity \( E^{-1}_k \dot{E}_k \).

We do this by defining

\[
Q_\delta = \chi_\delta Q
\]

(3.4.6)

where \( \chi_\delta \in \mathcal{C}^\infty \) satisfies

\[
\chi_\delta(x) = \begin{cases} 
0 & \text{for } x \in [-r_1, -r_1 + \delta] \cup [r_1 - \delta, r_1] \\
1 & \text{for } x \in [-r_1 + 2\delta, r_1 - 2\delta].
\end{cases}
\]

(3.4.7)

**Proposition 3.4.2.** For any \( \epsilon > 0 \) and any \( t \in (0, 2] \), there exists \( \delta > 0 \) such that

\[
|E^{-1}_k \langle (Q_\delta - Q)u_{n_k}(t), u_{n_k}(t) \rangle| < \epsilon
\]

(3.4.8)

for all \( k \), where \( (n_k) \) is a \( t \)-dependent subsequence of the positive integers with lower density bounded below by \( 1 - \epsilon \).

**Proof.** First we show that it suffices for each \( t \) to establish the spectral projector estimates

\[
\|\eta_1[\lambda, \lambda+1](\sqrt{-\Delta})\|_{L^2(M_t) \to L^\infty(M_t)} = O(\lambda^{1/2})
\]

(3.4.9)

and

\[
\|\eta_1[\lambda, \lambda+1](\sqrt{-\Delta})\|_{L^2(M_t) \to L^\infty(M_t)} = O(\lambda^{3/2}).
\]

(3.4.10)

Here \( \eta_t = \eta \circ I^{-1}_t \) where \( \eta : M_t \to \mathbb{R} \) is an fixed smooth cutoff function supported and equal to 1 in a neighbourhood of \( \partial M_t \cap \text{spt}(\phi) \) such that \( \eta \) vanishes in a neighbourhood of the semidisk.

Applying \( \eta_1[\lambda, \lambda+1](\sqrt{-\Delta}) \) to \( \sum_{\lambda_j \in [\lambda, \lambda+1]} a_j u_j \) and using the estimate (3.4.10) then yields
for each \( x \in M_t \).

Setting \( a_j = \nabla u_j(x) \) then yields the estimate

\[
\eta_t(x)^2 \sum_{\lambda_j \in [\lambda, \lambda+1)} |\nabla u_j(x)|^2 \leq C\lambda^3. \quad (3.4.11)
\]

Similarly, we obtain

\[
\eta_t(x)^2 \sum_{\lambda_j \in [\lambda, \lambda+1)} |u_j(x)|^2 \leq C\lambda. \quad (3.4.12)
\]

The estimates (3.4.11) and (3.4.12) then allow us to control each term of (3.4.8) in an average sense.

For example, as only the the horizontal component \( 1 - \chi_\delta \) of the cutoff function in (3.4.8) is \( \delta \)-dependent, we can integrate by parts in the second order term in (3.4.8) without loss.

Then, by writing \( \eta_\delta \) to denote the cutoff function in the new second order term and choosing \( \delta > 0 \) sufficiently small so that \( \eta_\delta = \eta_\delta \), the contribution of these terms to (3.4.8) is controlled by

\[
E^{-1} \sum_{E_j \leq E} E_j^{-1} \int_M \eta_\delta(x)|\nabla u_j(x)|^2 \, dx
\]

\[
\approx E^{-1} \sum_{k=1}^{E^{1/2} - 1} \sum_{\lambda_j \in [k,k+1)} E_j^{-1} \int_M \eta_\delta(x)|\nabla u_j(x)|^2 \, dx
\]

\[
\leq E^{-1} \sum_{k=1}^{E^{1/2} - 1} k^{-2} \int_M \eta_\delta(x)\eta(x) \sum_{\lambda_j \in [k,k+1)} |\nabla u_j(x)|^2 \, dx
\]

\[
\leq CE^{-1} \left( \sum_{k=1}^{E^{1/2} - 1} k \right) \int_M \eta_\delta(x) \, dx
\]

\[
\leq \tilde{C}_\delta
\]

for sufficiently small \( \delta > 0 \), where \( \tilde{C}_\delta \to 0 \) as \( \delta \to 0 \).

Together with analogous estimates for lower order terms, we obtain the estimate

\[
\frac{1}{n} \sum_{j=1}^{n} E_j^{-1} |\langle (Q_\delta - Q)u_j, u_j \rangle| < C_\delta \quad (3.4.13)
\]

where \( C_\delta \to 0 \) as \( \delta \to 0 \).

By taking \( \delta \) sufficiently small that \( C_\delta < \epsilon^2 \), we ensure that the collection of \( j \) with \( E_j^{-1} |\langle (Q_\delta - Q)u_j, u_j \rangle| \geq \epsilon \) has upper density at most \( \epsilon \).
The estimate (3.4.9) follows from Proposition 8.1 in [20]. Note that the finite propagation speed of the operator \( \cos(t\sqrt{-\Delta}) \), the post-composition with a cutoff near a flat boundary, and the small-time nature of the argument together imply that \( M_t \) can be treated as the half-plane, which certainly satisfies the geometric assumptions of the cited result.

By inserting the gradient operator in the dual estimate, it remains to control \( \|1_{(\lambda,\lambda+1)}(\sqrt{-\Delta})\nabla \eta\|_{L^1(M_t)} \rightarrow L^2(M_t) \) in order to give us (3.4.10).

The argument in the proof of Proposition 8.1 in [20] allows us to replace the spectral projector by a smooth spectral projector \( \rho_{ev}^\lambda(\sqrt{-\Delta}) \) where \( \rho_{ev}^\lambda(s) = \rho(s-\lambda) + \rho(-s-\lambda) \) and \( \rho \in S(\mathbb{R}) \) has non-negative Fourier transform supported in \([\epsilon/2,\epsilon]\) for some sufficiently small \( \epsilon \).

So it suffices to estimate the \( L^1 \rightarrow L^2 \) norm of the operator

\[
\rho_{ev}^\lambda(\sqrt{-\Delta})\nabla = \frac{1}{\pi} \int_\mathbb{R} \cos(t\sqrt{-\Delta})(e^{-it\lambda\hat{\chi}}(t) + e^{it\lambda\hat{\chi}}(-t))\nabla \eta(x) \, dt. \tag{3.4.14}
\]

This integral is supported close to \( t = 0 \), and hence by finite propagation speed, the kernel of the wave equation solution operator \( \cos(t\sqrt{-\Delta}) \) on \( M \) is identical to that of the half-plane.

Moreover, the kernel of the wave equation solution operator on the half-plane can be obtained from the free space wave kernel by the reflection principle, and their \( L^1 \rightarrow L^2 \) norms are identical.

This implies that it suffices to prove the estimate with the kernel for \( \cos(t\sqrt{-\Delta}) \) replaced by the free space wave kernel.

So the kernel to be estimated is

\[
\frac{1}{4\pi^3} \int_\mathbb{R} \int_\mathbb{R} \int_\mathbb{R} e^{i(x-y)\cdot \xi}(e^{-it\lambda\hat{\chi}}(t) + e^{it\lambda\hat{\chi}}(-t))\xi \cos(|\xi|t)\eta(x) \, dy \, d\xi \, dt
= \nabla_x(K_\lambda(x,y))\eta(x)
= \nabla_x(\lambda^{(n-1)/2}a_\lambda(x,y)e^{i\lambda\psi(x,y)})\eta(x)
= O(\lambda^{(n+1)/2})
\]

where \( K_\lambda, a_\lambda, \psi \) are as in Lemma 5.13 from [44], which we make use of in our penultimate line.

Duality then completes the proof of (3.4.10) and the proposition.

Proposition 3.4.1 allows us to use Theorem 3.1.4 and Proposition 3.4.2 to control the flow speed of a full density subsequence of the eigenfunctions for any fixed \( t \).

**Proposition 3.4.3.** For each \( t \in (0,2\] \) there exists a full density subsequence \((n_k)\) of the positive integers such that

\[
\liminf_{k \to \infty} E_{n_k}^{-1}(t)\dot{E}_{n_k}(t) \geq -\frac{\dot{A}(t)}{A(t)(1-d(t))} \tag{3.4.15}
\]

where \( d(t) \) denotes the proportion of the phase space volume that is in the completely integrable region \( S^*M_t \setminus U_t \).

**Proof.** From Proposition 3.4.1, Proposition 3.4.2 and Theorem 3.1.4, for all \( \epsilon > 0 \) we may choose a \( \delta > 0 \) and a subsequence of eigenfunctions with lower density bounded below by
1 − ε such that we have the estimate (3.4.8) and such that we have a unique semiclassical measure μ with μ|_{U_t} = aμ_L|_{U_t} for some non-negative constant α. We immediately have

$$a = \frac{\mu(U_t)}{1 - d(t)} \leq \frac{1}{1 - d(t)}. \quad (3.4.16)$$

The definition of semiclassical measures then implies

$$\liminf_{k \to \infty} E^{-1}_{n_k}(t) \dot{E}_{n_k} \geq -\frac{1}{2} \int S^* M \sigma(Q_\delta) d\mu - \epsilon \geq -\frac{1}{2} \int S^* M 4\phi_t \xi_2^2 d\mu - \epsilon \geq -\frac{1}{\pi(1 - d(t)) A(t)} \int_M \int_0^{2\pi} \phi_t(x) \sin^2(\theta) dx \ d\theta - \epsilon \geq -\frac{\dot{A}(t)}{(1 - d) A(t)} - \epsilon.$$

We can then apply Lemma D.1 to obtain a full density subsequence of eigenfunctions satisfying the estimate (3.4.17).

Moreover, we can strengthen the above to an almost-uniform result.

**Proposition 3.4.4.** For any ε > 0, there exists a full density subsequence \( n_k \) of positive integers and a family of sets \( B_{n_k} \subset (0, 2] \) with \( m(B_{n_k}) \to 0 \) such that

$$E^{-1}_{n_k}(t) \dot{E}_{n_k}(t) > -\frac{\dot{A}(t)}{A(t)(1 - d(t))} - \epsilon \quad (3.4.17)$$

for each \( t \in (0, 2] \setminus B_{n_k} \).

**Proof.** For each δ > 0 we define the subset \( S_\delta \subset \mathbb{N} \) as the collection of \( n \in \mathbb{N} \) such that

$$m\{t \in (0, 2] : E^{-1}_{n}(t) \dot{E}_{n}(t) \leq -\frac{\dot{A}(t)}{A(t)(1 - d(t))} - \epsilon\} > \delta. \quad (3.4.18)$$

If every \( S_\delta \) were of zero density, we could write \( S'_\delta := \mathbb{N} \setminus S_\delta \) and use Lemma D.1 to assemble a full density set satisfying the claims of the proposition.

Now suppose that \( S_\delta \) has positive upper density for some \( \delta > 0 \).

Since for every \( n \in S_\delta \), the sets \( B_n = \{t : E^{-1}_{n}(t) \dot{E}_{n}(t) \leq -\frac{\dot{A}(t)}{A(t)(1 - d(t))} - \epsilon\} \) have measure bounded below and are subsets of a set with finite measure, there must exist a further positive density subset \( \hat{S}_\delta \subset S_\delta \) such that

$$\bigcap_{n \in \hat{S}_\delta} B_n \neq \emptyset. \quad (3.4.19)$$

This can be seen for instance by applying the bounded convergence theorem to the function

$$\frac{1}{n} \sum_{j=1}^{n} 1_{B_j}. \quad (3.4.20)$$
The existence of a \( t \) in this intersection contradicts Proposition 3.4.3 and hence completes the proof.

The almost-uniform result in Proposition 3.4.4 can for our purposes be treated as a uniform bound on speed for large \( E_j \), in light of the following weaker bound for \( t \) in the sets \( B_j \) of diminishing measure for which (3.4.17) does not hold.

**Proposition 3.4.5.** There exists a positive constant \( C \) such that for every \( t \in (0, 2] \) and every \( j \), we have

\[
\dot{E}_j(t) = -\frac{1}{2} \langle Qu_j, u_j \rangle \geq -CE_j(t) \quad (3.4.21)
\]

**Proof.** Integration by parts and Cauchy–Schwartz on the left-hand side provides us with a lower bound of

\[
-\hat{C} \int_M |(\partial_\tau u_j)u_j| + |(\partial_y u_j)u_j| + |u_j|^2 \, dx \, dy \\
\geq -\hat{C}((-\Delta u_j, u_j) + (-\Delta u_j, u_j)^{1/2} + 1) \\
\geq -CE_j
\]

for a positive constant \( C \) that is uniform in time. \qed

**Corollary 3.4.6.** For any \( \epsilon > 0 \), there exists \( \delta > 0 \) and a full density subsequence \((n_k)\) such that we have

\[
-\int_S \dot{E}_{n_k}(t) \, dt \leq E_{n_k}(t_1) \left( \frac{A(t_1)}{A(t_1)(1-d(t_1))} + \epsilon \right) (t_2 - t_1) \quad (3.4.22)
\]

for any measurable set \( S \subseteq (0, 2] \) with measure greater than \( \delta \).

**Proof.** This follows from Proposition 3.4.4 and Proposition 3.4.5 by removing finitely many elements from the subsequence constructed in Proposition 3.4.4. \qed

### 3.5 Main results

We are now in a position to prove the main result of the chapter.

**Theorem 3.5.1.** The weak Percival’s conjecture 3.1.3 holds for the mushroom billiard \( M_t \) for any fixed inner and outer radii, and almost all “stalk lengths” \( t \in (0, 2] \).

We prove Theorem 3.5.1 by establishing the claim for \( t \in \mathcal{G} \subset (0, 2] \) satisfying a certain spectral nonconcentration property, and then prove that this set is of full measure.

For a given \( c > 0 \) we define \( c \)-clusters to be the connected components of \( \cup_{i \in \mathbb{N}} [\alpha_i^2 - c, \alpha_i^2 + c] \), indexed \( C_k \) in increasing order.

**Definition 3.5.2.** We call \( t \in (0, 2] \) good if for every \( \epsilon > 0 \), there exists some \( c > 0 \) with

\[
\liminf_{n \to \infty} \frac{\# \{ j \in \mathbb{N} : E_j(t) \in \cup_{k=1}^n C_k \}}{\# \{ j \in \mathbb{N} : \alpha_j^2 \in \cup_{k=1}^n C_k \}} < 1 + \epsilon^2 \quad (3.5.1)
\]

We denote the set of good times by \( \mathcal{G} \).
We begin by proving the claim for fixed $t \in \mathcal{G}$. We write $N_{\text{semidisk}}(C), N_{\text{mushroom}}(C)$ to denote the number of quasi-eigenvalues and eigenvalues respectively contained in a given $c$-cluster $C$.

The assumption (3.5.1) then implies the following Proposition.

**Proposition 3.5.3.** Suppose $t$ and $0 < c < 2/r^2$ are such that (3.5.1) holds, and all but finitely many $c$-clusters contain at least as many eigenvalues as quasi-eigenvalues. Then there exists a subset $J$ of quasi-eigenvalues with upper density at least $(1 - \epsilon)$ such that each quasi-eigenvalue in $J$ is contained in a cluster $C$ with

$$N_{\text{semidisk}}(C) \leq N_{\text{mushroom}}(C) \leq (1 + \epsilon)N_{\text{semidisk}}(C) \quad (3.5.2)$$

**Proof.** Index the $c$-clusters $C_k$ in increasing order.

Let

$$S = \{k : N_{\text{semidisk}}(C_k) \leq N_{\text{mushroom}}(C_k) \leq (1 + \epsilon)N_{\text{semidisk}}(C_k)\}$$

and

$$F = \{k : N_{\text{mushroom}}(C_k) < N_{\text{semidisk}}(C_k)\}.$$ 

From the defining property (3.5.1) of $t \in \mathcal{G}$, we have:

$$\liminf_{n \to \infty} \frac{\sum_{k \leq n} N_{\text{mushroom}}(k)}{\sum_{k \leq n} N_{\text{semidisk}}(k)} < 1 + \epsilon^2 \quad (3.5.3)$$

The definition of $S$ implies

$$\liminf_{n \to \infty} \left(1 - \frac{\sum_{k \leq n} N_{\text{semidisk}}(k)1_F(k)}{\sum_{k \leq n} N_{\text{semidisk}}(k)}\right) + \epsilon \frac{\sum_{k \leq n} N_{\text{semidisk}}(k)(1 - 1_F(k) - 1_S(k))}{\sum_{k \leq n} N_{\text{semidisk}}(k)} < 1 + \epsilon^2.$$  

The second term on the left-hand side is $o(1)$ from the finiteness assumption. Hence we obtain that the lower density of $\mathbb{N} \setminus (S \cup F)$ is bounded above by $\epsilon$. As $F$ is finite, and consequently of density 0, we can conclude that the upper density of $S$ is bounded below by $1 - \epsilon$ as required. 

**Theorem 3.5.4 (Main Theorem).** For each $t \in \mathcal{G}$, there exists $B_t \subset \mathbb{N}$ of density $d(t)$ such that any semiclassical measure associated to the eigenfunctions $(u_n)_{n \in B_t}$ is supported inside the completely integrable region.

**Proof.** First, we fix $\epsilon > 0$ and choose $c > 0$ small enough so that the inequality (3.5.1) holds.

Proposition 3.2.2 implies that we may choose the $\epsilon_1, \epsilon_2$ in applications of Proposition 3.3.1 to the increasing sequence of $c$-clusters to decay faster than any polynomial in the energy infima of these $c$-clusters.

By Weyl’s law, this ensures that for all but possibly finitely many $c$-clusters, we have (3.3.6) with $\delta$ decaying faster than any polynomial in energy. We remove the exceptional $c$-clusters, without any loss in density of our subset.
In light of Proposition 3.5.3, we can then select a subset of the remaining $c$-clusters such that the included subset of quasi-eigenvalues has upper density exceeding $1 - \epsilon$ and such that (3.5.2) holds for each cluster.

We can now apply Proposition 3.3.1 on a cluster-by-cluster basis, with parameter $\delta \to 0$ faster than any polynomial in energy.

From the $L^2$ boundedness of pseudodifferential operators with compactly supported symbols, this implies that for each $\epsilon$, we get a subsequence of eigenfunctions $u_{jk}$ such that any associated semiclassical measure $\mu$ satisfies

$$\mu(D_t \setminus U_t) \geq 1 - \epsilon^{1/4}.$$  \hspace{1cm} (3.5.4)

Moreover, by comparison to Proposition 3.2.3 and Weyl’s law for the mushroom, we obtain a lower bound of $d(t) - h(\epsilon)$ for the upper density of this eigenfunction subsequence with $h(\epsilon) \to 0$ as $\epsilon \to 0$. So for each $\epsilon > 0$, we can find a $c$ and a subsequence of $(u_n)$ with upper density at least $d(t) - h(\epsilon)$ which concentrates in the completely integrable region up to $\epsilon^{1/4}$ of its semiclassical mass.

We now take a sequence $\epsilon_j \to 0$ and denote the corresponding eigenvalue window widths by $c_j$. We write $B_{j,t}$ to denote the corresponding concentrating eigenfunction subsequences.

Lemma D.2 then allows us to obtain a subsequence $B_t$ of lower density at least $d(t)$ such that any associated semiclassical measure $\mu$ satisfies

$$\mu(D_t \setminus U_t) = 1.$$  \hspace{1cm} (3.5.5)

To show that the upper density of $B_t$ cannot exceed $d(t)$, we choose a function $\chi_\epsilon \in C^\infty_c(\mathbb{R}^2 \times \mathbb{R}^2)$ supported in the interior of $M$ such that the following properties are satisfied.

• $0 \leq \chi_\epsilon \leq 1$
• $\chi_\epsilon|_{D_t \setminus U_t} = 0$
• $\int_{D_t} \chi_\epsilon \, d\mu_L > (1 - \epsilon)\mu_L(U_t)$.

Applying the local Weyl law (Lemma 4 from [49]) to the corresponding semiclassical pseudodifferential operator $\chi_\epsilon(x, hD)$, we obtain

$$\frac{1}{n} \sum_{j \in [1, n] \cap B_t} \langle \chi_\epsilon(x, E_j^{-1/2}D)u_j, u_j \rangle + \frac{1}{n} \sum_{j \in [1, n] \cap B_t^c} \langle \chi_\epsilon(x, E_j^{-1/2}D)u_j, u_j \rangle > (1 - \epsilon)\mu_L(U_t)$$  \hspace{1cm} (3.5.6)

for sufficiently large $n$.

The localisation property (3.5.5) implies that the first summand is $o(1)$ in $n$. Hence we have

$$\frac{1}{n} \sum_{j \in [1, n] \cap B_t^c} \langle \chi_\epsilon(x, E_j^{-1/2}D)u_j, u_j \rangle > (1 - 2\epsilon)\mu_L(U_t)$$  \hspace{1cm} (3.5.7)

for sufficiently large $n$.

From Theorem 3.1.4 and the bound $a \leq \mu_L(U_t)^{-1}$ that is immediate from semiclassical measures being probability measures, it follows that a full density subset of the remaining summands must be bounded above by 1.
This implies that
\[
\frac{\# \{ j \leq n : j \in B_t^c \}}{n} > (1 - 2\epsilon)\mu_L(U_t) \tag{3.5.8}
\]
for sufficiently large \( n \).

Rearranging and passing to the limit \( n \to \infty \) and then \( \epsilon \to 0 \), we obtain the required upper bound of
\[
\limsup_{n \to \infty} \frac{\# \{ j \leq n : j \in B_t \}}{n} \leq 1 - \mu_L(U_t) = d(t). \tag{3.5.9}
\]

Hence \( B_t \) is a sequence of eigenfunctions of upper density \( d(t) \) with semiclassical mass supported in the completely integrable region.

**Proposition 3.5.5.** Let \( A_t = \mathbb{N} \setminus B_t \). Then for each \( t \in \mathcal{G} \), a full density subsequence of \( (u_n)_{n \in A_t} \) equidistributes in \( U_t \).

**Proof.** From Theorem 3.5.4, the sequence of eigenfunctions \( (u_n)_{n \in B_t} \) has all semiclassical mass in the completely integrable region and \( B_t \) has natural density \( d(t) \).

Applying the local Weyl law again with the function \( \chi_\epsilon \) from the proof of Theorem 3.5.4, we obtain
\[
\frac{1}{n} \sum_{j \in [1, n] \cap B_t^c} \langle \chi_\epsilon(x, E_j^{-1/2}D)u_j, u_j \rangle > (1 - 2\epsilon)\mu_L(U_t) \tag{3.5.10}
\]
for sufficiently large \( n \).

Then, splitting the summation into the set
\[
A_{\epsilon, t} = \{ j \in B_t^c : \langle \chi_\epsilon(x, E_j^{-1/2}D)u_j, u_j \rangle < 1 - \sqrt{\epsilon} \} \tag{3.5.11}
\]
and its complement, the upper bound of 1 for a full density subset of the summands in (3.5.10) then implies:
\[
\liminf_{n \to \infty} d_n(A_{\epsilon, t}) < 2\sqrt{\epsilon}\mu_L(U_t) \tag{3.5.12}
\]
using the notation \( d_n \) from Lemma D.1. Thus we obtain a subset of upper density exceeding \( 1 - O(\sqrt{\epsilon}) \) of \( A_t \) with at least \( \mu(U_t) > 1 - O(\sqrt{\epsilon}) \) for any corresponding semiclassical measure.

An application of Lemma D.2 then gives us a subsequence of \( A_t \) with full upper density and all semiclassical mass in \( U_t \).

Together with Theorem 3.1.4, this implies that we can find a subsequence \( (u_{n_k}) \) of \( A_t \) with full upper density such that every associated semiclassical measure is of the form \( 1_{U_t} \cdot \mu_L(U_t)^{-1}\mu_L \).

We now show that the set \( (0, 2] \setminus \mathcal{G} \) has Lebesgue measure 0.

**Proposition 3.5.6.** If \( (0, 2] \setminus \mathcal{G} \) has positive Lebesgue measure, then there exists some \( \epsilon > 0 \) and some interval \( \mathcal{I} = [t_1, t_2] \subset (0, 2] \) such that
\[
\frac{1}{|\mathcal{I}|} \int_{\mathcal{I}} \liminf_{n \to \infty} \left( \frac{\# \{ j \in \mathbb{N} : E_j(t) \in \bigcup_{i=1}^n [\alpha_i^2 - c, \alpha_i^2 + c] \}}{n} \right) dt > 1 + \epsilon \tag{3.5.13}
\]
for all \( c > 0 \). Moreover, we can find such \( \mathcal{I} \) with arbitrarily small length.
Proof. By the monotone convergence property of measures, if \( m((0, 2] \setminus \mathcal{G}) > 0 \) then there must exist \( \epsilon > 0 \) and a positive measure set \( S \subseteq (0, 2] \) on which we have
\[
\liminf_{n \to \infty} \left( \frac{\# \{ j \in \mathbb{N} : E_j(t) \in \bigcup_{i=1}^{n} [\alpha_i^2 - c, \alpha_i^2 + c] \}}{n} \right) > 1 + 2\epsilon
\] (3.5.14)
for all \( t \in S \) and for all \( 0 < c < 2/r_2^2 \). From the regularity of the Lebesgue measure, we can find an open set \( S \subseteq U \subseteq (0, 2] \) with \( m(U) < m(S) + \delta \) for an arbitrarily small \( \delta \). We then have
\[
\frac{1}{|S|} \int_S \liminf_{n \to \infty} \left( \frac{\# \{ j \in \mathbb{N} : E_j(t) \in \bigcup_{i=1}^{n} [\alpha_i^2 - c, \alpha_i^2 + c] \}}{n} \right) dt > 1 + 2\epsilon \] (3.5.15)
and
\[
\frac{1}{|U \setminus S|} \int_{U \setminus S} \liminf_{n \to \infty} \left( \frac{\# \{ j \in \mathbb{N} : E_j(t) \in \bigcup_{i=1}^{n} [\alpha_i^2 - c, \alpha_i^2 + c] \}}{n} \right) dt \geq 1. \] (3.5.16)
from our pointwise bounds on the integrands. By choosing \( \delta \) sufficiently small, we are thus guaranteed the estimate
\[
\frac{1}{|U|} \int_U \liminf_{n \to \infty} \left( \frac{\# \{ j \in \mathbb{N} : E_j(t) \in \bigcup_{i=1}^{n} [\alpha_i^2 - c, \alpha_i^2 + c] \}}{n} \right) dt > 1 + \epsilon. \] (3.5.17)
Writing the open set \( U \) as a countable union of disjoint open intervals, the average of the integrand over one such interval must exceed \( 1 + \epsilon \), as claimed.

To complete the proof, we observe if we partition \( I \) into arbitrarily many intervals of equal length, at least one of them must also satisfy (3.5.13).

To culminate the argument, we seek out a contradiction coming from the upper bound (3.4.17) on speed of eigenvalue variation and the lower bound (3.5.13) on the average proportion of eigenvalues lying in \( c \)-clusters.

**Proposition 3.5.7.** For any \( \epsilon > 0 \), there exists \( c > 0 \) such that
\[
\limsup_{m \to \infty} \frac{1}{m} \sum_{j=1}^{m} \frac{|\{ t \in I : E_j(t) \in \bigcup_i [\alpha_i^2 - c, \alpha_i^2 + c] \}|}{|I|} < d(t_1) + \epsilon
\] (3.5.18)
for any sufficiently small interval \( I \).

**Proof.** Note that we have the flow speed bound (3.4.17) for a full density subsequence of eigenvalues, so if we can establish the claimed inequality for each summand with a sufficiently large index that obeys the flow speed bound, density will allow us to draw the desired conclusion.

We now suppose \( E_j \) is a large eigenvalue that lies in this full density subsequence.

Writing \( X = (A(t_1)^{-1} - A(t_2)^{-1}) \) for brevity, Weyl’s law applied to the mushroom gives
\[
E_j(t_1) - E_j(t_2) > (4\pi X - 2\delta)j
\] (3.5.19)
for \( \delta > 0 \) and all sufficiently large \( j \).

Weyl’s law for the semidisk (recalling that we constructed the completely integrable
region quasimodes from semidisk eigenfunctions) gives us an upper bound of
\[
\left( \frac{\pi r^2 X}{2} + \delta \right) j \tag{3.5.20}
\]
for the number of quasi-eigenvalues in \( [E_j(t_2), E_j(t_1)] \) and hence an upper bound of
\[
2c \left( \frac{\pi r^2 X}{2} + \delta \right) j \tag{3.5.21}
\]
for the length of \( [E_j(t_2), E_j(t_1)] \) that lies within \( \cup_i [\alpha_i^2 - c, \alpha_i^2 + c] \).

Now suppose that \( E_j(t) \) spends proportion \( q_j \) of \( t \in I \) in \( \cup_i [\alpha_i^2 - c, \alpha_i^2 + c] \).

From Proposition 3.4.5, it follows that the \( q_j \) are uniformly bounded above by some \( 1 - \delta \).

This means that we apply Corollary 3.4.6 to find a lower bound for the time taken by an eigenvalue \( E_j \) in our full density subsequence to traverse the set \( [E_j(t_2), E_j(t_1)] \cup \cup_i [\alpha_i^2 - c, \alpha_i^2 + c] \). Heuristically, we can think of this as dividing the size of this set by an upper bound for the speed of the eigenvalue’s variation.

Precisely, we have
\[
(1 - q_j)(t_2 - t_1) > \frac{jX(4\pi - \pi r^2 c) - j\delta(1 + 2c)}{E_j(t_1) \left( \frac{A(t_1)}{A(t_1)(1 - d(t_1))} + \delta \right)} \tag{3.5.22}
\]
\[
= \frac{j}{E_j(t_1)} \cdot \left( \frac{X(4\pi - \pi r^2 c) - \delta(1 + 2c)}{A(t_1)(1 - d(t_1)) + \delta} \right) \tag{3.5.23}
\]
\[
> \left( \frac{4\pi}{A(t_1) + \delta} \right)^{-1} \cdot \left( \frac{X(4\pi - \pi r^2 c) - \delta(1 + 2c)}{A(t_1)(1 - d(t_1)) + \delta} \right) \tag{3.5.24}
\]
\[
> \frac{XA(t_2)^2(1 - d(t_1)) - \epsilon}{A(t_1)} \tag{3.5.25}
\]
where the final two lines follow from Weyl’s law and passing to sufficiently small \( \delta \) and \( c \) respectively.

Additionally, since \( A(t) \) is a linear polynomial in \( t \), we have
\[
X = \frac{A(t_2) - A(t_1)}{A(t_1)A(t_2)} = \frac{(t_2 - t_1)\dot{A}(t_1)}{A(t_1)A(t_2)} \tag{3.5.26}
\]
which implies that
\[
1 - q_j > \frac{A(t_1)}{A(t_2)}(1 - d(t_1)) - \frac{\epsilon}{2} \tag{3.5.27}
\]
\[
q_j < d(t_1) + (d(t_1) - 1) \left( \frac{A(t_1)}{A(t_2)} - 1 \right) + \frac{\epsilon}{2} \tag{3.5.28}
\]
\[
< d(t_1) + \epsilon \tag{3.5.29}
\]
for sufficiently small \( |I| \), using the uniform continuity of \( A \).

Thus we have the required inequality for all sufficiently small intervals \( I \) and all sufficiently large \( j \) in a full density subsequence on which (3.4.17) holds.
**Proposition 3.5.8.** For any $\epsilon > 0$ there exists $c > 0$, such that

$$
\frac{1}{|I|} \int |I| \liminf_{n \to \infty} \left( \frac{\# \{ j \in \mathbb{N} : E_j(t) \in \bigcup_{i=1}^{n} [\alpha_i^2 - c, \alpha_i^2 + c] \} }{n} \right) dt < 1 + \epsilon
$$

for all sufficiently small $|I|$.

**Proof.** By Fatou’s lemma, it suffices to show that we can find $c$ such that

$$
\frac{1}{|I|} \int |I| \frac{\# \{ j : E_j(t) \in \bigcup_{i=1}^{n} [\alpha_i^2 - c, \alpha_i^2 + c] \} }{n} dt < 1 + \frac{\epsilon}{2}
$$

for sufficiently large $n$. This quantity is bounded above by

$$
\frac{1}{n} \sum_{j : E_j(t_2) < \alpha_n^2 + c} \frac{|\{ t \in I : E_j(t) \in \bigcup_{i=1}^{n} [\alpha_i^2 - c, \alpha_i^2 + c] \}|}{|I|} = \frac{1}{n} \sum_{j : E_j(t_2) < \alpha_n^2 + c} q_j.
$$

The sum is controlled by the previous proposition, giving us an upper bound of

$$
\frac{1}{n} \cdot \max \{ j : E_j(t_2) < \alpha_n^2 + c \} (d(t_1) + \delta)
$$

for sufficiently large $n$. From Weyl’s law, we have

$$
\max \{ j : E_j(t_2) < \alpha_n^2 + c \} < (\alpha_n^2 + c) \left( \frac{A(t_2)}{4\pi} + \delta \right)
$$

for sufficiently large $n$. By taking $\delta, c, \text{ and } I$ sufficiently small, we then obtain

$$
\frac{1}{n} \sum_{j : E_j(t_2) < \alpha_n^2 + c} q_j < \left( \frac{\alpha_n^2}{n} \cdot \frac{A(t_2) d(t_2)}{4\pi} \right) + \epsilon.
$$

Inverting the estimate (3.2.11) provides an upper bound of $1 + \frac{\epsilon}{4}$ for the first summand on the right-hand side for all sufficiently large $n$, thus completing the proof.

**Corollary 3.5.9.** The set $G$ has full measure in $(0, 2]$.

**Proof.** This is an immediate consequence of Propositions 3.5.6 and 3.5.8.

This also completes the proof of Theorem 3.5.1.
KAM Theory

The KAM theory studied by Kolmogorov, Arnold and Moser in the 60’s led to an improved understanding of the classical dynamics of a Hamiltonian $H$ that is a small perturbation of a completely integrable Hamiltonian $H^0$. In particular, they established that the Lagrangian invariant tori corresponding to all but a $o(1)$ measure subset of frequencies survive this perturbation as the size of the perturbation tends to zero. We denote the union of these tori by $\Lambda$.

The paper [36] uses a local version of the KAM theorem to construct a Birkhoff normal form for Gevrey class Hamiltonians $H$ about $\Lambda$. This normal form generalises the notion of “action-angle” variables of a completely integrable Hamiltonian as discussed in [3], and as a consequence of the normal form construction, Popov obtains an effective stability result for the Hamiltonian flow near the union of remaining invariant tori. The natural setting for the estimates is that of Gevrey regularity. This work generalises Popov’s earlier work in [34] and [35] where he constructs a Birkhoff normal form for real analytic Hamiltonians.

The paper [37] uses semiclassical Fourier integral operators to construct a corresponding quantum normal form for a class of semiclassical pseudodifferential operators $\mathcal{P}_h$, and uses this normal form to obtain a family of quasimodes microlocalised near the union of preserved invariant tori of a Hamiltonian associated to $\mathcal{P}_h$. In particular, this result is applicable to the Schrödinger operator.

This can be regarded as an extension of the main result in [10], which establishes the existence of quasimodes microlocalised near the Lagrangian tori of a completely integrable Hamiltonian on a compact smooth manifold. These results were obtained by a similar quantum normal form.

In this chapter we construct a family of Birkhoff normal forms corresponding to the convex family of Gevrey smooth Hamiltonians $H = H^0 + tQ$, real-analytic in the parameter $t \in (-1, 1)$. We shall complete the construction of quasimodes in Chapter 5 and use them to obtain results about the localisation of exact eigenfunctions.

Our treatment runs along the same lines as that of Popov [36], yet we shall be fairly explicit in our presentation to ensure that the presence of the parameter $t$ does not affect any of the essential details in [36].
4.1 Notations

We begin by introducing some notation that will be frequently used throughout Chapters 4 and 5.

The following function spaces will be the natural ones for much of this topic.

Definition 4.1.1. For \( \rho \geq 1 \) and \( X \subset \mathbb{R}^n \) open, the Gevrey class of order \( \rho \) is given by

\[
G^{\rho}(X) := \{ f \in C^\infty(X) : \sup_{x \in X} |\partial_x^hf(x)|L^{-|\alpha|}\alpha!^{-\rho} < \infty \}. \tag{4.1.1}
\]

If \( f \in G^{\rho}_L(X) \), the supremum in (4.1.1) is denoted by \( \|f\|_L \). We will frequently suppress the \( L \) in our notation. Equipped with this norm, \( G^{\rho}_L(X) \) is a Banach space.

Gevrey regularity is generally weaker the real analyticity (they coincide when \( \rho = 1 \) as can be seen by using the Cauchy–Hadamard theorem to characterise analytic functions by the growth of their Taylor coefficients) and importantly, there exist bump functions in the Gevrey class for \( \rho > 1 \).

An important property of the Gevrey class that follows from Taylor’s theorem is that if a Gevrey function has vanishing derivatives, then locally it is super-exponentially small.

Proposition 4.1.2. Suppose \( f \in G^\rho(X) \), and \( \rho > 1 \). Then there exist positive constants \( c, C, \eta \) and \( r_0 \) only dependent on the Gevrey constant \( L \), the norm \( \|f\|_L \), and the set \( X \) such that

\[
f(x_0 + r) = \sum_{|\alpha| \leq \eta |r|^{1/(1-\rho)}} f_\alpha(x_0) r^\alpha + R(x_0, r) \tag{4.1.2}
\]

where \( f_\alpha = (\partial^\alpha f)/\alpha! \) and

\[
|\partial_x^\beta R(x_0, r)| \leq C^{1+|\beta|} \beta!^\rho e^{-c |r|^{1/(\rho-1)}} \quad \forall 0 < |r| \leq \min(r_0, d(x_0, \mathbb{R}^n \setminus X)). \tag{4.1.3}
\]

In the study of KAM systems, we will need to work on functions on products of domains in Euclidean spaces with different degrees of Gevrey regularity in each variable.

To this end, we introduce the class of anisotropic Gevrey spaces.

Definition 4.1.3. Suppose \( X \) and \( Y \) are open subsets of Euclidean spaces. Suppose that \( \rho_1, \rho_2 \geq 1 \) and \( L_1, L_2 > 0 \). Then

\[
G^{\rho_1, \rho_2}_{L_1, L_2}(X \times Y) = \{ f \in C^\infty(X \times Y) : \sup_{(x, y) \in X \times Y} |\partial_x^\alpha \partial_y^\beta f|L_1^{-|\alpha|}L_2^{-|\beta|}\alpha!^{-\rho_1}\beta!^{-\rho_2} < \infty \}. \tag{4.1.4}
\]

If \( f \in G^{\rho_1, \rho_2}_{L_1, L_2} \), then we denote the supremum in (4.1.4) by \( \|f\|_{L_1, L_2} \). Equipped with this norm, \( G^{\rho_1, \rho_2}_{L_1, L_2} \) is a Banach space. This definition extends in the natural way to \( k \geq 3 \) variables. Furthermore, some of these variables might lie in complex domains.

For the statement of the KAM theorem in the next section, we record the following notation.

Definition 4.1.4. If \( D \subset \mathbb{R}^n \) and \( s, r > 0 \) we write

\[
T^n + s := \{ z \in \mathbb{C}^n/2\pi\mathbb{Z}^n : |\text{Im}(z)| \leq s \} \tag{4.1.5}
\]

and

\[
D_{s, r} := \{ \theta \in \mathbb{C}^n/2\pi\mathbb{Z}^n : |\text{Im}(\theta)| < s \} \times \{ I \in \mathbb{C}^n : |I| < r \}, \tag{4.1.6}
\]
where $|·|$ denotes the sup-norm on $\mathbb{C}^n$ induced by the 2-dimensional $\ell^\infty$ norm on $\mathbb{C}$.

These domains arise from considering the analytic extension of real analytic Hamiltonians in action-angle variables. In this topic we frequently bound derivatives of analytic functions using Cauchy estimates, which requires the use of shrinking sequences of domains.

For simplicity of nomenclature, we call an analytic function of several complex variables real analytic if its restriction to a function of $n$ real variables is real-valued.

As a final notational convenience, we use $|·|$ to denote the $\ell^1$ norm when applied to elements of $\mathbb{Z}^n$ throughout this chapter, as well as the matrix norm induced by the sup norm on $\mathbb{C}^n$.

4.2 Formulation of the KAM theorem

Let $D^0 \subset \mathbb{R}^n$ be a bounded domain, and consider a completely integrable Hamiltonian $H^0(I) = H^0(\theta, I) : T^n \times D^0 \to \mathbb{R}$ in action-angle coordinates. To begin with, we shall assume that this Hamiltonian is real analytic.

In addition, we assume the non-degeneracy condition $\det \left( \frac{\partial^2 H}{\partial I^2} \right) \neq 0$. This assumption implies that the map relating the action variable $I$ to the frequency $\omega = \nabla H^0(I)$ is locally invertible.

In fact, we assume that $I \mapsto \nabla H^0(I)$ (4.2.1) is a diffeomorphism from $D^0$ to $\Omega^0 \subset \mathbb{R}^n$. The inverse to this map is given by $\nabla g^0$, where $g^0$ is the Legendre transform of $H^0$.

We now let $D \subset D^0$ be a subdomain, and denote by $\Omega = \nabla H^0(D)$ the corresponding frequency set.

The phase space $T^n \times D$ is then foliated by the family of Lagrangian tori $\{T^n \times \{I\} : I \in D\}$ that are invariant under Hamiltonian flow associated to $H^0$.

The KAM theorem asserts that small perturbations of $H(\theta, I) = H^0(I) + H^1(\theta, I)$ on $T^n \times D$ still possess a family of Lagrangian tori which fill up phase space up to a set of Liouville volume $o(1)$ in the size of the perturbation.

More precisely, if $\Omega := \{\omega : \omega = \nabla H^0\}$ is the set of frequencies for the quasi-periodic flow of $H^0$, the frequencies satisfying:

$$|\langle \omega, k \rangle| \geq \frac{\kappa}{|k|^\tau}$$

(4.2.2)

for all nonzero $k \in \mathbb{Z}^n$ and fixed $\kappa > 0$ and $\tau > n - 1$ also correspond to Lagrangian tori for the perturbed Hamiltonian $H$, provided $\|H - H^0\| < \epsilon(\kappa)$ in a suitable norm.

Such frequencies are said to be non-resonant, and we denote the set of non-resonant frequencies by $\Omega^*_\kappa$, suppressing the dependence on $\tau$ from our notation. These sets are obtained by taking the intersection of the sets

$$\{\omega \in \Omega : |\langle \omega, k \rangle| \geq \frac{\kappa}{|k|^\tau} \}$$

(4.2.3)
over all nonzero \( k \in \mathbb{Z}^n \), and hence \( \bigcap_{k>0} \Omega^*_k \) is closed and perfect like the Cantor set. Unlike the Cantor set however, \( \bigcap_{k>0} \Omega^*_k \) is of full measure in \( \Omega \), as can be seen from the observation that
\[
m(\{ \omega \in \mathbb{R}^n : |\langle k, \omega \rangle| < \frac{\kappa}{|k|^{r+1}} \}) = O(\frac{\kappa}{|k|^{r+1}}).
\]
(4.2.4)

**Remark 4.2.1.** Such sets occur frequently in perturbation theory as a method of handling the infamous “small divisor” problem that we shall encounter shortly in Section 4.4.

We work with the sets
\[
\Omega_\kappa := \{ \omega \in \Omega^*_\kappa : \text{dist}(\omega, \partial \Omega) \geq \kappa \}
\]
(4.2.5)

which has positive measure for sufficiently small \( \kappa \).

It is convenient to work with the set of points of Lebesgue density in \( \Omega_\kappa \), which we denote by
\[
\tilde{\Omega}_\kappa := \{ \omega \in \Omega : \frac{m(B(\omega, r) \cap \Omega_\kappa)}{m(B(\omega, r))} \to 1 \text{ as } r \to 0 \}.
\]
(4.2.6)

From the Lebesgue density theorem we have that \( m(\tilde{\Omega}_\kappa) = m(\Omega_\kappa) \). We also note that a smooth function vanishing on \( \Omega_\kappa \) is necessarily flat on \( \tilde{\Omega}_\kappa \).

The construction of the Birkhoff normal form stems from Theorem 4.2.2, which is a version of the KAM theorem localised around the frequency \( \omega \) which is taken as an independent parameter. This version is particularly useful for our application as it makes it an easier task to check the regularity of the invariant tori with respect to the frequency parameter.

To illustrate the setup of this theorem, we set
\[
\Omega' = \{ \omega \in \Omega : \text{dist}(\omega, \partial \Omega) \leq \kappa/2 \}, \quad D' = \nabla g^0(\Omega').
\]
(4.2.7)

Taking \( z_0 \in D' \) we let \( I = z - z_0 \) lie in a small ball of radius \( R \) about \( 0 \). That is, \( R \) is chosen such that \( B_R(z_0) \subset D \).

Taylor expanding gives us the expression
\[
H^0(z) = H^0(z_0) + \langle \nabla z H^0(z_0), I \rangle + \int_0^1 (1 - t) \langle \nabla^2 z H^0(z_0 + tI), I \rangle dt.
\]
(4.2.8)

We now take \( \omega \in \Omega^0 \) to be the corresponding frequency \( \nabla H^0(z_0) \). The inverse of the frequency map is
\[
\psi_0(\omega) = \nabla g^0(\omega),
\]
(4.2.9)

where \( g^0 \) is the Legendre transform of \( H^0 \).

Hence we can write
\[
H^0(z) = H^0(\psi_0(\omega)) + \langle \omega, I \rangle + \langle P^0(I; \omega) I, I \rangle
\]
(4.2.10)

where \( P^0 \) is the quadratic remainder term in (4.2.8).
Expanding about the point \( z_0 = \nabla g^0(\omega) \), we can write our perturbation \( H^1 \) locally as
\[
H^1(\theta, z) = H^1(\theta, \nabla g^0(\omega) + I) = P^1(\theta, I; \omega).
\]
(4.2.11)

This leads us to consider perturbed real analytic Hamiltonians in the form
\[
H(\theta, I; \omega) = H^0(\psi_0(\omega)) + \langle \omega, I \rangle + P(\theta, I; \omega) =: N(I; \omega) + P(\theta, I; \omega).
\]
(4.2.12)

where
\[
N(I; \omega) = H^0(\psi_0(\omega)) + \langle \omega, I \rangle\]
(4.2.13)
and
\[
P(\theta, I; \omega) = \langle P_0(I; \omega), I, I \rangle + P_1(\theta, I; \omega).
\]
(4.2.14)

The traditional formulations of the KAM theorem assert the existence of a Cantor family of tori that persist under small perturbations of a single Hamiltonian \( H^0 \) with domain \( D \). In the framework laid out above, we now have a Cantor family of Hamiltonians parametrised by \( \omega \in \Omega_\kappa \). Note that each of these Hamiltonians is only linear in \( I \).

The essence of our version of the frequency localised KAM theorem in Theorem 4.2.2 is then that for sufficiently small \( P \), we can find a symplectic change of variables that transforms \( H \) to a linear normal form in \( I \) with remainder quadratic in \( I \) for \( \omega \in \Omega_\kappa \). This establishes the persistence of the Lagrangian torus with frequency \( \omega \). In Section 4.8 we will pass from this result to a version of the KAM theorem on \( \mathbb{T}^n \times D \) in Theorem 4.8.1 that establishes the existence of a Cantor family of invariant tori for the original Hamiltonian \( H \).

The proof of Theorem 4.2.2 is based on a rapidly converging iterative procedure introduced by Kolmogorov [28].

To work with Gevrey smooth Hamiltonians, we fix \( L_0 \geq 1 \) and \( A_0 > 1 \), and assume that \( H^0 \in G_{L_0}^\rho \), \( g^0 \in G_{L_0}^\rho(\Omega^0) \) with the estimates
\[
\|H^0\|_{L_0}, \|g^0\|_{L_0} \leq A_0.
\]
(4.2.15)

For \( L_2 \geq L_1 \geq 1 \) with \( L_2 \geq L_0 \), we now consider the analytic family of Gevrey perturbation \( H^1 \in G_{L_1, L_2}^\rho \) with the perturbation norm
\[
\epsilon_H := \kappa^{-2} \|H^1\|_{L_1, L_2}.
\]
(4.2.16)

The estimate (4.2.15) implies that there is a constant \( C(n, \rho) \) dependent only on \( n \) and \( \rho \) such that taking
\[
R \leq \frac{C(n, \rho) \kappa}{A_0 L_0^2}
\]
(4.2.17)
is sufficient to ensure that \( B_R(z_0) \subset D \) for any \( z_0 \in D' \).

At this point we introduce the notational convention for this chapter that \( C \) represents an arbitrary positive constant, dependent only on \( n, \tau, \rho \) and \( L_0 \). Similarly, \( c \) will represent a positive constant strictly less than 1, also only dependent on \( n, \tau, \rho \) and \( L_0 \). We will be explicit when we stray from this convention.

The estimates (4.2.15) and (4.2.16), together with Proposition A.3 in [36]
show that our constructed functions \( P^0 \) and \( P^1 \) are in the Gevrey classes 
\[
\mathcal{G}_{CL_0,CL_2}(B_R \times \overline{\Omega}) \subset \mathcal{G}_{CL_2,CL_2}(B_R \times \overline{\Omega}) \quad \text{and} \quad \mathcal{G}_{L_1,L_2,CL_2,L_2}(\mathbb{T}^n \times B_R \times \overline{\Omega} \times (-1,1))
\]
respectively, where the \( C \) in \( \mathcal{G}_{CL_0,CL_2} \) does not depend on \( L_0 \) or \( L_2 \).

Additionally we have the estimate
\[
\|P^1\|_{L_1,CL_2,CL_2,CL_2} \leq \kappa^{-2} \epsilon_H. \tag{4.2.18}
\]

Dropping the factors in our Gevrey constants dependent only on \( n, \tau, \rho, L_0 \) for brevity of notation, we are in a position to state the local KAM theorem in terms of the weighted norm
\[
\langle P \rangle_r := r^2 \|P^0\|_{L_2,L_2} + \|P^1\|_{L_1,L_2,L_2,L_2} \tag{4.2.19}
\]
for \( 0 < r < R \).

Our proofs in this chapter shall largely follow those of Popov [36], but we shall be fairly explicit in order to keep the presentation relatively self-contained as well as to demonstrate regularity in the analytic parameter \( t \in (-1,1) \).

**Theorem 4.2.2.** Suppose \( 0 < \zeta \leq 1 \) is fixed and \( \kappa < L_2^{-1-\zeta} \). Then there exists \( N(n, \rho, \tau) > 0 \) and \( \epsilon > 0 \) independent of \( \kappa, L_1, L_2, R, \Omega \) such that whenever the Hamiltonian
\[
H(\theta, I; \omega, t) = H^0(\psi_0(\omega)) + \langle \omega, I \rangle + \langle P^0(I, \omega)I, I \rangle + P^1(\theta, I; \omega, t) \tag{4.2.20}
\]
and \( 0 < r < R \) are such that
\[
\langle P \rangle_r < \epsilon \kappa r L_1^{-N} \tag{4.2.21}
\]
we can find
\[
\phi \in G^{p(\tau+1)+1,1}(\Omega \times (-3/4, 3/4), \Omega)
\]
and
\[
\Phi = (U, V) \in G^{p,\rho(\tau+1)+1,1}(\mathbb{T}^n \times \Omega \times (-3/4, 3/4), \mathbb{T}^n \times B_R)
\]
such that

1. For all \( \omega \in \Omega_\kappa \) and all \( t \in (-3/4, 3/4) \), the map \( \Phi_{\omega,t} = \Phi(\cdot; \omega, t) : \mathbb{T}^n \to \mathbb{T}^n \times B_R \) is a \( G^p \) embedding, with image \( \Lambda_{\omega,t} \) an invariant Lagrangian torus with respect to the Hamiltonian \( H_{\phi(\omega,t),t}(\theta, I) = H(\theta, I; \phi(\omega,t), t) \). The Hamiltonian vector field restricted to this torus is given by
\[
X_{H_{\phi(\omega,t),t}} \circ \Phi_{\omega,t} = D\Phi_{\omega,t} \cdot \mathcal{L}_\omega \tag{4.2.22}
\]
where
\[
\mathcal{L}_\omega = \sum_{j=1}^n \omega_j \frac{\partial}{\partial \theta_j} \in T\mathbb{T}^n. \tag{4.2.23}
\]

2. There exist positive constants \( A \) and \( C \) dependent only on \( n, \tau, \rho, L_0 \) such that
\[
|\partial_\theta^\alpha \partial_\omega^\beta (U(\theta; \omega, t) - \theta)| + r^{-1} |\partial_\theta^\alpha \partial_\omega^\beta V(\theta; \omega, t)| + \kappa^{-1} |\partial_\omega^\beta (\phi(\omega,t) - \omega)| 
\leq A(\mathcal{L}_1)^{[\alpha]}(\mathcal{L}_1^{\tau+1}/\kappa)^{[\beta]} \alpha_\beta^p \rho^{p(\tau+1)+1}(\langle P \rangle_r / \kappa r)^{L_1^N} \tag{4.2.24}
\]
uniformly in \( \mathbb{T}^n \times \Omega \times (-3/4, 3/4) \).
We remark that at the endpoint \( t = 0 \), this result is trivial by taking \( \phi(\omega, 0) = \omega, U(\theta, \omega, 0) = \theta \) and \( V(\theta, \omega, 0) = \nabla g^0(\omega) \).

KAM theory can be viewed as a collection of techniques for dealing with perturbative problems, rather than just the specific theorem describing the dynamics of a perturbed completely integrable Hamiltonian system. In this spirit we shall first present an application to Arnold’s theorem on analytic diffeomorphisms of the circle before moving on to the proof of Theorem 4.2.2. Arnold’s theorem is simpler than the KAM theorem from a technical standpoint, nonetheless it serves as an excellent model to provide the intuition behind its proof. An exposition of this problem can be found in [47], from which this account is drawn.

### 4.3 Arnold’s theorem on analytic circle diffeomorphisms

Suppose \( f : \mathbb{R} \to \mathbb{R} \) is the lift of an orientation preserving analytic diffeomorphism \( \tilde{f} \) of \( T \). Then we have

\[
f(x + 1) = f(x) + 1 \quad \text{and} \quad f'(x) > 0
\]

for each \( x \in \mathbb{R} \).

The simplest examples of such functions are the rotation maps \( R_\theta(x) := x + \theta \), whose dynamics are well understood. In particular, for rational \( \theta \), the corresponding dynamical system obtained by iterating \( \tilde{R}_\theta \) is a periodic map, whilst for irrational \( \theta \) the orbits are dense and equidistributed on \( T \).

Denjoy’s theorem [19, p. 301] shows that these are in fact the only orientation preserving diffeomorphisms up to conjugation by a homeomorphic and 1-periodic change of variables \( \chi : \mathbb{R} \to \mathbb{R} \).

The question as to whether arbitrary orientation preserving diffeomorphisms were *analytically diffeomorphic* to rotation maps remained wide open until the invention of KAM theory. Arnold used KAM techniques to answer this question in the affirmative for small perturbations of suitably irrational rotation maps.

To make this precise, we first introduce the notion of suitable irrationality or nonresonance. This is a one-dimensional analogue to the nonresonance condition (4.2.2).

**Definition 4.3.1.** For \( \kappa, \rho > 0 \), we define

\[
\Omega_{\kappa, \rho} := \{ \theta \in \mathbb{R} : |\theta - m/n| > \frac{\kappa}{|n|^\rho} \text{ for all } m, n \in \mathbb{Z}, n \neq 0 \}.
\]

Intuitively, \( \Omega_{\kappa, \rho} \) consists of irrational numbers that can’t be approximated well by rationals. Dirichlet’s theorem on Diophantine approximation asserts that every irrational number can be approximated by a rational number \( m/n \) up to an order \( n^{-2} \) error, but it turns out that is the best approximation we can hope for from a typical irrational number.

**Proposition 4.3.2.** For \( \rho > 2 \) and any interval \([a, b]\), the set

\[
\bigcup_{\kappa > 0} ([a, b] \setminus \mathbb{Q}) \cap \Omega_{\kappa, \rho}
\]

is of full measure.

A proof of this fact can be found in [2, p. 116].
As we are interested in small \( \kappa \), we assume that \( \kappa \leq 1 \). We also require the notion of rotation number for an arbitrary circle diffeomorphism, which we now introduce.

**Proposition 4.3.3.** For \( f \) the lift of a circle diffeomorphism, the limit

\[
\theta = \lim_{n \to \infty} \frac{f^n(x) - x}{n}
\]

(4.3.4)

exists and is independent of \( x \). We say that \( \theta \) is the rotation number of the diffeomorphism \( \tilde{f} \).

The proof of Proposition 4.3.3 can be found in [19, p. 296].

It is easy to show that rotation numbers are preserved by homeomorphisms. Moreover, we have the following.

**Corollary 4.3.4.** If \( f(x) = x + \theta + \eta(x) \) is the lift of a circle diffeomorphism with rotation number \( \theta \), then \( \eta(x) \) vanishes at some \( x_0 \in [0, 1] \).

The benefit of working with real analytic functions is that have the Cauchy estimates from Proposition A.1 at our disposal. To this end, we define

\[
S_\sigma = \{ z \in \mathbb{C} : |\text{Im}(z)| \leq \sigma \}
\]

(4.3.5)

for \( \sigma > 0 \), and

\[
H_\sigma = \{ f : S_\sigma \to \mathbb{C} : f \text{ holomorphic and 1-periodic on } S_\sigma \text{ and real-valued on } \mathbb{R} \}.
\]

(4.3.6)

**Theorem 4.3.5** (Arnold’s Theorem). Suppose that \( \theta \in \Omega_{\kappa, \rho} \) and \( f \in H_\sigma \) for some \( \sigma > 0 \). Then there exists \( \epsilon > 0 \), dependent only on \( \kappa, \rho \) and \( \sigma \), such that for any \( f(x) = x + \theta + \eta(x) \) with rotation number \( \theta \) and with \( \eta \) satisfying \( \| \eta \|_\sigma < \epsilon \), there exists a real analytic diffeomorphism \( \chi \) such that \( \chi^{-1} \circ f \circ \chi = R_\theta \).

The proof proceeds in three steps.

1. First we linearise the equation defining the sought diffeomorphism \( \chi \), and use basic Fourier analysis to solve this linearised problem and explicitly bound the solution.

2. The next key step is to note that upon applying the coordinate change constructed from the linear problem, the function \( f(x) \) transforms to \( x + \theta + \tilde{\eta}(x) \), with \( \tilde{\eta} = O(\| \eta \|^2) \). These two steps are a Banach space analogue of Newton’s method for root approximation.

3. Finally, we iterate the process above. The primary challenge is to ensure the the composition of diffeomorphisms converges. This turns out to be a consequence of the rapid (quadratic) convergence.

If the perturbation \( \eta(x) \) in the statement of 4.3.5 is small, we expect that the sought diffeomorphism \( \chi(x) \) is of the form

\[
\chi(x) = x + \mu(x)
\]

(4.3.7)

with \( \mu \) small.
Substituting this into the equation

\[ f(\chi(x)) = \chi(R_\theta(x)) \]  

we arrive at

\[ \mu(x + \theta) - \mu(x) = \eta(x + \mu(x)). \]  

(4.3.8)

Since \( \eta \) and \( \mu \) are small, it is natural to drop higher order terms and instead consider the equation

\[ \mu(x + \theta) = \eta(x) \]  

(4.3.9)

which is now linear in the unknown function \( \mu \).

Additionally, since the functions \( \eta \) and \( \mu \) are periodic by construction, equation 4.3.10 can almost be formally solved up to a constant by simply equating nonzero Fourier coefficients on both sides. Note that the zero-th Fourier coefficient of the left-hand side of (4.3.10) vanishes, and so the resulting Fourier series will be formally equal to \( \eta(x) - \hat{\eta}(0) \).

Explicitly, with

\[ \mu(x) := \sum_{n \neq 0} \hat{\mu}(n)e^{2\pi inx} = \sum_{n \neq 0} \left( \int_0^1 e^{-2\pi inx} \mu(x) \, dx \right) e^{2\pi inx} \]  

(4.3.11)

we obtain

\[ \hat{\mu}(n) = \frac{\hat{\eta}(n)}{e^{2\pi in\theta} - 1} \]  

(4.3.12)

for nonzero \( n \).

For general irrational \( \theta \), the denominators in the series (4.3.12) can be very small. This is a common difficulty in perturbation theory, and one of the major innovations in KAM theory was the usage of Diophantine sets like \( \Omega_{\kappa,\rho} \) to control these denominators.

Indeed, we have

**Lemma 4.3.6.** If \( \theta \in \Omega_{\kappa,\rho} \), we have

\[ |e^{2\pi in\theta} - 1| \geq 4\kappa|n|^{-(\rho-1)}. \]  

(4.3.13)

**Proof.** \( \square \)

An immediate consequence of Lemma 4.3.6 is that the formal Fourier series (4.3.11) for \( \mu \) converges.

Using the estimate on Fourier coefficients of an analytic function from Proposition A.2, we obtain

\[ \|\mu\|_{\sigma-\delta} = \sup \left| \sum_{n \neq 0} \hat{\eta}(n)e^{2\pi inx} \right| \frac{1}{e^{2\pi in\theta} - 1} \]  

\[ \leq \sum_{n \neq 0} \frac{|n|^\rho - 1}{4\kappa} e^{2\pi \sigma|n|} e^{2\pi |n|(|\sigma-\delta|)} \|\eta\|_{\sigma} \]  

\[ \leq \frac{\Gamma(\rho)}{\kappa(2\pi)^\rho} \|\eta\|_{\sigma}. \]  

(4.3.14)

for \( \delta \) sufficiently small dependent only on \( \rho \), where (4.3.16) follows from estimating the sum in \( n \) by the integral defining the Gamma function.
This estimate allows us to show that $\chi$ is invertible on a suitable domain.

**Proposition 4.3.7.** If $\max\left(\delta, \|\eta\|_\sigma^{\frac{\kappa}{\rho^2+1}}\right)$ is sufficiently small (dependent only on the fixed quantities $\rho$ and $\kappa$), then $\chi$ has an analytic inverse on $\chi(S_{\sigma-2\delta})$.

Moreover, $\chi(S_{\sigma-2\delta}) \supseteq S_{\sigma-3\delta}$.

**Proof.** The analytic inverse function theorem implies that it suffices to show that $\|\text{Id} - \chi'\|_{\sigma-2\delta} = \|\mu'\|_{\sigma-2\delta} < 1$, which follows directly from the Cauchy estimate from Proposition A.1 and (4.3.14).

Moreover, since the assumptions of this proposition imply that $\|\mu\|_{\sigma-2\delta} < \delta$, we have $\chi(S_{\sigma-2\delta}) \supseteq S_{\sigma-3\delta}$ by a straightforward degree theoretic argument. \qed

**Proposition 4.3.8.** Under the conditions of Proposition 4.3.7, we have

$$\chi^{-1}(z) = z - \mu(z) + \nu(z),$$

(4.3.17)

with the estimate

$$\|\nu\|_{\sigma-4\delta} \leq \frac{C(\rho, \kappa)}{\delta^{2\rho+1}} \|\eta\|_\sigma^2$$

(4.3.18)

**Proof.** From (4.3.17) and Proposition 4.3.7, we have

$$z = (\chi^{-1} \circ \chi)(z)$$

(4.3.19)

$$= z + \mu(z) - \mu(z + \mu(z)) + \nu(z + \mu(z))$$

(4.3.20)

for all $z \in S_{\sigma-2\delta}$.

Hence

$$\nu(z) = \mu(\chi^{-1}(z)) \left( \int_0^1 \mu'(\chi^{-1}(z) + t\mu(\chi^{-1}(z))) \, dt \right)$$

(4.3.21)

for $z \in S_{\sigma-3\delta}$.

As in the second part of Proposition 4.3.7, we obtain $\chi(S_{\sigma-3\delta}) \supseteq S_{\sigma-4\delta}$.

Since we have $\|\mu\|_{\sigma-2\delta} < \delta$, it follows that

$$\chi^{-1}(z) + t\mu(\chi^{-1}(z)) \in S_{\sigma-2\delta}$$

(4.3.22)

for all $t \in [0,1]$ and all $z \in S_{\sigma-4\delta}$.

Thus

$$\|\nu\|_{\sigma-4\delta} \leq \|\mu\|_{\sigma-3\delta} \cdot \|\mu'\|_{\sigma-2\delta}$$

(4.3.23)

$$\leq \delta^{-1} \left( \frac{\Gamma(\rho)\|\eta\|_\sigma}{\kappa(2\pi\delta)^\rho} \right)^2$$

(4.3.24)

$$= \frac{C(\rho, \kappa)}{\delta^{2\rho+1}} \|\eta\|_\sigma^2.$$ 

(4.3.25)

\qed

The two estimates (4.3.14) and (4.3.18) allow us to bound the new error term $\tilde{\eta}$.

**Proposition 4.3.9.** Under the conditions of Proposition 4.3.7, we can write

$$\tilde{f}(z) = (\chi^{-1} \circ f \circ \chi)(z) = z + \theta + \tilde{\eta}(z)$$

(4.3.26)
with the estimate
\[ \|\tilde{\eta}\|_{\sigma-\delta} \leq \frac{C(\rho, \kappa)}{\delta^{2\rho+1}} \|\eta\|_{\sigma}^2 \] (4.3.27)

**Proof.** From (4.3.7), (4.3.17), and (4.3.10), we have

\[
(\chi^{-1} \circ f \circ \chi)(z) = z + \theta + \mu(z) + \eta(z + \mu(z)) - \mu(z + \theta + \mu(z) + \eta(z + \mu(z))) + \nu(z + \theta + \mu(z) + \eta(z + \mu(z))).
\] (4.3.28)

Hence

\[
\tilde{\eta}(z) = \tilde{f}(z) - z - \theta = \tilde{\eta}_0 + (\eta(z + \mu(z)) - \eta(z)) + (\mu(z + \theta) - \mu(z + \theta + \mu(z) + \eta(z + \mu(z)))) + \nu(z + \theta + \mu(z) + \eta(z + \mu(z))).
\] (4.3.30)

It remains to bound the right-hand side terms in (4.3.31) individually. First, from (4.3.14) and the conditions of Proposition 4.3.7 we bound

\[
\eta(z + \mu(z)) - \eta(z) = \mu(z) \int_0^1 \mu'(z + t\mu(z)) \, dt \]
(4.3.32)
\[
\Rightarrow \|\eta(z + \mu(z)) - \eta(z)\|_{\sigma-\delta} \leq \delta^{-1} \|\mu\|_{\sigma-\delta} \|\eta\|_{\sigma-2\delta} \leq \frac{C(\rho, \kappa)}{\delta^{\rho+1}} \|\eta\|_{\sigma}^2.
\] (4.3.34)

Similarly, we can bound

\[
\mu(z + \theta) - \mu(z + \theta + \mu(z) + \eta(z + \mu(z))) = -(\mu(z) + \eta(z + \mu(z))) \int_0^1 \mu'(z + \theta + t\mu(z) + \eta(z + \mu(z))) \, dt
\]
(4.3.35)
\[
\Rightarrow \|\mu(z + \theta) - \mu(z + \theta + \mu(z) + \eta(z + \mu(z)))\|_{\sigma-\delta} \leq \|\mu'\|_{\sigma-2\delta} (\|\mu\|_{\sigma-\delta} + \|\eta\|_{\sigma-3\delta}) \leq \frac{C(\rho, \kappa)}{\delta^{2\rho+1}} \|\eta\|_{\sigma}^2.
\] (4.3.39)

Since \(\|\eta\|_{\sigma} < \delta\) can be assumed, we have

\[
z + \theta + \mu(z) + \eta(z + \mu(z)) \in S_{\sigma-\delta}
\] (4.3.40)

whenever \(z \in S_{\sigma-\delta}\). This allows us to apply the estimate (4.3.18) for \(\nu\) directly.

The only remaining term in (4.3.31) is the zero-th Fourier coefficient \(\tilde{\eta}(0)\). To bound this, we observe that Corollary 4.3.4 implies the existence of an \(x_0 \in [0, 1]\) with \(\tilde{\eta}(x_0) = 0\). So setting \(z = x_0\) in (4.3.31) and rearranging yields

\[
|\tilde{\eta}(0)| \leq \frac{C(\rho, \kappa)}{\delta^{\rho+1}} \|\eta\|_{\sigma}^2 + \frac{C(\rho, \kappa)}{\delta^{2\rho+1}} \|\eta\|_{\sigma}^2 + \frac{C(\rho, \kappa)}{\delta^{2\rho+1}} \|\eta\|_{\sigma}^2.
\] (4.3.41)

Putting the estimates (4.3.32),(4.3.35),(4.3.18), and (4.3.41) into (4.3.31) completes the proof.

What we have completed at this stage is frequently referred to as the KAM step, and
the remainder of the proof of Theorem 4.3.5 is an induction formed by iterating the KAM step with a carefully chosen decreasing sequence of constants $\sigma_n$ and $\delta_n$.

The sequence of lifts of circle diffeomorphisms $(f_n)$ is defined inductively by

$$f_0(x) := f(x), \quad f_{n+1}(x) := \chi_n^{-1} \circ f_n \circ \chi_n.$$  \hfill (4.3.42)

We also define

$$\eta_0(x) := \eta(x), \quad \eta_{n+1}(x) := f_{n+1}(x) - x - \theta$$  \hfill (4.3.43)

and

$$\mu_n(x) = \chi_n(x) - x.$$  \hfill (4.3.44)

By construction, we have that $\mu_n$ satisfies

$$\mu_n(x + \theta) - \mu_n(x) = \eta_n(x) - \hat{\eta}_{n+1}(0).$$  \hfill (4.3.45)

The corresponding sequences of $\delta_n, \sigma_n$ are defined by

$$\delta_n := \frac{\sigma}{36(1 + n^2)}$$  \hfill (4.3.46)

and

$$\sigma_0 := \sigma, \quad \sigma_{n+1} = \sigma_n - 6\delta_n$$  \hfill (4.3.47)

for all $n \geq 0$.

The key feature of this choice is that $\lim_{n \to \infty} \sigma_n > \sigma/2 > 0$, so the decreasing sequence of domains $S_{\sigma_n}$ have width positively bounded below.

To control the errors, we introduce the sequence

$$\epsilon_0 := \|\eta\|_\sigma, \quad \epsilon_n := \epsilon_0^{(3/2)^n}.$$  \hfill (4.3.48)

We can now formulate the induction of the KAM step.

**Proposition 4.3.10.** For sufficiently small $\epsilon_0$, dependent only on $\sigma, \rho, \kappa$, we have $\eta_{n+1} \in H_{\sigma_{n+1}}$ and $\|\eta_{n+1}\|_{\sigma_{n+1}} \leq \epsilon_{n+1}$. Moreover, we have the estimates

$$\|\mu_n\|_{\sigma_n - \delta_n} \leq \frac{C(\rho, \kappa)\epsilon_n}{\delta_n^\rho}$$  \hfill (4.3.49)

and

$$\|\nu_n\|_{\sigma_n - 4\delta_n} \leq \frac{C(\rho, \kappa)\epsilon_n^2}{\delta_n^{2\rho+1}}$$  \hfill (4.3.50)

**Proof.** The estimates (4.3.49) and (4.3.50) follow from (4.3.14) and (4.3.18) respectively, provided that that $\|\eta_n\|_{\sigma_n} \leq \epsilon_n$ and $\delta_n^{-\rho-1}\epsilon_n < c(\rho, \kappa)$ for all $n \geq 0$, which we verify inductively.

If $\|\eta_n\|_{\sigma_n} \leq \epsilon_n$, then from Proposition 4.3.9, we have

$$\|\eta_{n+1}\|_{\sigma_{n+1}} \leq \frac{C(\rho, \kappa)\epsilon_n^2}{\delta_n^{2\rho+1}}$$  \hfill (4.3.51)

$$\leq C(\sigma, \rho, \kappa)n^{k\rho+2}\epsilon_n^2.$$  \hfill (4.3.52)
Hence it suffices to choose $\epsilon_0$ small enough that
\[
\epsilon_0^{(3/2)n} \leq \frac{C(\sigma, \rho, \kappa)}{n^{8\rho+4}} \tag{4.3.53}
\]
which is possible due to the rapid decay of the left-hand side.

Similarly, $\delta_n^{\rho-1} \epsilon_n < c(\rho, \kappa)$ can be ensured by taking $\epsilon_0$ sufficiently small, dependent only on $\rho, \kappa$ and $\sigma$, due to the rapid decay of $\epsilon_n$.

We can now complete the proof of Theorem 4.3.5 by analysing the convergence of $\chi^n(x) := \chi_0 \circ \chi_1 \circ \ldots \circ \chi_n(x)$.

**Proof of Theorem 4.3.5.** We first note that the definitions

From (4.3.44), it follows that
\[
\chi^n(x) = x + \sum_{k=0}^{n} \mu_{n-k}(h_k(x)) \tag{4.3.54}
\]
where the sequence of functions $h_k$ on $S_{\delta_n}$ is defined recursively by
\[
h_0(x) := x \tag{4.3.55}
\]
and
\[
h_m(x) = x + \sum_{k=0}^{m-1} \mu_{n-k}(h_k(x)). \tag{4.3.56}
\]

This gives us the estimate
\[
\|\chi^n(z) - z\|_{\sigma_n-2\delta_n} \leq \sum_{n=0}^{\infty} \frac{C(\rho, \kappa)\epsilon_n}{\delta_n^{\rho}} = \tilde{C} < \infty \tag{4.3.57}
\]
with convergence following from the rapid decay of $\epsilon_n$.

To prove convergence of the $\chi^n(z)$ on the limiting strip $S_{\sigma_\infty}$, we write
\[
\chi^{n+1}(z) - \chi^n(z) = \mu_n(z) \int_0^1 (\chi^n)'(z + t\mu_n(z)) \, dt. \tag{4.3.58}
\]

Using the Cauchy estimate from Proposition A.1 and (4.3.57), we obtain
\[
\|\chi^{n+1}(z) - \chi^n(z)\|_{\sigma_{n+1}} \leq \|\mu_n\|_{\sigma_n-2\delta_n} \cdot (1 + \|\chi^n)'(z) - 1\|_{\sigma_n-2\delta_n}) \tag{4.3.59}
\]
\[
\leq \|\mu_n\|_{\sigma_n-2\delta_n} \cdot (1 + \frac{3}{\delta_n}\|\chi^n(z) - z\|_{\sigma_n-2\delta_n}) \tag{4.3.60}
\]
\[
\leq \left(1 + \frac{3\tilde{C}}{\delta_n}\right) \cdot \frac{C(\rho, \kappa)\epsilon_n}{\delta_n^{\rho}} \tag{4.3.61}
\]
which is summable, again due to the rapid decay of $\epsilon_n$.

This implies that the $\chi^n(z)$ converge uniformly to an analytic limit function $\chi(z)$ on $S_{\sigma_\infty}$.

From the Cauchy estimate from Proposition A.1 and (4.3.57), we have
\[
\|\chi'(z) - 1\|_{(1-r)\sigma_\infty} \leq \frac{\tilde{C}}{r\sigma_\infty} \tag{4.3.62}
\]
for arbitrary \( r \in (0, 1/2) \).

Having fixed \( r \), we can choose \( \epsilon_0 \) sufficiently small so that \( \tilde{C} < r\sigma_\infty \) from the rapid decay of \( \epsilon_n \), and so as in Proposition 4.3.7, we obtain that \( \chi \) is invertible on \( S_{(1-r)\sigma_\infty} \) with \( \chi(S_{(1-r)\sigma_\infty}) \supseteq S_{(1-2r)\sigma_\infty} \).

Now, we have

\[
(f \circ \chi)(z) = \lim_{n \to \infty} (f \circ \chi^n)(z) \quad (4.3.63)
\]

\[
= \lim_{n \to \infty} \chi^n(z + \theta + \eta_n(z)) \quad (4.3.64)
\]

\[
= \chi(z + \theta) \quad (4.3.65)
\]

for all \( z \in S_{(1-2r)\sigma_\infty} \). Since \( \chi \) is invertible on this domain, we can apply \( \chi^{-1} \) to both sides, demonstrating that \( \chi \) is the sought diffeomorphism and we are done. \( \Box \)

In the next section, we return to the problem of dynamics in a small perturbation of a completely integrable Hamiltonian system. The largest technical difference from Arnold’s theorem is the involvement of functions that are only of Gevrey regularity and are not necessarily analytic. This necessitates the approximation of this sequence of Gevrey functions by a sequence of analytic functions in order to exploit the Cauchy estimates that are essential to the method of proof.

We begin by proving the KAM step.

### 4.4 The KAM step

We now prepare for the proof of Theorem 4.2.2 by first proving the result that will comprise the steps of the iterative argument. Given a Hamiltonian in the form

\[
H(\theta, I; \omega, t) = e(\omega; t) + \langle \omega, I \rangle + P(\theta, I; \omega, t) = N(I; \omega, t) + P(\theta, I; \omega, t), \quad (4.4.1)
\]

we aim to construct a \( t \)-dependent symplectic map \( \Phi \) and a \( t \)-dependent frequency transformation \( \phi \) such that for \( F = (\Phi, \phi) \), we have

\[
(H \circ F)(\theta, I; \omega, t) = N_+(I; \omega, t) + P_+(\theta, I; \omega, t) \quad (4.4.2)
\]

where \( N_+(I, \omega, t) = e_+ + \langle I, \omega \rangle \) and with \( |P_+| \) controlled by \( |P|^r \) for some \( r > 1 \).

This construction is analogous to that in [38], although our application requires working with families of Hamiltonians that are also real analytic in the additional parameter \( t \).

**Theorem 4.4.1.** Suppose \( c, h, v, s, r, \eta, \sigma, K \) are positive constants such that

\[
s, r < 1, \quad v < 1/6, \quad \eta < 1/8, \quad \sigma < s/5, \quad \epsilon \leq c \eta r \sigma^{r+1}, \quad \epsilon \leq c v h r, \quad h \leq \kappa/2K^{\tau+1}. \quad (4.4.3)
\]

where \( c \) is a constant dependent only on \( n \) and \( \tau \).

Suppose \( H(\theta, I; \omega, t) = N(I; \omega, t) + P(\theta, I; \omega, t) \) is real analytic on \( D_{s,r} \times O_h \times (-1, 1) \), and \( |P|_{s,r,h} \leq \epsilon \). Here, \( D_{s,r} \) is as in Definition 4.1.4 and

\[
O_h := \{ \omega \in \mathbb{C}^n : \text{dist}(\omega, \Omega_\kappa) < \eta \}. \quad (4.4.4)
\]
Then there exists a real analytic map
\[ \mathcal{F} = (\Phi, \phi) : D_{s-5\sigma,\eta r} \times O_{(1/2-3v)}h \times (-1, 1) \to D_{s,r} \times O_h \] (4.4.5)
where the maps
\[ \Phi : D_{s-5\sigma,\eta r} \times O_h \times (-1, 1) \to D_{s,r} \] (4.4.6)
and
\[ \phi : O_{(1/2-3v)}h \times (-1, 1) \to O_h \] (4.4.7)
are such that
\[ H \circ \mathcal{F} = e_+ + (\omega, t) + \langle \omega, I \rangle + P_+(\theta, I; \omega, t) = N_+ (I; \omega, t) + P_+(\theta, I; \omega, t) \] (4.4.8)
and we have the new remainder estimate
\[ |P_+|_{s-5\sigma,\eta r, (1/2-2v)}h \leq C \left( \frac{\epsilon^2}{K\eta\sigma^{\tau+1}} + (\eta^2 + K^n e^{-K\sigma}) \epsilon \right). \] (4.4.9)
Moreover \( \Phi \) is symplectic for each \((\omega, t)\) and has second component affine in \( I \). Finally, we have the following uniform estimates on the change of variables.
\[ |W(\Phi - id)|, |W(D\Phi - Id)W^{-1}| \leq \frac{C\epsilon}{K\eta\sigma^{\tau+1}} \] (4.4.10)
and
\[ |\phi - id|, vh|D\phi - Id| \leq \frac{C\epsilon}{r} \] (4.4.11)
where \( W = \text{diag}(\sigma^{-1}Id, r^{-1}Id) \). All estimates are uniform and analytic in the parameter \( t \in (-1, 1) \).

Proof. We first linearise \( P \) about \( I = 0 \) and truncate its Fourier series to order \( K \). If we define
\[ Q = P(\theta, 0; \omega, t) + I \cdot \nabla I P(\theta, 0; \omega, t), \] (4.4.12)
then the Cauchy estimate from Proposition A.1 for analytic functions together with Taylor’s theorem yields
\[ |Q|_{s,r} \leq C\epsilon \] (4.4.13)
and
\[ |P - Q|_{s,2\eta r} \leq C\eta^2 \epsilon. \] (4.4.14)
Defining
\[ R(\theta, I; \omega, t) := \sum_{|k| \leq K} \langle Q(\cdot, I; \omega, t), e^{i(k, \cdot)}e^{i(k, \theta)} \rangle, \] (4.4.15)
the Fourier truncation result, Proposition A.3 yields
\[ |R - Q|_{s-\sigma, r} \leq CK^{-n} e^{-K\sigma} \epsilon. \] (4.4.16)
When we apply this KAM step, we can assume \( K \) is sufficiently large (dependent on \( \sigma \)), so that we will in fact have
\[ |R|_{s-\sigma, r} \leq C\epsilon. \] (4.4.17)
All estimates thus far are uniform in \((\omega, t) \in O_h \times (-1, 1)\), and the function \( R(\theta, I; \omega, t) \) is still analytic in all variables.
Recalling the assumption $h \leq \kappa/(2K^{\tau+1})$, we next extend the nonresonance estimate (4.2.2) from the Cantor set $\Omega$ to the estimate

$$|\langle k, \omega \rangle| \geq \frac{\kappa}{2|k|^\tau}$$

(4.4.18)

for $\omega \in O_h$ and $|k| \leq K$. This of course motivates our Fourier truncation of $Q$.

We now make the ansatz that our sought transformation $\Phi$ can be obtained as the time-1 Hamiltonian flow generated by an undetermined function $F$ dependent on the parameters $\omega,t$. From Proposition 1.1.6, such transformations are known to be symplectic, and so it suffices to show that we can choose $F$ in a way that yields a transformation bringing $H$ to normal form up to an error term that is quantifiably smaller than $\epsilon$.

To this end, we analyse the expression

$$(N(I; \omega,t) + R(\theta,I; \omega,t)) \circ \Phi_F^\tau.$$  (4.4.19)

Recalling (1.1.16), Taylor’s theorem yields

$$(N(I; \omega,t) + R(\theta,I; \omega,t)) \circ \Phi_F^1 = N + \{N,F\} + \int_0^1 \{(1-\tau)\{N,F\}, F\} \circ \Phi_F^\tau d\tau + R + \int_0^1 \{R,F\} \circ \Phi_F^\tau d\tau.$$  (4.4.20)

The integral in (4.4.20) is of second order in $F$ and $R$, and is thus an ideal candidate to absorb into remainder term $F^+$, which will also contain the error from approximating $P$ by $R$. In other words, we need to choose $F$ in such a way that $N + \{N,F\} + R$ is in linear normal form. This amounts to solving

$$\{F,N\} + (N_+ - N) = R$$  (4.4.21)

for $F$ and then finding a suitable frequency transformation so that $N_+ = e_+(\omega;t) + \langle \omega, I \rangle$.

Taking the Fourier expansion

$$F = \sum_{k \in \mathbb{Z}^n} F_k e^{i\langle k, \theta \rangle},$$  (4.4.22)

we formally obtain

$$\{F,N\} = \sum_{j=1}^n \omega_j \frac{\partial F}{\partial \theta_j} = \sum_{k \in \mathbb{Z}^n} i\langle k, \omega \rangle F_k e^{i\langle k, \theta \rangle}.$$  (4.4.23)

Since $\omega$ is a nonresonant frequency, the factors $\langle k, \omega \rangle$ are nonzero for nonzero $k$ which allows us to choose

$$F_k = \frac{R_k}{i\langle k, \omega \rangle}$$  (4.4.24)

and define $F$ by the resulting Fourier series. This sum is finite, thanks to the earlier truncation of $R$.

To equate the zero-th Fourier coefficients in (4.4.21), we take $N_+ := N + \hat{R}(0)$.

**Remark 4.4.2.** The small denominators $\langle k, \omega \rangle$ occurring in (4.4.24) are a frequent and
problematic feature in the formal series that arise in perturbation theory. The use of non-resonance conditions such as (4.4.18) to control such denominators in the KAM theorem was a key advance for the field.

Using (4.4.24) and the estimate (4.4.18), we obtain the bound

\[
|F|_{r,s-2\sigma} \leq \sum_{|k|\leq K} \left| R_k \right| \left| \frac{e^{i\langle k,\theta \rangle}}{|\langle k,\omega \rangle|} \right|
\]

where the last line follows from (4.4.17). Furthermore, \( F \) is analytic in all variables so we may use the Cauchy estimate again to obtain

\[
|\partial_\theta F|_{s-3\sigma,r} \leq \frac{C\epsilon}{\kappa \sigma r + 1}
\]

and

\[
|\partial_I F|_{s-2\sigma,r/2} \leq \frac{C\epsilon}{\kappa r \sigma^2}.
\]

We can combine these estimates as

\[
\max\{r^{-1} |\partial_\theta F|, \sigma^{-1} |\partial_I F|\} \leq \frac{C\epsilon}{\kappa r \sigma^2 + 1}
\]

uniformly on \( D_{s-3\sigma,r/2} \times O_h \times (-1,1) \). At this point we can also estimate

\[
N_+ - N = \hat{R}(0) \leq |R|_{s-\sigma,r} \leq C\epsilon.
\]

Now the bounds on the \((\theta, I)\)-derivatives of \( F \) control the Hamiltonian flow \( \Phi = (U, V) \). Indeed, the estimates (4.4.3),(4.4.29) and (4.4.30) imply that

\[
|\partial_\theta F| \leq \eta r \leq r/8, \quad |\partial_I F| \leq \sigma
\]

and consequently that the time-1 flow is well-defined as a map

\[
D_{s-4\sigma,r/4} \to D_{s-3\sigma,r/2},
\]

with the component bounds

\[
|U - id| \leq |\partial_\theta F| \leq \frac{C\epsilon}{\kappa r \sigma^2 + 1}, \quad |V - id| \leq |\partial_I F| \leq \frac{C\epsilon}{\kappa r \sigma^2 + 1}.
\]

By construction, we have that \( F \) is affine linear in \( I \). Consequently, \( \partial_I F \) and \( V \) are both \( I \)-independent, and \( \partial_\theta F \) and \( U \) are affine linear in \( I \). To complete the estimates of \( \Phi \), we use the Cauchy estimate again to yield

\[
|\partial_I U - Id| \leq \frac{C\epsilon}{\kappa r \sigma^2 + 1}, \quad |\partial_\theta U| \leq \frac{C\epsilon}{\kappa \sigma r + 1}, \quad |\partial_\theta V - Id| \leq \frac{C\epsilon}{\kappa r \sigma^2 + 1}.
\]
uniformly on $D_{s-5\sigma,r/8} \supseteq D_{s-5\sigma,\eta r}$.

It remains to bound the new error term $P_+$ given by (4.4.20) and to construct the frequency map $\phi$ that transforms $N_+$ to normal form. First we treat the error term $P_+$.

By using the Cauchy estimate, we can bound

$$\{R, F\} \leq |\partial_1 R| |\partial_0 F| + |\partial_0 R| |\partial_1 F| \leq \frac{C\epsilon^2}{k\tau \sigma^{\tau + 1}}$$ (4.4.37)

uniformly on $D_{s-3\sigma,r/2}$.

In exactly the same way, recalling that $N_+ - N = \hat{R}(0)$, we use (4.4.32) to estimate

$$\{N_+ - N, F\} \leq \frac{C\epsilon^2}{k\tau \sigma^{\tau + 1}}.$$ (4.4.38)

Together with the mapping property (4.4.34), the discussion of $P_+$ following (4.4.20), and the bounds (4.4.16) and (4.4.14) we obtain

$$\left|\int_0^1 \{((1 - \tau)(R + \{N, F\}) + \tau R, F\} \circ \Phi_{s-5\sigma,\eta r} \right|_{s-5\sigma,\eta r} \leq \frac{C\epsilon^2}{k\tau \sigma^{\tau + 1}},$$ (4.4.39)

and

$$\left|\{P - R\} \circ \Phi|_{s-5\sigma,\eta r} \leq |P - R|_{s-4\sigma,2\eta r} \leq C(\eta^2 + K^n e^{-K\sigma})$$ (4.4.40)

which proves (4.4.9).

Finally, we have

$$N_+ = N + \hat{R}(0) = e(\omega,t) + \langle \omega, I \rangle + \hat{R}(0,0;\omega,t)$$ (4.4.43)

which we need to re-write in normal form $e_+(\omega) + \langle \omega, I \rangle$ by finding a suitable frequency transformation $\phi$. Noting that $R$ is linear in $I$ by construction, this amounts to inverting the map $\omega \mapsto \omega + q(\omega; t)$, where

$$q(\omega; t) = \hat{R}_I(0,0;\omega, t)$$ (4.4.44)

which is bounded by $C\epsilon/r \leq v h$ from the Cauchy estimate, (4.4.32), and our assumption (4.4.3).

An application of a version of the implicit function theorem, Proposition A.4, then constructs $\phi : O_{1/2-3v} h \times (-1,1) \to O_{1-4v} h$ inverse to $q$ which satisfies the claimed estimates.

As in [38],[34], Theorem 4.4.1 can be used to prove the KAM theorem for real analytic Hamiltonians $H(\theta, I; \omega,t)$. However, in order to treat the more general class of Gevrey smooth Hamiltonians $H \in G^{n,1}((\mathbb{T}^n \times D \times \Omega) \times (-1,1))$, we require the approximation result Proposition 4.5.1. This method was used to prove Theorem 2.1 in [36] without the presence of the parameter $t$. 

\[\square\]
4.5 Approximations of Gevrey functions

It is convenient to extend the $P^j$ to Gevrey functions $\tilde{P}^j \in G^{\rho,1}_{C_L1,C_L2}(\mathbb{T}^n \times \mathbb{R}^{2n} \times (-1,1))$ where $C$ depends only on $n$ and $\rho$ by making use of a Gevrey formulation of the Whitney extension theorem, from Theorem 4.7.1. We thus obtain the estimate

$$\|\tilde{P}^j\| \leq AL_1^{n+1}\|P^j\|$$  \hspace{1cm} (4.5.1)

where $A$ also only depends on $n$ and $\rho$. We can then cut-off $\tilde{P}^j$ without loss to have $(I,\omega)$ support in $B_1 \times B_\tilde{R} \subset \mathbb{R}^{2n}$, where $1 \ll \tilde{R}$ is such that $\Omega^0 \subset B_{\tilde{R}-1}$. From here, we suppress the tilde in our notation, as well as the factor $C$ in our Gevrey constant.

**Proposition 4.5.1.** Suppose $P \in G^{\rho,1}_{C_L1,C_L2}(\mathbb{T}^n \times \mathbb{R}^{2n} \times (-1,1))$ satisfies $\text{supp}(I,\omega)(P) \subset B_1 \times B_\tilde{R}$. If $u_j, w_j, v_j$ are positive real sequences monotonically tending to zero such that

$$v_j L_2, w_j L_2 \leq u_j L_1 \leq 1, \quad v_0, w_0 \leq L_2^{-1-\zeta}$$  \hspace{1cm} (4.5.2)

where $1 \leq L_1 \leq L_2$ and $0 < \zeta \leq 1$ are fixed, then we can find a sequence of real analytic functions $P_j : U_j \to \mathbb{C}$ such that

$$|P_{j+1} - P_j|_{U_{j+1}} \leq C(\tilde{R}^n + 1)L_1^n \exp \left( -\frac{3}{4}(\rho - 1)(2L_1 u_j)^{-1/(\rho-1)} \right) \|P\|,$$  \hspace{1cm} (4.5.3)

$$|P_0|_{U_0} \leq C(\tilde{R}^n + 1)\left( 1 + L_1^n \exp \left( -\frac{3}{4}(\rho - 1)(2L_1 u_0)^{-1/(\rho-1)} \right) \right),$$  \hspace{1cm} (4.5.4)

and

$$|\partial_\alpha^n(P - P_j)(\theta, I; \omega, t)| \leq C(1 + \tilde{R}^n)L_1^n L_2 \exp \left( -\frac{3}{4}(\rho - 1)(2L_1 u_j)^{-1/(\rho-1)} \right)$$  \hspace{1cm} (4.5.5)

in $\mathbb{T}^n \times B_1 \times B_\tilde{R} \times (-1,1)$ for $|\alpha| \leq 1$, where

$$U_j^m := \{(\theta, I; \omega, t) \in \mathbb{C}^n / 2\pi \mathbb{Z}^n \times \mathbb{C}^n \times \mathbb{C}^n \times \mathbb{C} :$$

$$|\text{Re}(\theta)| \leq \pi, |\text{Re}(I)| \leq 2, |\text{Re}(\omega)| \leq \tilde{R} + 1, |\text{Re}(t)| \leq 1,$$

$$|\text{Im}(\theta)| \leq 2u_j, |\text{Im}(I)| \leq 2v_j, |\text{Im}(\omega_k)| \leq 2w_j, |\text{Im}(t)| \leq (2L_2)^{-1} \}$$  \hspace{1cm} (4.5.6)

and

$$U_j := U_j^1$$  \hspace{1cm} (4.5.7)

where we have identified $[-\pi, \pi]^n$ with $\mathbb{T}^n$ for simplicity of notation.

**Proof.** We first extend $P$ to functions $F_j : U_j^2 \to \mathbb{C}$ that are almost analytic in $(\theta, I, \omega)$ and analytic in $t$. The Gevrey estimate on $t$-derivatives of $P$ imply that the Taylor expansions in $t$ have radius of convergence $L_2^{-1}$, and so the expression

$$F_j(\theta + i\tilde{\theta}, I + i\tilde{I}, \omega + i\tilde{\omega}, t + i\tilde{t}) := \sum_{\mathcal{M}_j} \partial_\alpha^n \partial_\beta^\gamma \partial_\delta^\epsilon P(\theta, I; \omega, t) \frac{(i\tilde{\theta})^\alpha (i\tilde{I})^\beta (i\tilde{\omega})^\gamma (i\tilde{t})^\delta}{\alpha! \beta! \gamma! \delta!}$$  \hspace{1cm} (4.5.8)

is convergent on $U_j^2$ where the index set is given by

$$\mathcal{M}_j = \{ (\alpha, \beta, \gamma, \delta) : \alpha_k \leq N_1, \beta_k \leq N_2, \gamma_k \leq N_3 \}$$  \hspace{1cm} (4.5.9)
where
\[ N_1 = [(2L_1 u_j)^{-1/(\rho - 1)} + 1, \quad N_2 = (2L_2 v_2)^{-1/(\rho - 1)} + 1, \quad N_3 = (2L_2 w_j)^{-1/(\rho - 1)} + 1. \] (4.5.10) 

A simple consequence of Stirling’s formula is that for \( s \in (0, 1] \) and \( m \in \mathbb{Z} \cap [1, t^{-1/(\rho - 1)} + 1], \)
we have the estimate
\[ s^m m! \rho^{-1} \leq C(\rho) m^{\frac{\rho - 1}{\rho}} e^{-m(\rho - 1)} \] (4.5.11)
which will be used repeatedly in estimating the almost analyticity of \( F_j \). By using the Gevrey estimates of \( P \) to control each summand on \( F_j \), we obtain the bound
\[ |F_j|_{L^2} \leq \|P\| \cdot \sum_{M_j} (2^{-\delta}(2L_1 u_j)^{n}(2L_2 v_j)^{|\beta|}(2L_2 w_j)^{|\gamma|}(2L_2 t)^{\delta}(\alpha!\beta!\gamma!)^{\rho-1}. \] (4.5.12)

Now, for \( k \) with \( \alpha_k > 0 \), we can apply (4.5.11) to obtain
\[ (2L_1 u_j)^{\alpha_k \alpha_k! \rho^{-1}} \leq C \alpha_k^{\frac{\rho - 1}{\rho}} e^{-\alpha_k(\rho - 1)}. \] (4.5.13)

Bounding the factors \( (2L_2 v_j)^{|\beta|} \beta_k! \rho^{-1} \) and \( (2L_2 w_j)^{|\gamma|} \gamma_k! \rho^{-1} \) in the same way, we arrive at the bound
\[ |F_j|_{L^2} \leq 2\|P\| \left( 1 + C(\rho) \sum_{m=1}^{\infty} m^{1-\rho} e^{-m(\rho - 1)} \right)^{3n} = C\|P\|. \] (4.5.14)

We now consider the Cauchy–Riemann operator \( \tilde{\partial}_{z_k} = \frac{1}{2}(\partial_{\theta_k} + \partial_{\bar{\theta}_k}) \) where \( z_k = \theta_k + i\bar{\theta}_k \).
Applying \( \tilde{\partial}_{z_k} \) to \( F_j \), we obtain
\[ 2\tilde{\partial}_{z_k} = \sum_{M_j, \alpha_k = N_1} \partial_{\theta_k}^{\alpha_k} \partial_{\bar{\theta}_k}^{\beta_k} P(\theta, I; \omega, t) \frac{(i\bar{\theta})^{\alpha}(i\bar{\omega})^{\beta}(i\bar{t})^{\gamma}}{\alpha!\beta!\gamma!}\delta! \] (4.5.15)

From the Gevrey estimates for \( P \), we can estimate each summand by
\[ |F_j|_{L^2} \leq \|P\| \|L_1(2L_1 u_j)^{|\alpha|}(2L_2 v_j)^{|\beta|}(2L_2 w_j)^{|\gamma|}(2L_2 t)^{\delta}(\alpha!\beta!\gamma!)^{\rho-1}(\alpha_k + 1)\rho. \] (4.5.16)

Using the fact that \( (2L_1 u_j)^{-1/(\rho - 1)} \leq N_1 \leq (2L_1 u_j)^{-1/(\rho - 1)} + 1 \) together with (4.5.11), we can bound
\[ (2L_1)^{\alpha_k \alpha_k! \rho^{-1}}(\alpha_k + 1) \rho \leq C(2L_1 u_j)^{-\frac{3\rho - 1}{2\rho - 2}} \exp(-\rho - 1)(2L_1 u_j)^{-1/(\rho - 1)}. \] (4.5.17)

Hence
\[ |\tilde{\partial}_{z_k} F_j|_{L^2} \leq C L_1(2L_1 u_j)^{-\frac{3\rho - 1}{2\rho - 2}} \exp(-\rho - 1)(2L_1 u_j)^{-1/(\rho - 1)}\|P\| \] (4.5.18)
\[ \leq C L_1 \exp \left( -\frac{3}{4}(\rho - 1)(2L_1 u_j)^{-1/(\rho - 1)} \right) \|P\|. \] (4.5.19)

where the last line follows from absorbing the power term into the exponential. The constant \( C \) depends only on \( n \) and \( \rho \).
Applying the same method of estimation to $F_j$ with an additional differentiation yields
\[
\max(|\bar{\partial}_k \partial_l \bar{\partial}_z F_j|, |\partial_k \partial_l \bar{\partial}_z F_j|, |\bar{\partial}_k \partial_l \partial_z F_j|, |\partial_k \partial_l \partial_z F_j|) \leq CL_1 L_2 \exp \left( -\frac{3}{4}(\rho - 1)(2L_1 u_j)^{-1/(\rho - 1)} \right) \|P\|. \tag{4.5.20}
\]
if $\beta + \gamma \leq 1$, where we have also used $1 \leq L_1 \leq L_2$.

We can also bound $\bar{\partial}_{k+\alpha} F_j$ and $\bar{\partial}_{\omega+k} F_j$ in the same fashion as (4.5.18). Indeed, since $L_2 \leq \min((L_2 v_j)^{-1/\xi}, (L_2 w_j)^{-1/\xi})$, we obtain the stronger estimate
\[
|\bar{\partial}_{\omega} F_j| \leq C \exp(\frac{3}{4}(\rho - 1)(2L_1 u_j)^{-1/(\rho - 1)}) \|P\|. \tag{4.5.22}
\]
\[
|\partial_{\omega} F_j| \leq C \exp(\frac{3}{4}(\rho - 1)(2L_1 u_j)^{-1/(\rho - 1)}) \|P\|. \tag{4.5.23}
\]
\[
|\partial_{\omega} F_j| \leq C \exp(\frac{3}{4}(\rho - 1)(2L_1 u_j)^{-1/(\rho - 1)}) \|P\|. \tag{4.5.24}
\]
\[
|\partial_{\omega} F_j| \leq C \exp(\frac{3}{4}(\rho - 1)(2L_1 u_j)^{-1/(\rho - 1)}) \|P\|. \tag{4.5.25}
\]
where $C$ now depends on $n, \rho, \xi$ only.

Generalising, we now let $z = (\theta + i\hat{\theta}, I + i\hat{I}, \omega + i\hat{\omega}, t + i\hat{t}) \in \mathbb{C}^n \times \mathbb{C}^{2n} \times \mathbb{C}$. Then if $\alpha_k \leq 1$ for each $k$, applying the operator $\bar{\partial}_z^\alpha$ to $F_j$ amounts to restricting the index set $M_j$ to the multi-indices with $k$-th component maximal, for each $k$ with $\alpha_k = 1$. Hence we obtain
\[
|\bar{\partial}_z^\alpha F_j|_{U_j^2} \leq CL_n \exp(\frac{3}{4}(\rho - 1)(2L_1 u_j)^{-1/(\rho - 1)}) \|P\|. \tag{4.5.26}
\]
We also obtain the derivative bound
\[
|\partial_z^{\beta+\gamma} F_j|_{U_j^2} \leq CL_n [\beta + \gamma] \exp(\frac{3}{4}(\rho - 1)(2L_1 u_j)^{-1/(\rho - 1)}) \|P\|. \tag{4.5.27}
\]
for $|\alpha| \geq 1, |\beta + \gamma| \leq 1$ as in (4.5.20). Of course, if $\alpha_{3n+1} = 1$, the operator $\bar{\partial}_z^\alpha$ will annihilate $F_j$ because of the analyticity in $t$.

Having constructed a family of almost-analytic extensions $F_j$ of $P$, the next step is to approximate the $F_j$ by functions that are real analytic in $U_j^2$. The key tool here is Green’s formula.
\[
\frac{1}{2\pi i} \int_{\partial D} \frac{f(\eta)}{\eta - \zeta} d\eta + \frac{1}{2\pi i} \int_{\partial \overline{D}} \frac{\bar{f}(\eta)}{\eta - \zeta} d\eta \wedge d\bar{\eta} = f(\zeta) \tag{4.5.28}
\]
if $D$ is a bounded domain symmetric with respect to the real axis with a piecewise smooth positively oriented boundary $\partial D$, $f \in \mathcal{C}^1(\overline{D})$, and $\zeta \in D$. We also observe that the first summand in (4.5.28) is real analytic in $D$ if $f(\eta) = \overline{f(\overline{\eta})}$ on $\partial D$, a property satisfied by our $F_j$ by the nature of their construction. We define open rectangles in $\mathbb{C}$ given by
\[
D_k = \{z : |\text{Re}(z)| < a_k, |\text{Im}(z)| < b_k\} \tag{4.5.29}
\]
where
\[
a_k = \begin{cases} 
\pi & \text{if } 1 \leq k \leq n \\
2 & \text{if } n + 1 \leq k \leq 2n \\
\bar{R} + 1 & \text{if } 2n + 1 \leq k \leq 3n
\end{cases}
\]
and
\[
|z_1 - \eta_1| \geq 2u_j \quad \text{for} \quad k \neq 1.
\]

Additionally, we define the oriented union of line segments
\[
\Gamma = [-\pi - 2iu_j, \pi - 2iu_j] \cup [\pi + 2iu_j, -\pi + 2iu_j].
\] (4.5.30)

We also introduce the $2\pi$-periodic meromorphic function
\[
K(\eta, \zeta) = \lim_{N \to \infty} \sum_{|k| \leq N} \frac{1}{\eta - \zeta + 2\pi k} = \frac{1}{\eta - \zeta} + K_1(\eta, \zeta).
\] (4.5.31)

Writing $F_{j,0}(z) = F_j(z)$, we define
\[
F_{j,1}(z) := \frac{1}{2\pi i} \int_{\partial D_1} F_{j,0}(\eta_1, z_2, \ldots, z_{3n+1}) K(\eta_1, z_1) d\eta_1.
\] (4.5.32)

This function is analytic and $2\pi$-periodic for $z_1$ in the strip $|\text{Im}(z_1)| < 2u_j$. Moreover, it satisfies the identity $F_{j,1}(z) = F_{j,1}(\bar{z})$. If $z_1 \in D_1$, we can safely avoid poles by integrating about $\partial D_1$. Thus, periodicity gives us
\[
F_{j,1}(z) = \frac{1}{2\pi i} \int_{\partial D_1} F_{j,0}(\eta_1, z_2, \ldots, z_{3n+1}) K(\eta_1, z_1) d\eta_1.
\] (4.5.33)

Green’s formula (4.5.28) then implies
\[
F_{j,1}(z) = F_{j,0}(z) - \frac{1}{2\pi i} \int_{D_1} \bar{\partial}_{\eta_1} F_{j,0}(\eta_1, z_2, \ldots, z_{3n+1}) K(\eta_1, z_1) d\eta_1 \wedge d\bar{\eta}_1.
\] (4.5.34)

which extends to $\text{Re}(z_1) = \pi$ by continuity.

Defining the set $U_{j,1}^2 = U_j^2 \cap \{|\text{Im}(z_1)| \leq u_j\}$, we claim that for any multi-indices $\alpha, \beta, \gamma$ with $\alpha_1 = 0$ and $|\beta| + |\gamma| \leq 1$, we have the estimate
\[
|\partial_{z_k}^\alpha \bar{\partial}_{z_k}^\beta \partial_{z_k}^\gamma (F_{j,1} - F_{j,0})|_{U_{j,1}^2} \leq CL_1^n L_2^{[|\beta|+|\gamma|]} \exp\left(-\frac{3}{4}(\rho - 1)(2L_1 u_j)^{-1/(\rho-1)}\right) \|P\|. \quad (4.5.35)
\]

For $k \neq 1$, this estimate follows directly from (4.5.27) by differentiating $F_{j,1}$ under the integral. For $k = 1$, we first use (4.5.28) to write
\[
\frac{1}{2\pi i} \int_{D_1} \bar{\partial}_{\eta_1} F_{j,0}(\eta_1, z_2, \ldots, z_{3n+1}) \frac{d\eta_1}{\eta_1 - z_1} \quad (4.5.36)
\]
\[
= -z_1 \bar{\partial}_{z_1} F_j(z) + \frac{1}{2\pi i} \int_{D_1} \bar{\partial}_{\eta_1} F_{j,0}(\eta_1, z_2, \ldots, z_{3n+1}) - \bar{\partial}_{z_1} F_j(z) \frac{d\eta_1}{\eta_1 - z_1} \wedge d\bar{\eta}_1.
\]
\]

We can then differentiate under the integral as before and use (4.5.27) to establish (4.5.35). Together with (4.5.27), we arrive at
\[
|\partial_{z_k}^\alpha \bar{\partial}_{z_k}^\beta \partial_{z_k}^\gamma F_{j,1}|_{U_{j,1}^2} \leq CL_1^n L_2^{[|\beta|+|\gamma|]} \exp\left(-\frac{3}{4}(\rho - 1)(2L_1 u_j)^{-1/(\rho-1)}\right) \|P\| \quad (4.5.37)
\]

for $|\alpha| \geq 1$. 

for $k \neq 1$. 

\begin{align*}
&= -z_1 \bar{\partial}_{z_1} F_j(z) + \frac{1}{2\pi i} \int_{D_1} \bar{\partial}_{\eta_1} F_{j,0}(\eta_1, z_2, \ldots, z_{3n+1}) - \bar{\partial}_{z_1} F_j(z) \frac{d\eta_1}{\eta_1 - z_1} \wedge d\bar{\eta}_1. \\
&= -z_1 \bar{\partial}_{z_1} F_j(z) + \frac{1}{2\pi i} \int_{D_1} \bar{\partial}_{\eta_1} F_{j,0}(\eta_1, z_2, \ldots, z_{3n+1}) - \bar{\partial}_{z_1} F_j(z) \frac{d\eta_1}{\eta_1 - z_1} \wedge d\bar{\eta}_1.
\end{align*}
Since by construction $F_{u} \alpha_{5}$ and $F_{u} \alpha_{6}$ are always satisfied, as well as decreasing sequences of the parameters $\rho, \sigma_{j}, K_{j}$ such that the hypotheses of Theorem 4.4.1 are always satisfied, we obtain the estimate (4.5.35) in $U_{j}^{2}$ for $\alpha$ with $\alpha_{k} = 0$ for $k \leq m$.

For $n + 1 \leq m \leq 3n$, we no longer require $2\pi$ periodicity of our construction, and can instead define

$$F_{j,m}(z) := \frac{1}{2\pi i} \int_{\partial D_{m}} \frac{F_{j,0}(z_{1},\ldots,\eta_{m},\ldots,\eta_{3n+1})}{\eta_{m} - z_{m}} \, d\eta_{m} \quad (4.5.38)$$

for $z \in U_{j}^{2}$. We can then proceed as above to estimate $F_{j,m}$ in $U_{j,m}$, defined inductively by

$$U_{j,m}^{2} = U_{j,m-1}^{2} \cap \{|\text{Im}(z_{m})| \leq p_{m}\} \quad (4.5.39)$$

where $p_{m} = v_{j}, w_{j}$ for $n + 1 \leq m \leq 2n$ and $2n + 1 \leq m \leq 3n$ respectively.

For $2n + 1 \leq m \leq 3n$, the constant $C$ in the estimate (4.5.35) has to be multiplied by $(\bar{R} + 1)^{m-2n}$ to account for the measure of $D_{m}$, but this is the only dependence of $C$ on anything other than $n, \rho, \zeta$. We absorb this dependence into the $C$ for the rest of the proof.

We now set $P_{j} := F_{j,3n}$, which is analytic in $U_{j}$ by construction. We have shown that for $l \leq 1$ and for any $1 \leq k \leq 3n + 1$, we have

$$|\partial_{x_{k}}^{l}(P_{j} - F_{j})|_{U_{j}} \leq C L_{1}^{n} L_{2}^{l} \exp\left(-\frac{3}{4}(\rho - 1)(2L_{1}u_{j})^{-1/(\rho-1)}\right) \|P\|. \quad (4.5.40)$$

In particular, this implies that

$$|P_{j+1} - P_{j}|_{U_{j+1}} \leq |P_{j+1} - F_{j+1}|_{U_{j+1}} + |P_{j} - F_{j}|_{U_{j+1}} + |F_{j+1} - F_{j}|_{U_{j+1}} \quad (4.5.41)$$

$$\leq C L_{1}^{n} \exp\left(-\frac{3}{4}(\rho - 1)(2L_{1}u_{j})^{-1/(\rho-1)}\right). \quad (4.5.42)$$

Since by construction $F_{j}(z) = P(z)$ for real $z$, we also obtain

$$|\partial_{x_{k}}^{l}(P_{j}(x) - P(x))|_{U_{j} \cap \mathbb{R}^{3n+1}} \leq C L_{1}^{n} L_{2}^{l} \exp\left(-\frac{3}{4}(\rho - 1)(2L_{1}u_{j})^{-1/(\rho-1)}\right). \quad (4.5.43)$$

The final claimed estimate arises as follows

$$|P_{0}|_{U_{0}} \leq |F_{0}|_{U_{0}} + |P_{0} - F_{0}|_{U_{0}} \leq C \left(1 + L_{1}^{n} \exp\left(-\frac{3}{4}(\rho - 1)(2L_{1}u_{0})^{-1/(\rho-1)}\right)\right) \|P\|. \quad (4.5.44)$$

Our next goal is to set up an iterative scheme based on Theorem 4.4.1 that converges in the Gevrey class $G^{\rho,(\rho+1)+1,\rho(1+1)+1,1}(\mathbb{T}^{n} \times D \times \Omega \times (-1,1))$. This involves defining decreasing sequences of our parameters $s_{j}, r_{j}, h_{j}, \eta_{j}, \epsilon_{j}, \sigma_{j}, K_{j}$ such that the hypotheses of Theorem 4.4.1 are always satisfied, as well as decreasing sequences of the the parameters $u_{j}, v_{j}, w_{j}$ such that the hypotheses of the Proposition 4.5.1 are always satisfied.
4.6 The iterative scheme

We begin by examining how the parameters are defined in the first iteration. It is convenient to choose a weighted error \( 0 < E < 1/64 \) and a fixed \( 0 < \hat{\epsilon} \leq 1 \) such that

\[
\eta = E^{1/2}, \quad \epsilon = \hat{\epsilon} Ke^\sigma \tau^{+1}. \tag{4.6.1}
\]

We can then define \( K \) and \( h \) by

\[
K^n e^{-K\sigma} = E, \quad h = \frac{K}{2K^{\tau+1}}. \tag{4.6.2}
\]

The choice for \( K \) is motivated by the form of the error term (4.4.9) and the choice for \( h \) is motivated by (4.4.3).

We define the subsequent values of the parameters \( r, s, \sigma \) by

\[
r_+ = \eta r, \quad s_+ = s - 5\sigma, \quad \sigma_+ = \delta \sigma \tag{4.6.3}
\]

where \( \delta(\rho) \in (0, 1) \) will subsequently be chosen in a convenient way.

The KAM step together with the definitions of our parameters yields the estimate

\[
|P_+|_{s_+, r_+, (1/2-3\nu)h} \leq \frac{1}{2} c_1^{1/2} K r_+ \sigma_+^{\tau+1} E^{3/2} \tag{4.6.4}
\]

for some \( c_1 > 1 \) dependent only on \( n, \rho, \tau \).

We can then choose our subsequent weighted and unweighted error as

\[
E_+ = c_1^{1/2} E^{3/2}, \quad \epsilon_+ = \hat{\epsilon} K r_+ \sigma_+^{\tau+1} E_+. \tag{4.6.5}
\]

The constants \( \eta_+, K_+, h_+ \) can then be defined in terms of \( E_+ \) and \( \sigma_+ \) as before, and we have \( c_1 E_+ = (c_1 E)^{3/2} \) which will give us rapid convergence provided \( c_1 < 1 \). Moreover, if we have

\[
h_+ \leq (1/2 - 3\nu)h \tag{4.6.6}
\]

then we obtain

\[
|P_+|_{s_+, r_+, h_+} \leq \epsilon_+/2. \tag{4.6.7}
\]

For brevity of notation, we write \( D_j = D_{s_j, \sigma_j}, O_j = O_{h_j}, V_j = D_j \times O_j \times (-1, 1) \).

Now, in terms of \( \delta \) we choose our sequences \( s_j \) and \( \sigma_j \) to decay at a geometric rate, by setting

\[
s_j = s_0 \delta^j, \quad \sigma_j = \sigma_0 \delta^j, \quad s_0 = 5\sigma_0 / (1 - \delta) \tag{4.6.8}
\]

where \( \sigma_0 \ll 1 \) remains to be chosen to be convenient for the subsequent estimates.

A direct consequence of this definition is that

\[
s_{j+1} = s_j - 5\sigma_j, \quad \sigma_j = (1 - \delta)s_j/5. \tag{4.6.9}
\]
For the parameter sequences required in our approximation lemma, we set

\[ u_j = 2s_0\delta^j M(\rho), \quad v_j = 2r_0\delta^j M(\rho), \quad w_j = 2h_0\delta^j M(\rho) \quad (4.6.10) \]

where \( M(\rho) = (25(1-\delta)^{-2} + 2)^{1/2} \) and assume for now that these sequences satisfy \( (4.5.2) \).

We write \( U_j = U_j^1 \cap \{|I| < r\} \) where \( U_j^1 \) is defined as in Proposition 4.5.1. Now, by applying Proposition 4.5.1 to the terms \( P^0, P^1 \) from (4.2.14), we obtain sequences \( P^0_j, P^1_j \) of real analytic functions in \( U_j^1 \) that are good approximations to \( P^0 \) and \( P^1 \).

We set \( P_j(\theta, I; \omega, t) := (P^0_j(I; \omega, t)I) + P^1_j(\theta, I; \omega, t). \quad (4.6.11) \)

Proposition 4.5.1, together with the factors picked up during the Whitney extension of \( P^0, P^1 \) in (4.5.1) then yields the estimates

\[ |P_0|_{U_0} \leq C(n, \rho, \zeta)(\bar{R}^n + 1)L_1^{2n+1}(P)_r \quad (4.6.12) \]

and

\[ |P_j - P_{j-1}|_{U_j} \leq C(n, \rho, \zeta)(\bar{R}^n + 1)L_1^{2n+1}(P)_r e^{-B_0\sigma_j^{-1/(\rho-1)}} \quad (4.6.13) \]

\[ = C(n, \rho, \zeta)(\bar{R}^n + 1)L_1^{2n+1}(P)_r e^{-B_0\sigma_j^{-1/(\rho-1)}} \quad (4.6.14) \]

where \( B_0, \bar{B}_0 \) are constant multiples of \( L_1^{-1/(\rho-1)} \) by factors depend only on \( \rho \) and \( \delta \).

We now choose the constant

\[ \hat{\epsilon} := \langle R \rangle r L_1^{N-n-2}(akr)^{-1}. \quad (4.6.15) \]

and set

\[ \bar{\epsilon}_j := \hat{\epsilon} \kappa r_0 \sigma_j^{\tau+1} \exp(-B_0\sigma_j^{-1/(\rho-1)}) \quad (4.6.16) \]

where \( N(n, \tau, \rho) \) and \( a(n, \rho, \tau, \zeta, \bar{R}) \) are constants to be fixed later in such a way that we can bound \( P_0 \) and \( P_j - P_{j-1} \) by \( \epsilon_0 \) and \( \bar{\epsilon}_j \) respectively, yielding rapid convergence of the \( P_j \).

Note that we will have \( \hat{\epsilon} \leq 1 \) by taking the \( \epsilon \) in Theorem 4.2.2 sufficiently small.

We then define the weighted error \( E_j \) by

\[ E_j := c_1^{-1} \exp(-B\sigma_j^{-1/(\rho-1)}) \quad (4.6.17) \]

where \( c_1(n, \rho, \tau) \) is the constant from (4.6.4) and

\[ B := B_0(\delta^{-1/(\rho-1)} - 1)/2. \quad (4.6.18) \]

The desired recurrence \( E_{j+1} = \epsilon_j^{1/2} E_j^{3/2} \) together with the recurrence for \( \sigma_j \) then forces \( \delta = (2/3)^{\rho-1} \), which in turn forces \( B = B_0/4 = A_0(\rho)L_1^{-1/(\rho-1)} \).

We now define the remaining parameters using \( E_j \) as previously discussed. We set \( \eta_j = E_j^{1/2}, r_{j+1} = \eta_j r_j \) and write

\[ \epsilon_j = \bar{\epsilon} \kappa r_j \sigma_j^{\tau+1} E_j. \quad (4.6.19) \]

Our definition of \( E_j \) in (4.6.17) was chosen precisely so that we can obtain the inequality \( \epsilon_j \leq \epsilon_{j+1}/2 \) which allows us to conveniently handle the error terms of the analytic
approximations in our iterative scheme.

We define \( K_j \geq 1 \) implicitly by \( K_j^n e^{-K_j \sigma_j} = E_j \).

Setting \( x_j = K_j \sigma_j \), we get

\[
x_j^n e^{-x_j} = E_j \sigma_j^n = c_1^{-1} \sigma_j^n \exp(-B \sigma_j^{-1/(\rho-1)})
\]

(4.6.20)

and taking logarithms this becomes

\[
x_j - n \log(x_j) = B \sigma_j^{-1/(\rho-1)} - n \log(\sigma_j) + \log(c_1).
\]

(4.6.21)

For convenience in our numerous estimates that require \( \sigma_j \) to be small, we now set

\[
\sigma_0 = \sigma L_1^{-1}(\log(L_1 + e))^{-1/(\rho-1)}
\]

(4.6.22)

where \( \sigma \leq \tilde{\sigma}(n, \rho) \ll 1 \).

We can then bound the right-hand side of (4.6.21) using

\[
B \sigma_j^{-1/(\rho-1)} - n \log(\sigma_j) + \log(c_1) \geq A_0(L_1 \sigma_0)^{-1/(\rho-1)} + o(1)
\]

(4.6.23)

\[
\geq A_0 \sigma_0^{-1/(\rho-1)} \gg 1.
\]

(4.6.24)

Hence for each \( j \geq 1 \) we indeed obtain a unique \( x_j(\sigma) \) such that

\[
x_j \geq x_j - n \log(x_j) \geq B \sigma_j^{-1/(\rho-1)} \gg 1.
\]

(4.6.27)

This implies that we uniformly have

\[
x_j - n \log(x_j) = x_j(1 + o(1))
\]

(4.6.28)

as \( \sigma \to 0 \). On the other hand we can also bound this quantity above by again making use of (4.6.22) to arrive at

\[
x_j - n \log(x_j) \leq B \sigma_j^{-1/(\rho-1)}(1 - nA_0^{-1}(L_1 \sigma_j)^{1/(\rho-1)} \log(L_1 \sigma_j) + nA_0^{-1}(L_1 \sigma_0)^{1/(\rho-1)}(\log(L_1) + \log(c_1))))
\]

\[
= B \sigma_j^{-1/(\rho-1)}(1 + o(1))
\]

as \( \sigma \to 0 \), again uniformly with respect to \( j \). Hence

\[
B \sigma_j^{-1/(\rho-1)} \leq x_j \leq B \sigma_j^{-1/(\rho-1)}(1 + o(1))
\]

(4.6.29)

as \( \sigma \to 0 \).

Defining \( h_j = \kappa/(2K_j^{\rho+1}) \), we also fix \( v = 1/54 \).

Having set up the parameters, our next task is to establish the remaining hypotheses in (4.4.3), noting that the last is immediate from our definition of \( h_j \). First we observe
\[ \epsilon_j = \epsilon \rho_j \sigma_j^{\tau+1} E_j \leq \kappa \rho_j \sigma_j^{\tau+1} E_j \leq c \kappa \eta_j \rho_j \sigma_j^{\tau+1} \]

since \( \eta_j^2 = E_j \) and \( E_j = c_1^{-1} \exp(-B \sigma_j^{-1/(\rho-1)}) \) is \( o(1) \) as \( \sigma \to 0 \). This establishes the first remaining hypothesis in (4.4.3). To prove the only remaining hypothesis, we compute

\[ \frac{\epsilon_j}{r_j h_j} = \epsilon \kappa_0 \rho_j^{\tau+1} E_j / h_j \]

\[ \leq 2 E_j x_j^{\tau+1} \]

\[ \leq 2 c_1^{-1} \exp(-B \sigma_j^{-1/(\rho-1)})(B \sigma_j^{-1/(\rho-1)})^{\tau+1}(1 + o(1)) \]

\[ \leq c(\rho, \tau) \exp(-A_0 (L_1 \sigma_j)^{-1/(\rho-1)}) \]

\[ \leq c(\rho, \tau) \exp(-A_0 (\sigma \delta \tau)^{-1/(\rho-1)}) \]

which is \( o(1) \) as \( \sigma \to 0 \) and hence we have verified the hypotheses of Theorem 4.4.1.

Moreover, for \( \tilde{\sigma}(n, \rho, \tau) \) sufficiently small, we have

\[ \prod_{j=0}^{\infty} (1 + C \epsilon_j / r_j h_j) \leq \exp(\sum_{j=0}^{\infty} C \epsilon_j / r_j h_j) \leq 2 \]

where the constant \( C(n, \tau) \) in (4.6.38) comes from the estimates in Theorem 4.4.1.

From (4.6.29), we have \( (x_j / x_{j+1}) = (\sigma_{j+1} / \sigma_j)^{1/(\rho-1)}(1 + o(1)) \). Consequently, we have

\[ (h_{j+1}/h_j) = (x_j / x_{j+1})^{\tau+1}(\sigma_{j+1}/\sigma_j)^{\tau+1} \]

\[ = \delta^{(\tau+1)/\rho/(\rho-1)}(1 + o(1)) \]

\[ = (2/3)^{\rho/(\rho-1)}(1 + o(1)) \]

\[ < (4/9)^\rho \]

\[ < 1/2 - 3v \]

for \( \tilde{\sigma}(n, \rho, \tau) \) sufficiently small, which verifies \( h_+ < (1/2 - 3v)h \).

We now inductively establish the claim

\[ \tilde{\epsilon}_j < \frac{1}{2} \epsilon_{j+1}. \]

After doing this and choosing the constants \( a, N, \tilde{\sigma} \) as required, we will at last be in a position to state and prove the iterative result that arises from \( n \) applications of the KAM step.

Unravelling the definitions of \( \epsilon_j, \tilde{\epsilon}_j \), we obtain

\[ \tilde{\epsilon}_0 / \epsilon_1 = (r_0 \sigma_0^{\tau+1} \exp(-4B \sigma_0^{-1/(\rho-1)})) / (r_1 \sigma_1^{\tau+1} E_1) = C(n, \rho, \tau) \exp(-2B \sigma_0^{-1/(\rho-1)}) \leq 1/2 \]

for sufficiently small \( \tilde{\sigma} \ll 1 \).
Moreover
\[(\varepsilon_j/\varepsilon_{j+1})(\varepsilon_j/\varepsilon_{j-1}) = \exp(-\frac{4}{3}B\sigma_j^{-1/(\rho-1)}) \cdot \frac{\varepsilon_j}{\varepsilon_{j+1}} = C(n, \rho, \tau) \exp(-\frac{1}{3}B\sigma_j^{-1/(\rho-1)}) \leq 1\] (4.6.46)
for sufficiently small $\tilde{\sigma} \ll 1$. This establishes (4.6.44).

Next we show that the sequences $u_j, v_j, w_j$ defined in (4.6.10) satisfy the hypotheses (4.5.2) of Proposition 4.5.1. First we observe from the definition of $\sigma_0$ that
\[4s_0L_1 = 20\sigma_0(1 - \delta)^{-1}L_1 \leq 20(1 - \delta^{-1})\sigma \leq 1\] (4.6.47)
for sufficiently small $\tilde{\sigma} \ll 1$. At this point, we fix $\tilde{\sigma}(n, \rho, \tau) \ll 1$. For the sequence $w_j$, we estimate
\[h_0 = \frac{\kappa\sigma_{0}^{\tau+1}}{2\rho_{0}^{\tau+1}} \leq K\sigma_{0}^{\tau+1} \leq \kappa\sigma_0 \leq \kappa s_0 \leq L_2^{-1-\zeta}s_0.\] (4.6.48)
This implies that $w_jL_2 \leq u_jL_1$ and $w_0 \leq L_2^{-1-\zeta}$.

Setting
\[r_0 = c(n, \rho, \tau, \zeta)\tau\] (4.6.49)
and using (4.2.17), we obtain
\[r_0 \leq cL_2^{-1-\zeta} < L_2^{-1-\zeta}\] (4.6.50)
and
\[r_0L_2 < cL_1^{-\zeta} \leq s_0L_1\] (4.6.51)
by choosing the $c$ in (4.6.49) sufficiently small.

Finally, we need to choose $0 < a(n, \rho, \tau, \zeta, \bar{R}) \leq 1$ and $N(n, \rho, \tau)$ such that
\[|P_0|U_0 \leq \tilde{\epsilon}_0\] (4.6.52)
and
\[|P_j - P_{j-1}|U_j \leq \tilde{\epsilon}_j\] (4.6.53)
for $j \geq 1$. From (4.6.12), we obtain
\[|P_0|U_0 \leq C_0(\bar{R}^n + 1)(P)_1L_1^{2n+1} = \tilde{\epsilon}krC_0(\bar{R}^n + 1)L_1^{-N+3n+3}a.\] (4.6.54)
On the other hand, the definition of $\sigma_0$ in 4.6.22 implies that
\[\sigma_0^{\tau+1}\exp(-4B\sigma_0^{-1/(\rho-1)}) = \sigma_0^{\tau+1}\exp(-4A_0(L_1\sigma_0)^{-1/(\rho-1)})\] (4.6.55)
\[= L_1^{-\tau-1}\log(L_1 + e)^{-(\rho-1)(\tau+1)}(L_1 + e)^{-4A_0\sigma_0^{-1/(\rho-1)}}\] (4.6.56)
\[\geq C(n, \rho, \tau)L_1^{-4A_0\sigma_0^{-1/(\rho-1)}-\tau-2}.\] (4.6.57)
We now fix
\[a = CC_0^{-1}(\bar{R}^n + 1)^{-1}r_0^{-1} = CC_0^{-1}(\bar{R}^n + 1)^{-1}c\] (4.6.58)
and
\[N = 4A_0\sigma_0^{-1/(\rho-1)} + \tau + 3n + 5\] (4.6.59)
where the $C(n, \rho, \tau), C_0(n, \rho), c(n, \rho, \tau, \zeta), A_0$ are as in (4.6.57), (4.6.54). This yields
\[|P_0|U_0 \leq C_0(\bar{R}^n + 1)(P)_1L_1^{2n+1} \leq \tilde{\epsilon}_0.\] (4.6.60)
We can then insert this estimate into (4.6.13) to obtain
\begin{equation}
|P_j - P_{j-1}|_{U_j} \leq C_0(\hat{R}^n + 1)(P)_{\ast} L_{1}^{2n+1} \exp(-B_0 \sigma_j^{-1/(p-1)}) \leq \tilde{\epsilon}_j. \tag{4.6.61}
\end{equation}

We have now set up the necessary sequences for our iterative scheme in addition to the key ingredient of Theorem 4.4.1.

We define the Hamiltonian
\begin{equation}
H_j(\theta, I; \omega, t) = N_0(I; \omega) + P_j(\theta, I; \omega, t) = \langle \omega, I \rangle + P_j(\theta, I; \omega, t) \tag{4.6.62}
\end{equation}
which is real analytic in \(U_j\). For \(j \geq 0\) we denote by \(D_j\) the class of real-analytic diffeomorphisms from \(D_{j+1} \times O_{j+1} \times (-1, 1) \rightarrow D_j \times O_j\) of the form
\begin{equation}
\mathcal{F}(\theta, I; \omega, t) = (\Phi(\theta, I; \omega), \phi(\omega; t)) = (U(\theta; \omega, t), V(\theta, I; \omega, t), \phi(\omega; t)) \tag{4.6.63}
\end{equation}
where \(\Phi\) is affine in \(I\) and canonical for fixed \((\omega, t)\) and the variable \(t \in (-1, 1)\) is regarded as a parameter.

**Proposition 4.6.1.** Suppose \(P_j\) is real analytic on \(U_j\) for each \(j \geq 0\), and that we have the estimates
\begin{equation}
|P_0|_{U_0} \leq \tilde{\epsilon}_0 \tag{4.6.64}
\end{equation}
and
\begin{equation}
|P_j - P_{j-1}|_{U_j} \leq \tilde{\epsilon}_j \tag{4.6.65}
\end{equation}
for each \(j \geq 1\).

Then for each \(j \geq 0\), we can find a real-analytic normal form \(N_j(I; \omega, t) = e_j(\omega, t) + \langle \omega, I \rangle\) and a real analytic map \(\mathcal{F}^j\) given by
\begin{equation}
\mathcal{F}^{j+1} = \mathcal{F}_0 \circ \ldots \circ \mathcal{F}_j : D_{j+1} \times O_{j+1} \times (-1, 1) \rightarrow (D_0 \times O_0) \cap U_j \tag{4.6.66}
\end{equation}
with the convention that the empty composition is the identity and where the \(\mathcal{F}_j \in D_j\) are such that
\begin{equation}
H_j \circ \mathcal{F}^{j+1} = N_{j+1} + R_{j+1} \tag{4.6.67}
\end{equation}
\begin{equation}
|R_{j+1}|_{j+1} \leq \epsilon_{j+1} \tag{4.6.68}
\end{equation}
\begin{equation}
|\bar{W}_j(\mathcal{F}_j - \text{id})|_{j+1}, |\bar{W}_j(D \mathcal{F}_j - \text{Id})\bar{W}_j^{-1}| < \frac{C\epsilon_j}{r_j h_j} \tag{4.6.69}
\end{equation}
\begin{equation}
|\bar{W}_0(\mathcal{F}^{j+1} - \mathcal{F}^j)|_{j+1} < \frac{C\epsilon_j}{r_j h_j} \tag{4.6.70}
\end{equation}
where the constants \(C\) depend only on \(n\) and \(\rho\) and \(\bar{W}_j = \text{diag}(\sigma_j^{-1} \text{Id}, r_j^{-1} \text{Id}, h_j^{-1} \text{Id})\).

**Proof.** An immediate application of Theorem 4.4.1 provides us with \(\mathcal{F}_0 \in D_0\) such that \(H_0 \circ \mathcal{F}_0 = N_1 + R_1\), where \(|R_1|_1 \leq \epsilon_1\). We proceed by induction, assuming that we have \(H_{j-1} \circ \mathcal{F}^j = N_j + R_j\), where \(N_j(I; \omega, t) = e_j(\omega, t) + \langle \omega, I \rangle\) is a real analytic normal form, \(R_j\) is real analytic in \(D_j \times O_j\), and \(|R_j|_j \leq \epsilon_j\).

We now apply Theorem 4.4.1 to the Hamiltonian \(N_j + R_j\) in order to find \(\mathcal{F}_j \in D_j\). From (4.6.7), we have \((N_j + R_j) \circ \mathcal{F}_j = N_{j+1} + \tilde{R}_{j+1}\) with the estimate
\begin{equation}
|\tilde{R}_{j+1}|_{j+1} \leq \frac{1}{2} \hat{\epsilon}_{k_{j+1}} \sigma_{j+1} \frac{1}{2} \hat{E}_{j+1}^{3/2} = \frac{1}{2} \epsilon_{j+1}. \tag{4.6.71}
\end{equation}
Once we show that
\[ \mathcal{F}^{j+1} : D_{j+1} \times O_{j+1} \times (-1, 1) \to U_j \]  
(4.6.72)
we can also establish
\[ |(P_j - P_{j-1}) \circ \mathcal{F}^{j+1}|_{j+1} \leq |P_j - P_{j-1}|_{U_j} \leq \frac{1}{2} \epsilon_{j+1}. \]  
(4.6.73)

Using our inductive assumption, we can rewrite
\[
H_j \circ \mathcal{F}^{j+1} = (N_0 + P_{j-1}) \circ \mathcal{F}^{j+1} + (P_j - P_{j-1}) \circ \mathcal{F}^{j+1}
\]
(4.6.74)
\[
= (H_{j-1} \circ \mathcal{F}^j) \circ \mathcal{F}_j + (P_j - P_{j-1}) \circ \mathcal{F}^{j+1}
\]
(4.6.75)
\[
= (N_j + R_j) \circ \mathcal{F}_j + (P_j - P_{j-1}) \circ \mathcal{F}^{j+1}.
\]
(4.6.76)
This gives us
\[ H_j \circ \mathcal{F}^{j+1} = N_{j+1} + R_{j+1} \]  
(4.6.77)
with
\[ |R_{j+1}|_{j+1} \leq \epsilon_{j+1}. \]  
(4.6.78)

It remains to verify (4.6.72), which we can do via the estimate (4.6.69). From (4.6.38) we obtain
\[
|\tilde{W}_0 D \mathcal{F}^{j+1} \tilde{W}_j^{-1}|_{j+1} = \left| \left( \prod_{k=0}^{j-1} W_k D \mathcal{F}_k \tilde{W}_k^{-1} \right) (\tilde{W}_j D \mathcal{F}_j \tilde{W}_j^{-1}) \right|
\]
(4.6.79)
\[
= \left| \left( \prod_{k=0}^{j-1} W_k D \mathcal{F}_k \tilde{W}_k^{-1} \right) (\tilde{W}_j D \mathcal{F}_j \tilde{W}_j^{-1}) \right|
\]
(4.6.80)
\[
= \left( \prod_{k=0}^{j-1} |\tilde{W}_k \tilde{W}_k^{-1}| \right) \prod_{k=0}^{\infty} \left( 1 + \frac{C \epsilon_k}{r_k h_k} \right)
\]
(4.6.81)
\[
\leq 2 \left( \prod_{k=0}^{j-1} |W_k W_k^{-1}| \right)
\]
(4.6.82)
\[
\leq 2 \delta^j
\]
(4.6.83)
where
\[ |\tilde{W}_k \tilde{W}_k^{-1}| = \sup(s_{k+1}/s_k, r_{k+1}/r_k, h_{k+1}/h_k) = s_{k+1}/s_k = \delta \]  
(4.6.84)
follows from the definition of the sequences \( s_j, r_j, h_j \).

Writing \( x + iy = (\theta, I, \omega) \in D_{j+1} \times O_{j+1} \), we can Taylor expand \( \mathcal{F}^{j+1} \) about \( x \) to obtain
\[ \mathcal{F}^{j+1}(x + iy) = \mathcal{F}^{j+1}(x) + i \tilde{W}_j^{-1} \left( \int_0^1 W_0 D \mathcal{F}^{j+1}(x + isy) \tilde{W}_j^{-1} ds \right) \tilde{W}_j y. \]  
(4.6.85)
From (4.6.79), the integral expression is bounded by \( 2 \delta^j \). Since \( |\tilde{W}_j y| \leq ((s_j \sigma_j)^2 + 2)^{1/2} = (25(1 - \delta)^{-2} + 2)^{1/2} = M(\rho) \), and \( \mathcal{F}^{j+1}(x) \) is real, we can conclude that \( \mathcal{F}^{j+1} : D_{j+1} \times O_{j+1} \times (-1, 1) \to U_j \). This completes the proof. \( \square \)

The next step is to find Gevrey estimates for the \( S_j := \mathcal{F}^{j+1} - \mathcal{F}^j \) so that we can show that this iterative scheme does indeed converge in the Gevrey class. We drop the dependence on \( t \) from our notation, and remind ourselves that the resulting Gevrey estimates
will be uniform in $t$. To this end we introduce the domains

$$
\tilde{D}_j := \{ (\theta, I) \in D_j : \text{Im}(\theta) < s_j/2 \}, \hspace{1cm} \tilde{O}_j := \{ \omega \in \mathbb{C}^n : \text{dist}(\omega, \Omega_\kappa) < h_j/2 \} \quad (4.6.86)
$$

For multi-indices $\alpha, \beta$ with $|\beta| \leq m$, we also introduce the following notation for the $(m - |\beta|)$-th Taylor remainder in the frequency variable, centred at $\omega$.

$$
R_m^\beta(\partial_\theta^\alpha \partial_\omega^\beta S^j)(\theta, I, \omega') := \partial_\theta^\alpha \partial_\omega^\beta S^j - \sum_{|\gamma| \leq m - |\beta|} (\omega' - \omega)^\gamma \partial_\theta^\beta \partial_\omega^{\beta + \gamma} S^j(\theta, I, \omega)/\gamma! \quad (4.6.87)
$$

We then have the following Gevrey estimates.

**Lemma 4.6.2.**

$$
|W_0^{\alpha} \partial_\theta^\alpha \partial_\omega^\beta S^j(\theta, 0, \omega)| \leq \epsilon AC^{m+|\alpha|+|\beta|} \left| L_{\text{max}}^{m+|\alpha|+|\beta|(\tau+1)+1} \kappa^{-|\beta|} \alpha! \beta! \rho'! E_j^{1/2} \right| 
$$

for all $(\theta, 0; \rho, \tau, \xi, \zeta) \in \hat{D}_{j+1} \times \hat{O}_{j+1} \times (-1, 1)$, where $\rho' = \rho(\tau + 1) + 1$.

$$
|W_0(R_m^\beta(\partial_\theta^\alpha \partial_\omega^\beta S^j))(\theta, 0, \omega')| \leq \epsilon AC^{m+|\alpha|+1+1} \left| L_{\text{max}}^{m+|\alpha|+(m+1)(\tau+1)+1} \kappa^{-m-1} \alpha! \beta! \rho' \kappa^{m+1} \rho' E_j^{1/2} \right|
$$

for all $\theta \in \mathbb{T}^n$, $\omega, \omega' \in \Omega_\kappa$ and $|\beta| \leq m$, where the constants $A, C$ only depend on $n, \rho, \tau, \xi$.\\

**Proof.** For ease of notation, we define

$$
M_{j, \alpha, \beta}(\theta, 0, \omega) := |W_0^{\alpha} \partial_\theta^\alpha \partial_\omega^\beta S^j(\theta, 0, \omega)| \quad (4.6.90)
$$

and

$$
L_{j, \alpha, \beta}^m(\theta, 0, \omega') := |W_0(R_m^\beta(\partial_\theta^\alpha \partial_\omega^\beta S^j))(\theta, 0, \omega')|. \quad (4.6.91)
$$

Now the Cauchy estimate for the derivative of an analytic function together with (4.6.70) yields

$$
M_{j, \alpha, \beta}(\theta, 0, \omega) \leq \frac{2^{m+|\beta|}C\alpha! \beta! \epsilon j}{r_j h_j s_{j+1}^{m+1} h_{j+1}} = \frac{2^{m+|\beta|}C(n, \rho) \kappa E_j \sigma_j^{\gamma+1}}{h_j s_{j+1}^{m+1} h_{j+1}} \quad (4.6.92)
$$

in $\tilde{D}_{j+1} \times \tilde{O}_{j+1}$.\\

Now from the definitions at the beginning of this section, we have

$$
s_j = 5(1 - \delta)^{-1} A_0^{-1} L_1^{-1}(B \sigma_j^{-1/\rho} - (\rho - 1)) \quad (4.6.93)
$$

and

$$
h_j = \frac{\kappa \sigma_j^{\gamma+1}}{2 x_j^{\gamma+1}}. \quad (4.6.94)
$$

Since (4.6.29) implies

$$
h_{j+1}^{-1} \leq C(\tau, \rho) \kappa^{-1} L_1^{\gamma+1} (B \sigma_j^{-1/\rho})^{\rho(\tau+1)}, \quad (4.6.95)
$$

§4.6 The iterative scheme
we can bound
\[ M_{j,a,b} \leq \varepsilon AC^{\alpha+|\beta|} \xi_{1}^{\alpha+|\beta|} \kappa^{-|\beta|} \alpha! \beta! (B\sigma_{j}^{-1/(\rho-1)} (\rho-1)(\alpha-\tau)+\rho(\tau+1)|\beta|+1) \exp(-B\sigma_{j}^{-1/(\rho-1)}) \] (4.6.96)

By redefining \( A \) and \( C \) and absorption into the exponential, we arrive at
\[ M_{j,a,b} \leq \varepsilon AC^{\alpha+|\beta|} \xi_{1}^{\alpha+|\beta|} \kappa^{-|\beta|} \alpha! \beta! |\beta|^{1/2} E_{j}^{1/2} \] (4.6.97)

where \( A \) and \( C \) depend only on \( n, \rho, \tau \). This is precisely the claim (4.6.88).

To prove the second claimed estimate (4.6.89), we first assume that the \( \omega, \omega' \in \Omega_{\kappa} \) satisfy \( |\omega' - \omega| < h_{j+1}/8 \).

From the elementary inequality
\[ \frac{(\beta + \gamma)!}{\gamma!} \leq 2^{\beta+\gamma} \beta! \leq 2^{\beta+\gamma} \frac{(m+1)!}{(m-|\beta|+1)!} \] (4.6.98)

together with the established bound (4.6.88), we can estimate \( L_{j,a,b}^{m} \) by Taylor expansion.

\[
L_{j,a,b}^{m} \leq \sum_{|\gamma| \geq m-|\beta|+1} |\omega' - \omega|^{\gamma} M_{j,a,b+\gamma}(\theta, 0, \omega')/\gamma!
\]
\[
= CA!(m+1)! \sum_{|\gamma| \geq m-|\beta|+1} \left( \frac{4^{\beta} \xi_{1}^{\alpha+|\beta|} \kappa^{-|\beta|} \alpha! \beta! |\beta|^{1/2}}{(m-|\beta|+1)!} \right) |\omega' - \omega|^{\gamma} E_{j}^{1/2}
\]
\[
\leq CA!(m+1)! \sum_{|\gamma| \geq m-|\beta|+1} \left( \frac{4^{\beta} \xi_{1}^{\alpha+|\beta|} \kappa^{-|\beta|} \alpha! \beta! |\beta|^{1/2}}{(m-|\beta|+1)!} \right) |\omega' - \omega|^{\gamma} E_{j}^{1/2}
\]
\[
\leq \varepsilon AC^{m+|\alpha|+1} \xi_{1}^{\alpha+(m+1)(\tau+1)+1} \kappa^{-m-1} \left( \frac{4^{\beta} \xi_{1}^{\alpha+|\beta|+1} \kappa^{-|\beta|} \alpha! \beta! |\beta|^{1/2}}{(m-|\beta|+1)!} \right) |\omega' - \omega|^{\gamma} E_{j}^{1/2}
\]

where \( A \) and \( C \) depend only on \( n, \rho \) and \( \tau \).

For \( |\omega' - \omega| \geq h_{j+1}/8 \), we can obtain the same estimate by using (4.6.88) to estimate \( L_{j,a,b}^{m} \) term by term directly from it’s definition.

The Cauchy estimate from Proposition A.1 immediately implies

**Corollary 4.6.3.**

\[ |W_{0}(\partial_{\theta}^{\alpha} \partial_{\omega}^{\beta} \partial_{t}^{\gamma} S_{i})(\theta, 0; \omega, t)| \leq \varepsilon AC^{\alpha+|\beta|+|\gamma|} \xi_{1}^{\alpha+|\beta|+|\gamma|} \kappa^{-|\beta|} \alpha! \beta! |\beta|^{1/2} E_{j}^{1/2} \] (4.6.99)

for all \( (\theta, 0; \omega, t) \in \tilde{D}_{j+1} \times \tilde{O}_{j+1} \times (-3/4, 3/4) \), where \( \rho' = \rho(\tau+1)+1 \).

\[
|W_{0}(R_{0}^{\omega} \partial_{\theta}^{\alpha} \partial_{\omega}^{\beta} \partial_{t}^{\gamma} S_{i})(\theta, 0, \omega', t)|
\]
\[
\leq \varepsilon AC^{m+|\alpha|+|\beta|+1} \xi_{1}^{\alpha+(m+1)(\tau+1)+1} \kappa^{-m-1} \left( \frac{4^{\beta} \xi_{1}^{\alpha+|\beta|+1} \kappa^{-|\beta|} \alpha! \beta! |\beta|^{1/2}}{(m-|\beta|+1)!} \right) |\omega - \omega'|^{\gamma} E_{j}^{1/2}
\]

for all \( \theta \in \mathbb{T}^{n}, \omega, \omega' \in \Omega_{\kappa}, t \in (-3/4, 3/4) \) and \( |\beta| \leq m \), where the constants \( A, C \) only depend on \( n, \rho, \tau, \zeta \).

From Proposition 4.6.1 and Lemma 4.6.3, the rapid decay of \( E_{j} \) implies that the limit
\[
\partial_{\theta}^{\alpha} \partial_{\omega}^{\beta} \partial_{t}^{\gamma} \mathbf{F}(\theta, 0; \omega, t) := \lim_{j \to \infty} \partial_{\theta}^{\alpha} \partial_{\omega}^{\beta} \partial_{t}^{\gamma} (\mathbf{F}(\theta, 0; \omega, t) - (\theta, 0, \omega))
\] (4.6.101)
exists for each \((\theta, \omega, t) \in \mathbb{T}^n \times \Omega_\kappa \times (-3/4, 3/4)\), and each triple of multi-indices \(\alpha, \beta, \gamma\). Convergence is uniform, and the limit is smooth in \(\theta\) and \(t\) and continuous in \(\omega\), with \(\partial_\theta^\alpha \partial_t^\gamma (\mathcal{H}^\beta) = \partial_\theta^\alpha \partial_t^\gamma \mathcal{H}^\beta\), justifying the notation in (4.6.101).

### 4.7 Whitney extension

We now need to use the jet \(\mathcal{H} = (\partial_\theta^\alpha \partial_t^\gamma \mathcal{H}^\beta)\) of continuous functions \(\mathbb{T}^n \times \Omega_\kappa \times (-3/4, 3/4) \to \mathbb{T}^n \times D \times \Omega\) to obtain a Gevrey function on \(\mathbb{T}^n \times \Omega \times (-3/4, 3/4)\) by using a Gevrey version of the Whitney extension theorem.

To this end, we define

\[
(R^m_\omega \partial_\theta^\alpha \partial_t^\gamma \mathcal{H})_\beta(\theta, \omega', t) := \partial_\theta^\alpha \partial_t^\gamma \mathcal{H}^\beta(\theta, \omega', t) - \sum_{|\delta| \leq m-|\beta|} (\omega' - \omega)^\delta \partial_\theta^\alpha \partial_t^\gamma \mathcal{H}^{\beta+\delta}(\theta; \omega, t)/\gamma! 
\]

(4.7.1)

In this notation, the results of Corollary 4.6.3 yield

\[
|W_0 \partial_\theta^\alpha \partial_t^\gamma \mathcal{H}^\beta(\theta; \omega, t)| \leq \epsilon AL_1(CL_1)^{|\alpha|}(CL_1^{\tau+1}/\kappa)^{|\beta|} C^\gamma \alpha! \beta! \gamma! 
\]

(4.7.2)

and

\[
|W_0(R^m_\omega \partial_\theta^\alpha \partial_t^\gamma \mathcal{H})_\beta(\theta, \omega', t)| \leq \epsilon AL_1(CL_1)^{|\alpha|}(CL_1^{\tau+1}/\kappa)^m C^\gamma \alpha! \beta! (m+1)!^\gamma! 
\]

(4.7.3)

for \(|\beta| \leq m\), and \((\theta, \omega, \omega', t) \in \mathbb{T}^n \times \Omega_\kappa \times \Omega_\kappa \times (-3/4, 3/4)\), where \(A\) and \(C\) depend only on \(n, \rho, \tau\).

These estimates allow us to apply the following consequence of Theorem 6.

**Theorem 4.7.1.** Suppose \(K \subset \mathbb{R}^n\) is compact, and \(1 \leq \rho < \rho'\). If the jet \((f^{\alpha, \beta, \gamma})\) of functions \(f^{\alpha, \beta, \gamma} : \mathbb{T}^n \times K \times (-3/4, 3/4) \to \mathbb{R}\) is continuous on \(\mathbb{T}^n \times K \times (-3/4, 3/4)\) and is smooth in \((\theta, t) \in \mathbb{T}^n \times (-3/4, 3/4)\) for each fixed \(\omega \in K\) where

\[
\partial_\theta^\alpha \partial_t^\gamma (f^{\alpha, \beta, \gamma}) = f^{\alpha+\alpha', \beta, \gamma+\gamma'} 
\]

(4.7.4)

and we have the estimates

\[
|f^{\alpha, \beta, \gamma}(\theta; \omega, t)| \leq AC_1^{\alpha} C_2^{\beta} C_3^{\gamma} \alpha! \beta! \gamma! 
\]

(4.7.5)

and

\[
|(R^m_\omega \partial_\theta^\alpha \partial_t^\gamma f)_\beta(\theta, \omega', t)| \leq AC_1^{\alpha} C_2^{m+1} C_3^{\gamma} \alpha! \beta! (m+1)!^\gamma! 
\]

(4.7.6)

then there exist positive constants \(A_0, C_0\), dependent only on \((n, \rho, \tau)\) (in particular, independent of the set \(K\)) such that we can extend \(f\) to \(\tilde{f} \in G^{\rho, \gamma, \beta}(\mathbb{T}^n \times \mathbb{R}^n \times (-3/4, 3/4))\) such that \(\partial_\theta^\alpha \partial_t^\gamma \tilde{f} = f^{\alpha, \beta, \omega}\) on \(\mathbb{T}^n \times K \times (-3/4, 3/4)\) and

\[
|\partial_\theta^\alpha \partial_t^\gamma \tilde{f}(\theta, \omega)| \leq A_0 A \max(C_1, 1) C_0^{\alpha+|\beta|+|\gamma|+n} C_1^{\alpha+n} C_2^{\beta} C_3^{\gamma} \alpha! \beta! \gamma! 
\]

(4.7.7)

The proof of Theorem 4.7.1 is an application ([36] Theorem 3.7) of a version of the Whitney extension theorem in anisotropic spaces of non quasi-analytic functions.

We are now ready to prove Theorem 4.2.2.
Proof of Theorem 4.2.2. From the estimates (4.7.2) and (4.7.3) (taking \( \gamma = 0 \)), we can apply Theorem 4.7.1 to extend the jet \( J \) to a Gevrey function

\[ \mathcal{H} = (\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3) : \mathbb{T}^n \times \mathbb{R}^n \times (-3/4, 3/4) \rightarrow \mathbb{T}^n \times \mathbb{R}^n \times \mathbb{R}^n. \]  

Moreover, we have the estimate

\[ |\overline{W}_0 \partial_0^a \partial_2^b \mathcal{H}(\theta; \omega, t)| \leq \epsilon A L_0^{n+2} (CL_1)^{|a|} (CL_1^r + 1/\kappa)^{|b|} \alpha^p \beta^q \nu \]  

where \( \epsilon = (P)_r L_1^{N-n-2} (a \kappa r)^{-1} \leq 1 \) and \( A(n, \rho, \tau), C(n, \rho, \tau), \alpha(n, \rho, \tau, \zeta, \tilde{R}) \) are positive constants.

We define

\[ \mathcal{F} = (\Phi, \phi) = (U, V, \phi) := (\mathcal{H}_1(\theta; \omega, t) + \theta, \mathcal{H}_2(\theta; \omega, t), \mathcal{H}_3(\omega, t) + \omega) \]  

Recalling that \( r_0 = cr \), where \( c(n, \rho, \tau, \zeta) \ll 1 \), and \( h_0 \leq \kappa \sigma_0^{r+1} < \kappa \), we can rewrite the Gevrey estimate (4.7.9) as

\[ |\partial_0^a \partial_2^b (U(\theta; \omega, t) - \theta)| + r^{-1} |\partial_0^a \partial_2^b V(\theta; \omega, t)| + \kappa^{-1} |\partial_2^a (\phi(\omega, t) - \omega)| \leq \frac{A(P)_r L_1^N (CL_1)^{|a|} (CL_1^r + 1/\kappa)^{|b|} \alpha^p \beta^q \nu}{\kappa r} \]  

where \( A \) and \( C \) depend only on \( n, \rho, \tau, \zeta, \tilde{R} \) and \( N \) depends only on \( n, \rho, \) and \( \tau \).

If we now choose \( \epsilon < A^{-1} \) in the statement of Theorem 4.2.2, then we obtain

\[ |V(\theta; \omega)| \leq A \epsilon r < r \leq R \]  

and

\[ \phi(\omega) \in \Omega_{\kappa} \]  

for \( \omega \in \Omega_{\kappa} \). Thus we have established the estimate (4.2.24).

It remains to check that \( \{ \Phi(\theta; \omega, t) : \theta, \omega \in \mathbb{T}^n \} \) is an embedded invariant Lagrangian torus for the Hamiltonian \( H(\theta, I; \phi(\omega, t), t) \) when \( \omega \in \Omega_{\kappa} \).

From the estimate

\[ |H_{j-1} \circ \mathcal{F}^j - N_j| \leq C_{\epsilon_j} \]  

that holds on \( D_j \times O_j \times (-3/4, 3/4) \), the Cauchy estimate yields

\[ |W_j(\text{J}(D\Phi)^T \nabla H_{j-1} \circ \mathcal{F}^j - J\nabla N_j)| \leq \frac{C_{\epsilon_j}}{r_j \sigma_j} \]  

for each \( j \geq 0 \) on

\[ \bigcap_{j\geq0} D_j \times O_j \times (-3/4, -3/4) = \{ |\text{Im}(\theta)| < s_0/2 \} \times \{ 0 \} \times \Omega_{\kappa} \times (-3/4, 3/4) \]  

where

\[ W_j := \text{diag}(\sigma_j^{-1}, r_j^{-1}). \]  

Recalling that \( J\nabla H = X_H \) and \( JMTJ = M^{-1} \) for symplectic matrices \( M \) (1.1.7), we
can rewrite the above estimate as

\[ |W_j(D\Phi^j)^{-1}X_{H_{j-1}} \circ F^j - X_N| \leq \frac{C \epsilon_j}{r_j \sigma_j}. \] (4.7.19)

From the uniform estimate (4.6.79), we then obtain

\[ |X_{H_{j-1}} \circ F^j - D\Phi^j \cdot X_N| \leq \frac{C \epsilon_j}{r_j \sigma_j} \] (4.7.20)

on \( \mathbb{T}^n \times \{0\} \times \Omega \times (-3/4, 3/4) \) where \( N = \langle \omega, I \rangle \) which differs from \( N_j \) by a term dependent only on \( \omega \).

Since \( \nabla H_j \rightarrow \nabla H \) uniformly, the rapid decay of \( \epsilon_j \) allows us to conclude that

\[ X_{H(\cdot;\phi(\omega,t),t)} \circ \Phi = D\Phi \cdot X_N. \] (4.7.21)

Hence \( \{ \Phi(\theta;\omega,t) : \theta \in \mathbb{T}^n \} \) is an embedded invariant Lagrangian torus for the Hamiltonian \( H(\theta, I; \phi(\omega,t), t) \) with frequency \( \omega \in \Omega \). This completes the proof. \( \square \)

**Remark 4.7.2.** A similar result is also obtained for real analytic Hamiltonians in [36].

### 4.8 Birkhoff normal form

We obtain a Birkhoff normal form for near-integrable Hamiltonians using a version of the KAM theorem that is a consequence of Theorem 4.2.2. The Gevrey index \( \rho(\tau + 1) + 1 \) frequently appears in these results, and so we introduce \( \rho' := \rho(\tau + 1) + 1 \).

**Theorem 4.8.1.** Fix \( 0 < \zeta \leq 1 \) and let \( H^0(I) \) be a real-valued non-degenerate \( G^\rho \) smooth Hamiltonian defined on \( D^0 \) and let \( D \) be a subdomain with \( \overline{D} \subset D^0 \). We define \( \Omega = \nabla H^0(D) \) and fix \( L_2 \geq L_1 \geq 1 \) and \( \kappa \leq L_2^{1-\zeta} \) such that \( L_2 \geq L_0 \) and \( \Omega \kappa \neq \emptyset \). Then there exists \( N = N(n, \rho, \tau) \) and \( \epsilon > 0 \) independent of \( \kappa, L_1, L_2 \) and \( D \subset D^0 \) such that for any \( H \in G^{\rho, \rho'}_{L_1, L_2, L_4}(\mathbb{T}^n \times D \times (-1,1)) \) with norm

\[ \epsilon_H := \kappa^{-2}\|H - H^0\|_{L_1, t_2, t_2} \leq \epsilon L_1^{-N} \] (4.8.1)

there exists a map

\[ \Phi = (\bar{U}, \bar{V}) \in G^{\rho, \rho'}(\mathbb{T}^n \times \Omega \times (-3/4, 3/4), \mathbb{T}^n \times D) \] (4.8.2)

such that

1. For each \( \omega \in \Omega \), and each \( t \in (-3/4, 3/4) \), \( \Lambda_\omega = \{ \Phi(\theta;\omega,t) : \theta \in \mathbb{T}^n \} \) is an embedded invariant Lagrangian torus of \( H \), and \( X_H \circ \Phi(\cdot;\omega,t) = D\Phi(\cdot;\omega,t) \cdot L_\omega. \)

2. There exist constants \( A, C > 0 \) independent of \( \kappa, L_1, L_2 \) and \( D \subset D^0 \) such that

\[ |\partial_\theta^\alpha \partial_\omega^\beta (\bar{U}(\theta;\omega,t) - \theta)| + \kappa^{-1} |\partial_\theta^\alpha \partial_\omega^\beta (\bar{V}(\theta;\omega,t) - \nabla q^0(\omega))| \leq A(C L_1)^{[\alpha]}(C L_1^{1/\kappa})^{[\beta]} \alpha^{\rho'} \beta^{\rho'} L_1^{N/2} \epsilon_H^{1/2} \] (4.8.3)

uniformly in \( \mathbb{T}^n \times \Omega \times (-3/4, 3/4) \).
Proof. We begin by choosing the $\epsilon > 0$ guaranteed by Theorem 4.2.2 sufficiently small so that $r = R := \kappa \sqrt{\epsilon_H}$ satisfies (4.2.17) for

$$\sqrt{\epsilon_H} \leq \epsilon(1 + \|P^0\|_{L_0, L_0})^{-1} L_1^{-N}$$

where $N(n, \rho, \tau)$ is as in Theorem 4.2.2.

Writing

$$H(\theta, I; \omega, t) = H^0(\psi_0(\omega)) + \langle \omega, I \rangle + P(\theta, I; \omega, t)$$

with

$$P(\theta, I; \omega, t) = \langle P^0(I; \omega)I, I \rangle + P^1(\theta, I; \omega, t)$$

as in (4.2.12), we can make use of Theorem 4.2.2.

Indeed, we have

$$\langle P \rangle_r = r^2\|P^0\|_{L_2, L_2} + \|P^1\|_{L_1, L_2, L_2}$$

and hence the assumption (4.2.21) is satisfied for the Hamiltonian $H(\theta, I; \omega, t)$.

Upon application of Theorem 4.2.2, we obtain the family of transformations

$$\Phi = (U, V) \in G^{p, \rho, \theta_1}(T^n \times \Omega \times (-3/4, 3/4), T^n \times B_R).$$

We can now define $\bar{\Phi} : T^n \times \Omega \times (-3/4, 3/4) \to T^n \times \mathbb{R}^n$ by

$$\bar{\Phi}(\theta; \omega, t) = (U(\theta; \omega, t), V(\theta; \omega, t) + (\nabla g^0)(\phi(\omega; t))).$$

By cutting off in $\omega$, we may assume that $V = 0$ outside of the domain $\Omega'$ defined in (4.2.7).

Since $\nabla g^0(\phi(\omega; t)) \in B_R$ and $R \leq \kappa/4$, we obtain that $\bar{\Phi}$ maps $T^n \times \Omega \times (-3/4, 3/4)$ into $T^n \times D$ as required.

Moreover, from Theorem 4.2.2, it follows that $\{\bar{\Phi}(\theta; \omega, t) : \theta \in T^n\}$ is an invariant Lagrangian torus for the Hamiltonian $H(\cdot, \cdot; t)$ with frequency $\omega$. The estimates for $\bar{\Phi}$ follow readily from (4.2.24).

We can now use Theorem 4.8.1 to obtain the Birkhoff normal form as done by Popov in [36].

**Theorem 4.8.2.** Suppose the assumptions of Theorem 4.8.1 hold. Then there exists $N(n, \rho, \tau) > 0$ and $\epsilon > 0$ independent of $\kappa, L_1, L_2, D$ such that for any $H \in G^{p, \rho, 1}_{L_1, L_2, L_4}(T^n \times D \times (-1, 1))$ with

$$\epsilon_H \leq \epsilon L_1^{-N-2(\tau+2)}$$

where $\epsilon_H$ is as in (4.8.1), there is a family of $G^{p, \rho, \theta}$ maps $\omega : D \times (-1/2, 1, 2) \to \Omega$ and a family of maps $\chi \in G^{p, \rho, \theta}(T^n \times D \times (-1/2, 1, 2), T^n \times D)$ that are diffeomorphisms and exact symplectic transformations respectively for each fixed $t \in (-1/2, 1, 2)$. Moreover, we can choose the maps $\omega$ and $\chi$ such that family of transformed Hamiltonians

$$\tilde{H}(\theta, I; t) := (H \circ \chi)(\theta, I; t)$$
is of Gevrey class $G^{p,q}(\mathbb{T}^n \times D \times (-1/2, 1/2))$ and can be decomposed as

$$K(I,t) + R(\theta,I;t) := \tilde{H}(0,I;t) + (\tilde{H}(\theta,I;t) - \tilde{H}(0,I;t))$$ (4.8.16)

such that:

1. $\mathbb{T}^n \times \{I\}$ is an invariant Lagrangian torus of $\tilde{H}(\cdot, t)$ for each $I \in E_{\kappa}(t) = \omega^{-1}(\tilde{\Omega}_{\kappa}; t)$ and each $t \in (-1/2, 1/2)$.

2. $\partial_\theta^\beta (\nabla K(I;t) - \omega(I;t)) = \partial_\theta^\beta R(\theta,I;t) = 0$ for all $(\theta,I;t) \in \mathbb{T}^n \times E_{\kappa}(t) \times (-1/2, 1/2), \beta \in \mathbb{N}^n$.

3. There exist $A,C > 0$ independent of $\kappa,L_1,L_2$, and $D \subset D^0$ such that we have the estimates

$$|\partial_\theta^\alpha \partial_\gamma^\beta \partial_t^\delta \phi(\theta,I;t)| + |\partial_\theta^\alpha \partial_\gamma^\beta (\omega(I;t) - \nabla H^0(I))| + |\partial_\theta^\alpha \partial_\gamma^\beta (\tilde{H}(\theta,I;t) - H^0(I))|$$

$$\leq AkC[\alpha+|\beta|+|\delta|]L_1^{|\beta|}(L_1^{k+1}/\kappa)^{|\beta|} \alpha! \beta! \delta! L_1^{N/2} \sqrt{\epsilon H}$$ (4.8.17)

uniformly in $\mathbb{T}^n \times D \times (-1/2, 1/2)$ for all $\alpha, \beta$, where $\phi \in G^{p,q}(\mathbb{T}^n \times D \times (-1/2, 1/2))$ is such that $(\theta,I) + \phi(\theta,I;t)$ generates the symplectic transformation $\chi$ in the sense of Proposition 1.1.7.

**Remark 4.8.3.** For our purposes, high regularity in the $t$-parameter is not required, so we have dropped from analyticity to $G^{p,q}$ regularity in $t$ at this point in order to simplify the proceeding arguments. I expect that analyticity in $t$ could be preserved by using a stronger variant of the Komatsu implicit function theorem than Corollary B.2

**Proof.** We begin by taking $\epsilon,N$ as in Theorem 4.8.1 and noting that $\epsilon H \leq \epsilon L_1^{-N-2}$ by assumption. This implies that the factor $(AC\kappa)L_1^{N/2} \sqrt{\epsilon H}$ occurring in the Gevrey estimate (4.8.3) can be bounded above by $AC\sqrt{\epsilon}$. Hence, taking $\epsilon$ small enough that both the conclusion to Theorem 4.8.1 holds as well as $AC\sqrt{\epsilon} < 1/2$, we can first apply the Cauchy estimate from Proposition A.1 to (4.8.3) in $t$, and then apply a variant of the Komatsu implicit function theorem, Corollary B.2, to obtain a solution $\theta(\gamma;\omega,t) : \mathbb{T}^n \times \Omega \times (-1/2, 1/2) \to \mathbb{T}^n$ to the implicit equation

$$\tilde{U}(\theta;\omega,t) = \gamma.$$ (4.8.18)

Moreover, this solution satisfies the Gevrey estimate

$$|\partial_\gamma^\alpha \partial_\omega^\beta \partial_t^\delta (\theta(\gamma;\omega,t) - \gamma)| \leq AkC[\alpha+|\beta|+|\delta|]L_1^{|\beta|}(L_1^{k+1}/\kappa)^{|\beta|} \alpha! \beta! \delta! L_1^{N/2} \sqrt{\epsilon H}$$ (4.8.19)

uniformly on $\mathbb{T}^n \times \Omega \times (-1/2, 1/2)$.

We set $F(\gamma;\omega,t) := \tilde{V}(\theta(\gamma;\omega,t);\omega,t)$. In terms of $(\gamma;\omega,t)$, the Lagrangian torus $\Lambda_\omega$ is now given by $(\gamma,F(\gamma;\omega,t) : \gamma \in \mathbb{T}^n)$ for each $\omega \in \Omega_{\kappa}$ and each $t \in (-1/2, 1/2)$. Moreover, Proposition B.4 on the composition of Gevrey functions gives us the estimate

$$|\partial_\gamma^\alpha \partial_\omega^\beta \partial_t^\delta (F(\gamma;\omega,t) - \nabla g^0(\omega))| \leq AkC[\alpha+|\beta|+|\delta|]L_1^{|\beta|}(L_1^{k+1}/\kappa)^{|\beta|} \alpha! \beta! \delta! L_1^{N/2} \sqrt{\epsilon H}.$$ (4.8.20)

We next construct functions $\psi \in G^{p,q}(\mathbb{R}^n \times \Omega \times (-1/2, 1/2))$ and $R \in G^{p,q}(\Omega \times (-1/2, 1/2))$ such that the function

$$Q(x;\omega,t) := \psi(x;\omega,t) - \langle x,R(\omega,t) \rangle$$ (4.8.21)
is 2π-periodic in $x$ and satisfies
\[ \nabla_x \psi(x; \omega, t) = F(p(x), \omega, t) \]  \hspace{1cm} (4.8.22)
in $\mathbb{R}^n \times \Omega_\kappa \times (-1/2, 1/2)$ where $p : \mathbb{R}^n \to \mathbb{T}^n$ is the canonical projection as well as the estimate
\[ |\partial_\omega^\alpha \partial_t^\beta \partial_x^\gamma Q(x; \omega, t)| + |\partial_\omega^\alpha \partial_t^\beta (R(\omega, t) - \nabla g^0(\omega))| \leq A \kappa C |\alpha| |\beta| |\gamma| L_1^{n+1} L_1^{N/2} \sqrt{\epsilon H} \]  \hspace{1cm} (4.8.23)
for $(x; \omega, t) \in \mathbb{R}^n \times \Omega \times (-1/2, 1, 2)$.

We do this by first integrating the canonical 1-form $I dx$ over the chain
\[ c_x := \{(sx, F(p(sx); \omega, t)) : 0 \leq s \leq 1\} \subset \mathbb{R}^n \times D. \]  \hspace{1cm} (4.8.25)
We define
\[ \tilde{\psi}(x; \omega, t) := \int_{c_x} \sigma = \int_0^1 (F(p(sx); \omega, t), x) \, ds \]  \hspace{1cm} (4.8.26)
in $\mathbb{R}^n \times \Omega \times (-1/2, 1, 2)$. From the estimate (4.8.20) it follows that $\tilde{\psi}(x; \omega, t) \sim \nabla g^0(\omega, x)$ is bounded above by the right hand side of (4.8.23) in $[0, 4\pi]^n \times \Omega \times (-1/2, 1, 2)$. Hence if we define $R_j(\omega, t) = (2\pi)^{-1} \tilde{\psi}(2\pi e_j; \omega, t)$, then $R - \nabla g^0$ satisfies the required estimates in (4.8.23).

Since for $\omega \in \Omega_\kappa$ we know that $\Lambda_\omega$ is a Lagrangian torus, it follows that the integral of the canonical 1-form over any closed chain in $\Lambda_\omega$ is homotopy invariant. This means that such an integral is a homomorphism from the fundamental group of $\Lambda_\omega$ to $\mathbb{R}$. Hence
\[ \tilde{\psi}(x + 2\pi m; \omega, t) \equiv \tilde{\psi}(x; \omega, t) = (2\pi m, R(\omega, t)) \]  \hspace{1cm} (4.8.27)
and so the function
\[ \tilde{Q}(x; \omega, t) := \tilde{\psi}(x, \omega) - \langle x, R(\omega, t) \rangle \]  \hspace{1cm} (4.8.28)
both satisfies the Gevrey estimate in (4.8.23) and is 2π-periodic in $x$ for $(\omega, t) \in \Omega_\kappa \times (-1/2, 1, 2)$.

To obtain the sought $Q$ in (4.8.21) from $\tilde{Q}$, we use an averaging trick. Choosing $f \in G_\kappa^0(\mathbb{R}^n)$ for some positive constant $C$ such that $f$ is supported in $[\pi/2, 7\pi/2]^n$ and
\[ \sum_{k \in \mathbb{Z}^n} f(x + 2\pi k) = 1 \]  \hspace{1cm} (4.8.29)
for each $x \in \mathbb{R}^n$, it then follows that
\[ Q(x; \omega, t) := \sum_{k \in \mathbb{Z}^n} f(x + 2\pi k) \tilde{Q}(x + 2\pi k; \omega, t) \]  \hspace{1cm} (4.8.30)
is 2π-periodic in $x$ for every $\omega \in \Omega$, and coincides with $\tilde{Q}$ for $\omega \in \Omega_\kappa$. Moreover, $Q$ satisfies the same Gevrey estimate (4.8.23) as $\tilde{Q}$. We define
\[ \psi(x; \omega, t) := Q(x; \omega, t) + \langle x, R(\omega, t) \rangle. \]  \hspace{1cm} (4.8.31)
Note that by multiplying $Q$ and $R - \nabla g^0$ by a cut-off function $h \in G_\kappa^0$ which is equal to 1 in a $\omega$-neighbourhood of $\Omega_\kappa$ and vanishes for dist($\omega, \mathbb{R}^n \setminus \Omega$) ≤ $\kappa/2$ where $C > 0$ is
independent of $\Omega \subset \Omega^0$, we can assume that $\psi(x; \omega, t) = \langle x, \nabla g^0(\omega) \rangle$ for $\text{dist}(\omega, \mathbb{R}^n \setminus \Omega) \leq \kappa/2$. This cutoff preserves the Gevrey estimates on $\psi$.

Now since $\epsilon_H L_1^{N+2(\tau+2)} \leq \epsilon$, we have that $K A(CL_1)(CL_1^{+1}/\kappa)L_1^{N/2} \sqrt{\epsilon_H} \leq A C^2 \sqrt{\tau}$. By taking $\epsilon$ sufficiently small we have that $\omega \mapsto \nabla_x \psi(x; \omega, t)$ is a diffeomorphism for any fixed $x \in \mathbb{R}^n$ from the Gevrey estimate (4.8.23). Hence we have a $G^\rho$-foliation of $\mathbb{T}^n \times D$ by Lagrangian tori $\Lambda_\omega = \{(p(x), \nabla_x \psi(x, \omega)) : x \in \mathbb{R}^n\}$ where $\omega \in \Omega$.

In the sought coordinate change, the action $I(\omega, t)$ of the Lagrangian torus $\Lambda_\omega$ will be given by $R(\omega, t)$. Hence from (4.8.23) and Proposition B.1, it follows that for $\epsilon$ sufficiently small, the map

$$(\omega, t) \mapsto (I(\omega, t), t) = (R(\omega, t), t)$$

is a $G^\rho$-diffeomorphism and we have the Gevrey estimate

$$|\partial_\theta^\alpha \partial_I^\beta (\omega(I, t) - \nabla H^0(I))| \leq A K C^{(\alpha+\beta)+[\beta]} (L_1^{+1} / \kappa)^{[\beta]} \alpha! \beta! \rho I^{N/2} \sqrt{\epsilon_H} \tag{4.8.34}$$

uniformly for $(\theta, I, t) \in \mathbb{T}^n \times D \times (-1/2, 1, 2)$.

We construct the sought symplectomorphism $\chi$ using the generating function $\Phi(x, I, t)$, setting

$$\Phi(x, I, t) = \psi(x, \omega(I, t); t)$$

and noting that we have the required $2\pi$-periodicity of $\phi(x, I, t) := \Phi(x, I, t) - \langle x, I \rangle$, and from Proposition B.4, we also have the estimate

$$|\partial_\theta^\alpha \partial_I^\beta (\Phi(x, I, t) - \langle x, I \rangle)| \leq A K C^{(\alpha+\beta)+[\beta]} (L_1^{+1} / \kappa)^{[\beta]} \alpha! \beta! \rho I^{N/2} \sqrt{\epsilon_H} \tag{4.8.36}$$

We can then apply Corollary B.2 to solve the implicit equation

$$\partial_\gamma \Phi(\gamma, I, t) = \theta$$

for $\gamma$ with the estimate

$$|\partial_\theta^\alpha \partial_I^\beta (\gamma(\theta, I, t) - \theta)| \leq A K C^{(\alpha+\beta)+[\beta]} (L_1^{+1} / \kappa)^{[\beta]} \alpha! \beta! \rho I^{N/2} \sqrt{\epsilon_H} \tag{4.8.38}$$

This completes the construction of a symplectomorphism $\chi$ satisfying

$$\chi(\partial_\gamma \Phi(\theta, I, t), I) = (\theta, \partial_\theta \Phi(\theta, I, t)).$$

It follows that

$$(\theta, F(\theta; \omega, t)) = \chi(\partial_\gamma \Phi(\theta, I(\omega), t), I(\omega)) = \chi(\theta, I(\omega), t)$$

for $\omega \in \Omega_\kappa$ and so

$$\Lambda_\omega = \{\chi(\theta, I(\omega), t) : \theta \in \mathbb{T}^n\}. \tag{4.8.41}$$

for $(\omega, t) \in \Omega_\kappa \times (-1/2, 1, 2)$.

We now set $H, K, R$ as in the theorem statement in terms of the symplectomorphism $\chi$. Since $H$ is constant on $\Lambda_\omega$ for each $\omega \in \Omega_\kappa$, it follows that $R(\cdot, I; t)$ is identically zero for each $I = I(\omega)$ with $\omega \in \Omega_\kappa$. Hence $R$ is flat at $I \in E_\kappa(t)$, since each point in $E_\kappa(t)$ is of positive density in $I(\Omega_\kappa)$.

Finally, the Gevrey estimate in (4.8.17) for $\tilde{H}(\theta, I, t) - H(I, t)$ follows from Proposition...
Remark 4.8.4. In addition to the quantum normal form we will construct in Chapter 5, Theorem 4.8.2 can also provide a short proof of effective stability of the Hamiltonian flow near $\Lambda$ (see [36] Corollary 1.3).

To conclude this chapter, we compute the integrable term $K$ of a Birkhoff normal form for the Hamiltonian $H(\theta, I) = H^0(I) + tH^1(\theta, I)$ to second order in $t$.

The key observation is that the initial KAM step leaves the completely integrable term unchanged if the perturbation has average zero over every torus $T^n \times \{I\}$ in the action-angle variables of the completely integrable Hamiltonian.

In Section 4.4.1, we proved a version of the KAM step that is localised by a frequency parameter. In this setting, the assertion is that

$$N_+(I; \omega, t) = N(I; \omega, t)$$

(4.8.42)

if

$$\int_{T^n} P(\theta, I; \omega, t) \, d\theta = 0$$

(4.8.43)

for each $I \in D$ and each $t \in (-1, 1)$. This can be seen directly from (4.4.43) for example.

An analogous fact holds for a version of the KAM step without parameters as in Section 3 of [16], and this version is most convenient for the proof of our final result in this chapter.

Using this result, we are able to change variables using two initial KAM step symplectic maps in order to increase the order in $t$ of the perturbation size before applying Theorem 5.1.1. This improves the estimate in (5.1.1).

As in the proof of Theorem 4.2.2, tracking the value of the various decreasing sequences of positive constants is necessary to prove convergence of the iterative scheme generated by this step. Fortunately, we shall only require two applications of the KAM step of Galavotti.

We first state the the result for real-analytic Hamiltonians before using the approximation techniques of Section 4.5 to generalise to the Gevrey setting.

Proposition 4.8.5. Suppose $H(\theta, I; t) = H^0(I) + H^1(\theta, I; t)$ is a real analytic Hamiltonian in $T^n \times D \times B_5$ that has an analytic extension to

$$W_{s,r}(D) := \{ (\theta, I) \in \mathbb{C}^n/2\pi \mathbb{Z} \times \mathbb{C}^n : |\text{Im}(\theta)| < s, \text{dist}(I, D) < r \}. $$

(4.8.44)

Suppose further that the conditions

$$\left| \frac{\partial H^0}{\partial I} \right| \leq E,$$  

(4.8.45)

$$\left| \left( \frac{\partial^2 H^0}{\partial I^2} \right)^{-1} \right| \leq \eta,$$  

(4.8.46)

and

$$\left( \left| \frac{\partial H^1}{\partial I} \right| + r^{-1} \left| \frac{\partial H^1}{\partial \theta} \right| \right) \leq \epsilon$$  

(4.8.47)

are satisfied.

Taking

$$V_{C,N} = \{ I \in D : |\langle \nabla H^0(I), k \rangle| \geq \frac{1}{C|k|^n} \quad \forall 0 < |k| \leq N \},$$

(4.8.48)
then for sufficiently small $\epsilon > 0$ we consider the set
\[
\tilde{D}_+ = \{ I \in D : \text{dist}(I, \partial D) > \tilde{r}/2 \text{ and } I \in V_{C,N} \}
\]
and the smoother set
\[
D_+ := \bigcup_{I \in \tilde{D}_+} B(I, \tilde{r}/2)
\]
as a new action domain, where $N$ is chosen in a way that depends on $C, \epsilon$ and $s$ and $\tilde{r}$ is proportional to $r$ and is otherwise dependent on $n, E, C, n$ with $\tilde{r} < r/2$.

There exists a family of real analytic symplectic maps
\[
\chi : \mathbb{T}^n \times D_+ \times B_5 \to \mathbb{T}^n \times D
\]
that extend analytically to a new domain of holomorphy
\[
W_{s_+,r_+}(D_+)
\]
such that
\[
(H \circ \chi)(\theta, I; t) = H^0(I) + (2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, I; t) \, d\theta + H^2(\theta, I; t).
\]
with
\[
\sup_{D_{s_+,r_+}} |H^2| = O(\epsilon^{3/2})
\]
with constant depending only on $n, C$ and $E$.

This result is essentially a restatement of the KAM step in Section 3 of [16], in which more explicit details are given. We remark that the constant $C$ describes the family of nonresonant frequencies preserved by the iterative scheme, with $V_{C,\infty}$ having large measure for large $C$.

We now consider the family of real analytic Hamiltonians of the form
\[
H(\theta, I; t) := H^0(I) + t H^1(\theta, I)
\]
still satisfying the hypotheses of Proposition 4.8.5.

Applying Proposition 4.8.5, we obtain
\[
(H \circ \chi)(\theta, I; t) = H^0(I) + t \cdot (2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, I) \, d\theta + H^2(\theta, I; t).
\]
We will require a stronger estimate on the error term $H^2(\theta, I; t)$ that $O(\epsilon^{3/2})$, due to the presence of a square root in (4.8.17). Hence we need another iteration of Proposition 4.8.5. A second iteration of Proposition 4.8.5 yields a family of symplectic maps $\tilde{\chi} : \mathbb{T}^n \times D_{++} \to \mathbb{T}^n \times D$ that extend analytically to $D_{s_+,r_+}$ and such that
\[
(H \circ \tilde{\chi})(\theta, I; t) = H^0(I) + t \cdot (2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, I) \, d\theta + (2\pi)^{-n} \int_{\mathbb{T}^n} H^2(\theta, I; t) \, d\theta + H^3(\theta, I; t)
\]
with $|H^3| = O(t^{9/4})$.

This implies that the completely integrable part of $(H \circ \tilde{\chi})(0, I; t)$ is given by $H^0(I) + t \cdot (2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, I) \, d\theta + O(t^{3/2})$, valid in the domain $\mathbb{T}^n \times D_{++}$, where $D_{++}$ is an open
subset of $D$ given by

\[ D_{++} := \bigcup_{I \in \tilde{D}_{++}} B(I, \tilde{r}_+/2) \]  

(4.8.58)

where

\[ \tilde{D}_{++} = \{ I \in D_+ : \text{dist}(I, \partial D_+) > \tilde{r}_+/2 \text{ and } I \in V_{C_+, N_+} \} \]  

(4.8.59)

where $C_+$ and $N_+$ are defined inductively in [16]. In particular, we can ensure that the set $E_\kappa(t) = \omega_{\kappa} - 1(\tilde{\Omega}_\kappa; t)$ is contained in $D_{++}$ by fixing $\tau = n$, taking $C$ in Proposition 4.8.5 sufficiently large. Thus we have

**Proposition 4.8.6.** Suppose $H(\theta, I; t) = H_0(I) + tH_1(\theta, I)$ is a real analytic Hamiltonian in $T^n \times D \times B_\delta$ that has an analytic extension to

\[ W_{s, r}(D) := \{ (\theta, I) \in C^n/2\pi\mathbb{Z} \times C^n : |\text{Im}(\theta)| < s, \text{dist}(I, D) < r \}. \]  

(4.8.60)

Suppose further that the conditions

\[ \left| \frac{\partial H_0}{\partial I} \right| \leq E, \]  

(4.8.61)

\[ \left| \left( \frac{\partial^2 H_0}{\partial I^2} \right)^{-1} \right| \leq \eta, \]  

(4.8.62)

and

\[ \left( \left| \frac{\partial H_1}{\partial I} \right| + r^{-1} \left| \frac{\partial H_1}{\partial \theta} \right| \right) \leq \epsilon \]  

(4.8.63)

are satisfied.

Then for sufficiently small $t \in (-\delta, \delta)$, there exists a subdomain $\tilde{D} \subset D$ containing $E_\kappa(t)$ and a family of real analytic symplectic maps

\[ \chi : T^n \times \tilde{D} \times (-\delta, \delta) \to T^n \times D \]  

(4.8.64)

that analytically extend to a new domain of holomorphy

\[ W_{s_+, \tilde{r}_+}(\tilde{D}) \]  

(4.8.65)

such that

\[ (H \circ \chi)(\theta, I; t) = \tilde{H}_0(I; t) + \tilde{H}_1(\theta, I; t). \]  

(4.8.66)

with

\[ \partial_t \tilde{H}_0(I; 0) = (2\pi)^{-n} \int_{T^n} H_1(\theta, I) \, d\theta \]  

(4.8.67)

and

\[ \sup_{D_{s_+, \tilde{r}_+}} |\tilde{H}_1| = O(t^{9/4}) \]  

(4.8.68)

with constant depending only on $n, C$ and $E$.

We can now generalise this result to the Gevrey setting.

**Proposition 4.8.7.** Suppose $H(\theta, I; t) = H_0(I) + tH_1(\theta, I)$ satisfies the assumptions of Theorem 4.8.1. Then for sufficiently small $t \in (-\delta, \delta)$, there exists a subdomain $\tilde{D} \subset D$
containing \( E_\kappa(t) \) and a \( G^{\rho,\rho} \) family of symplectic maps

\[
\chi : \mathbb{T}^n \times \tilde{D} \times (-\delta, \delta) \to \mathbb{T}^n \times D
\]

such that

\[
(H \circ \chi)(\theta, I; t) = \tilde{H}^0(I; t) + \tilde{H}^1(\theta, I; t).
\]

with

\[
\partial_t \tilde{H}^0(I; 0) = (2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, I) \, d\theta
\]

and

\[
\|\tilde{H}^1\|_{CL_1,CL_2,CL_2} = O(t^{\rho/4})
\]

with constant independent of \( \kappa \) and with \( C \) dependent only on \( n \) and \( \rho \).

**Proof.** This result is established via the approximation of Gevrey functions by real-analytic functions. First, we use Theorem 4.7.1 to extend \( H^0 \) and \( H^1 \) to the domain \( \mathbb{T}^n \times \mathbb{R}^n \times (-1, 1) \), before cutting off in \( I \) to a ball \( B_R \) with \( D^0 \subset B_{R-1} \) as done at the start of Section 4.5.

From the same methods used in the proof of Proposition 4.5.1, we may then construct sequences of real analytic functions \( P_j^0 \) and \( P_j^1 \) on shrinking \( j \) dependent complex domains \( U_j \) containing \( \mathbb{T}^n \times \mathbb{R}^n \times (-1, 1) \) with a corresponding sequence \( u_j \to 0 \) such that

\[
|P_{j+1}^k - P_j^k|_{U_{j+1}} \leq C(D^0, L_1, L_2) \exp \left( -\frac{3}{4}(\rho - 1)(2L_1 u_j)^{-1/(\rho-1)} \right) \|H^k\|
\]

and

\[
|\partial_x^\alpha (P_j^k - H^k)(\theta, I; t)| \leq C(D^0, L_1, L_2) \exp \left( -\frac{3}{4}(\rho - 1)(2L_1 u_j)^{-1/(\rho-1)} \right)
\]

in \( \mathbb{T}^n \times B_R \times (-1, 1) \) for \( |\alpha| \leq 1 \). These sequences \( P_j^k \) are convergent in \( G^{\rho,\rho,1}(\mathbb{T}^n \times \mathbb{R}^n \times (-1, 1)) \), as is shown in [26] Proposition 2.2. (This fact can be readily obtained by applying Cauchy estimates to (4.8.73)).

Now for each \( j \in \mathbb{N} \), we can apply Proposition 4.8.51 to obtain a real analytic symplectic map

\[
\chi_j : \mathbb{T}^n \times D_+ \to \mathbb{T}^n \times D_+
\]

defined in shrinking holomorphy domains that comprise the first KAM step for \( H_j^0 + H_j^1 \).

Note that for an individual KAM step, the symplectic map \( \chi_j \) is defined using a generating function \( \Phi_j \) that is a weighted sum of finitely many Fourier components of \( H_j^1 \) (see [16] Equation 3.14).

This implies that as \( P_j^0 + tP_j^1 \to H^0 + tH^1 \) in \( G^{\rho,\rho,1}(\mathbb{T}^n \times D_+ \times (-1, 1)) \), the generating functions \( \Phi_j \) converges to some

\[
\Phi \in G^{\rho,\rho,1}(\mathbb{T}^n \times D_+ \times (-1, 1))
\]

in the \( G^{\rho,\rho,1} \) sense.

From the Komatsu implicit function theorem, Corollary B.2, it follows that the corresponding symplectic maps \( \chi_j \) converge to some

\[
\chi^1 \in G^{\rho,\rho,1}(\mathbb{T}^n \times D_+ \times (-1, 1))
\]
in the Gevrey sense.

Similarly, the symplectic maps \( \tilde{\chi}_j \) that arise from comprise a single KAM step for the Hamiltonians
\[
(P^0_j + tP^1_j) \circ \chi_j
\]
can also be seen to converge to some
\[
\chi^2 \in G^{\rho,\rho,1}(\mathbb{T}^n \times D_{++}, \mathbb{T}^n \times D_+).
\]

It follows that the family of symplectic maps \( \chi_j \circ \tilde{\chi}_j \) whose existence is asserted by applying Proposition 4.8.6 to \( H^0_j + H^1_j \) converge to \( \chi := \chi^1 \circ \chi^2 \) in the \( G^{\rho,\rho,1} \)-sense.

Moreover, if we write
\[
(P^0_j + tP^1_j) \circ \chi_j \circ \tilde{\chi}_j = \tilde{H}^0_j(I; t) + \tilde{H}^1_j(\theta, I; t).
\]
in the notation of Proposition 4.8.6, we have that \( \tilde{H}^0_j, \tilde{H}^1_j \) are convergent sequences in \( G^{\rho,\rho,1} \), and so it follows that their limits \( \tilde{H}^0, \tilde{H}^1 \) satisfy
\[
\partial_t \tilde{H}^0(I; 0) = (2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, I) d\theta
\]
and
\[
\| \tilde{H}^1 \|_{CL_1, CL_2, CL_2} = O(t^{9/4})
\]
as required. \( \square \)

By applying Proposition 4.8.7 to \( H(\theta, I; t) = H^0(I) + tH^1(\theta, I) \) with \( t \) sufficiently small, we can then invoke Theorem 4.8.2 on the Hamiltonian
\[
\tilde{H}(\theta, I; t) = (H \circ \chi)(\theta, I; t)
\]
with an improved error term.

**Proposition 4.8.8.** Suppose the assumptions of Theorem 4.8.1 hold for the Hamiltonian
\[
H(\theta, I; t) = H^0(I) + tH^1(\theta, I; t) \in G^{\rho,\rho,1}(\mathbb{T}^n \times D \times (-1, 1)).
\]

Then there exists \( N(n, \rho, \tau) > 0 \) and \( \epsilon > 0 \) independent of \( L_1, L_2, D \) such that for any \( H \in G^{\rho,\rho,1}_{L_1, L_2, L_2}(\mathbb{T}^n \times D \times (-1, 1)) \) with
\[
\kappa^{-2t} \| H^1 \|_{L_1, L_2, L_2} = \epsilon_H \leq \epsilon L_1^{-N-2(\tau+2)}
\]
there is a subdomain \( \tilde{D} \subset D \) containing \( E_n(0) \) and a family of \( G^{\rho,\rho} \) maps \( \omega : \tilde{D} \times (-1/2, 1, 2) \to \Omega \) and a family of maps \( \chi \in G^{\rho,\rho,\rho}(\mathbb{T}^n \times \tilde{D} \times (-1/2, 1, 2), \mathbb{T}^n \times \tilde{D}) \) that are diffeomorphisms and exact symplectic transformations respectively for each fixed \( t \in (-1/2, 1, 2) \). Moreover, we can choose the maps \( \omega \) and \( \chi \) such that family of transformed Hamiltonians
\[
\tilde{H}(\theta, I; t) := (H \circ \chi)(\theta, I; t)
\]
is of Gevrey class \( G^{\rho,\rho,\rho}(\mathbb{T}^n \times \tilde{D} \times (-1/2, 1, 2)) \) and can be decomposed as
\[
K(I; t) + R(\theta, I; t) := \tilde{H}(0, I; t) + (\tilde{H}(\theta, I; t) - \tilde{H}(0, I; t))
\]
such that:

1. \( T^n \times \{ I \} \) is an invariant Lagrangian torus of \( \tilde{H}(\cdot, \cdot; t) \) for each \( I \in E_n(t) = \omega^{-1}(\tilde{\Omega}_\kappa) \) and each \( t \in (-1/2, 1, 2) \).

2. \( \partial_I^3 (\nabla K(I; t) - \omega(I; t)) = \partial_I^3 R(\theta, I; t) = 0 \) for all \((\theta, I; t) \in T^n \times E_n(t) \times (-1/2, 1, 2), \beta \in \mathbb{N}^n \).

3. There exist \( A, C > 0 \) independent of \( \kappa, L_1, L_2, \) and \( D \subset D^0 \) such that we have the estimates

\[
|\partial_\theta^\alpha \partial_I^\beta \partial_t^\delta \phi(\theta, I; t)| + |\partial_\theta^\alpha \partial_I^\beta (\omega(I; t) - \nabla \tilde{H}^0(I))| + |\partial_\theta^\alpha \partial_I^\beta (\tilde{H}(\theta, I; t) - \tilde{H}^0(I))| \leq AC|\alpha| + |\beta| + |\delta| L_1^{\alpha + |\beta| + |\delta|} (L_1^{-1}/\kappa)^{1/2} L_1^{1/2} |t|^{9/8}
\]

uniformly in \( T^n \times \tilde{D} \times (-1/2, 1, 2) \) for all \( \alpha, \beta \), where \( \phi \in G^\rho, \rho' \) \((T^n \times \tilde{D} \times (-1/2, 1, 2)) \) is such that \((\theta, I) + \phi(\theta, I; t)\) generates the symplectic transformation \( \chi \) in the sense of Proposition 1.1.7 and \( \tilde{H}^0, \tilde{H}^1 \) are as in Proposition 4.8.7.

4. \[
\partial_t K(I; t) = (2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, I; t) + o(1)
\]

uniformly in \( T^n \times \tilde{D} \times (-1/2, 1, 2) \).

**Proof.** The only new claim in this Proposition is (4.8.89), which follows from (4.8.88) and the expression (4.8.71) for \( \tilde{H}^0 \). Note that the exponent 9/4 in (4.8.88) comes from (4.8.72) and the square root in (4.8.17). \( \square \)
Quantum Birkhoff normal form in KAM systems

In this chapter, we continue from the construction of a Birkhoff normal form in Chapter 4 and construct a quantum Birkhoff normal form for the quantisations of Gevrey KAM Hamiltonians, following the work of Popov in [37]. The main consequence of this construction for our purposes is the construction of exponentially accurate quasimodes localising onto the invariant KAM tori.

5.1 Quantum Birkhoff normal form

We let $M$ be a compact $G^\rho$-smooth manifold of dimension $n \geq 2$ and consider the family of formally self-adjoint Schrödinger type semiclassical pseudodifferential operator

$$P_h := h^2 \Delta_g + V(x, hD)$$  \hspace{1cm} (5.1.1)

acting on half densities $f|dx|^{1/2}$ where each $V$ is a self adjoint Gevrey symbol in the class $S_\ell(T^*M)$ from Definition 2.2.5 where $\ell = (\rho, \mu, \eta)$ with $\rho(\tau+n)+1 > \mu > \rho' = \rho(\tau+1)+1$ and $\nu = \rho(\tau+n+1)$.

The paradigmal example here is the semiclassical Schrödinger operator

$$P_h = h^2 \Delta + V(x)$$  \hspace{1cm} (5.1.2)

for a smooth and compactly supported potential $V(x) \in G^\rho$ that is bounded below.

We suppose further that there exists an exact symplectomorphism $\chi_1 : \mathbb{T}^n \times D \to U \subset \mathbb{T}^*M$ such that the transformed Hamiltonian $H(\phi, I) = P_0 \circ \chi_1$ can be placed in a $G^\rho_{\phi'}$ Birkhoff normal form (4.8.16) about a family of invariant tori $\Lambda := \{ \Lambda_\omega : \omega \in \Omega \}$ with $\Omega \in \Omega_n$.

From Theorem 4.8.2, we have shown that this is the case if $P_0 = H^0 + tH^1$ is an analytic one-parameter family of small perturbations of a completely integrable and non-degenerate Hamiltonian $H^0$.

From this point on, we specialise from (5.1.1) to considering one-parameter families of operators of the form

$$P_h(t) := H^0(x, hD) + tH^1(x, hD)$$  \hspace{1cm} (5.1.3)

where the Hamiltonian $H = H^0 + tH^1$ satisfies the assumptions of Theorem 4.8.2.

In this case, the family of maps $\chi_1(t)$ can be taken to be the transformation into “action-angle" coordinates, the existence of which is guaranteed by the Liouville–Arnold
quantum analogue to the Birkhoff normal form of the Hamiltonian $H$ as the quantisation of the symplectomorphism that transforms $H$ into its Birkhoff normal form.

As a consequence of Theorem 5.1.1, we will obtain a family of Gevrey quasimodes microlocalised (in the sense of the Gevrey microsupport introduced in Definition 2.2.8) onto a family $\Lambda$ of the nonresonant invariant Lagrangian tori constructed in Chapter 4.

We sketch the details of the proof of Theorem 5.1.1 in this chapter, following the argument of Popov [37].

The construction of $U_h(t)$ can be broken into multiple steps. We begin by constructing a family of semiclassical Fourier integral operators $T_h$ that conjugate $P_h(t)$ to a family of pseudodifferential operators $P_h^0(t) : C^\infty(\mathbb{T}^n;L)$ with principal symbol equal to $K_0(I;t) + R_0(\theta,I;t)$, the Birkhoff normal form of $H$, and vanishing subprincipal symbol. This arises as the quantisation of the symplectomorphism that transforms $H$ into its Birkhoff normal form.

The symbol of the operator $P_h^0(t)$ satisfies the property (5.1.5) to $O(h^2)$, and to improve this, we replace the conjugating Fourier integral operator $T_h$ with $T_h A_h$ for a suitable elliptic pseudodifferential operator $A_h$ whose symbol is determined iteratively on the family of Cantor-like set $\{(\theta,I;t) \in \mathbb{T}^n \times \mathbb{R}^n \times (-1,1) : I \in E_\kappa(t)\}$ by solving equations of the form

$$\langle \nabla K_0, \partial_\theta \rangle f(\theta,I;t) = g(\theta,I;t)$$  

(5.1.7)
referred to in the literature as homological equations. In this manner the “flatness condition” of (5.1.6) is obtained for $j > 0$, where the $j = 0$ statement was established by Theorem 4.8.2.

5.2 Conjugation by a h-FIO

In this section we first conjugate $P_h(t)$ by an Fourier integral operator to a semiclassical pseudodifferential operator on $\mathbb{T}^n \times D$.

We do this by quantising the $G^\rho$ symplectic maps $\chi_1 : \mathbb{T}^n \times D \to T^*M$ and $\chi_0 : \mathbb{T}^n \times D \to \mathbb{T}^n \times D$ that transform the unperturbed Hamiltonian $H^0$ to action-angle variables and transform the perturbed Hamiltonian to Birkhoff normal form respectively.

We define

$$C_1 = \{ (\chi_1(y, \eta), y, \eta) : (y, \eta) \in \mathbb{T}^n \times D \}$$

and the flipped graph

$$C_1' = \{ (x, y, \xi, \eta) : (x, \xi) = \chi_1(y, -\eta) \}$$

which is a Lagrangian submanifold of $T^*M$.

Because $\chi_1$ is exact-symplectic, we can quantise this map as a semiclassical Fourier integral operator which is a semiclassical Lagrangian distribution associated with $C_1'$ [35] [13].

In this construction, the Maslov class of the tori $\{ \Lambda_\omega : \omega \in \Omega_\kappa \}$ (as defined in Section 3.4 of [14]) can be identified with elements of $\vartheta \in H^1(T^n; \mathbb{Z})$ via the symplectic map $\chi_0 \circ \chi_1 : \mathbb{T}^n \times D \to T^*M$.

Following [35] and [10], we can then associate a smooth line bundle $L$ over $\mathbb{T}^n$ with the class $\vartheta$, such that smooth sections $f \in C^\infty(\mathbb{T}^n, L)$ can be canonically identified with smooth functions $\tilde{f} \in C^\infty(\mathbb{R}^n, \mathbb{C})$ satisfying the quasiperiodicity condition

$$\tilde{f}(x + 2\pi p) = \exp \left( \frac{i\pi}{2} \langle \vartheta, p \rangle \right) \tilde{f}(x)$$

for all $p \in \mathbb{Z}^n$.

We now construct this semiclassical Fourier integral operator microlocally. First we need to parametrise $C_1'$ locally by phase functions. To this end, we consider a fixed $\zeta_0 = (x_0, y_0, \xi_0, \eta_0) \in C_1' \subset T^*(M \times \mathbb{T}^n)$. We fix an analytic chart $U_0$ about $x$. Then the implicit function theorem shows that there exists a unique $\phi \in C^\omega(U_1, U_2)$, where $U_1$ is a local chart in the torus and $U_1 \times U_2$ is a rectangular neighbourhood of $(y_0, \xi_0)$ such that

$$C_1 = \{ (\partial_{\xi} \phi, \xi, y, \partial_{\eta} \phi) \}$$

and

$$\det (\partial^2_{y\xi} \phi) \neq 0$$

and

$$\phi(y_0, \xi_0) = x_0 \cdot \xi_0 - f(\zeta_0)$$

where $f$ is the function on $C_1'$ such that $df = i^* \alpha$, where $i : C_1' \to T^*(M \times \mathbb{T}^n)$ is the inclusion and $\alpha$ is he canonical one form.
Then $\Psi(x, y, \xi) = x \cdot \xi - \phi(y, \xi)$ parametrises $C'_1$ locally in the sense that on the set $O_\Psi = \{(x, y, \xi) : \partial_\xi \Psi = 0\}$, we have

- $\text{rank}(\partial_{(x,y,\xi)} \partial_\xi \Psi) = n$
- $(x, y, \xi) \mapsto (x, y, \partial_x \Psi, \partial_y \Psi)$ is a local diffeomorphism.

Additionally, we have $\Psi(x_0, y_0, \xi_0) = f(\zeta_0)$.

We are now ready to define semiclassical Fourier integral operators associated to $C'_1$ mapping

$$C^\infty(T^n, \Omega^{1/2} \otimes \mathbb{L}) \to C^\infty_0(M, \Omega^{1/2}). \quad (5.2.7)$$

We fix $\sigma > 1, \ell = (\sigma, \sigma, 2\sigma - 1)$ and choose $a \in S_\ell((U_0 \times U_1) \times U_2)$ as in Definition 2.2.5.

Motivated by equation (5.2.3), we extend $a$ to

$$\tilde{a}(x, y + 2\pi p, \xi; h) = e^{-\frac{i}{2\pi} \theta \cdot p} a(x, y, \xi; h) \text{ for } (x, y, \xi) \in U \times \mathbb{R}^n \quad (5.2.8)$$

and extend $\Psi$ periodically to $U_0 \times (U_1 + 2\pi \mathbb{Z}^n) \times U_2$.

Then given a section $a \in C^\infty(T^n, \mathbb{L})$, we define

$$T_h u(h) := (2\pi h)^{-n} \int_{\mathbb{R}^n} \int_{U_1} e^{i\Psi (x,y,\xi)/h} \tilde{a}(x, y, \xi; h) \tilde{u}(y) \, dy \, d\xi \quad (5.2.9)$$

noting the invariance of this integral under the translations

$$U_1 \mapsto U_1 + 2\pi p \text{ where } p \in \mathbb{Z}^n. \quad (5.2.10)$$

The class of semiclassical Fourier integral operators is then the set of finite sums of operators given microlocally by 5.2.9. As in the standard theory of Hörmander [24] [14], the above definition is independent of our choices of parametrising phase functions.

The principal symbol of $T_h$ is $e^{i\varphi(\zeta)} \Upsilon(\zeta)$, where $\Upsilon$ is a smooth section in

$$\Omega^{1/2}(C'_1) \otimes M_{C'_1} \otimes \pi^*_2(L') \quad (5.2.11)$$

where $\pi_2$ is the canonical projection onto the torus.

We can trivialise the half-density bundle by pulling back the canonical half-density on $T^n \times D$ by the corresponding canonical projection, and we can canonically identify $\pi^*_2(L')$ with the dual $M'_{C'_1}$ of the Maslov bundle as is done in [10].

This allows us to canonically identify the principal symbol for a semiclassical Fourier interal operator in this setting with a function $b \in C^\infty(C'_1)$.

Additionally, we have

$$a_0(\partial_\xi \phi(y, \xi), y, \xi, -\partial_y \phi(y, \xi)) = a_0(\partial_\xi' (y, \xi), y, \xi)|\det(\partial_{\xi y}^2 \phi)|^{-1/2} \quad (5.2.12)$$

where $a_0$ is the leading term in the amplitude corresponding to a given microlocal
Theorem 5.2.1. We can choose a semiclassical Fourier integral operator $\tilde{T}$ of the pullback of the invariant tori of $\mathcal{T}$ and all of the involved pseudodifferential operators have symbols in $S^1$. Suppose

1. $Q_h = T^*_1T_{1h}$ is a semiclassical pseudodifferential operator in $C^\infty(T^n, \mathbb{L})$;
2. $Q_h$ has vanishing subprincipal symbol;
3. $P^1_h = T^*_1P_hT_{1h}$ is a semiclassical pseudodifferential operator in $C^\infty(T^n, \mathbb{L})$;
4. $T_{1h}$ is microlocally invertible in a neighbourhood of the pullback of the union of invariant tori $\Lambda \subset T^*M$ by $\chi_1$, with $P^1_h = T^{-1}_{1h}P_hT_{1h} + h^2R_h$ where $R_h$ is a semiclassical pseudodifferential operator in $C^\infty(T^n, \mathbb{L})$;
5. The principal symbol of $P^1_h$ is equal to $H \circ \chi_1$ and its subprincipal symbol vanishes; and all of the involved pseudodifferential operators have symbols in $S^\nu(T^n \times D)$.

Theorem 5.2.2. Suppose $T_{2h}(t) : L^2(T^n, \mathbb{L}) \to L^2(T^n, \mathbb{L})$ is a semiclassical Fourier integral operator associated to the canonical relation graph of the Birkhoff normal form transformation $\chi(t) : (y, I) \mapsto (x, \xi)$ with kernel given by

$$(2\pi h)^{-n} \int e^{i((x-y) - \phi(x, I; t))/h}b(x, I; h) dI$$

where

1. $\phi(x, I; t) = \Phi(x, I; t) - \langle x, I \rangle$;
2. $\Phi \in C^{1,0,s}(T^n \times D \times (-1, 1))$ is a generating function of the Birkhoff normal form canonical transformation;
3. $b \in S^\nu_\ell(T^n \times D)$ where $\ell = (\sigma, \mu, \sigma + \mu - 1)$ and $\mu > s = \tau' + 1 > \sigma > 1$;
4. The principal symbol is equal to 1 in a neighbourhood of $T^n \times D$.

Then $T_{2h}(t)^*T^*_1P_hT_{1h}T_{2h}(t)$ is a family of semiclassical pseudodifferential operator with symbol in $S^\nu_\ell$ smoothly varying in the parameter $t$. The principal symbol of $\tilde{P}(t)$ is $H(t) = H(t) \circ \chi(t)$ and its subprincipal symbol vanishes.

Moreover, the conjugating operator $T_h := T_{1h}T_{2h}(t)$ is microlocally invertible in a neighbourhood of the pullback of the family $\Lambda$ of invariant tori by $\chi_1 \circ \chi(t) : T^n \times D \to T^*M$, and we have

$$T^*_h = T^{-1}_h + O(h^2)$$

for this microlocal inverse.

Proof of these two results can be found in [35] (Section 1). In light of Theorem 5.2.2, we make the following definition.

$$\tilde{P}_h(t) = T^*_1(t)P_h(t)T_{1h}(t).$$
5.3 Conjugation by an elliptic h-PDO

From Section 5.2, we have conjugated \( P_h(t) \) to a family of self-adjoint semiclassical operators \( \tilde{P}_h(t) \) with symbol \( \tilde{p} \in S_\ell(T^n \times D) \) satisfying the flatness condition (5.1.6) to order \( h^2 \), where \( \ell = (\rho, \rho', \rho + \rho' - 1) \). That is to say, the formal summation of \( \tilde{p} \)

\[
\sum_{j=0}^{\infty} \tilde{p}_j(\theta, I; t) h^j
\]

satisfies

\[
\tilde{p}_0(\theta, I; t) = K_0(I; t) + R_0(\theta, I; t)
\]

and

\[
\tilde{p}_1(\theta, I; t) = 0.
\]

The next step of the proof of Theorem 5.1.1 is the improvement of the order of the flatness condition by composition with a suitable elliptic semiclassical pseudodifferential operator \( A_h(t) = \text{Id} + O(h) \) with symbol \( a(t) = \infty \sum_{j=1} a_j(t) \).

\[
a(t) = \sum_{j=1}^{\infty} a_j(t).
\]

To motivate the method, we suppose that a quantum Birkhoff normal form \( P_0^h \) exists in the sense of Theorem 5.1.1. Our current operator \( \tilde{P}_h \) is equal to \( P_0^h \) up to order \( h^2 \) by construction. Hence, we have

\[
T_h(t) A_h(t) \tilde{P}_h(t) = T_h(t) \tilde{P}_h(t) A_h(t) + T_h(t) [A_h(t), \tilde{P}_h(t)]
\]

for some semiclassical pseudodifferential operator \( B_h(t) \) in the symbol class \( S_\ell(T^n \times D) \). From (2.1.12), the symbol of the commutator is equal to

\[
- \left( \partial_\theta^6 a_1 \partial_I^2 \tilde{p}_0 \right) h^2 = - \mathcal{L}_\omega(\theta, I; t) a_1
\]

where \( \mathcal{L}_\omega = \langle \omega, \partial_\theta \rangle a_1(\theta, I; t) \). Thus to improve the order of the flatness condition, it suffices to choose \( a_1 \) solving the homological equation

\[
\mathcal{L}_\omega(\theta, I; t) a_1 = b_0
\]

where \( b_0 \) denotes the principal symbol of \( B_h(t) \). Indeed, if (5.3.8) is solvable, then we have

\[
T_h(t) A_h(t) P_h(t) = P_0^h(t) T_h(t) A_h(t) + O(h^3).
\]

Extending this idea, it is shown by Popov [35] that we can choose higher order terms of the symbol \( a \) in an iterative fashion by the solution of such a homological equation for each power of \( h \) that we gain. The consequence is the following theorem.

**Theorem 5.3.1.** There exists \( a, K^0, r \in S_\ell(T^n \times D) \) where \( \ell = (\rho, \mu, \nu) \) such that

\[
a(\theta, I; t, h) \sim \sum_{j=0}^{\infty} a_j(\theta, I; t) h^j
\]
\[ K^0(I; t, h) \sim \sum_{j=0}^{\infty} K_j(I; t) h^j \]  

(5.3.11)

and

\[ r(\theta, I; t, h) \sim \sum_{j=0}^{\infty} r_j(\theta, I; t) h^j \]  

(5.3.12)

where \( a_0 = 1, r_0 = R_0, K_1 = 0, \) and

\[ \tilde{p} \circ a - a \circ K^0 \sim r. \]  

(5.3.13)

where each \( r_j(\theta, I; t) \) is flat in \( I \) on \( \mathbb{T}^n \times E_\kappa(t) \).

The symbol \( K^0 \) in the statement of theorem corresponds to the sought symbol \( K^0 \) in Theorem 5.1.1, while the symbol \( R^0 \) is then constructed by solving \( a \circ R^0 = r \), which is possible by ellipticity.

We shall detail the technicalities of the solution of the homological equation in Gevrey classes in Section 5.4, and consequently outline the full construction of the full symbol of \( a \). For now, we are in a position to explain how Theorem 5.1.1 follows from Theorem 5.3.1 and Theorem 5.2.2.

**Proof of Theorem 5.1.1.** From the definition (5.2.15) of \( \tilde{P}_h(t) \), it follows that

\[ P_h(t)T_h(t)A_h(t) = T_h(t)T_h^{-1}(t)P_h(t)T_h(t)A_h(t) \]  

(5.3.14)

\[ = T_h(t)\tilde{P}_h(t)A_h(t) \]  

(5.3.15)

\[ = T_h(t)A_h(t)(K^0_h(t) + R^0_h(t)) + O(h^\infty). \]  

(5.3.16)

Thus, for

\[ V_h(t) := T_h(t)A_h(t), \]  

(5.3.17)

\( V_h(t) \) satisfies all of the desired properties of \( U_h(t) \) in Theorem 5.1.1, besides the condition of being microlocally unitary.

This \( V_h \) will not in general be unitary, so we define the semiclassical pseudodifferential operator \( W_h = V_h^*V_h \) with formal symbol

\[ \sum_{j=0}^{\infty} w_j(\theta, I) h^j \in FS_\ell(\mathbb{T}^n \times D). \]  

(5.3.18)

Then \( w_0 = 1 \) and we have

**Lemma 5.3.2.** For each \( j \), \( p^0_j(I) \) is real valued on \( E_\kappa(t) \) and \( w_j(\theta, I) \) is \( \theta \)-independent for \( I \in E_\kappa(t) \).

A proof of this lemma is contained in [35].

If we now define \( Q_h(t) = W_h^{-1/2}(t) \), Lemma 5.3.2 implies that its symbol \( q \in S_\ell(\mathbb{T}^n \times D) \) satisfies

\[ q(\theta, I; t, h) = q(0, I; t, h) \]  

(5.3.19)

to infinite order at \( \mathbb{T}^n \times E_\kappa(t) \), and choosing \( U_h(t) = V_h(t) \circ Q_h(t) \) completes the construction of a unitary operator satisfying the requirements in Theorem 5.1.1. \( \square \)
5.4 Homological equations and the proof of Theorem 5.3.1

In this section, we outline the proof of Theorem 5.3.1. The key ingredient is the following result which solves the homological equation (5.4.3).

**Theorem 5.4.1.** Suppose \( f(\cdot, \cdot; t) \in G^{\rho,\mu}(\mathbb{T}^n \times D) \) satisfies the estimate

\[
|\partial_{\theta}^\alpha \partial_{I}^\beta f(\theta, I; t)| \leq d_0 C |\alpha| + |\mu| |\beta| + q
\]

uniformly in the smooth parameter \( t \in (-1, 1) \) for some \( q > 0 \) and some \( C \geq 1 \) and that for each \( I \in D \), we have

\[
\int_{\mathbb{T}^n} f(\theta, I; t) d\theta = 0.
\]

Then for any smooth family \( \omega(\cdot, \cdot; t) \in G_{\lambda}^{\rho}(D, \Omega) \) there is a solution \( u(\cdot, \cdot; t) \in G^{\rho,\mu}(\mathbb{T}^n \times D) \) to the equation

\[
L_{\omega} u(\theta, I; t) = f(\theta, I; t) \quad (\theta, I) \in \mathbb{T}^n \times E_{\kappa}(t)
\]

\[
u(0, I; t) = 0 \quad I \in D
\]

where \( L_{\omega} = \langle \omega(I; t), \frac{\partial}{\partial \theta} \rangle \). Moreover, \( u \) is smooth in the parameter \( t \) and satisfies the estimate

\[
|\partial_{\theta}^\alpha \partial_{I}^\beta u(\theta, I; t)| \leq A_0 C^{n+\tau+|\alpha|+|\mu|+1} \Gamma(\rho|\alpha| + |\mu| |\beta| + \rho(n + \tau + 1) + q)
\]

where \( A \) depends only on \( n, \rho, \tau \) and \( \mu \).

**Proof.** We begin by taking the Fourier expansions of \( f \) and \( u \) in the \( \theta \)-variable. That is, we have

\[
f(\theta, I; t) = \sum_{k \in \mathbb{Z}^n} e^{i<k,\theta>} f_k(I; t)
\]

and

\[
u(\theta, I; t) = \sum_{k \in \mathbb{Z}^n} e^{i<k,\theta>} u_k(I; t).
\]

where \( u_k \) is to be determined.

From (5.4.2), we have that \( f_0 = 0 \). We set \( u_0 = 0 \) and assemble the solution \( u(\theta, I; t) \) to the homological equation by choosing suitable Fourier coefficients.

We first fix \( \delta > 0 \) so that

\[
(2 + \rho) \delta \leq \mu - \rho(\tau + 1) - 1
\]

and choose \( \psi \in G^{1+\delta}(\mathbb{R}) \) such that \( 0 \leq \psi \leq 1 \) and

\[
\psi(x) = \begin{cases} 
0 & \text{if } x \geq \kappa/2 \\
1 & \text{if } x \leq \kappa/4.
\end{cases}
\]

We note that this is possible from the construction of Gevrey bump functions in [39].

We then define

\[
u_k(I; t) := [(\omega(I; t), k) + i\kappa |k|^{-\tau} \psi(|\omega(I; t), k||k|^{-\tau})]^{-1} f_k(I; t)
\]
for \( I \in D \) and nonzero \( k \). The definition of \( E_\kappa(t) = \omega^{-1}(\Omega_\kappa; t) \) immediately implies that
\[
g_k(I; t) = \langle \omega(I; t), k \rangle \quad \text{for} \quad I \in E_\kappa(t). \tag{5.4.11}
\]
Moreover, since we either have
\[
|\langle \omega(I; t), k \rangle||k|^\tau \geq \kappa/4 \tag{5.4.12}
\]
or
\[
\kappa|k|^{-\tau}\psi(|\langle \omega(I; t), k \rangle||k|^\tau|) = \kappa|k|^{-\tau} \tag{5.4.13}
\]
for any fixed \( k \in \mathbb{Z}^n \) and \( I \in D \), it follows that
\[
|\langle \omega(I; t), k \rangle + i\kappa|k|^{-\tau}\psi(|\langle \omega(I; t), k \rangle||k|^\tau|)| \geq \kappa|k|^{-\tau} \geq \kappa|k|^{-\tau}/4. \tag{5.4.14}
\]
The rapid decay of Fourier coefficients together with the estimate (5.4.14) imply that the Fourier series (5.4.7) is convergent and that \( u \) is smooth in \( t \). It remains to estimate the derivatives of \( u \).

Since \( \omega(\cdot; t) \in C^\rho \) \( = C^{\rho(\tau+1)+1} \), we have
\[
|\partial_\tau^p \omega(I; t, k)| \leq C_1^{(|\alpha|)}|k|\alpha!(|\alpha| - 1)!\rho(\tau+1) \tag{5.4.15}
\]
for \( \tilde{C}_1 = L_0 \max(c(n, \rho, \tau), \|\omega\|_{L_0}) \). We also have that the function \( \psi_k(x) = \kappa|k|^{-\tau}\psi(x||k|^\tau|) \) satisfies
\[
|\partial_\tau^p \psi_k(x)| \leq \tilde{C}_2^{p+1}p^{1+\delta}|k|^{\tau(p-1)}. \tag{5.4.16}
\]
Using the Faa di Bruno formula, we can now bound the derivatives of the second term in
\[
\langle \omega(I; t), k \rangle + i\kappa|k|^{-\tau}\psi(|\langle \omega(I; t), k \rangle||k|^\tau|). \tag{5.4.17}
\]
Indeed, we obtain
\[
|\partial_\tau^p (\kappa|k|^{-\tau}\psi(|\langle \omega(I; t), k \rangle||k|^\tau|))| \leq \sum_{\alpha} \sum_{\rho=1} C_1^{(|\alpha|)}\tilde{C}_2^{p+1}|k|^{\tau p + p - \tau}\rho^{\delta}p! \prod_{j=1}^p (|\beta_j| - 1)!\rho(\tau+1) \tag{5.4.18}
\]
\[
\prod_{j=1}^n \left( \frac{\alpha_j + p - 1}{p - 1} \right) \leq 2^{(|\alpha|)}2^{n(p-1)}. \tag{5.4.21}
\]
Hence, by taking \( C_1 \) sufficiently large (dependent on \( n, \rho, \tau, \|\omega\| \)) and utilising the inequal-
we arrive at an estimate for $g_k$

$$|\partial^\alpha g_k| \leq C_1 |\alpha|!^{1+\delta} \max_{1 \leq \rho \leq |\alpha|} (|k|^{\tau \rho + \rho - \tau} (|\alpha| - p)^{\rho(\tau + 1)})$$ (5.4.24)

for nonzero $k \in \mathbb{Z}^n$ and nonzero multi-indices $\alpha$.

We can then apply the Faa di Bruna formula again together with the estimate (5.4.24) to obtain

$$|\partial^\alpha (g_k(I; t))^{-1}| \leq \sum_{p=1}^{\alpha} \frac{\alpha!}{\rho^{p+1}} \prod_{j=1}^p |\partial^\beta g(I; t)|$$

$$\leq \sum_{p=1}^{\alpha} \frac{\alpha!}{\rho^{p+1}} \prod_{j=1}^p |\partial^\beta g|$$

$$\leq \tilde{C}_0 |\alpha|!^{1+\delta} \max_{1 \leq j \leq |\alpha|} (|k|^{\tau j + j + \tau} (|\gamma| - j)^{\rho(\tau + 1)})$$ (5.4.25)

for sufficiently large $\tilde{C}_0$, dependent only on $n, \rho, \tau, \|\omega\|_{L_0}$ and our choice of cutoff function $\psi$.

We introduce the following quantity that simplifies the notation of the estimates to come

$$\langle k \rangle_m := 1 + \sum_{j=1}^n |k_j|^m$$ (5.4.26)

for multi-indices $k$.

For $m = 1$, we have $|k| \leq \langle k \rangle_1$, and for nonzero $k$ we additionally have $\langle k \rangle_1 \leq 2|k|$. By splitting $\langle k \rangle_m$ into its $n+1$ summands and integrating (5.4.1) by parts, the estimate $|k|^m \leq n^m \langle k \rangle_m$ yields

$$|k^\beta \langle k \rangle_m \partial^\rho f_k(I; t)| \leq (n+1)d_0 C^{[\beta]+\mu[\alpha]+m} \Gamma(\rho|\beta + \mu|\alpha + \rho m + q)$$ (5.4.27)

for $m \in \mathbb{N}$ and any multi-indices $\alpha, \beta$. This estimate is uniform in $I \in D$.

We will need to apply (5.4.27) multiple times with different choices of $m$, so we define

$$m(j) = [(\tau + \delta)j + \tau] + j + n + 1$$ (5.4.28)

for $j \in \mathbb{N}$. By choosing $\delta < 1 + |\tau| - \tau$, we have the following bound for $m(j)$.

$$\tau j + j + \tau + n + \delta < m(j) \leq (\tau + \delta + 1)j + \tau + n + 1.$$ (5.4.29)

Setting $W(k) := \langle k \rangle_1^{-n-\delta}$, we proceed to bound

$$A_{k,\alpha,\beta} := |\langle k \rangle_1^{n+\delta} k^{\alpha} \partial^\beta u_k(I; t)|$$ (5.4.30)

using (5.4.25). This bound will be made $k$-independent, and so summation will complete the proof.
The Leibniz rule implies
\[ A_{k,\alpha,\beta} \]
\[ \leq (k)_{1}^{n+\delta} \sum_{\gamma \leq \beta} \left( \frac{\beta}{(\beta - \gamma)!} \right) C_{0}^{1+1+\delta} \max_{0 \leq j \leq |\gamma|} |k|^{\nu j+\tau j} (|\gamma - j|)^{\nu(\tau + 1)j} (k)^{-1} \left| k^{\alpha} (k)_{m(j)} \partial_{I}^{\beta-\gamma}(f_{k}(I)) \right| \]
\[ \leq A \sum_{\gamma \leq \beta} \left( \frac{\beta}{(\beta - \gamma)!} \right) C_{0}^{1+1+\delta} \max_{0 \leq j \leq |\gamma|} (|\gamma - j|)^{\nu(\tau + 1)j} (k)^{-1} \left| k^{\alpha} (k)_{m(j)} \partial_{I}^{\beta-\gamma}(f_{k}(I)) \right| \]
\[ \leq A \sum_{\gamma \leq \beta} \left( \frac{\beta}{(\beta - \gamma)!} \right) C_{0}^{1+1+\delta} \max_{0 \leq j \leq |\gamma|} (|\gamma - j|)^{\nu(\tau + 1)j} \left| k^{\alpha} (k)_{m(j)} \partial_{I}^{\beta-\gamma}(f_{k}(I)) \right| \]
\[ \leq A \sum_{\gamma \leq \beta} \left( \frac{\beta}{(\beta - \gamma)!} \right) C_{0}^{1+1+\delta} \max_{0 \leq j \leq |\gamma|} C^{\nu|\alpha|+\mu|\beta-\gamma|+m(j)} (|\gamma - j|)^{\nu(\tau + 1)j} \Gamma(s) \]
where
\[ s := \rho|\alpha| + \mu|\beta - \gamma| + \rho m(j) + q \leq \rho|\alpha| + \mu|\beta - \gamma| + \rho(\tau + 1)j + \rho(\tau + n + 1) + 1. \] (5.4.33)
and \( A, C_{0} \) are constants only dependent on \( n, \rho, \tau, \delta \).

Stirling’s formula implies that \( \gamma^{\delta} \leq \Gamma(\delta |\gamma|) \) and \( (|\gamma - j|)^{\nu(\tau + 1)} \leq \Gamma(\rho(\tau + 1)j). \)

Since \( s \geq 1 + |\beta - \gamma| \geq 1 \), we can estimate
\[ B_{j,\beta,\gamma} \]
\[ := \frac{\beta}{(\beta - \gamma)!} \gamma^{\delta} (|\gamma| - j)^{\nu(\tau + 1)j} \Gamma(s) \] (5.4.34)
\[ \leq \frac{\beta}{(\beta - \gamma)!} (C_{2}C_{3})^{\nu} \Gamma(s + \rho(\tau + 1)j + \delta |\gamma|) \] (5.4.35)
\[ \leq s(s + 1) \ldots (s + |\gamma| - 1)(C_{2}C_{3})^{\nu} \Gamma(s + \rho(\tau + 1)j + \delta |\gamma|) \] (5.4.36)
\[ \leq (C_{2}C_{3})^{\nu} \Gamma(s + \rho(\tau + 1)j + (1 + \delta) |\gamma|) \] (5.4.37)
\[ \leq (C_{2}C_{3})^{\nu} \Gamma(s + \rho(\tau + 1)j + 1 + q - \delta |\gamma|) \] (5.4.38)
where the last inequality follows from replacing the first \( |\gamma| \) factors in the previous line with larger factors that can be absorbed into the Gamma function. From (5.4.8), we have \( \mu > 2 \delta + \rho(\tau + 1) + 1 \). Together with (5.4.38), we can further estimate
\[ B_{j,\beta,\gamma} \leq \frac{(C_{2}C_{3})^{\nu} \Gamma(\rho|\alpha| + \mu|\beta| + \rho(\tau + n + 1) + 1 + q - \delta |\gamma|)}{(C_{2}C_{3})^{\nu} \Gamma(\rho|\alpha| + \mu|\beta| + \rho(\tau + n + 1) + 1 + q - \delta |\gamma|)} \] (5.4.39)
for nonzero \( \gamma \).

Combining our estimates for \( A_{k,\alpha,\beta} \) and \( B_{j,\beta,\gamma} \), and redefining \( A(n, \rho, \tau) \) appropriately,
we obtain
\[ A_{k,\alpha,\beta} \leq Ad_0 \sum_{\gamma \leq \beta} C_0^{[\gamma]+1} C^{[\alpha]+\mu[\beta-\gamma]+(\tau+\delta+1)[\gamma]+(\tau+n+1)} \max_j (B_j,\beta,\gamma) \]
\[ \leq Ad_0 C^{[\alpha]+\mu[\beta]+\tau+n+1} \Gamma(\rho[\alpha]+\mu[\beta]+\rho(\tau+n+1)+1+q)) \left( 1 + \sum_{\gamma \neq 0} \frac{C_0^{[\gamma]+1}(C_2 C_3)^{[\gamma]}}{\Gamma(\delta[\gamma])} \right) \]
\[ \leq Ad_0 C^{[\alpha]+\mu[\beta]+\tau+n+1} \Gamma(\rho[\alpha]+\mu[\beta]+\rho(\tau+n+1)+1+q)). \] (5.4.41)

Hence, we can conclude that
\[
|\partial_\rho^a \partial_\tau^b u| \leq \sum_{k \in \mathbb{Z}^n} |k^\alpha \partial_\tau^\beta u_k(I)|
\]
\[
\leq \left[ \sum_{k \in \mathbb{Z}^n \gamma \leq \beta} \langle k \rangle_{\gamma+n+\delta}^{-1} \right] \cdot Ad_0 C^{[\alpha]+\mu[\beta]+\tau+n+1} \Gamma(\rho[\alpha]+\mu[\beta]+\rho(\tau+n+1)+1+q)) \]
\[ = Ad_0 C^{[\alpha]+\mu[\beta]+\tau+n+1} \Gamma(\rho[\alpha]+\mu[\beta]+\rho(\tau+n+1)+1+q)). \] (5.4.42)

where \( A \) depends only on \( n, \rho, \tau \) as required.

\[ \square \]

Completion of proof of Theorem 5.3.1. To prove Theorem 5.3.1, we need to find \( a, K^0 \in S_\ell(\mathbb{T}^n \times D) \) such that \( r = p \circ a - a \circ K^0 \) is flat on \( E_\ell(t) \), where \( \ell = (\rho, \mu, \nu) \) by continuing the iterative procedure outlined in the discussion preceding the theorem.

This can be done by using the composition formula (2.1.12) and induction using Theorem 5.4.1 to assemble the symbols \( a, K^0 \) in terms of their semiclassical expansions (5.3.10) and (5.3.11).

We begin by setting \( a_0 = 0, K_1 = 0, r_0 = R_0 \) and recalling that \( p_0(I; t) = K_0(I; t) + R_0(\theta, I; t) \) and \( p_1 = 0 \) from the construction of the semiclassical pseudodifferential operator \( \tilde{P}_h \) in Theorem 5.2.2.

The composition formula (2.1.12) then implies that
\[ r_0(\theta, I; t) = R_0(\theta, I; t) \] (5.4.43)
and
\[ r_1(\theta, I; t) = a_1(\theta, I; t)R_0(\theta, I; t) \] (5.4.44)
which are both flat in \( I \) at \( E_\ell(t) \) since \( R_0 \) is.

For \( j \geq 2 \), the composition formula (2.1.12) yields
\[ r_j(\theta, I; t) = \frac{-i}{(\mathcal{L}_\omega a_j-1)(\theta, I; t)} + p_j(\theta, I; t) - K_j(I; t) + F_{j1}(\theta, I; t) - F_{j2}(\theta, I; t) \] (5.4.45)
where
\[ F_{j1}(\theta, I; t) = \sum_{s=1}^{j-2} \sum_{r+|\gamma|=j-s} \frac{1}{\gamma!} \partial_\gamma^r p_r(\theta, I; t) \partial_\tau^\alpha a_s(\theta, I; t) \] (5.4.46)
and
\[ F_{j2}(\theta, I; t) = \sum_{s=1}^{j-2} a_s(\theta, I; t) K^0_{j-s}(I; t). \] (5.4.47)

We solve the equations \( r_j = 0 \) by using the expression (5.4.45) and Theorem 5.4.1.
where $\vartheta$ that this class is independent of choice of torus by the local constancy of the Maslov class.

Suppose now that Theorem 5.5.2. $L > 0$ of nonresonant actions. For an arbitrary constant $G$ defines a parametrised by $a, K$ Theorem 5.4.1. This allows us to complete the construction of the symbols of $a, K^0$ and $r$ as required.

\section{5.5 Quasimode construction}

Having completed the proof of Theorem 5.1.1, we now complete Popov’s construction of a family of Gevrey quasimodes for $P_h(t)$ that are semiclassically supported on a family of nonresonant tori. Moreover, these quasimodes are smooth in the parameter $t \in (-1, 1)$.

\begin{definition}
A $G^p$ family of quasimodes $Q(t)$ for $P_h(t)$ is a family
\[
\{(u_m(x; t, h), \lambda_m(t, h)) : m \in \mathcal{M}_h(t)\} \subset C^\infty(M \times D_h(m)) \times C^\infty(D_h(m))
\]
parametrised by $h \in (0, h_0]$ where

- $\mathcal{M}_h(t) \subset \mathbb{Z}^n$ is a $h$-dependent finite index set;
- $D_h(m) = \{t \in (0, 1) : m \in \mathcal{M}_h(t)\}$
- each $u(\cdot; t, h)$ is uniformly of class $G^p$;
- \[\|P_h(t)u_m(\cdot; t, h) - \lambda_m(t; h)u_m(\cdot; t, h)\|_{L^2} \leq C e^{-c/h^{1/p}} \quad \forall m \in \mathcal{M}_h(t)\];
- \[|\langle u_m(\cdot; t, h), u_l(\cdot; t, h) \rangle - \delta_{ml}| \leq C e^{-c/h^{1/p}} \quad \forall m, l \in \mathcal{M}_h(t)\].
\end{definition}

\begin{theorem}
Suppose now that $t \in (-1, 1)$ is fixed and $S \subset E_n(t)$ is a closed collection of nonresonant actions. For an arbitrary constant $L > 1$, we define the index set
\[
\mathcal{M}_h := \{ m \in \mathbb{Z}^n : \text{dist}(S, h(m + \vartheta/4)) < Lh\}
\]
where $\vartheta \in \mathbb{Z}^n$ is the Maslov class of any Lagrangian tori $\{\chi(\mathbb{T}^n \times \{I\})\}$ with $I \in S$. Note that this class is independent of choice of torus by the local constancy of the Maslov class.

Then
\[
\{(u_m(x; t, h), \lambda_m(t; h)) : m \in \mathcal{M}_h(t)\} := (U_h(t)e_m, K^0(h(m + \vartheta/4); t, h)
\]
defines a $G^p$ family of quasimodes for $P_h(t)$ that has Gevrey microsupport on the family of tori
\[
\Lambda_S = \bigcup_{t \in S} \Lambda_{\omega(t; t)} = \bigcup_{t \in S} \chi(\mathbb{T}^n \times \{I\}) \subset T^* M
\]
where \( \{ e_m \}_{m \in \mathbb{Z}^n} \) is the orthonormal basis of \( L^2(\mathbb{T}^n; \mathbb{L}) \) associated to the quasiperiodic functions

\[
\tilde{e}_m(x) := \exp(i\langle m + \vartheta/4, x \rangle)
\]  

\[ (5.5.5) \]

**Proof.** From the definition of the functions \( e_m \), it follows that

\[
P^0_h(t)(e_m)(\theta) = \sigma(P^0_h(t))(\theta, h(m + \vartheta/4))e_m(\theta)
\]

\[ (5.5.6) \]

\[
= (K^0(h(m + \vartheta/4); t, h) + R^0(\theta, h(m + \vartheta/4); t, h))e_m(\theta)
\]

\[ (5.5.7) \]

\[ = (\lambda_m(t; h) + R^0(\theta, h(m + \vartheta/4)))e_m(\theta). \]  

\[ (5.5.8) \]

From the definition (5.5.2) of the index set \( M_h(t) \) and from (4.1.2), it thus follows that

\[
P_h(t)(U_h(t)e_m) = U_h(t)P^0_h(t)e_m = O(e^{-c/h^{1/\rho}}) \]  

\[ (5.5.9) \]

upon an application of Theorem 5.1.1. This establishes that the \( U_h(t)e_m \) are exponentially accurate quasimodes.

The almost-orthogonality of the \( U_h(t)e_m \) then follows from the fact that \( U_h(t) \) is almost unitary from Theorem 5.1.1, and the \( e_m \) are exactly orthogonal by construction.

This completes the proof. \( \square \)

These quasimodes are as numerous as we could hope for, indeed the index set \( M_h(t) \) satisfies the local Weyl asymptotic

\[
\lim_{h \to 0} (2\pi h)^n \# M_h = m(\mathbb{T}^n \times S) = \mu(\Lambda_S)
\]  

\[ (5.5.10) \]

where \( m \) denotes the \((2n)\)-dimensional Lebesgue measure and \( \mu \) denotes the symplectic measure \( d\xi dx \). To see this, we can denote by \( U \) the union of \( n \)-cubes centred at the lattice points in \( M_h \) with side length \( h \). The containment

\[
S \subset U \subset \{ I : \text{dist}(I, S) < \tilde{L}h \}
\]  

\[ (5.5.11) \]

for a constant \( \tilde{L} \) then yields the claim by monotone convergence of measures, noting that since \( S \) is closed we have

\[
S = \overline{S} = \cap_{h > 0} \{ I : \text{dist}(I, S) < \tilde{L}h \}.
\]

\[ (5.5.12) \]

In the special case of \( S = \{ I \} \), we have a family of \( G^0 \) quasimodes with microsupport on an individual torus \( \chi(\mathbb{T}^n \times \{ I \}) \).
Chapter 6

Eigenvalue Localisation Results

6.1 Introduction

In this section, we consider a one-parameter family of elliptic semiclassical pseudodifferential operators obtained by quantising a one-parameter family of KAM Hamiltonian systems. As an application of the quasimode construction in Chapter 5, we prove that the quantum dynamics generated by this family of operators can only be quantum ergodic for a Lebesgue null subset of the possible parameter values.

Suppose that $M$ is an $n$-dimensional compact boundaryless Riemannian manifold of regularity $G^\rho$. The family of Hamiltonian systems in question is then given by

$$H(x, \xi; t) = \|\xi\|^2_g + V(x, \xi) + tQ(x, \xi) \quad (6.1.1)$$

where $H(x, \xi; 0) \in G^\rho(T^*M)$ is a completely integrable system satisfying the nondegeneracy condition (4.2.1) for a suitable choice of action-angle coordinates $(\theta, I) \in \mathbb{T}^n \times D$, and $Q \in G^\rho(T^*M)$ is a non-negative and compactly supported symbol with positive quantisation. In particular, these symbols live in $S_\ell(T^*M)$ in the notation of Definition 2.2.5 with $\ell = (\rho, \mu, \eta)$, with $\rho(\tau + n) + 1 > \mu > \rho' = \rho(\tau + 1) + 1$ and $\nu = \rho(\tau + n + 1)$ as in Chapter 5.

The quantisation of this Hamiltonian is given by the elliptic operator

$$P_h(t) := h^2 \Delta_g + V(x) + tQ(x, hD). \quad (6.1.2)$$

The operator $P_h(t)$ then acts on $L^2(M)$ with domain $H^2_h(M)$, the semiclassical Sobolev space defined as in [50] Chapter 8. The inverse of $P_h(t)$ is compact and self-adjoint, and so the eigenfunctions of $P_h(t)$ form an orthonormal basis for $L^2(M)$ with an unbounded purely discrete spectrum consisting of the eigenvalues $0 \leq E_1(t; h) < \ldots$.

If the Hamiltonian flow corresponding to the Hamiltonian $H(x, \xi, t) = \sigma(P_h) = |\xi|^2_g + V(x) + tQ(x, \xi)$ were ergodic on the energy surface $\Sigma_E = H^{-1}(E) \subset T^*M$ for a regular value $E$ of the Hamiltonian, then the quantum ergodicity theorem [23] asserts that $P_h(t)$ is quantum ergodic on $\Sigma_E$ in the sense of Definition 1.3.2.

However, by rewriting (6.1.1) locally in the action angle variables $(\theta, I)$ for the completely integrable Hamiltonian $H(x, \xi; 0)$, we see that for fixed $Q \in G^\rho_{L_1, L_2} (\mathbb{T}^n \times D)$ with $L_2 \geq L_1 \geq 1$, the results of Section 4.8 imply that $H(x, \xi; t)$ generates KAM dynamics for every $t \in [0, \delta]$, for sufficiently small $\delta > 0$, dependent on $Q$. As KAM dynamics are far from ergodic dynamics in character, I expect that $P_h(t)$ is typically not
quantum ergodic, and that there could even exist sequences of eigenfunctions for $P_h(t)$ with semiclassical mass entirely supported on individual invariant tori.

As a first result in this area of investigation, we prove that $P_h(t)$ is not quantum ergodic by arguing by contradiction. A key assumption is the existence of what we shall call a slow torus for our family of perturbations.

**Definition 6.1.1.** A slow torus in the energy band $[a, b]$ for the Hamiltonian

$$ H(\theta, I; t) = H^0(I) + tH^1(\theta, I) $$

in action-angle coordinates, is a Lagrangian invariant torus $\Lambda_{\omega_0}$ with nonresonant frequency $\omega_0 \in \tilde{\Omega}_\kappa$ and energy $H^0(I(\omega_0)) = H^0(\nabla g^0(\omega_0)) \in [a, b]$ in the notation of Theorem 4.8.2 that satisfies

$$ \left(2\pi\right)^{-n} \int_{T^n} H^1(\theta, I(\omega_0)) \, d\theta < \inf_{E \in [a, b]} \frac{1}{\mu_E(\Sigma_E)} \int_{\Sigma_E} Q(x, \xi) \, d\mu_E $$

at $t = 0$.

**Remark 6.1.2.** It should be mentioned that the Lagrangian tori that foliate the completely integrable space $D_t \setminus U_t$ in Chapter 3 are analogously slow with respect to the perturbation parameter $t$, making the spectral non-concentration arguments in Section 6.3 and Section 5.3 rather similar in spirit.

However, in the KAM setting we do not have a result analogous to Theorem 3.1.4 that precisely describes the restriction of semiclassical measures to the non-integrable region. This makes it more difficult to prove results on eigenfunction localisation in KAM systems compared to mixed systems, at least via the methods used in this thesis.

The assumption (6.1.4) holds in a small energy band $[a, b]$ whenever the completely integrable Hamiltonian system $H(x, \xi; 0)$ possesses a nonresonant torus on some energy surface $\Sigma_E$ on which the perturbation $Q$ is smaller than it is on average on $\Sigma_E$. We call such a torus a slow torus to emphasise the key feature that the quasi-eigenvalues associated to such a torus will increase at a slower rate than the typical increase of eigenvalues at the same energy. The intuition behind this stems from the Hadamard variational formula (6.2.1), and the fact that the associated quasimodes localise onto $\Lambda_{\omega_0}$ in phase space.

This condition is rather mild, and will typically be satisfied for some torus provided that $Q$ is not constant on energy surfaces. An example of a Hamiltonian system with a slow torus could be explicitly constructed by taking an axisymmetric surface of revolution, and choosing $Q(x, \xi)$ to be small near a collar containing a stable geodesic.

We are now in a position to state our main theorem.

**Theorem 6.1.3.** Suppose $M$ is a compact boundaryless $G^\rho$-smooth Riemannian manifold, and the Hamiltonian $H(x, \xi; t)$ is regular on the energy interval $[a, b]$ and can be written as

$$ H(x, \xi; t) = H^0(I) + tH^1(\theta, I) \in G_{L_1, L_2, L_2}^{0, 0, 1}(\mathbb{T}^n \times D \times (-1, 1)) $$

in a subdomain $\mathbb{T}^n \times D \times (-1, 1) \subset \mathbb{T}^n \times D^0 \times (-1, 1)$ of local action-angle variables for the non-degenerate completely integrable Hamiltonian $H^0 \in G_{L_0}^{0}(D^0)$ with Legendre transform $g^0 \in G_{L_0}^{0}(\Omega^0)$. 

Suppose further that \( a < b \) and there exists a slow torus \( \Lambda_{\omega_0} \) in the energy band \([a,b]\).

Then there exist \( \epsilon, \delta > 0 \) such that for almost all \( t \in [0, \delta] \) the quantisation \( P_h(t) \) of \( H(x,\xi; t) \) is quantum ergodic over the energy surface \( \Sigma_E \) for at most a Lebesgue measure \((1 - \epsilon)(b - a)\) subset of \( E \in [a,b] \).

We prove Theorem 6.1.3 by using an argument based on controlling the flow speed of eigenvalues in the parameter \( t \) as in Chapter 3.

We begin by using the slow torus condition and choosing \( \epsilon_1 > 0 \) sufficiently small so that

\[
(2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, I(\omega_0)) \, d\theta < \inf_{E \in [a,b]} \frac{1}{\mu_E(\Sigma_E)} \int_{\Sigma_E} Q(x, \xi) \, d\mu_E - 2\epsilon_1 \quad (6.1.6)
\]

is satisfied at \( t = 0 \).

We next observe that it suffices to prove the theorem for a small interval \([a,b]\) containing the energy \( H^0(I(\omega_0)) \), and so we can scale our interval \([a,b]\) by a small factor \( \lambda \) to ensure that the condition

\[
\sup_{E \in [a,b]} \frac{1}{\mu_E(\Sigma_E)} \int_{\Sigma_E} Q \, d\mu_E - \inf_{E \in [a,b]} \frac{1}{\mu_E(\Sigma_E)} \int_{\Sigma_E} Q \, d\mu_E =: Q_+(0) - Q_-(0) < \epsilon_2 < \epsilon_1 \quad (6.1.7)
\]

is satisfied.

Now Theorem 4.8.8 applies, and we obtain a family of symplectic maps \( \chi \in G^{\rho, \varphi'}(\mathbb{T}^n \times D \times (-1/2, 1/2), \mathbb{T}^n \times D) \) and a family of diffeomorphisms \( \omega \in G^\rho(\Omega \times (-1/2, 1/2), \Omega) \) such that

\[
H(\chi(\theta, I; t); t) = K(I; t) + R(\theta, I; t) \quad (6.1.8)
\]

where \( R \) is flat in \( I \) at the nonresonant actions \( I \in E_\kappa(t) \).

Defining the action map \( I \in G^\rho(\Omega \times (-1/2, 1/2)) \) implicitly by

\[
\omega := \omega(I(\omega; t); t) \quad (6.1.9)
\]

we can specify the action coordinates of a nonresonant torus at \( t \in [0, \delta] \) in the Birkhoff normal form furnished by \( \chi(\cdot, \cdot; t) \).

We begin by expanding the nonresonant slow torus \( \Lambda_{\omega_0} \) to a positive measure family of slow tori.

**Proposition 6.1.4.** For any \( 0 < c < \min(\frac{b - a}{2}, \text{dist}(H^0(I(\omega_0))), [a, b]^c) \), there exists \( r > 0 \) and \( \delta > 0 \) such that for any \( \omega \in \Omega := B(\omega_0, r) \cap \Omega_\kappa \), the torus \( \Lambda_\omega = \chi(\mathbb{T}^n \times \{I(\omega, t)\}) \) has energy

\[
K(I(\omega; t); t) \in [K(I(\omega_0; t); t) - c, K(I(\omega_0; t); t) + c] \quad (6.1.10)
\]

for all \( t \in [0, \delta] \) and the torus \( \Lambda_{\omega_0}(t) \) has energy

\[
K(I(\omega_0; t); t) \in [a + c, b - c] \quad (6.1.11)
\]

for all \( t \in [0, \delta] \).

In particular, the family of tori

\[
\Lambda(t) := \bigcup_{\omega \in \Omega} \Lambda_\omega \quad (6.1.12)
\]
is a positive measure family of tori entirely contained within the energy band \([a, b]\).

Moreover, \(r\) and \(\delta\) can be chosen small enough to ensure

\[
(2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, I(\omega; t)) \, d\theta < (2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, I(\omega)) \, d\theta + \epsilon_1 < Q_- - \epsilon_1.
\]

for each \(\omega \in \overline{\Omega}\) and each \(t \in [0, \delta]\).

We can also choose \(\delta > 0\) small enough to ensure that

\[
Q_+(t) - Q_-(t) < 2\epsilon_2.
\]

for all \(t \in [0, \delta]\).

**Proof.** From the regularity of \(\chi, I,\) and \(K\) established in Theorem 4.8.2, it follows that we can take \(r < Lc\) for some constant \(L > 0\) to ensure that (6.1.10) is satisfied at \(t = 0\) for sufficiently small \(c > 0\). Similarly, we can ensure that

\[
(2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, I(\omega)) \, d\theta < (2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, I(\omega_0)) \, d\theta + \epsilon_1/2
\]

holds for \(|I - I(\omega_0)|\) sufficiently small. (6.1.13) is satisfied at \(t = 0\). In particular this implies that

\[
(2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, I(\omega)) \, d\theta < (2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, I(\omega_0)) \, d\theta + \epsilon_1/2
\]

for all \(\omega \in \overline{\Omega} = B(\omega_0, r) \cap \tilde{\Omega}_\kappa\) upon taking \(r\) sufficiently small.

The regularity of \(\chi, I,\) and \(K\) in the parameter \(t\) then allow us to then deduce that (6.1.10) and (6.1.13) are satisfied for \(t \in [0, \delta]\), for sufficiently small \(\delta > 0\) and for each \(\omega \in \overline{\Omega}\).

Finally, the estimate (6.1.14) for small \(\delta\) follows from the regularity of

\[
\frac{1}{\mu_E(\Sigma_E)} \int_{\Sigma_E} Q \, d\mu_E
\]

in \(t\).

From the Birkhoff normal form (6.1.8), we construct a quantum Birkhoff normal form as in Chapter 5. The constructions of Section 5.5 then provide us with a family of quasi-modes localising onto \(\Lambda(t)\) by taking \(S(t) = \{I(\omega; t) : \omega \in \overline{\Omega}\}\) and defining the index set \(\mathcal{M}_h(t)\) as in (5.5.2).

We can now define \(t\)-dependent energy bands

\[
\mathcal{I}(t) := [K_0(I(\omega_0, t), t) - c, K_0(I(\omega_0, t), t) + c] \subseteq [a, b]
\]

and the union of \(h^\alpha\)-width quasi-eigenvalue windows

\[
W(t; h) := \mathcal{I}(t) \cap \bigcup_{m \in \mathcal{M}_h(t)} [K^0(h(m + \vartheta/4), t; h) - h^\alpha, K^0(h(m + \vartheta/4), t; h) + h^\alpha]
\]

where \(\alpha > 2n\) and \(K_0 = K\) and \(K^0\) are as in Theorem 5.1.1.
§6.2 Eigenvalue and quasieigenvalue variation

For the sake of brevity, we introduce the notation
\[ \mu_m(t; h) := K^0(h(m + \vartheta/4), t; h) \] (6.1.20)
for the quasi-eigenvalues under consideration.

The number of exact eigenvalues lying in the union of quasi-eigenvalue windows \( W(t; h) \) is given by
\[ N(t; h) := \# \{ E_k(t; h) \in W(t; h) \}. \] (6.1.21)

Next we obtain asymptotic estimates for the number of eigenvalues and the number of quasi-eigenvalues contained in the interval \( I(t) \) as \( h \to 0 \).

**Proposition 6.1.5.** We have the asymptotic estimate
\[ \# M_h(t) \sim (2\pi h)^n \mu(\mathbb{T}^n \times \{ I(\omega, t) : \omega \in \mathbb{N} \}). \] (6.1.22)
for each \( t \in [0, \delta] \).

Furthermore, we have
\[ \limsup_{h \to 0} (2\pi h)^n \# \{ k \in \mathbb{N} : E_k(t; h) \in [a, b] \text{ for some } t \in [0, \delta] \} \leq \mu(\{(x, \xi) : H(x, \xi; 0) \in [a - M\delta, b]\}) \] (6.1.23)
where \( M = \| Q(x, hD) \|_{L^2 \to L^2} \).

Here \( \mu \) denotes the symplectic measure \( d\xi dx \) on \( T^* M \).

**Proof.** The estimate (6.1.22) is a consequence from (5.5.10), and (6.1.23) follows from (6.2.2) and an application of the Weyl law (2.3.3).

A crucial fact is that if we scale the interval \([a, b]\) and \( c \) about \( K_0(I(\omega_0, 0), 0) \) by a factor of \( \lambda > 0 \), then the \( r > 0 \) chosen in Proposition 6.1.4 can be chosen to be scaled by the factor \( \lambda > 0 \) as well. Doing this for small \( \lambda > 0 \) will make the quantity
\[ \sup_{t \in [0, \delta]} Q_+(t) - \inf_{t \in [0, \delta]} Q_-(t) < 3\epsilon_2 = o(1). \] (6.1.25)
However we will still have
\[ \limsup_{h \to 0} \frac{\# \{ k \in \mathbb{N} : E_k(t; h) \in [a, b] \text{ for some } t \in [0, \delta] \}}{\inf_{t \in [0, \delta]} \# M_h(t)} \] (6.1.26)
uniformly bounded in \( t \) and \( \lambda \).

### 6.2 Eigenvalue and quasieigenvalue variation

For each fixed \( h > 0 \), the operators \( P_h(t) \) comprise an holomorphic family of type A in the sense of [27] and so we can choose eigenvalues and corresponding eigenprojections holomorphic in the parameter \( t \). Thus if at each time \( t \) we order our eigenpairs \( E_k(t; h) \) in order of increasing energy, by holomorphy it follows that \( E_k \) will be continuous, piecewise smooth, and have multiplicity 1 for all but finitely many \( t \in [0, \delta] \). On this cofinite set, we
Eigenvalue Localisation Results

have

\[ \dot{E}_k(t; h) = \langle \mathcal{P}_h(t) u_k(t; h), u_k(t; h) \rangle = \langle Q(x, hD) u_k(t; h), u_k(t; h) \rangle \] (6.2.1)

from \((u_k)\) being an orthonormal basis. We will control (6.2.1) using an assumption of quantum ergodicity.

Suppose, for the sake of contradiction, that for every \(\epsilon > 0\) there exists a positive measure subset \(B_\epsilon\) of \([0, \delta]\) such that for each \(t \in B_\epsilon\), \(P_h(t)\) is quantum ergodic on a set of energies \(E_t \subset [a, b]\) with \(m(E_t) > (b - a)(1 - \epsilon/2)\).

Analogously to Section 3.4, we can use quantum ergodicity to deduce bounds on the speed at which the eigenvalues monotonically increase in \(t\) under the assumption of eigenfunction equidistribution.

We first note that we have a global in time bound

\[ E'_k(t) \leq M = \|Q\|_{L^2 \to L^2} < \infty \] (6.2.2)

differentiation of the expression

\[ E_k(t) = \langle P_h(t) u_k(t), u_k(t) \rangle. \] (6.2.3)

Now, our first result estimate on eigenvalue variation is one localised to individual energy surfaces on which we have quantum ergodicity.

**Proposition 6.2.1.** For each \(t \in B_\epsilon\), each \(E \in E_t\), and each \(\hat{\epsilon} > 0\), there exists an \(h\)-dependent subset \(S_E(t; h) \subset \{ k \in \mathbb{N} : E_k(t; h) \in [E - h, E + h] \}\) and \(h_E > 0\) such that

\[ \#S_E(t; h) > 1 - \hat{\epsilon} \] (6.2.4)

and

\[ \dot{E}_k(t; h) \in [Q_-, \epsilon_2, Q_+ + \epsilon_2] \] (6.2.5)

for all \(h < h_E\) and all \(k \in S_E(t; h)\).

**Proof.** This claim follows immediately from the Hadamard variational formula (6.2.1), Definition (1.3.2), and (6.1.7).

Proposition 6.2.1, together with the definition of \(B_\epsilon\) allows us to obtain a similar result over the whole energy band \([a, b]\).

**Proposition 6.2.2.** For each \(t \in B_\epsilon\), there exists an \(h\)-dependent subset \(S(t; h) \subset \{ k \in \mathbb{N} : E_k(t; h) \in [a, b] \}\) and \(h_0 > 0\) such that

\[ \#S(t; h) > \frac{1 - C(n, Q, b - a)\epsilon}{\epsilon_1^2} \] (6.2.6)

and

\[ \dot{E}_k(t; h) \in [Q_-, \epsilon_2, Q_+ + \epsilon_2] \] (6.2.7)

for all \(h < h_0\) and all \(k \in S(t; h)\).
§6.2 Eigenvalue and quasieigenvalue variation

Proof. Noting that this result is pointwise in time, we omit the parameter \( t \) in our notation and we define

\[
f(E, k; h) := 1_{|E - E_k(h)| < h} \cdot h^{n-1} \left| \langle Q(x, hD)u_k(h), u_k(h) \rangle - \frac{1}{\mu_E(\Sigma_E)} \int_{\Sigma_E} Q d\mu_E \right|^2.
\]

From Definition 1.3.2 of quantum ergodicity, we have

\[
\sum_{k \in \mathbb{N}} f(E, k; h) \to 0 \quad \text{(6.2.9)}
\]
as \( h \to 0 \) for each \( E \in \mathcal{E} \). Hence, for any \( \dot{\epsilon} > 0 \), we can find an \( h_0 > 0 \) such that

\[
\int_a^b \left( \sum_{k \in \mathbb{N}} f(E, k; h) \right) dE < \dot{\epsilon}(b - a)(1 - \epsilon) + N(b - a)\epsilon \quad \text{(6.2.10)}
\]

where \( N \) is an upper bound for

\[
h^{n-1} \sum_{E_k(h) \in [E - h, E + h]} \left| \langle Q(x, hD)u_k(h), u_k(h) \rangle - \frac{1}{\mu_E(\Sigma_E)} \int_{\Sigma_E} Q d\mu_E \right|^2 \leq 4 \max \left( \sup_{H^{-1}([a,b])} (Q), \|Q(x, hD)\|_{L^2 \to L^2} \right)^2 \cdot h^{n-1} \# \{ k \in \mathbb{N} : E_k(h) \in [E - h, E + h] \}
\]

which exists due to the uniform Weyl law in \( h \)-energy bands proven in [32]. Taking \( \dot{\epsilon} = N\epsilon \), we obtain

\[
\int_a^b \left( \sum_{k \in \mathbb{N}} f(E, k; h) \right) dE < 2N(b - a)\epsilon \quad \text{(6.2.12)}
\]

for \( 0 < h < h_0 \).

Now, by interchanging the order of integration and noting that

\[
\left| \frac{1}{\mu_{E_k(h)}(\Sigma_{E_k(h)})} \int_{\Sigma_{E_k(h)}} Q d\mu_{E_k(h)} - \frac{1}{\mu_E(\Sigma_E)} \int_{\Sigma_E} Q d\mu_E \right| = O(h) \quad \text{(6.2.13)}
\]

uniformly for \( E_k(h) \in [E - h, E + h] \), we have

\[
4h^n \sum_{E_k(h) \in [a,b]} \left| \langle Qu_k(h), u_k(h) \rangle - \frac{1}{\mu_{E_k(h)}(\Sigma_{E_k(h)})} \int_{\Sigma_{E_k(h)}} Q d\mu_{E_k(h)} \right|^2 \leq 2N(b - a)\epsilon \quad \text{(6.2.14)}
\]

for \( h < h_0 \), where \( h_0 \) has been redefined.

Denoting by \( S(t; h) \) the collection of all indices \( k \in \mathbb{N} \) with

\[
\dot{E}_k(t; h) \in [Q_- - \epsilon_2, Q_+ + \epsilon_2],
\]

we can deduce from (6.2.14) that

\[
2h^n \sum_{k \in S(t; h)^c} \epsilon_2^2 \leq N(b - a)\epsilon \quad \text{(6.2.16)}
\]
which implies
\[
\frac{\#S(t;h)}{\#\{E_k(t;h) \in [a,b]\}} > 1 - \frac{C(n,Q,b-a)\epsilon}{\epsilon^2} \quad (6.2.17)
\]
from Weyl’s law, as required.

From the outer regularity of the Lebesgue measure, we can find an open interval 
\[ J = (t_1, t_2) \subseteq [0,\delta] \] such that
\[
\frac{m(B \cap J)}{m(J)} > 1 - \tilde{\epsilon}. \quad (6.2.18)
\]
for any \( \tilde{\epsilon} > 0 \). We can strengthen the pointwise bound on eigenvalue flow speed from 
Proposition 6.2.2 for \( t \in B \) to an almost-uniform bound on the interval \( J \).

**Proposition 6.2.3.** There exists a subset \( \tilde{B} \subseteq B \cap J \) and a \( h_0 > 0 \) such that
\[
\frac{m(\tilde{B})}{m(B \cap J)} > 1 - \tilde{\epsilon} \quad (6.2.19)
\]
and for any \( h < h_0 \) and any \( t \in \tilde{B} \), there exists a subset \( Z(t,h) \subseteq \{ k \in \mathbb{N} : E_k(t,h) \in [a,b] \} \) such that
\[
\frac{\#Z(t,h)}{\#\{k \in \mathbb{N} : E_k(t,h) \in [a,b]\}} > 1 - C\epsilon \quad \forall 0 < h_* < h_0 \quad (6.2.20)
\]
and
\[
\dot{E}_k(t,h) \in [Q_- - \epsilon_2, Q_+ + \epsilon_2] \quad \forall k \in Z(t,h). \quad (6.2.21)
\]
where \( C \) is the same constant as in (6.2.6).

**Proof.** We set
\[
G(h) := \{ t \in B \cap J : \frac{\#\{k \in \mathbb{N} : E_k(t,h) \in [a,b] \text{ and } \dot{E}_k(t,h) \in [Q_- - \epsilon_2, Q_+ + \epsilon_2]\}}{\#\{k \in \mathbb{N} : E_k(t,h) \in [a,b]\}} \geq 1 - C\epsilon \quad \forall 0 < h_* < h \}. \quad (6.2.22)
\]
From Proposition 6.2.2, we know that
\[
B \cap J = \bigcup_{h>0} G(h) = \bigcup_{n \in \mathbb{N}^+} G(1/n) \quad (6.2.23)
\]
and so the claim follows from monotone convergence.

On the other hand, we also have an upper bound for the variation of the quasi-eigenvalues.

**Proposition 6.2.4.** For all sufficiently small \( \delta > 0 \) and all \( t \in [0,\delta] \), we have
\[
\limsup_{h \to 0} \partial_t \mu_m(t;h) \leq Q_- - \epsilon_1/2. \quad (6.2.24)
\]
for all \( m \in \cup_{t \in [0,\delta]} \mathcal{M}_h(t) \) uniformly in \( t \).

**Proof.** From Proposition 4.8.8, we have
\[
K_0(I,t) = H^0(I) + t \cdot (2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta,I) \ d\theta + O(t^{3/8}) \quad (6.2.25)
\]
for any \(I \in D\). Hence we have
\[
\partial_t(K_0(h(m + \vartheta/4); t)) < (2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, h(m + \vartheta/4)) d\theta + \epsilon_1/2 \tag{6.2.26}
\]
for all \(t \in [0, \delta]\), taking \(\delta\) sufficiently small.

From the definition of \(M_h(t)\), we know that \(|h(m + \vartheta/4) - I(\omega; t)| < Lh\) for some \(\omega \in \overline{\Omega}\), and so from the regularity of \(I\) in \(t\) it follows that
\[
\partial_t(K_0(h(m + \vartheta/4); t)) < (2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, I(\omega; t)) d\theta + \epsilon_1/2 + O(h) \tag{6.2.27}
\]
for some \(\omega \in \overline{\Omega}\). This allows us to use (6.1.13).

Indeed, we have
\[
\partial \mu_m(t; h) = \partial_t(K_0(h(m + \vartheta/4); t, h)) \tag{6.2.28}
\]
\[
= \partial_t(K_0(h(m + \vartheta/4); t)) + O(h) \tag{6.2.29}
\]
\[
\Rightarrow \limsup_{h \to 0} \partial_t \mu_m(t; h) \leq Q_--\epsilon_1/2. \tag{6.2.30}
\]

In particular, we can choose \(B > 0\) and \(h_0 > 0\) such that
\[
\partial_t \mu_m(t; h) < B < Q_--\epsilon_1/4. \tag{6.2.31}
\]

for all \(t \in J\) and all \(h < h_0\).

**Remark 6.2.5.** We have abused notation slightly here by writing \(\mu_m(t; h)\) even when \(m \notin M_h(t)\). That is, we track the behaviour of \(K_0(h(m + \vartheta/4), t; h)\) even for \(t \in [0, \delta]\) such that this does not correspond to a quasi-eigenvalue in our family. This is a necessity due to the rough nature of the set \(\{I(\omega; t) : \omega \in \overline{\Omega}\}\) of nonresonant actions. Indices \(m \in \mathbb{Z}^n\) will typically be elements of \(M_h(t)\) for only \(O(h)\)-sized \(t\)-intervals at a time.

### 6.3 Spectral non-concentration

We can now complete proof of Theorem 6.1.3 by proving a spectral non-concentration result that follows from the results of Section 6.2.

**Proposition 6.3.1.** There exists sufficiently small \(\epsilon\), such that
\[
\frac{N(t_*; h)}{\#M_h(t_*)} < 1/2 \tag{6.3.1}
\]
for some \(t_* \in J\) and all \(0 < h < h_0\), where \(N\) is as in (6.1.21).

**Proof.** We begin by defining
\[
A(t; h) = \{k \in \mathbb{N} : E_k(t; h) \in W(t; h)\}. \tag{6.3.2}
\]
where \(W(t; h)\) is as in (6.1.19). It then suffices to show that
\[
\frac{\#A(t; h)}{\#M_h(t)} < 1/2 \tag{6.3.3}
\]
can be made arbitrarily small for some \( t \in J \) by taking \( \epsilon \) arbitrarily small in the definition of \( \mathcal{B}_t \) given in the introduction of Section 6.2.

We do this by averaging in \( t \) and exploiting Proposition 6.2.3. First we define

\[
B_k(h) := \{ t \in J : k \in Z(t; h) \}
\]

where \( Z \) is as defined in Proposition 6.2.3, and consider the problem of bounding

\[
\sum_{k \in \mathbb{N}} \int_J 1_{B_k}(t) \, dt
\]

noting that only finitely many terms of the sum are nonzero. This quantity controls the average amount of time that the individual eigenvalues that meet \( W(t; h) \) for some \( t \in J \) spend travelling at the approximately the ergodic rate given by the small interval

\[
[\tilde{Q}_-, \tilde{Q}_+] := [Q_- - \epsilon_2, Q_+ + \epsilon_2].
\]

By definition, this can be re-written as

\[
\int_{\tilde{B}} \#Z(t; h) \, dt < (1 - C\epsilon) \cdot \int_{\tilde{B}} \#\{ k \in \mathbb{N} : E_k(t; h) \in [a, b] \} \, dt
\]

\[
> (1 - C\epsilon)(1 - \bar{\epsilon})m(J) \cdot \#\{ k \in \mathbb{N} : E_k(0; h) \in [a, b - M\delta] \}
\]

where \( \tilde{B} \) is the set constructed in Proposition 6.2.3.

Hence the average amount of time spent by eigenvalues that meet \( W(t; h) \) for some \( t \in J \) is bounded below by

\[
(1 - C\epsilon)(1 - \bar{\epsilon})m(J) \cdot \frac{\#\{ k \in \mathbb{N} : E_k(0; h) \in [a, b - M\delta] \}}{\#\{ k \in \mathbb{N} : E_k(0; h) \in [a - M\delta, b] \}} =: (1 - \eta)m(J).
\]

where \( \eta \to 0 \) if \( \epsilon, \bar{\epsilon}, \) and \( \delta \) do.

It follows that at least \((1 - \eta^{1/2})\) proportion of the eigenvalues \( E_k(t; h) \) that meet \( W(t; h) \) for some \( t \in J \) have speed \( E_k'(t; h) \in [\tilde{Q}_-, \tilde{Q}_+] \) for at least \( m(J)(1 - \eta^{1/2}) \) time.

We denote the collection of such indices \( k \) by \( \mathcal{F} \).

Taking \( E(t; h) := E_k(t; h) \) for some \( k \in \mathcal{F} \), we have

\[
E(t_2; h) - E(t_1; h) > m(J)(1 - \eta^{1/2})\tilde{Q}_-.
\]

On the other hand, we now bound \( E(t_2; h) - E(t_1; h) \) above. To do this, we define \( \tilde{E}(t; h) = E(t; h) - Bt \) and \( \tilde{\mu}_m(t; h) = \mu_m(t; h) - Bt \) where \( B \) was the upper bound in (6.2.31).

Then the transformed quasi-eigenvalue windows \( \tilde{\mu}_m(t; h) \) are non-increasing. From this it follows that if \( \tilde{E}(s; h) \in [\tilde{\mu}_m(s; h) - h^\alpha, \tilde{\mu}_m(s; h) + h^\alpha] \) and \( m \in \mathcal{M}_h(s) \) for some \( s \in J \), then \( \tilde{E}(s'; h) - \tilde{E}(s; h) < 2h^\alpha \), where \( s' \) is the final time \( t \in J \) such that \( m \in \mathcal{M}_h(t) \) and \( \tilde{E}(t; h) \in [\tilde{\mu}_m(t; h) - h^\alpha, \tilde{\mu}_m + h^\alpha] \). This implies that \( E(s'; h) - E(s; h) < 2h^\alpha + B(s' - s) \).

Generalising this idea, we can cover \( \{ t \in J : E(t; h) \in W(t; h) \} \) as a finite union of almost-disjoint intervals \( \bigcup_i J_i \) with \( I_j = [s_j, s'_j] \) defined as follows:

1. We define \( s_0 := \inf\{ t \in J : E(t; h) \in W(t; h) \} \), and we choose an \( m(0) \in \mathcal{M}_h(s_0) \) such that \( E(t; h) \in [\mu_m(0)(t; h) - h^\alpha, \mu_m(0)(t; h) + h^\alpha] \) and \( m(0) \in \mathcal{M}_h(t) \) for all sufficiently small \( t - s_0 > 0 \).
2. We then define \( s_0 := \sup \{ t \in J : E(t; h) \in [\mu_{m(0)}(t; h) - h^\alpha, \mu_{m(0)}(t; h) + h^\alpha] \} \).

3. If \( t \in \{ J : t > s_{j-1} \) and \( E(t; h) \in W(t; h) \} \) is empty, we terminate the inductive process, otherwise we proceed inductively by defining \( s_j := \inf \{ t \in J : t > s_{j-1} \) and \( E(t; h) \in W(t; h) \} \) and choosing a corresponding \( m(j) \in \mathcal{M}_h(s_j) \) such that \( E(t; h) \in [\mu_{m(j)}(t; h) - h^\alpha, \mu_{m(j)}(t; h) + h^\alpha] \) and \( m(j) \in \mathcal{M}_h(t) \) for all sufficiently small \( t - s_{j-1} > 0 \).

4. We then define \( s'_j := \sup \{ t \in J : E(t; h) \in [\mu_{m(j)}(t; h) - h^\alpha, \mu_{m(j)}(t; h) + h^\alpha] \} \).

From the Weyl asymptotics, this procedure must terminate after finitely many iterations.

**Remark 6.3.2.** In the case that \( E(t; h) \) is still in a quasi-eigenvalue window after the window corresponding to \( \mu_{m(j)} \), we will have \( s_{j+1} = s'_j \). This is the only kind of overlap possible between the intervals \( I_j \). We also remark that the \( m(j) \) are necessarily distinct, by the nature of this construction.

For each such interval \( I_j = [s_j, s'_j] \), we have that \( E(s'_j; h) - E(s_j; h) < 2h^\alpha + B(s'_j - s_j) \).

As there can be at most \( O(h^{-n}) \) intervals \( I_j \), we obtain:

\[
\sum_j E(s'_j; h) - E(s_j; h) < B \sum_j (s'_j - s_j) + O(h^{\alpha-n}). \tag{6.3.11}
\]

For such eigenvalues, we thus obtain the upper bound

\[
E(t_2; h) - E(t_1; h) < \sum_j (E(s'_j; h) - E(s_j; h)) + \left( m(J)(1 - \eta^{1/2}) - \sum_j (s'_j - s_j) \right) \tilde{Q}_+ + m(J)\eta^{1/2}M
\]

\[
< (B - \tilde{Q}_+) \sum_j (s'_j - s_j) + m(J)\eta^{1/2}M + m(J)(1 - \eta^{1/2})\tilde{Q}_+ \tag{6.3.12}
\]

in the limit \( h \to 0 \).

Rearranging (6.3.12) and using (6.3.10), we arrive at

\[
(\tilde{Q}_+ - B) \sum_j (s'_j - s_j) < m(J)\eta^{1/2}M + m(J)(1 - \eta^{1/2})(\tilde{Q}_+ - \tilde{Q}_-). \tag{6.3.13}
\]

By taking \( \epsilon, \tilde{\epsilon} \) and \( \delta \) small and then passing to sufficiently small \( h > 0 \) we can thus bound

\[
\frac{1}{|J|} \int_J 1_{A_k}(t; h) \, dt \tag{6.3.14}
\]

by an arbitrarily small positive constant \( \gamma \) for all \( k \in \mathcal{F} \).
Hence there exist constants $c, C > 0$ such that for sufficiently small $\eta$ small and then passing to sufficiently small $h > 0$.

This average being less than 1/2 implies the existence of a $t_\ast \in J$ satisfying the claims of the proposition (possibly dependent on the arbitrarily small parameter $h > 0$).

We proceed as in the beginning of Proposition 3.3.1. We denote by $U$, the $h$-dependent span of all eigenfunctions with eigenvalues in $\mathcal{I}(t_\ast, h)$.

**Proposition 6.3.3.** For sufficiently small $h > 0$, the projections

$$w_m(t_\ast, h) = \pi_U(v_m(t_\ast, h))$$

(6.3.19)

are linearly independent.

**Proof.** First, we show that the estimate from Definition 5.5.1 on the error of quasimodes implies that the projections $\pi_U(v_m(t_\ast, h))$ are small. In particular, for $m \in \mathcal{M}_h(t_\ast)$, we have

$$\| (P_h(t_\ast) - \mu_m(t_\ast, h)) \sum_{j \in \mathbb{N}} \langle v_m(t_\ast, h), u_j(t_\ast, h) \rangle u_j \|^2 < Ce^{-ch^{-1/\rho}}$$

$$\Rightarrow \sum_{|E_j - \mu_m| > h^{\alpha}} |E_j(t, h) - \mu_m(t, h)|^2 \| \langle v_m(t_\ast, h), u_j(t_\ast, h) \rangle \|^2 < Ce^{-ch^{-1/\rho}}$$

$$\Rightarrow \sum_{E_j \notin \mathcal{I}(t_\ast, h)} \| \langle v_m(t_\ast, h), u_j(t_\ast, h) \rangle \|^2 < \frac{Ce^{-ch^{-1/\rho}}}{h^{2\alpha}}.$$ 

Hence there exist constants $c, C > 0$ such that for sufficiently small $h$, we have

$$\| \pi_U(v_m(t_\ast, h)) \|^2 \geq 1 - Ce^{-ch^{-1/\rho}}.$$ 

(6.3.20)

It follows that the $\pi_U(v_m(t, h))$ are almost orthogonal for distinct $m, k \in \mathcal{M}_h(t)$.

$$|\langle \pi_U(v_m(t_\ast, h)), \pi_U(v_k(t_\ast, h)) \rangle| \leq |\langle v_m(t_\ast, h), v_k(t_\ast, h) \rangle| + |\langle \pi_U(\perp v_m(t_\ast, h)), \pi_U(\perp v_k(t, h)) \rangle|$$

$$< Ce^{-ch^{-1/\rho}} + \sqrt{(1 - \| \pi_U(v_m(t, h)) \|^2)(1 - \| \pi_U(v_k(t, h)) \|^2)}.$$ 

Hence there exist constants $c, C > 0$ such that we have

$$|\langle \pi_U(v_m(t_\ast, h)), \pi_U(v_k(t_\ast, h)) \rangle - \delta_{k,m}| \leq Ce^{-ch^{-1/\rho}}.$$ 

(6.3.21)

for all sufficiently small $h$. 
If we enumerate the quasimodes $v_m(t_*, h)$ by positive integers rather than $m \in \mathbb{Z}^n$, we can then form the Gram matrix $M(h) \in \text{Mat}(\#\mathcal{M}_h(t_*), \mathbb{R})$, with entries given by

$$M_{ij}(h) = \langle \pi_U(v_i(t_*, h)), \pi_U(v_j(t_*, h)) \rangle. \quad (6.3.22)$$

Since

$$\|M - I\|_{HS}^2 \leq C(\#\mathcal{M}_h(t_*))^2 e^{-c h^{-1}/\rho}. \quad (6.3.23)$$

for some constants $c, C > 0$, it follows as in Proposition 3.3.1 that we can invert $M = I + (M - I)$ as a Neumann series. Since $M$ is nonsingular, we can therefore conclude that the collection of functions

$$\{\pi_U(v_m(t_*, h)) : m \in \mathcal{M}_h(t_*)\} \quad (6.3.24)$$

is linearly independent.

We are now in a position to complete the proof of Theorem 6.1.3.

**Completion of proof of Theorem 6.1.3.** Having chosen $\delta > 0$ in Proposition 6.2.4, we have shown in Proposition 6.3.1 that if there does not exist an $\epsilon > 0$ satisfying the claim of Theorem 6.1.3, then there exists a $t_* \in [0, \delta]$ at which we have the spectral non-concentration result (6.3.1).

On the other hand, we have shown in Proposition 6.3.3 that the projections $\pi_U(v_m(t_*, h))$ are $\#\mathcal{M}_h(t_*)$ linearly independent vectors in a vector space of dimension $\dim(U) = N(t_*, h) < \#\mathcal{M}_h(t_*)/2$.

This contradiction completes the proof.
Appendix A

Estimates for analytic functions

In this appendix we prove several elementary but important estimates for analytic functions.

**Proposition A.1.** Suppose \( \tilde{\Omega}_j \subset \mathbb{C} \) are open sets and \( \Omega_j \subset \tilde{\Omega}_j \) are such that \( \text{dist}(\Omega_j, \mathbb{C} \setminus \tilde{\Omega}_j) < r_j \).

Define

\[
\Omega = \prod_{j=1}^n \Omega_j \tag{A.1}
\]

and

\[
\tilde{\Omega} = \prod_{j=1}^n \tilde{\Omega}_j. \tag{A.2}
\]

If the analytic function \( f : \tilde{\Omega}^n \to \mathbb{C} \) satisfies

\[
\|f\|_\Omega = A < \infty \tag{A.3}
\]

then we have

\[
\|\partial^\alpha f\|_\Omega \leq Ar^{-\alpha} \alpha! \tag{A.4}
\]

for each multi-index \( \alpha \).

**Proof.** From the Cauchy integral formula, we have

\[
f(z) = \frac{1}{(2\pi i)^n} \oint_{\partial B(z_1, r_1)} \oint_{\partial B(z_2, r_2)} \cdots \oint_{\partial B(z_n, r_n)} \frac{f(w)}{w - z} \, dw_1 \, dw_2 \cdots dw_n. \tag{A.5}
\]

which yields

\[
\partial^\alpha f(z) = \frac{\alpha!}{(2\pi i)^n} \oint_{\partial B(z_1, r_1)} \oint_{\partial B(z_2, r_2)} \cdots \oint_{\partial B(z_n, r_n)} \frac{f(w)}{(w - z)^{\alpha+1}} \, dw_1 \, dw_2 \cdots dw_n. \tag{A.6}
\]

upon repeated differentiation, where \( 1 \) denotes the multi-index \( (1, 1, \ldots, 1) \). Hence

\[
\|\partial^\alpha f\|_\Omega \leq Ar^{-\alpha} \alpha! \tag{A.7}
\]

as required. \( \square \)

**Proposition A.2.** Suppose \( f \) is a \( 1 \)-periodic analytic function on the complex strip

\[
S_\sigma = \{ z : |\text{Im}(z)| < \sigma \} \subset \mathbb{C}^n \tag{A.8}
\]
with the estimate
\[ \|f\|_{S_\sigma} = A < \infty. \]  
(A.9)

Then for \( m \in \mathbb{Z}^n \), the \( m \)-th Fourier coefficient
\[ \hat{f}(m) = \int_{[0,1]^n} e^{-2\pi i m \cdot x} f(x) \, dx \]  
(A.10)
satisfies the estimate
\[ |\hat{f}(m)| \leq e^{-2\pi |m|\sigma} A \]  
(A.11)

Proof. From Cauchy’s theorem and the periodicity of \( f \), we can replace the integral (A.10) with
\[ \int_0^1 e^{-2\pi i m \cdot (x - i\delta q)} f(x - i\delta q) \, dx \]  
(A.12)
for any \( 0 < \delta < \sigma \), where \( q_j = \text{sgn}(m_j) \).

The desired estimate then follows immediately from the triangle inequality and letting \( \delta \to \sigma \). \( \square \)

As a consequence of Proposition A.2, it is a straightforward matter to control the error incurred by truncating the Fourier series of an analytic function.

**Proposition A.3.** Suppose \( f \) is a 1-periodic analytic function on the complex strip
\[ S_\sigma = \{ z : |\text{Im}(z)| < \sigma \} \subset \mathbb{C}^n \]  
(A.13)
with the estimate
\[ \|f\|_{S_\sigma} = A < \infty. \]  
(A.14)

Then we have
\[ \left\| f(x) - \sum_{|m| \leq K} \hat{f}(m)e^{2\pi i m \cdot x} \right\|_{S_{\sigma - \delta}} \leq C(n)K^ne^{-K\delta}\|f\|_{S_\sigma}. \]  
(A.15)
for any \( 0 < \delta < \sigma \).

We also have an implicit function theorem for real analytic functions. Defining
\[ O_h = \{ \omega \in \mathbb{C}^n : \text{dist}(\omega, \Omega) < h \} \]  
(A.16)
where distances in \( \mathbb{C}^n \) are taken with the sup-norm, we have the following.

**Proposition A.4.** Suppose \( f : O_h \times (-1, 1) \to \mathbb{C}^n \) is real analytic, and we have the estimate
\[ |f|_h < \infty, \]  
(A.17)
then for any \( 0 < v < 1/6 \) such that
\[ |f - id|_h \leq vh \]  
(A.18)
the function has a real analytic inverse \( g : O_{(1/2-3v)h} \times (-1, 1) \to O_{(1-4v)h} \) that satisfies the estimate
\[ \max(|g - id|_{(1/2-3v)h}, 3vh|D\phi - Id|_{(1/2-3v)h}) \leq |f - id|_h \]  
(A.19)
uniformly in $t \in (-1, 1)$. The matrix norm in (A.19) is the norm induced by equipping $\mathbb{C}^n$ with the sup-norm.

Proposition A.4 can be proven in the same way as in Lemma 3.4 of [36]. The only difference is that we need to work on domains of the form $O_{\lambda h} \times B^C_{\Gamma}$, and invert maps of the form

$$\tilde{f}(\omega, t) := (f(\omega, t), t)$$

(A.20)

for given $f$ satisfying the assumptions of the proposition uniformly in $t$. 
Estimates for analytic functions
Appendix B

Properties of anisotropic Gevrey classes

In this appendix, we collect several results on anisotropic Gevrey classes from the appendix of [36]. The first of these is an implicit function theorem of Komatsu.

Proposition B.1. Suppose that $F \in G_{L_1,L_2}^{\rho,\rho'}(X \times \Omega^0, \mathbb{R}^n)$ where $X \subset \mathbb{R}^n$, $\Omega^0 \subset \mathbb{R}^m$ and $L_1 \|F(x,\omega) - x\|_{L_1,L_2} \leq 1/2$. Then there exists a local solution $x = g(y,\omega)$ to the implicit equation

$$F(x,\omega) = y$$

(B.1)

defined in a domain $Y \times \Omega$. Moreover, there exist constants $A,C$ dependent only on $\rho,\rho',n,m$ such that $g \in G_{CL_1,CL_2}^{\rho,\rho'}(Y \times \Omega, X)$ with $\|g\|_{CL_1,CL_2} \leq A\|F\|_{L_1,L_2}$.

A consequence of this theorem is established by Popov in [36].

Corollary B.2. Suppose $F \in G_{L_1,L_2}^{\rho,\rho'}(T^n \times \Omega, T^n)$ where $\Omega^0 \subset \mathbb{R}^m$ and $L_1 \|F(\theta,\omega) - \theta\|_{L_1,L_2} \leq 1/2$. Then there exists a local solution $x = g(y,\omega)$ to the implicit equation

$$F(x,\omega) = y$$

(B.2)

defined on $T^n \times \Omega$. Moreover, there exist positive constants $A,C$ dependent only on $\rho,\rho',n,m$ such that $g \in G_{CL_1,CL_2}^{\rho,\rho'}(T^n \times \Omega, T^n)$ with $\|g\|_{CL_1,CL_2} \leq A\|F\|_{L_1,L_2}$.

Finally, we have two results on the composition of functions of Gevrey regularity, which can also be found in [36].

Proposition B.3. Let $X \subset \mathbb{R}^n$, $Y \subset \mathbb{R}^m$, and $\Omega \subset \mathbb{R}^k$ be open sets. Suppose $g \in G_{L_1}^{\rho,\rho'}(\Omega,Y)$ with $\|g\|_{L_1} = A_1$ and $f \in G_{B,L_2}^{\rho,\rho'}(X \times Y)$ with $\|f\|_{B,L_2} = A_2$. Then the composition $F(x,\omega) := f(x,g(\omega))$ is in $G_{B,L}^{\rho,\rho'}(X \times \Omega)$, where

$$L = 2^{l} + \rho' L_1 \max(1,A_1L_2)$$

with $l = \max(k,m,n)$. Moreover we have the Gevrey norm estimate

$$\|F\|_{B,L} \leq A_2$$

Proposition B.4. Let $X \subset \mathbb{R}^n$, $Y \in \mathbb{R}^m$, and $\Omega \subset \mathbb{R}^k$ be open sets. Suppose $g \in G_{B_1,L_1}^{\rho,\rho'}(X \times \Omega,Y)$ with $\|g\|_{B_1,L_1} = A_1$ and $f \in G_{B_2,L_2}^{\rho,\rho'}(Y \times \Omega)$. Then the composition
$F(x, \omega) := f(g(x, \omega), \omega)$ is in $G_{B,L}^{\rho, \rho'}(X \times \Omega)$, where

$$B = 4^l(4l)^\rho B_1 \max(1 + A_1 B_2)$$

and

$$L = L_2 + 4^l(4l)^\rho L_1 \max(1, A_1 B_2)$$

with $l = \max(k, m, n)$. Moreover we have the Gevrey norm estimate

$$\|F\|_{B,L} \leq A_2.$$
Appendix C

Whitney extension theorem

In this appendix, we prove a version of the Whitney extension theorem for anisotropic Gevrey classes. The proof is adapted from the work of Bruna [5] in the case without in the non-anisotropic case.

Definition C.1.

\[ C^\infty_{M, \tilde{M}}(X \times Y) = \{ f \in C^\infty(X \times Y, \mathbb{R}) : \sup_{(x,y) \in X \times Y} \sup_{\alpha, \beta} \left( \frac{|(\partial^\alpha_x \partial^\beta_y f)(x, y)|}{L_1^{\alpha} L_2^{\beta} M_\alpha \tilde{M}_\beta} \right) < \infty \text{ for some } L_j > 0 \} \]  

(C.1)

where \( X, Y \) are open sets in Euclidean spaces of possibly differing dimension, \( \alpha, \beta \) are multi-indices of the appropriate dimension, and \( M \) and \( \tilde{M} \) are positive sequences satisfying

1. \( M_0 = 1 \)
2. \( M_k^2 \leq M_{k-1} M_{k+1} \)
3. \( M_k \leq A^k M_j M_{k-j} \)
4. \( M_{k+1} \leq A^k M_{k+1} \)
5. \( M_{k+1}/(k M_k) \) is increasing
6. \( \sum_{k \geq 0} M_k/M_{k+1} \leq A p M_p/M_{p+1} \text{ for } p > 0 \)

where \( A > 0 \) is a positive constant.

In the Gevrey case of interest to us, \( M_k = k!^{\rho_1}, \tilde{M}_k = k!^{\rho_2} \).

For fixed \( L_j > 0 \), the supremum in (C.1) defines a norm which equips a subspace of \( C^\infty_{M, \tilde{M}}(X \times Y) \) with a Banach space structure.

The space \( C^\infty_{M, \tilde{M}}(X \times Y) \) is then the inductive limit of these spaces as \( L = L_1 = L_2 \to \infty \), which identifies it a Silva space.

For \( f \in C^\infty_{M, \tilde{M}}(X \times Y) \), and \( z = (z_1, z_2) \in X \times Y, x \in X \) we define

Definition C.2.

\[ (T^m_x f)(z) := \sum_{|\alpha| \leq m} \frac{(\partial^\alpha_x f)(x, z_2)}{\alpha!} (z_1 - x)^\alpha \]  

(C.2)

Definition C.3.

\[ (R^m_x f)(z) := f(z) - (T^m_x f)(z). \]  

(C.3)

To slightly generalise this notation, for a jet \( f^{\alpha, \beta} \) of continuous functions, we write
Definition C.4.

\[(R^nf)_{\alpha,\beta}(z) := f^{\alpha,\beta}(z) - (T^m_x |\alpha| f^{\alpha,\beta})(z)\]  

(C.4)

We can now pose the central question:

Given a compact set \(K \subset X\), under what conditions is it true that an arbitrary continuous jet \(\tilde{f}^{\alpha,\beta} : K \times X \to \mathbb{R}\) is the jet of a function \(f \in \mathcal{C}_{M,K}^\infty (X \times Y)\)?

We assume without loss of generality here that the set \(X\) is a full Euclidean space \(\mathbb{R}^d\), rather than just an open subset thereof.

This question is the anisotropic non quasi-analytic analogue of Whitney’s extension theorem from classical analysis, which deals with the \(\mathcal{C}^\infty\) case.

We begin by finding necessary conditions for the existence of such an extension, before proving that these conditions are indeed sufficient.

**Proposition C.5.** Suppose \(f \in \mathcal{C}_{M,K}^\infty (X \times Y)\) with Gevrey constants \(L_1, L_2\). Then there exists a constant \(A\) dependent only on the dimensions of \(X, Y\) and on \(M, \tilde{M}\) such that the jet \(f^{\alpha,\beta} = \partial_x^{\alpha} \partial_y^{\beta} f\) satisfy

\[|f^{\alpha,\beta}| \leq AL_1^{[\alpha]} L_2^{[\beta]} M_{[\alpha]} \tilde{M}_{[\beta]}\]  

(C.5)

and

\[|(R^nf)_{k,l}(z)| \leq A\tilde{L}_1^{n+1} M_{n+1} L_2^{[l]} \tilde{M}_{[l]} \cdot |z_1 - x|^{n+1} \]  

(C.6)

for all non-negative integers \(m, n\) and all multi-indices \(|k| \leq m, |l| \leq n\), where \(\tilde{L}_1 = CL_1\) with the \(C\) dependent only on the dimension of \(X\).

**Proof.** The first estimate (C.5) follows immediately from the definition of \(\mathcal{C}_{M,K}^\infty (X \times Y)\).

We prove the second claim (C.6) by making use of the estimate (C.5) on the jet \(f^{\alpha,\beta} = \partial_x^{\alpha} \partial_y^{\beta} f\) and Taylor expansion.

\[R^nf(z) = \sum_{|\alpha| = n+1} \frac{n+1}{\alpha!} (z_1 - x)^\alpha \int_0^1 (1-t)^n f^{\alpha,0}(x + t(z_1 - x), z_2) \, dt\]  

(C.7)

\[\leq \left( \sup_{|\alpha| = n+1} \sup_{z \in X \times Y} |f^{\alpha,0}(z)| \right) \cdot \sum_{|\alpha| = n+1} \frac{|z_1 - x|^{\alpha}}{\alpha!}\]  

(C.8)

\[\leq \left( \sup_{|\alpha| = n+1} \sup_{z \in X \times Y} |f^{\alpha,0}(z)| \right) \cdot \frac{C^{n+1} |z_1 - x|^{n+1}}{(n+1)!}\]  

(C.9)

Hence

\[|(R^nf)_{k,l}(z)| = |(R^nf)_{k,l}(z)| \leq A\tilde{L}_1^{n+1} M_{n+1} L_2^{[l]} \tilde{M}_{[l]} \cdot |z_1 - x|^{n+1} \]  

(C.11)

as required.

Subsequently, for simplicity of notation, we omit the tilde in \(\tilde{L}_1\) with the understanding that we are allowed to absorb constants that are dependent only on the dimensions of \(X, Y\) and on the sequences \(M, \tilde{M}\).
Theorem C.6. Suppose \((f^{\alpha,\beta}) : K \times Y \to \mathbb{R}\) is a jet of continuous functions smooth in \(y\) that satisfies
\[
\partial_y^\gamma (f^{\alpha,\beta}) = f^{\alpha,\beta+\gamma}
\] (C.12)
as well as the conditions (C.5) and (C.6) on \(K \times Y\). Then there exists a function \(f \in C^\infty_{\tilde{M}}(X \times Y)\) such that \(\partial_x^{\alpha,\beta} f = f^{\alpha,\beta}\) on \(K \times Y\).

Moreover, there exist constants \(C_0, C_1\) dependent only on the dimensions of \(X\) and \(Y\) and the weight sequences \((M_k), \tilde{M}_k\) such that
\[
\|f\|_{C^1_{L_1, L_2}} \leq C_0 A. \quad \text{(C.13)}
\]

Before proving Theorem C.6, we need to collect some lemmas, the proofs of which can be found in [5].

Proposition C.7. Suppose \(K \subset \mathbb{R}^d\) is compact. Then there exists a collection of closed cubes \(\{Q_j\}_{j \in \mathbb{N}}\) with sides parallel to the axes such that

1. \(\mathbb{R}^d \setminus K = \bigcup_j Q_j\);
2. \(\text{int}(Q_j)\) are disjoint;
3. \(\delta_j := \text{diam}(Q_j) \leq d_j := d(Q_j, K) \leq 4\delta_j\);
4. For \(0 < \lambda < 1/4\), \(d(z, K) \sim \delta_j\) for \(z \in Q_j^* := (1 + \lambda)Q_j\);
5. Each \(Q_i^*\) intersects at most \(D = (12)^{2d}\) cubes \(Q_j^*\);
6. \(\delta_i \sim \delta_j\) if \(Q_i^* \cap Q_j^* \neq \emptyset\).

Proposition C.8. For each \(\eta > 0\), there exists a family of functions \(\phi_i \in C^\infty_{\tilde{M}}(\mathbb{R}^d)\) such that

1. \(0 \leq \phi_i\);
2. \(\text{supp}(\phi_i) \subset Q_i^*\);
3. \(\sum_i \phi_i(z) = 1\) for \(z \in \mathbb{R}^d\);
4. \(|\partial^\alpha_i \phi_i(z)| \leq Ah(B\eta d(z, K))|\eta|^{\alpha} |M_{\alpha}|\) for \(z \in Q_i^*\).

where \(A, B > 0\) are constants and
\[
h(t) := \sup_k \frac{k!}{t^k M_k}. \quad \text{(C.14)}
\]

Proposition C.9. Suppose \(T \in \mathcal{L}(E, F)\) is a continuous linear surjection between Silva spaces. Then for any bounded set \(B \subset F\), there exists a bounded set \(C \subset E\) with \(T(C) = B\).

We also require an anisotropic version of Carleman’s theorem, which is the special case of C.6 with \(K = \{0\}\), and Gevrey analogue of Borel’s theorem from classical analysis.

Proposition C.10. Let \((g_\alpha)_{\alpha \in \mathbb{N}_d}\) be a multisequence of functions in \(C^\infty_{\tilde{M}}(Y)\) such that
\[
|\partial_y^\alpha g_\alpha(y)| \leq KL_1^{\alpha} |L_2||M_{\alpha}| \tilde{M}_{|\alpha|}. \quad \text{(C.15)}
\]
for some constant \(K > 0\).
Then there exists a function \( f \in C_{M,\tilde{M}}^\infty(X \times Y) \) such that \( g_\alpha(y) = \partial_x^\alpha f(0, y) \) for all \( y \in Y \). Moreover, \( \|f\|_{C^1 L^2} \leq AK \) for some constants \( A, C > 0 \) independent of \( f, L_1, \) and \( L_2 \).

**Proof.** We adapt the solution of [33] of the classical Carleman problem to this setting. Key is that the assumptions on \( M \) imply that the hypotheses of [33] are satisfied. Hence as in the proof of [33] Theorem 2.1 (ai), we can construct compactly supported \( \chi_p(x) \in C_{M_p}(\mathbb{R}) \) for each non-negative integer \( p \) such that

\[
\chi_p^{(k)}(0) = \delta(k, p) \tag{C.16}
\]

and

\[
\|\chi_p\|_{L^{2+A^{-1}}} \leq \frac{1}{M_p} \left( \frac{Ae}{L} \right)^p \tag{C.17}
\]

for some dimensional constant \( A \) and any \( L > 0 \).

Hence we can define

\[
\chi_\alpha(x) := \prod_{j=1}^d \chi_{\alpha_j}(x_j) \tag{C.18}
\]

for \( \alpha \in \mathbb{N}^d \) which satisfies

\[
\chi^{(\beta)}_\alpha(0) = \delta(\beta, \alpha). \tag{C.19}
\]

Moreover, we have the estimate

\[
|\chi^{(\beta)}_\alpha| = \prod_{j=1}^d |\chi^{(\beta_j)}_{\alpha_j}| \tag{C.20}
\]

\[
\leq \prod_{j=1}^d \frac{1}{M_{\alpha_j}} \left( \frac{Ae}{L} \right)^{\alpha_j} (L(2 + A^{-1}))^{\beta_j} M_{\beta_j} \tag{C.21}
\]

\[
\leq \left( \frac{Aec(d, M)}{L} \right)^{|\alpha|} M_{|\alpha|}^{-1} (L(2 + A^{-1}))^{|\beta|} M_{|\beta|}. \tag{C.22}
\]

By taking \( L = 2CL_1 = 2Aec(d, M)L_1 \), we can estimate

\[
|\partial_x^k \partial_y^l (\chi_\alpha(x)g_\alpha(y))| \leq K \left( (C/L)^{|\alpha|} M_{|\alpha|}^{-1} (L(2 + A^{-1}))^{|k|} M_{|k|} \right) \cdot \left( L_1^{[\alpha]} L_2^{[\alpha]} M_{|\alpha|} \tilde{M}_{|\beta|} \right) \tag{C.23}
\]

\[
\leq K \cdot 2^{-|\alpha|} (2CL_1(2 + A^{-1}))^{[\alpha]} L_2^{[\alpha]} M_{|\alpha|} \tilde{M}_{|\beta|}. \tag{C.24}
\]

Where \( A, C, \) and \( K \) are constants independent of \( f, L_1, \) and \( L_2 \).

Hence we have that \( \|\chi_\alpha(x)g_\alpha(y)\|_{2CL_1(2+A^{-1}),L_2} \leq K \cdot 2^{-|\alpha|} \). It follows that

\[
f(x, y) := \sum_{\alpha \in \mathbb{N}^d} \chi_\alpha(x)g_\alpha(y) \tag{C.25}
\]

converges in the \( C_{M,\tilde{M}}^\infty(X \times Y) \) sense, and satisfies \( \partial_x^\alpha f(0, y) = g_\alpha(y) \) as required.  

Equipped with these tools, we are ready to prove Theorem C.6.

**Proof of Theorem C.6.** We begin by estimating the difference in Taylor expansions about different points in \( K \).
Using the identity

\[(T^a_x f)(z) - (T^a_y f)(z) = \sum_{|\alpha| \leq n} (z_1 - x)^\alpha (R^\alpha_y f)_{\alpha,0}(x, z_2)\]  

(C.26)

we can estimate

\[|\partial_z^{k,l}((T^a_x f)(z) - (T^a_y f)(z))| \leq AL_1^{n+1}M_{n+1}L_2^{\ell_2} M_\ell \frac{|z_1 - x| + |z_1 - y|^{n-|\ell|+1}}{(n - |\ell| + 1)!}.\]  

(C.29)

We now invoke Proposition C.10.

For \(x \in X\) consider the map \(T_x : C^\infty_{M,\tilde{M}}(X \times Y) \rightarrow G_x\) given by \((T_x f)_{\alpha}(y) := f^{\alpha,0}(x, y)\) where the space \(G_x\) consists of all multisequences of analytic functions \(f_{\alpha} : Y \rightarrow \mathbb{R}\) satisfying \(|f_{\alpha}| \leq AL_1^{[\alpha]}L_2^{[\beta]} M_{[\alpha]} M_{[\beta]}\) for some \(A > 0\).

From the assumed estimate (C.5) on \(f^{\alpha,\beta}\), Proposition C.10 applies, and for each \(x \in K\), we can find a function \(f_x \in C^\infty_{M,\tilde{M}}(X \times Y)\) such that

\[\partial_z^{\alpha,\beta} f_x(x, z_2) = f^{\alpha,\beta}(x, z_2)\]  

(C.30)

for each \(\alpha, \beta\). Moreover, the conclusion of Proposition C.10 implies that there exist constants \(B = C_0 A, K_1 = C_1 L_1, K_2 = L_2 > 0\) such that the estimate

\[|\partial_z^{\alpha,\beta} f_x(z)| \leq BK_1^{[\alpha]}K_2^{[\beta]} M_{[\alpha]} M_{[\beta]}\]  

(C.31)

holds uniformly, where \(C_j\) depend only on the dimensions of \(X\) and \(Y\) and the weight sequences \(M_k, \tilde{M}_k\).

Hence we can bound

\[|\partial_z^{k,l}(f_x(z) - (T^m_x f_x)(z))| \leq |(R_x^{m,n} f_x)_{k,l}(z)|\]  

(C.32)

using the same calculation as in Proposition C.5. We obtain

\[|\partial_z^{k,l}(f_x(z) - (T^a_y f)(z))| \leq A(C_1 L_1)^{n+1}M_{n+1}L_2^{\ell_2} M_\ell \frac{|z_1 - x|^{n-|\ell|+1}}{(n - |\ell| + 1)!}.\]  

(C.34)

The upshot of this estimate is that we can replace \(T^a_x f\) and \(T^a_y f\) in (C.29) with \(f_x\) and \(f_y\) respectively.

That is, we have

\[|\partial_z^{k,l}(f_x(z) - f_y(z))| \leq A(C_1 L_1)^{n+1}M_{n+1}L_2^{\ell_2} M_\ell \frac{|z_1 - x| + |z_1 - y|^{n-|\ell|+1}}{(n - |\ell| + 1)!}.\]  

(C.35)

We now fix \(k, l\) and vary \(n \geq k\) in order to optimise the upper bound (C.35).
By defining the quantity
\[ h(t) := \sup_{k \geq 0} \frac{k!}{t^k M_k} \] (C.36)
as in [5] we obtain
\[ |\partial_z^{k,l}(f_x(z) - f_y(z))| \leq A(C_1 L_1)^{|k| M_{|k|} L_2^{|l|} M_{|l|}} h((C_1 L_1)(|z_1 - x_1| + |z_1 - y_1|))^{-1}. \] (C.37)
by using property (3) following Definition C.1.

The next step in the construction is to use Proposition C.8 to piece together the functions \( f_x \) using a \( C_M^\infty \) partition of unity subordinate to the cover arising from the decomposition of \( X \setminus K \) by cubes in Proposition C.7.

Taking the collection \( \{Q_j\}_{j \in N} \) of cubes in \( X = \mathbb{R}^d \) constructed by Proposition C.7, we choose \( x_j \in K \) such that \( d(x_j, Q_j) = d(Q_j, K) \).

Note that the conclusion of Proposition C.7 implies that
\[ |z - x_j| \sim d(z, K) \] (C.38)
for all \( z \in Q_j \).

Now taking \( \phi_j \) as in Proposition C.8, we define:
\[ \tilde{f}(z) := \begin{cases} f(z) & \text{if } z_1 \in K \\ \sum_i \phi_i(z_1) f_{x_i}(z) & \text{if } z \in X \setminus K. \end{cases} \] (C.39)

Note that since the partition of unity \( \{\phi_j\} \) is locally finite, the function \( \tilde{f}(z) \) is smooth in \( (X \setminus K) \times Y \).

It remains to check that \( \tilde{f} \) is smooth elsewhere, and moreover that \( \tilde{f} \in C_M^\infty(X \times Y) \).

To this end, for \( x \in K \) and \( z_1 \in X \setminus K \), we estimate
\[ \partial_z^{\alpha,\beta}(\tilde{f}(z) - f_x(z)) = \sum_{k \leq \alpha} \left( \sum_i \partial_z^{\alpha}(\phi_i)(z_1) \cdot \partial_z^{\alpha-\beta}(f_{x_i}(z) - f_x(z)) \right). \] (C.40)

First we estimate the \( k = 0 \) term. If \( z_1 \in \text{spt}(\phi_i) = Q_i^s \), we have
\[ d(z_1, x_i) \sim d(z_1, K) \leq d(z_1, x) \] (C.41)
and hence we have
\[ |\sum_i \phi_i(z_1) \cdot \partial_z^{\alpha,\beta}(f_{x_i}(z) - f_x(z))| \leq A(C_1 L_1)^{|\alpha| M_{|\alpha|} L_2^{|\beta|} M_{|\beta|}} h((C_1 L_1)(|z_1 - x|))^{-1}. \] (C.42)
from (C.37).

We now estimate the terms with \( |k| > 0 \). For \( x \in X \setminus K \), we choose \( \bar{x} \in K \) with \( d(x, \bar{x}) = d(x, K) \).

Since \( \sum_i \partial^k \phi_i = 0 \), we have
\[ \sum_i (\partial^k \phi_i)(z_1) \cdot \partial_z^{\alpha-\beta}(f_{x_i}(z) - f_x(z)) = \sum_i (\partial^k \phi_i)(z_1) \cdot \partial_z^{\alpha-\beta}(f_{x_i}(z) - f_{\bar{x}}(z)). \] (C.43)
Now as before, we exploit the fact that $d(z_1, x_i) \sim d(z_1, K)$ to bound

$$|\partial_{x}^{\alpha} f_x(z) - f(z)| \leq A(C_1 L_1)^{1-|\alpha|} M_{[\alpha]} L_2^{[\beta]} M_{[\beta]} h((C_1 L_1) d(z_1, K))^{-1}. \quad (C.44)$$

Since $\log(M_j)$ is an increasing convex sequence with first term 0, it is also superadditive, and we have $M_{|k|} M_{[\beta]} \leq M_{|k|+|\beta|}$. Hence for $|k| \geq 1$, we can use property (4) in Proposition C.8 to conclude that

$$\sum_{i} (\partial_{x}^{\alpha} f_i(z_1) - f_x(z)) \leq A M_{[\alpha]} M_{[\beta]} (C_1 L_1)^{1-|\alpha|} L_2^{[\beta]} \eta^{[\beta]} h((C_1 L_1) d(z_1, K))^{-1}. \quad (C.45)$$

where $\eta$ remains to be chosen.

Equation (15) from [5] implies the existence of a constant $c > 0$ such that

$$\frac{h(t)}{h(ct)} \leq A$$

for some $A > 0$.

Hence we choose $\eta = (C_1 L_1)/cB$ to arrive at the estimate

$$\sum_{i} (\partial_{x}^{\alpha} f_i(z_1) - f_x(z)) \leq A(C_1 L_1)^{1-|\alpha|} L_2^{[\beta]} M_{[\alpha]} M_{[\beta]} \eta^{[\beta]} h((C_1 L_1) |z_1 - x|)^{-1}. \quad (C.47)$$

Combining (C.42) and (C.47), we arrive at

$$|\partial_{z}^{\alpha} (\tilde{f}(z) - f_x(z))| \leq A L_2^{[\beta]} M_{[\alpha]} \tilde{M}_{[\beta]} ((C_1 L_1) + \eta)^{[\alpha]} h((C_1 L_1) |z_1 - x|)^{-1} \quad (C.48)$$

for $z \in (X \setminus K) \times Y$.

The estimate (C.48) is key to proving $\tilde{f} \in C^\infty(X \times Y)$ (and that the derivatives coincide with the values given by the jet $f^{\alpha, \beta}$), as well as the subsequent deduction of $C^\infty_{M, \tilde{M}}$ regularity.

We write

$$\tilde{f}^{\alpha, \beta}(z) := \begin{cases} \partial_{z}^{\alpha} \tilde{f}(z) & \text{if } z_1 \in X \setminus K, \\ f^{\alpha, \beta}(z) & \text{if } z_1 \in K. \end{cases} \quad (C.49)$$

The smoothness of each $\tilde{f}^{\alpha, \beta} : X \times Y \to \mathbb{R}$ readily follows from the fact that each $f^{\alpha, \beta} : K \times Y \to \mathbb{R}$ is smooth in $y$, together with the estimate

$$|\tilde{f}^{\alpha, \beta}(z) - \partial_{z}^{\alpha} T^m f(z)| = o(|z_1 - x|^{-|\alpha|}). \quad (C.50)$$

For $z$ with $z_1 \in K$, the estimate (C.50) comes immediately from (C.6) on $K \times Y$. Otherwise, it is a consequence of the estimate (C.48), the defining property (C.30) of the functions $f_x$, and the fact that the function $h(t)$ increases faster than any polynomial in $t^{-1}$ as $t \to 0$.

Finally, we need to check $C^\infty_{M, \tilde{M}}$ regularity. That is, we need to verify that the Gevrey estimate

$$\|f\|_{C^1 L_1, L_2} \leq C_0 A. \quad (C.51)$$

for some constants $C_0, C_1$ dependent only on the dimensions of the spaces $X$ and $Y$ and the weight sequences $M_k, \tilde{M}_k$. 

In light of (C.5), it only remains to prove (C.51) on $(X \setminus K) \times Y$, and by multiplication by a cutoff function we may assume $d(z_1, K)$ is bounded.

Then, by applying (C.48) with $x = \bar{z}_1$ we can further reduce the problem to verifying (C.51) for $f_x$, uniformly in $x \in K$. However this was established earlier in (C.31).

Hence, the proof is complete. \qed
In this section, we prove the following abstract lemma that we have used several times to assemble full density subsequences along which a given function has limit 0.

**Lemma D.1.** If there exists a function \( g : \mathbb{N} \to \mathbb{R}^+ \) and a family of subsets \( S_j \subset \mathbb{N} \) such that

\[
\liminf_{n \to \infty} \frac{\# \{k \leq n : k \in S_j\}}{n} > d - \epsilon_j
\]

and

\[
\limsup_{n \in S_j \to \infty} g(n) < \epsilon'_j
\]

where \( \epsilon_j, \epsilon'_j \downarrow 0 \), then there exists a subset \( S \subset \mathbb{N} \) such that

\[
\liminf_{n \to \infty} \frac{\# \{k \leq n : k \in S\}}{n} \geq d
\]

and

\[
\lim_{n \in S \to \infty} g(n) = 0.
\]

**Proof.** For ease of notation, we define

\[
d_n(A) = \frac{\# \{k \leq n : k \in A\}}{n}
\]

for \( A \subseteq \mathbb{N} \) and \( n \in \mathbb{N} \).

We have \( g(n) < 2\epsilon'_j \) for cofinitely many elements of \( S_j \), and we denote these sets by \( S'_j \). Now let

\[
B_j = \{k \in \mathbb{N} : g(k) \geq 2\epsilon'_j\} \subseteq \mathbb{N} \setminus S'_j.
\]

Since each \( d_n \) respects the partial ordering of set inclusion and is additive with respect to disjoint unions, we can construct a strictly increasing sequence \( (N_j)_{j \in \mathbb{N}} \) such that \( N_1 = 1 \) and \( d_n(B_j) < 1 - d + 2\epsilon_j \) for all \( n \geq N_j \).

We define

\[
B = \bigcup_{j \in \mathbb{N}} B_j \cap [N_j, \infty).
\]

If \( n \in [N_j, N_{j+1}) \), then any \( k \in [1, n] \cap B \) must lie in \( B_i \) for some \( i \leq j \) and hence in \( B_j \).

This implies that for \( n \in [N_j, N_{j+1}) \) we have \( d_n(B) \leq d_n(B_j) < 1 - d + 2\epsilon_j \) and consequently, that \( \limsup_{n \to \infty} d_n(B) \leq 1 - d \).

We now take \( S := \mathbb{N} \setminus B \), with the required density bound.
\[
\liminf_{n \to \infty} d_n(S) \geq d. \tag{D.8}
\]

To complete the proof we observe that if \( n \in [N_j, \infty) \cap S \), then \( n \in \mathbb{N} \setminus B_i \) for each \( i \leq j \), and hence \( g(n) < 2\epsilon_j \). This establishes (D.4). \(\square\)

For sequences without well-defined natural densities, an analogue of Lemma D.1 holds for the notion of upper density, with an easier proof.

**Lemma D.2.** If there exists a function \( g : \mathbb{N} \to \mathbb{R}^+ \) and a family of subsets \( S_j \subset \mathbb{N} \) such that
\[
\limsup_{n \to \infty} \frac{\#\{k \leq n : k \in S_j\}}{n} > d - \epsilon_j \tag{D.9}
\]
and
\[
\limsup_{n \in S_j \to \infty} g(n) < \epsilon_j' \tag{D.10}
\]
where \( \epsilon_j, \epsilon_j' \downarrow 0 \), then there exists a subset \( S \subset \mathbb{N} \) such that
\[
\limsup_{n \to \infty} \frac{\#\{k \leq n : k \in S\}}{n} \geq d \tag{D.11}
\]
and
\[
\lim_{n \in S \to \infty} g(n) = 0. \tag{D.12}
\]

**Proof.** For each \( j \in \mathbb{N} \), we have \( d_n(S_j) > d - \epsilon_j \) for infinitely many \( N \). We define \( N_0 = 0 \) and inductively choose an increasing sequence \( (N_k) \) of positive integers such that \( d_{N_k}(S_k) > d - \epsilon_k \) and such that \( g(n) < \epsilon_j' \) for \( n > N_k \) in \( S_k \).

We can then construct the set
\[
S := \bigcup_{k=1}^{\infty} \left( \bigcup_{j \geq k} S_j \right) \cap (N_{k-1}, N_k]. \tag{D.13}
\]

Since \( d_{N_k}(S) \geq d_{N_k}(S_k) > d - \epsilon_k \), we thus obtain
\[
\limsup_{n \to \infty} d_n(S) \geq d. \tag{D.14}
\]

Moreover, for \( n \in S \) with \( n > N_k \), we have \( g(n) < \epsilon_j' \), and hence \( g(n) \to 0 \) along \( S \). This completes the proof. \(\square\)
Bibliography


