ANALYSIS & PDE

Volume 16 No. 1 2023

SEÁN GOMES

GENERIC KAM HAMILTONIANS ARE NOT QUANTUM ERGODIC

msp

ANALYSIS AND PDE
Vol. 16 (2023), No. 1, pp. 119–171
DOI: 10.2140/apde.2023.16.119

GENERIC KAM HAMILTONIANS ARE NOT QUANTUM ERGODIC

SEÁN GOMES

We show that under generic conditions, the quantisation of a 1-parameter family of KAM perturbations $P(x, \xi; t)$ of a completely integrable and Kolmogorov nondegenerate Gevrey smooth Hamiltonian is not quantum ergodic for a full-measure subset of parameter values $t \in (0, \delta)$.

1. Introduction

1A. *Hamiltonian dynamics.* Let M be a compact boundaryless Riemannian G^ρ smooth manifold of dimension $n \geq 2$, and let $P(x, \xi) \in \mathcal{C}^\infty(T^*M)$ be a completely integrable Hamiltonian with $P(x, \xi) \to \infty$ as $|\xi| \to \infty$. Complete integrability is the assumption that there exist n functionally independent conserved quantities of the Hamiltonian flow that are pairwise in involution.

The Liouville–Arnold theorem asserts that we can locally choose symplectomorphisms

$$\chi : \mathbb{T}^n \times D \to T^*M \tag{1A.1}$$

such that the transformed Hamiltonian

$$H^0(\theta, I) = (P \circ \chi)(\theta, I) \tag{1A.2}$$

is independent of θ. It follows that the Hamiltonian flow is quasiperiodic and constrained to n-dimensional Lagrangian tori, given in local coordinates by

$$\dot{I} = 0, \quad \dot{\theta} = \nabla H(I). \tag{1A.3}$$

Under the Kolmogorov nondegeneracy condition $\det(\nabla_I^2 H) \neq 0$, we can locally index the invariant Lagrangian tori Λ_ω by the frequency $\omega = \nabla_I H$ of their quasiperiodic motion.

Gomes was supported by an Australian postgraduate award and research training program stipend.
MSC2010: 58J51.
Keywords: quantum ergodicity, KAM Hamiltonians.

If we now consider a smooth one-parameter family of perturbed Hamiltonians given by $H(\theta, I; t)$ in action-angle coordinates with $H(\theta, I; 0) = H^0(I)$, a natural question is whether or not an invariant tori Λ_ω with quasiperiodic motion of frequency ω still exists in the perturbed Hamiltonian dynamics.

This question was resolved positively by Kolmogorov [1954], Arnold [1983], and Moser [1966]. In particular, they established that the Lagrangian invariant tori corresponding to all but an $o(1)$-symplectic-measure subset of frequencies survive this perturbation as the size of the perturbation tends to zero.

In particular persisting tori are those with frequencies ω in a set Ω_κ determined by the Diophantine condition (3B.2), where $\tau > n-1$ is fixed and the choice of κ then dictates the measure of the union of preserved tori.

Popov [2004b] used a local version of the KAM theorem to construct a Birkhoff normal form for Gevrey class Hamiltonians H about Λ. This normal form generalises the notion of "action-angle" variables of a completely integrable Hamiltonian as discussed in [Arnold 1989]. As a consequence of the normal form construction, Popov obtained 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 earlier work in [Popov 2000a; 2000b], where a Birkhoff normal form is constructed for real-analytic Hamiltonians.

1B. ***Quantum ergodicity.*** We now consider the quantisation of a KAM Hamiltonian system given by a family of self-adjoint and uniformly elliptic semiclassical pseudodifferential operators

$$\mathcal{P}_h(t) = \sum_{j=0}^{m} P_j(x, hD; t) h^j, \tag{1B.1}$$

with real-valued full symbol in the Gevrey class $S_\ell(T^*M)$ from Definition B.5, analytic in the parameter t, where $\ell = (\rho, \mu, \nu)$, with $\rho(\tau+n)+1 > \mu > \rho' = \rho(\tau+1)+1$ and $\nu = \rho(\tau+n+1)$. Furthermore, we assume $\mathcal{P}_h(t)$ acts on half-densities in $\mathcal{C}^\infty(M; \Omega^{1/2})$ with principal symbol $P_0(x, \xi; t)$ completely integrable and nondegenerate at $t = 0$, and with vanishing subprincipal symbol. The operator $\mathcal{P}_h(t)$ then has an orthonormal basis of eigenfunctions $u_j(t; h)$ and corresponding real eigenvalues $E_j(t; h) \to \infty$ for each fixed t, h.

The Bohr correspondence principle asserts that aspects of the classical dynamics should be reflected in the spectral theory of $\mathcal{P}_h(t)$ in the semiclassical limit $h \to 0$. A rigorous manifestation of this correspondence principle is the celebrated quantum ergodicity theorem, due to [Shnirelman 1974; Colin de Verdière 1985; Zelditch 1987], which asserts that billiards with ergodic geodesic flow have eigenfunctions satisfying a quantum notion of equidistribution, made precise using the machinery of pseudodifferential operators.

We work with a semiclassical formulation of quantum ergodicity. Let $d\mu_E$ denote the measure on the energy surface $\Sigma_E = p^{-1}(E)$ induced by the symplectic measure $|d\xi \wedge dx|$ on T^*M by

$$|d\mu_E \wedge dE| = |d\xi \wedge dx|. \tag{1B.2}$$

If a Hamiltonian $p(x, \xi) \in \mathcal{C}^\infty(T^*M)$ generates an ergodic Hamiltonian flow on every energy surface Σ_E with $E \in [a, b]$ and $dp|_{p^{-1}([a,b])} \neq 0$, then for any semiclassical pseudodifferential operator A of semiclassical order 0, the quantum ergodicity theorem states that

$$h^n \sum_{E_j(h)\in[a,b]} \left| \langle A_h u_j(h), u_j(h)\rangle - \frac{1}{\mu_{E_j}(\Sigma_{E_j})} \int_{\Sigma_{E_j}} \sigma(A)\, d\mu_{E_j} \right|^2 \to 0. \tag{1B.3}$$

The quantum ergodicity theorem is originally due to Shnirelman [1974], Zelditch [1987], and Colin de Verdière [1985]. The semiclassical formulation of the quantum ergodicity theorem (1B.3) is a straightforward consequence of the sharper formulation in [Helffer et al. 1987], or [Dyatlov and Guillarmou 2014], in which the statement is localised to $O(h)$ energy bands. From (1B.3), a standard diagonal argument introduced in [Colin de Verdière 1977] shows that

$$\lim_{h\to 0} \left| \langle A_h u_j(h), u_j(h)\rangle - \frac{1}{\mu_{E_j}(\Sigma_{E_j})} \int_{\Sigma_{E_j}} \sigma(A)\, d\mu_{E_j} \right| = 0 \tag{1B.4}$$

uniformly for a family $\Lambda(h) \subset \{E_j(h) \in [a,b]\}$ of full density, in the sense that

$$\frac{\#\Lambda(h)}{\#\{E_j(h) \in [a,b]\}} \to 1. \tag{1B.5}$$

We say that a semiclassical pseudodifferential operator of the form (1B.1) is *quantum ergodic* if its eigenfunctions satisfy (1B.3).

In the appendix to [Marklof and O'Keefe 2005], Zelditch raises the question of *converse quantum ergodicity*: to what extent is it possible for nonergodic Hamiltonian systems such as those in the KAM regime to have quantum ergodic quantisations? In the extreme situation of quantum complete integrability, rigorous results on eigenfunction microlocalisation onto unions of Lagrangian tori have been established in [Toth and Zelditch 2003], which clearly rules out quantum ergodicity. In the intermediate regimes between complete integrability and ergodicity, fewer rigorous results on the question of converse quantum ergodicity are known. In the appendix to [Marklof and O'Keefe 2005], Zelditch shows that the "pimpled spheres", which are S^2 with a metric deformed polar cap, are not quantum ergodic, exploiting the periodicity of the flow in a strong way. In [Gutkin 2009] it is shown that the "racetrack billiard" is quantum ergodic but not ergodic, with phase space splitting into two disjoint invariant sets of equal measure.

As KAM dynamics are far from ergodic dynamics in character, the Bohr correspondence principle suggests that $\mathcal{P}_h(t)$ is typically *not* quantum ergodic, and that under generic conditions on the perturbation, there could exist sequences of eigenfunctions for $\mathcal{P}_h(t)$ with semiclassical mass entirely supported on individual invariant tori.

This localisation was proven for quasimodes in [Popov 2004a], where semiclassical Fourier integral operators were used to construct a quantum Birkhoff normal form for a class of semiclassical pseudodifferential operators $\mathcal{P}_h(t)$. This quantum Birkhoff normal form is used to obtain a family of quasimodes microlocalised near the union of KAM Lagrangian tori of a Hamiltonian associated to $\mathcal{P}_h$. A similar construction was previously made in [Colin de Verdière 1977], which establishes the existence of quasimodes microlocalised near the Lagrangian tori of a completely integrable Hamiltonian on a compact smooth manifold.

As pointed out in [Zelditch 2004], however, the passage from quasimode microlocalisation statements to microlocalisation statements for genuine eigenfunctions typically requires information on the spectral concentration of the operator in question.

One way in which this information can be obtained is by considering the spectral flow of $\mathcal{P}_h(t)$ in an analytic parameter t as in this paper. The Hadamard variational formula allows us to rule out spectral concentration for full measure t, given suitable information on the expectation of the quantum observable

$$\langle \partial_t \mathcal{P}_h(t) u_j(t; h), u_j(t; h)\rangle, \tag{1B.6}$$

which can be obtained from conditions like (1B.4). One can then draw conclusions about eigenfunction microlocalisation from those about quasimode microlocalisation.

In [Hassell 2010], this technique was exploited to obtain the existence of a sequence of Laplacian eigenfunctions on the Bunimovich stadium that does not equidistribute, at least for a full-measure set of aspect ratios. This strategy was also exploited in [Gomes 2018], where the author establishes a weak form of Percival's conjecture for the mushroom billiard.

It is the purpose of this paper to use the same technique to show that quantisations of KAM Hamiltonian systems in the sense of (1B.1) are typically *not* quantum ergodic, at least for full measure $t \in (0, \delta)$.

We follow [Popov 2004a] in working in the category of Gevrey regularity for our Hamiltonian P, due to the availability of explicit and full details of the quantum Birkhoff normal form construction in this setting.

1C. *Statement of results.* The following is the main result of this paper.

Theorem 1.1. *Suppose M is a compact boundaryless G^ρ manifold and $\mathcal{P}_h(t)$ is a family of self-adjoint elliptic semiclassical pseudodifferential operators acting on $\mathcal{C}^\infty(M; \Omega^{1/2})$ with fixed positive differential order such that*:

(i) *The operator $\mathcal{P}_h(t)$ has full symbol real-valued, analytic in t, and in the Gevrey class $S_\ell(T^*M)$ from Definition B.5, where $\ell = (\rho, \mu, \nu)$, with $\rho(\tau+n)+1 > \mu > \rho' = \rho(\tau+1)+1$ and $\nu = \rho(\tau+n+1)$.*

(ii) *The principal symbol $P_0(x, \xi; t)$ lies in $G^{\rho,1}(T^*M \times (-1, 1))$.*

(iii) *$P_0(x, \xi; 0)$ is a completely integrable and nondegenerate Hamiltonian.*

(iv) *The subprincipal symbol of $\mathcal{P}_h(t)$ vanishes.*

(v) *In an action-angle variable coordinate patch $\mathbb{T}^n \times D$ for the unperturbed Hamiltonian $P_0(x, \xi; 0)$, the KAM Hamiltonian can be written as $H(\theta, I; t) = P_0(\,\cdot\,, \cdot\,; t) \circ \chi$, and we define $H^0(I) := H(\theta, I; 0)$.*

(vi) *The KAM perturbation is such that*
$$\int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta$$
is nonconstant on some regular energy surface $\{I \in D : H^0(I) = E\}$ in the action-angle coordinate patch.

Then for any regular energy band $P_0^{-1}([a, b])$ with $E \in (a, b)$ for the energy surface in condition (vi), *there exists $\delta > 0$ such that the family of operators $\mathcal{P}_h(t)$ is not quantum ergodic in $[a, b]$ for full measure $t \in (0, \delta)$.*

Remark 1.2. Though we choose to work with Gevrey class Hamiltonians, it should be noted that we only require quasimodes for $\mathcal{P}_h(t)$ of order $O(h^{(3n+2)/2})$ to carry out the arguments in Section 2. In particular this implies that Theorem 1.1 should hold in the $\mathcal{C}^\infty$ setting, where $O(h^\infty)$ quasimodes are constructed in [Colin de Verdière 1977].

Remark 1.3. The condition (vi) is a rather mild one. Indeed for Hamiltonian perturbations of the form $H^0(I)+tH^1(\theta, I)$, it is equivalent to the functional independence of $H^0(I)$ and $\int_{\mathbb{T}^n} H^1(\theta, I)\,d\theta$. This holds for generic choice of H^1.

1D. ***Examples.*** The broad class of operators satisfying the assumptions of Theorem 1.1 are perturbations of completely integrable Schrödinger-type operators

$$\mathcal{P}_h = -h^2\Delta_g + V(x). \tag{1D.1}$$

In particular, Theorem 1.1 applies to the case of the semiclassical Laplace–Beltrami operator ($V=0$) on a manifold with perturbed metric (M, g_t), where (M, g_0) has completely integrable and nondegenerate geodesic flow.

The model example of a completely integrable geodesic flow is that of the flat torus

$$\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n. \tag{1D.2}$$

The Hamiltonian that generates the geodesic flow on $\mathbb{T}^n$ can be written as $|I|^2$, where $I \in \mathbb{R}^n$ is dual to the spatial variable $\theta \in \mathbb{T}^n$. This is clearly a nondegenerate and completely integrable Hamiltonian system. Similarly, in [Knörrer 1980], it is shown that the geodesic flow on an n-axial ellipsoid E is completely integrable and nondegenerate. Thus the Laplace–Beltrami operator for metric perturbations of both of these manifolds is covered by Theorem 1.1, provided the generic condition (vi) is satisfied.

For an explicit family of examples, one can consider $\mathbb{T}^2$, equipped with the metric

$$g = d\theta_1^2 + d\theta_2^2 + t\chi(\theta_1, \theta_2)\,d\theta_1\,d\theta_2$$

for $t > 0$ small and $\chi \in \mathcal{C}^\infty(\mathbb{T}^2)$ arbitrary. The Hamiltonian corresponding to $-h^2\Delta_g$ is

$$H(\theta, I) = I_1^2 + I_2^2 + t\chi(\theta_1, \theta_2)I_1I_2,$$

and we have that

$$\int_{\mathbb{T}^2} \partial_t H(\theta, I)\,d\theta = I_1I_2\int_{\mathbb{T}^2}\chi(\theta)\,d\theta \tag{1D.3}$$

is nonconstant over any energy surface $|I| = E > 0$; thus condition (vi) of Theorem 1.1 is satisfied.

1E. ***Outline of paper.*** In Section 3A, we introduce some definitions and notations that are prevalent throughout the paper.

In Section 2, we prove Theorem 1.1 by contradiction. We now outline the strategy of the proof. In Section 2B, under the assumptions of (vi) in Theorem 1.1, Proposition 2.5 makes use of the calculation in Section 3E to obtain an upper bound for the flow speed of a positive density family of the quasieigenvalues constructed in Section 4C. On the other hand, the assumption of quantum ergodicity of $\mathcal{P}_h(t)$ for large measure t yields an estimate for the variation of a large density subset of exact eigenvalues in (2B.22). The results in this section establish a gap (2B.23) between the flow speed of these quasieigenvalues and exact eigenvalues that ensures that individual eigenvalues cannot spend large measure $t \in (0, \delta)$ within $O(h^{n+1})$ distance of any of the quasieigenvalues. This is formalised in Section 2C, where it is deduced that there exists $t_* \in (0, \delta)$ at which there are very few actual eigenvalues within $O(h^{n+1})$ distance of the union of quasieigenvalue windows. An elementary spectral theory contradiction is arrived at from this spectral nonconcentration, completing the proof.

In Section 3, we construct a Gevrey class Birkhoff normal form for the family of Hamiltonians $P(x, \xi; t)$. The construction is that of [Popov 2004b], with our only additional concern being establishing the regularity of this Birkhoff normal form construction in the parameter t. In Section 3E, we compute the derivative of the integrable term $K(I; t)$ of the Birkhoff normal form in the parameter t. This is done by applying two KAM iterations to $P(x, \xi; t)$ prior to the application of the Birkhoff normal form construction of Theorem 3.10.

In Section 4, we recall the quantum Birkhoff normal form construction of [Popov 2004a], formulated in Theorem 4.1. This construction yields a Gevrey family of quasimodes microlocalising on the KAM Lagrangian tori of the Hamiltonian $P(x, \xi; t)$. For the spectral flow arguments in Section 2C we require that the associated quasieigenvalues are smooth in t, which is a statement entirely about the symbols of this quantum Birkhoff normal form.

In Appendix A, we introduce the anisotropic classes of Gevrey functions that are used throughout this paper as well as some of their basic properties.

In Appendix B, we introduce the semiclassical pseudodifferential calculus for Gevrey class symbols.

In Appendix C, we collect two elementary assertions about analytic functions.

In Appendix D, we state and prove a version of the Whitney extension theorem for the anisotropic class of Gevrey functions.

2. Proof of Theorem 1.1

2A. ***Introduction.*** We begin by assuming that $\mathcal{P}_h(t)$ is a family of operators satisfying the assumptions of Theorem 1.1.

Condition (vi) in Theorem 1.1 implies that there exists a nonresonant frequency $\omega_0 \in \widetilde{\Omega}_\kappa$ with associated Lagrangian torus Λ_{ω_0} such that the average of $\partial_t P_0(x, \xi; 0)$ over the torus Λ_{ω_0} differs from the average of $\partial_t P_0(x, \xi; 0)$ over the associated energy surface

$$\{(x, \xi) \in T^*M : P_0(x, \xi; 0) = H^0(I(\omega_0))\}. \tag{2A.1}$$

Moreover, we can ensure that Λ_{ω_0} lies in an arbitrarily small energy window $[a, b]$ about the regular energy E from the condition (vi). Without loss of generality, the hypotheses of Theorem 1.1 thus guarantee the existence of what we shall call a *slow torus*.

Definition 2.1. A *slow torus* in the energy band $[a, b]$ for the unperturbed Hamiltonian

$$H(\theta, I; 0) = H^0(I), \tag{2A.2}$$

written in action-angle coordinates, is a Lagrangian invariant torus Λ_{ω_0} with nonresonant frequency $\omega_0 \in \widetilde{\Omega}_\kappa$ and energy $H^0(I(\omega_0)) \in (a, b)$ in the notation of Theorem 3.10 that satisfies

$$(2\pi)^{-n}\int_{\mathbb{T}^n} \partial_t H(\theta, I(\omega_0); 0)\, d\theta < \inf_{E\in[a,b]} \frac{1}{\mu_E(\Sigma_E)} \int_{\Sigma_E} \partial_t P_0(x, \xi; 0)\, d\mu_E \tag{2A.3}$$

at $t = 0$.

We call such a torus a *slow torus* to draw intuition from the special case where $\partial_t \mathcal{P}_h(t)$ is a positive operator. In this case, as t evolves, the quasieigenvalues associated to such a torus increase as t evolves at a slower rate than the typical increase of eigenvalues at the same energy. The intuition behind this stems from the Hadamard variational formula (2B.8), and the fact that the associated quasimodes microlocalise onto Λ_{ω_0}. This intuition is confirmed in Section 3E, by a more careful analysis of the leading-order behaviour as $t \to 0$ of the integrable term in the Birkhoff normal form established in Theorem 3.10. Under the assumption of quantum ergodicity, this analysis implies a discrepancy (2B.23) in the spectral flow of genuine eigenvalues and quasieigenvalues attached to slow tori. Consequently, we obtain the spectral nonconcentration statement Proposition 2.10.

We begin by using the slow torus condition and choosing $c > 0$ sufficiently small so that

$$(2\pi)^{-n}\int_{\mathbb{T}^n} \partial_t H(\theta, I(\omega_0; 0); 0)\, d\theta < \inf_{E\in[a,b]} \frac{1}{\mu_E(\Sigma_E)} \int_{\Sigma_E} \partial_t P_0(x, \xi; 0)\, d\mu_E - 3c \tag{2A.4}$$

is satisfied.

As the quantum ergodicity condition (1B.3) is preserved upon passing to energy subintervals, we can assume that $[a, b]$ is an arbitrarily small energy window containing $H^0(I(\omega_0; 0))$. In particular, we can scale our interval $[a, b]$ by a small factor λ to ensure that the condition

$$\sup_{E\in[a,b]} \frac{1}{\mu_E(\Sigma_E)} \int_{\Sigma_E} \partial_t P_0\, d\mu_E - \inf_{E\in[a,b]} \frac{1}{\mu_E(\Sigma_E)} \int_{\Sigma_E} \partial_t P_0\, d\mu_E =: Q_+(0) - Q_-(0) < \epsilon < c \tag{2A.5}$$

is satisfied for any particular $\epsilon < c$. From the regularity of P_0, one can achieve this by taking

$$\lambda = O(\epsilon). \tag{2A.6}$$

Through the course of this section, we will track the size of various small quantities in terms of this ϵ, which we will eventually take small in the proof of Proposition 2.10.

Proposition 3.14 applies to H, and we obtain a family of symplectomorphisms

$$\chi \in G^{\rho,\rho',\rho'}\big(\mathbb{T}^n \times D \times \big(-\tfrac{1}{2}, \tfrac{1}{2}\big), \mathbb{T}^n \times D\big) \tag{2A.7}$$

and a family of diffeomorphisms

$$\omega \in G^{\rho',\rho'}\big(D \times \big(-\tfrac{1}{2}, \tfrac{1}{2}\big), \Omega\big) \tag{2A.8}$$

such that

$$H(\chi(\theta, I; t); t) = K(I; t) + R(\theta, I; t), \tag{2A.9}$$

where $\partial_I^\alpha R(\theta, I; t) = 0$ for nonresonant actions $I \in E_\kappa(t)$. Using the diffeomorphism (2A.8), we can define an action map $I \in G^{\rho', \rho'}\big(\Omega \times \big(-\frac{1}{2}, \frac{1}{2}\big)\big)$ implicitly by

$$\tilde{\omega} = \omega(I(\tilde{\omega}; t); t) \tag{2A.10}$$

and we can use this map to specify the action coordinates of a nonresonant torus with fixed frequency at any $t \in \big(-\frac{1}{2}, \frac{1}{2}\big)$ in the Birkhoff normal form furnished by $\chi(\,\cdot\,, \cdot\,; t)$.

We first obtain a positive-measure family of slow tori near Λ_{ω_0}.

Proposition 2.2. *There exists $r > 0$ and $\delta > 0$ such that for any $\omega \in \bar{\Omega} := B(\omega_0, r) \cap \widetilde{\Omega}_\kappa$, the torus $\Lambda_\omega = \chi(\mathbb{T}^n \times \{I(\omega, t)\})$ has energy*

$$K(I(\omega; t), t) \in [a, b] \tag{2A.11}$$

for all $t \in (0, \delta)$.

In particular, the family of tori

$$\Lambda(t) := \bigcup_{\omega \in \bar{\Omega}} \Lambda_\omega \tag{2A.12}$$

is a positive-measure family of KAM tori entirely contained within the energy band $[a, b]$.

Moreover, r and δ can be chosen small enough to ensure

$$\begin{aligned}(2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I(\omega; t); t)\, d\theta &< (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I(\omega; 0); t)\, d\theta + \epsilon \\ &< \inf_{t \in (0,\delta)} Q_-(t) - 2c \end{aligned} \tag{2A.13}$$

for each $\omega \in \bar{\Omega}$ and each $t \in (0, \delta)$.

We can also choose $\delta > 0$ small enough to ensure that

$$Q_+ - Q_- := \sup_{t \in (0,\delta)} Q_+(t) - \inf_{t \in (0,\delta)} Q_-(t) < 2\epsilon. \tag{2A.14}$$

In particular r, δ can be taken to be $O(\epsilon)$, with constant independent of t and h.

Proof. From the regularity of χ, I, and K established in Theorem 3.10, it follows that we can take $r = O(\lambda)$ to ensure that (2A.11) is satisfied at $t = 0$, where $\lambda = O(\epsilon)$ is as in (2A.6). Similarly, we can ensure that

$$(2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta < (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I(\omega_0); t)\, d\theta + \epsilon/2 \tag{2A.15}$$

holds for $|I - I(\omega_0)| = O(\lambda)$. Since (2A.4) is satisfied at $t = 0$, it follows that

$$(2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I(\omega; 0); 0)\, d\theta < Q_-(0) - 3c + \epsilon/2 \tag{2A.16}$$

for all $\omega \in \bar{\Omega} = B(\omega_0, r) \cap \widetilde{\Omega}_\kappa$ upon taking $r = O(\lambda)$.

The regularity of χ, I and K in the parameter t then allows us to deduce that (2A.11) and (2A.13) are satisfied for $t \in (0, \delta)$, for sufficiently small $\delta > 0$ and for each $\omega \in \bar{\Omega}$. In particular, we can take $\delta = O(\lambda) = O(\epsilon)$.

Finally, the estimate (2A.14) for small δ follows from the regularity of

$$\frac{1}{\mu_E(\Sigma_E)} \int_{\Sigma_E} \partial_t P_0 \, d\mu_E \tag{2A.17}$$

in t and E. ☐

We can now apply the quantum Birkhoff normal form construction outlined in Section 4. From Theorem 4.5, we obtain a family of quasimodes that microlocalise onto the family of KAM tori $\Lambda(t)$ introduced in (2A.12).

In particular, following Section 4C, we take $S(t) = \{I(\omega; t) : \omega \in \overline{\Omega}\}$ and define the index set $\mathcal{M}_h(t)$ as in (4C.2). Then for each $m \in \mathcal{M}_h(t) \subset \mathbb{Z}^n$, we have a quasimode v_m with corresponding quasieigenvalue μ_m as in (4C.3). We introduce notation for the union of h^{n+1}-width energy windows about the quasieigenvalue associated to tori in $\Lambda(t)$:

$$W(t; h) := \bigcup_{m \in \mathcal{M}_h(t)} [\mu_m(t; h) - h^{n+1}, \mu_m(t; h) + h^{n+1}]. \tag{2A.18}$$

We also introduce the index set

$$G(h) = \{j \in \mathbb{N} : E_j(t) \in [a, b] \text{ for some } t \in (0, \delta)\} \tag{2A.19}$$

of the eigenvalues that can possibly play a role in the spectral flow considerations in Section 2C.

To conclude this section, we collect asymptotic estimates for the number of eigenvalues and the number of quasieigenvalues that are in the energy window $[a, b]$ as $h \to 0$.

Proposition 2.3. *We have the asymptotic estimate*

$$\#\mathcal{M}_h(t) \sim (2\pi h)^{-n} \mu(\mathbb{T}^n \times \{I(\omega, t) : \omega \in \overline{\Omega}\}) \tag{2A.20}$$

for each $t \in (0, \delta)$.

Furthermore, we have

$$\limsup_{h \to 0} (2\pi h)^n \# G(h) \leq \mu(\{(x, \xi) : P_0(x, \xi; 0) \in [a - M\delta, b + M\delta]\}), \tag{2A.21}$$

where M is the uniform bound on spectral flow speed in (2B.11) *and $G(h)$ is as in* (2A.19).

Here μ denotes the symplectic measure $d\xi\, dx$ on T^*M.

Proof. The estimate (2A.20) is a consequence from (4C.8), and (2A.21) follows from (2B.11) and an application of the semiclassical Weyl law [Zworski 2012, Theorem 14.11]. ☐

From Proposition 2.3, it follows that we can bound

$$\frac{\# G(h)}{\inf_{t \in (0,\delta)} \# \mathcal{M}_h(t)} \tag{2A.22}$$

for $t \in (0, \delta(\epsilon))$ and $h < h_0(\epsilon)$. Moreover, this upper bound is uniform in ϵ. By the nature of their construction in Proposition 2.2, the quasieigenvalues $\mu_m(t; h)$ lie in $[a, b]$ for all $t \in (0, \delta)$.

It is convenient to introduce the subset $\widetilde{G}(h) \subset G(h)$ given by

$$\widetilde{G}(h) = \{j \in \mathbb{N} : E_j(t) \in [a, b] \text{ for all } t \in (0, \delta)\}. \tag{2A.23}$$

By choosing $\delta(\epsilon) > 0$ appropriately small, we can ensure that a large proportion of eigenvalues that lie in $[a, b]$ for some $t \in (0, \delta)$ lie in $[a, b]$ for all $t \in (0, \delta)$.

Proposition 2.4. *We can choose $\delta(\epsilon) = O(\epsilon^2)$ such that*

$$\frac{\#\widetilde{G}(h)}{\#G(h)} \geq 1 - C\epsilon \tag{2A.24}$$

for all $\epsilon < \epsilon_0$ and $h < h_0(\epsilon)$, where $C > 0$ is a constant.

Proof. We have the bound

$$\frac{\#G(h)}{\#\widetilde{G}(h)} \leq \frac{N_h([a + M\delta, b - M\delta])}{N_h([a - M\delta, b + M\delta])}, \tag{2A.25}$$

where $N_h(I)$ counts the semiclassical eigenvalues of the operator $\mathcal{P}_h(0)$ in I. Recalling that the interval $[a, b]$ is of scale $\lambda = O(\epsilon)$, it follows that for any choice of $\delta = O(\epsilon^2)$, the ratio of phase space volumes

$$\frac{\mu(P_0(x, \xi; 0) \in [a - M\delta, a + M\delta] \cup [b - M\delta, b + M\delta])}{\mu(P_0(x, \xi; 0) \in [a - M\delta, a + M\delta])} \tag{2A.26}$$

can be bounded by a constant multiple of ϵ for all sufficiently small ϵ. Application of the semiclassical Weyl asymptotics to (2A.25) completes the proof. □

2B. ***Eigenvalue and quasieigenvalue variation.*** We now turn our attention to the variation of quasieigenvalues and eigenvalues as $t \in (0, \delta)$ varies. The quasieigenvalues can be handled rather explicitly.

Proposition 2.5. *For any all sufficiently small $\delta(\epsilon) > 0$ and all $t \in (0, \delta)$, we have*

$$\limsup_{h \to 0} \partial_t \mu_m(t; h) \leq Q_- - c \tag{2B.1}$$

for all $m \in \bigcup_{t \in (0,\delta)} \mathcal{M}_h(t)$ uniformly in t.

Proof. From Proposition 3.14, we have

$$K_0(I; t) = H^0(I) + t \cdot (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta + O(t^{9/8}) \tag{2B.2}$$

for any $I \in D$. Hence we have

$$\partial_t (K_0(h(m + \vartheta/4); t)) < (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, h(m + \vartheta/4); 0)\, d\theta + \epsilon \tag{2B.3}$$

for all $t \in (0, \delta(\epsilon))$, taking δ sufficiently small. From the definition of $\mathcal{M}_h(t)$, 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

$$\partial_t (K_0(h(m + \vartheta/4); t)) < (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I(\omega; t); t)\, d\theta + \epsilon + O(h) \tag{2B.4}$$

for some $\omega \in \overline{\Omega}$. This allows us to compute

$$\begin{aligned}\partial_t \mu_m(t;h) &= \partial_t(K^0(h(m+\vartheta/4);t,h)) \\ &= \partial_t(K_0(h(m+\vartheta/4);t)) + O(h) \\ &< (2\pi)^{-n}\int_{\mathbb{T}^n} \partial_t H(\theta, I(\omega;t);t)\,d\theta + \epsilon + O(h) \\ &< Q_- - 2c + \epsilon + O(h) \end{aligned} \tag{2B.5}$$

which implies

$$\limsup_{h\to 0} \partial_t \mu_m(t;h) < Q_- - 2c + \epsilon, \tag{2B.6}$$

using (2B.4), (2A.13), and (2A.14). □

In particular, we can choose $B > 0$ and $h_0 > 0$ such that

$$\partial_t \mu_m(t;h) < B < Q_- - c \tag{2B.7}$$

for all $t \in (0,\delta)$ and all $h < h_0$.

Remark 2.6. We abused notation slightly here by writing $\mu_m(t;h)$ even when $m \notin \mathcal{M}_h(t)$. 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 quasieigenvalue 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 $\mathcal{M}_h(t)$ for only $O(h)$-sized t-intervals at a time.

Remark 2.7. This is the last part of the argument that involves placing an additional restriction on the size of $\delta > 0$.

We now consider the variation of eigenvalues. For each fixed $h > 0$, the operators $\mathcal{P}_h(t)$ comprise an holomorphic family of type A in the sense of [Kato 1966] and so we can choose eigenvalues and corresponding eigenprojections holomorphic in the parameter t. Thus if at each time t we order our eigenpairs $E_j(t;h)$ in order of increasing energy, by holomorphy it follows that E_j will be continuous and piecewise smooth in $(0,\delta)$. On the cofinite set where E_j is differentiable in t, we have

$$\partial_t E_j(t;h) = \langle \partial_t \mathcal{P}_h(t) u_j(t;h), u_j(t;h)\rangle, \tag{2B.8}$$

since (u_j) is an orthonormal basis. We will control (2B.8) using our assumption of quantum ergodicity.

To this end, we now suppose for the sake of contradiction that there exists a positive-measure set $\mathcal{B} \subset (0,\delta)$ such that $\mathcal{P}_h(t)$ is quantum ergodic in the sense of (1B.3) for every $t \in \mathcal{B}$.

Proposition 2.8. *For every $t \in \mathcal{B}$ and $\epsilon > 0$, there exists $h_0(t,\epsilon)$ such that, for all $h < h_0(t,\epsilon)$, we have*

$$\left|\langle \partial_t \mathcal{P}_h(t) u_j(t;h), u_j(t;h)\rangle - \int_{\Sigma_{E_j(t;h)}} \partial_t P_0 \, d\mu_{E_j(t;h)}\right| < \epsilon \tag{2B.9}$$

for a family of indices $S(t;h) \subset \{j \in \mathbb{N} : E_j(t;h) \in [a,b]\}$ with

$$\frac{\# S(t;h)}{\{j \in \mathbb{N} : E_j(t;h) \in [a,b]\}} > 1 - \epsilon. \tag{2B.10}$$

Proof. This is a direct application of (1B.4). □

We also note that we have a global-in-time bound

$$\partial_t E_j(t;h) \le M < \infty \tag{2B.11}$$

from differentiation of the expression

$$E_j(t;h) = \langle \mathcal{P}_h(t) u_j(t;h), u_j(t;h)\rangle \tag{2B.12}$$

and using a routine elliptic parametrix construction that is uniform in $t \in (0,1)$ to bound the quantity

$$\langle \partial_t \mathcal{P}_h(t) u_j(t;h), u_j(t;h)\rangle \tag{2B.13}$$

given that $E_j(t;h)$ lies in a fixed energy band $[a,b]$.

Recalling (2A.5), Proposition 2.8 implies that

$$\langle \partial_t \mathcal{P}_h(t) u_j, u_j\rangle \in [Q_- - \epsilon, Q_+ + \epsilon] \tag{2B.14}$$

for all $j \in S(t;h)$ such that E_j is smooth at t, and all $h < h_0(t,\epsilon)$.

Now, from the outer regularity of the Lebesgue measure, we may then choose a subinterval $J \subset (0,\delta)$ such that

$$\frac{m(\mathcal{B}\cap J)}{m(J)} > 1-\epsilon. \tag{2B.15}$$

We can then apply the monotone convergence theorem to upgrade Proposition 2.8 for $t \in \mathcal{B}$ to a statement that is uniform in a large-measure subset of J.

Proposition 2.9. *There exists a subset $\widetilde{\mathcal{B}} \subseteq \mathcal{B}\cap J$ and an $h_0(\epsilon) > 0$ such that*

$$\frac{m(\widetilde{\mathcal{B}})}{m(J)} > 1-2\epsilon \tag{2B.16}$$

and, for any $h < h_0(\epsilon)$ and any $t \in \widetilde{\mathcal{B}}$, there exists a subset

$$Z(t,h) \subset \{j \in \mathbb{N} : E_j(t,h) \in [a,b]\} \tag{2B.17}$$

such that

$$\frac{\#Z(t,h)}{\#\{j \in \mathbb{N} : E_j(t,h) \in [a,b]\}} > 1-2\epsilon \quad \text{for all } 0 < h < h_0 \tag{2B.18}$$

and

$$\langle \partial_t \mathcal{P}_h(t) u_j, u_j\rangle \in [Q_- - \epsilon, Q_+ + \epsilon] \quad \text{for all } j \in Z(t,h), \tag{2B.19}$$

for all $Z(t;h)$ such that E_j is smooth at t and all $h < h_0(\epsilon)$.

Proof. For fixed $\eta, \epsilon > 0$, we define

$$\mathcal{B}_\eta := \{t \in \mathcal{B}\cap J : h_0(t,\epsilon) > \eta\}, \tag{2B.20}$$

where $h_0(t,\epsilon)$ is as in Proposition 2.8. As $\mathcal{B}\cap J = \bigcup_{\eta>0} \mathcal{B}_\eta$, countable additivity implies that for sufficiently small $\eta_0 > 0$, we have

$$m(\mathcal{B}_{\eta_0}) > \frac{1-2\epsilon}{1-\epsilon} m(\mathcal{B}\cap J) > (1-2\epsilon)m(J). \tag{2B.21}$$

We now take $\widetilde{\mathcal{B}} = \mathcal{B}_{\eta_0}$ and $Z(t; h) = S(t; h)$ in the notation of Proposition 2.8, and (2B.19) follows from (2B.14). $\square$

In light of Proposition 2.9, we redefine Q_-, Q_+ to be the endpoints of the enlarged interval in (2B.19). Hence

$$\langle \partial_t \mathcal{P}_h(t) u_j, u_j \rangle \in [Q_-, Q_+] \quad \text{for all } j \in Z(t, h). \tag{2B.22}$$

In terms of the redefined Q_-, Q_+, we have

$$Q_- - B > c - \epsilon > 0, \tag{2B.23}$$

and so we have established a discrepancy between the typical speed of eigenvalue flow and the upper bound for the speed of quasieigenvalue flow.

2C. *Spectral nonconcentration.*

We can now complete the proof of Theorem 1.1 by proving a spectral nonconcentration result that follows from the results of Section 2B.

Proposition 2.10. *Under the quantum ergodicity assumption $m(\mathcal{B}) > 0$ imposed in Section 2B, for sufficiently small $\epsilon > 0$ there exists $t_* \in \widetilde{\mathcal{B}} \subset \mathcal{B}$ such that*

$$\frac{N(t_*; h)}{\#\mathcal{M}_h(t_*)} < \frac{1}{2} \tag{2C.1}$$

for a sequence $h_j \to 0$, where

$$N(t; h) := \#\{j \in \mathbb{N} : E_j(t; h) \in W(t; h)\} \tag{2C.2}$$

is the number of exact eigenvalues lying in the union $W(t, h)$ of the quasieigenvalue windows as introduced in (2A.18).

Proof. The method of proof is by averaging in t and using Proposition 2.9 to show that most individual eigenfunctions cannot lie in $W(t, h)$ for a significant proportion of $t \in J$. We begin by defining

$$A_j(h) = \{t \in J : E_j(t; h) \in [a, b]\}, \tag{2C.3}$$

$$B_j(h) = \{t \in J : j \in Z(t; h)\}, \tag{2C.4}$$

$$C_j(h) = \{t \in J : E_j(t; h) \in W(t; h)\}. \tag{2C.5}$$

From Proposition 2.9, for each $t \in \widetilde{\mathcal{B}}$ we have

$$\sum_{j \in \mathbb{N}} 1_{B_j} \geq (1 - 2\epsilon) \sum_{j \in \mathbb{N}} 1_{A_j} \tag{2C.6}$$

for $h < h_0(\epsilon)$. Integrating, we obtain

$$\sum_{j \in \mathbb{N}} \int_{\widetilde{\mathcal{B}}} 1_{B_j} \, dt \geq (1 - 2\epsilon) \sum_{j \in \mathbb{N}} \int_{\widetilde{\mathcal{B}}} 1_{A_j} \, dt. \tag{2C.7}$$

Hence

$$\begin{aligned} \sum_{j \in \mathbb{N}} \int_J 1_{B_j} \, dt &\geq (1 - 2\epsilon) \sum_{j \in G(h)} \left(\int_J 1_{A_j} \, dt - \int_{J \setminus \widetilde{\mathcal{B}}} 1_{A_j} \, dt \right) \\ &\geq (1 - 2\epsilon) \sum_{j \in G(h)} \left(\int_J 1_{A_j} \, dt - 2\epsilon m(J) \right), \end{aligned} \tag{2C.8}$$

which can be rewritten as

$$\sum_{j\in\mathbb{N}} m(B_j) \geq (1-2\epsilon) \sum_{j\in G(h)} (m(A_j) - 2\epsilon m(J)). \tag{2C.9}$$

From the definitions (2A.19) and (2A.23), we know that $m(A_j) > 0$ only if $j \in G(h)$ and $m(A_j) = m(J)$ if $j \in \widetilde{G}(h)$. Thus we can estimate

$$\begin{aligned}\frac{1}{\#G(h)} \sum_{j\in\mathbb{N}} m(B_j) &\geq (1-2\epsilon)\left(\frac{\#\widetilde{G}(h)}{\#G(h)} - 2\epsilon\right) m(J) \\ &\geq (1-2\epsilon)(1-O(\epsilon))m(J) \\ &=: (1-\eta)m(J),\end{aligned} \tag{2C.10}$$

where $\limsup_{h\to 0} \eta(\epsilon; h) = o_\epsilon(1)$. Consequently we have

$$m(B_j) \geq (1-\eta^{1/2})m(A_j) \tag{2C.11}$$

for a subfamily $\mathcal{F}(h) \subset \widetilde{G}(h)$ with

$$\frac{\#\mathcal{F}(h)}{\#G(h)} \geq 1 - \eta^{1/2} - O(\epsilon) \tag{2C.12}$$

in the limit $h \to 0$, where we have made use of Proposition 2.4.

Taking $E(t; h) := E_j(t; h)$ for some $j \in \mathcal{F}$, the bound from the Hadamard variational formula (2B.22) yields

$$E(t_2; h) - E(t_1; h) \geq ((1-\eta^{1/2})Q_- - M\eta^{1/2})m(J), \tag{2C.13}$$

where M is the uniform bound on eigenvalue flow speed for eigenvalues in $[a, b]$ and $J = [t_1, t_2]$.

On the other hand, we now bound $E(t_2; h) - E(t_1; h)$ above. To do this, we define

$$\widetilde{E}(t; h) = E(t; h) - Bt \quad \text{and} \quad \tilde{\mu}_m(t; h) = \mu_m(t; h) - Bt,$$

where B was the upper bound in (2B.7). Then the transformed quasieigenvalue windows

$$\widetilde{W}_m(t; h) = [\tilde{\mu}_m(t; h) - h^{n+1}, \tilde{\mu}_m(t; h) + h^{n+1}]$$

are nonincreasing. From this it follows that if $\widetilde{E}(s; h) \in [\tilde{\mu}_m(s; h) - h^{n+1}, \tilde{\mu}_m(s; h) + h^{n+1}]$ and $m \in \mathcal{M}_h(s)$ for some $s \in J$, then $\widetilde{E}(s'; h) - \widetilde{E}(s; h) < 2h^{n+1}$, where s' is the final time $t \in J$ such that $m \in \mathcal{M}_h(t)$ and $\widetilde{E}(t; h) \in [\tilde{\mu}_m(t; h) - h^{n+1}, \tilde{\mu}_m + h^{n+1}]$. This implies $E(s'; h) - E(s; h) < 2h^{n+1} + B(s' - s)$.

Generalising this idea, we can cover each $C_j(h)$ with a finite union of intervals $\bigcup_k I_k$ with $I_k = [s_k, s_k']$ defined as follows:

(i) 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 $E(t; h) \in [\mu_{m(0)}(t; h) - h^{n+1}, \mu_{m(0)}(t; h) + h^{n+1}]$ and $m(0) \in \mathcal{M}_h(t)$ for all sufficiently small $t - s_0 > 0$.

(ii) We then define $s_0' := \sup\{t \in J : E(t; h) \in [\mu_{m(0)}(t; h) - h^{n+1}, \mu_{m(0)}(t; h) + h^{n+1}]\}$.

(iii) If $\{t \in J : t > s'_{k-1}$ and $E(t; h) \in W(t; h)\}$ is empty, we terminate the inductive process; otherwise we proceed inductively by defining $s_k := \inf\{t \in J : t > s'_{k-1}$ and $E(t; h) \in W(t; h)\}$ and choosing a corresponding $m(k) \in \mathcal{M}_h(s_k)$ such that $E(t; h) \in [\mu_{m(k)}(t; h) - h^{n+1}, \mu_{m(k)}(t; h) + h^{n+1}]$ and $m(k) \in \mathcal{M}_h(t)$ for all sufficiently small $t - s_{k-1} > 0$.

(iv) We then define $s'_k := \sup\{t \in J : E(t; h) \in [\mu_{m(k)}(t; h) - h^{n+1}, \mu_{m(k)}(t; h) + h^{n+1}]\}$.

From the Weyl asymptotics, this procedure must terminate after finitely many iterations.

Remark 2.11. In the case that $E(t; h)$ is still in a quasieigenvalue window for t arbitrarily close to, but greater than s'_k, we will have $s_{k+1} = s'_k$. This is the only kind of overlap possible between the intervals I_k. We also remark that the $m(k)$ are necessarily distinct, by the nature of this construction.

For each such interval $I_k = [s_k, s'_k]$, we have that $E(s'_k; h) - E(s_k; h) \le 2h^{n+1} + B(s'_k - s_k)$. As there can be at most $O(h^{-n})$ intervals I_k, we obtain

$$\sum_k E(s'_k; h) - E(s_k; h) \le B \sum_k (s'_k - s_k) + O(h). \tag{2C.14}$$

For such eigenvalues, we thus obtain the upper bound

$$\begin{aligned} E(t_2; h) - E(t_1; h) &\le \sum_k (E(s'_k; h) - E(s_k; h)) + \left(m(J)(1-\eta^{1/2}) - \sum_k (s'_k - s_k)\right) Q_+ + m(J)\eta^{1/2} M \\ &\le (B - Q_+) \sum_k (s'_k - s_k) + m(J)(1-\eta^{1/2}) Q_+ + m(J)\eta^{1/2} M + O(h) \\ &\le (B - Q_+) m(C_j) + ((1-\eta^{1/2}) Q_+ + M\eta^{1/2}) m(J) + O(h) \end{aligned} \tag{2C.15}$$

in the limit $h \to 0$. Rearranging (2C.15) and using (2C.13), we arrive at

$$(Q_+ - B)\frac{m(C_j)}{m(J)} \le 2M\eta^{1/2} + (1-\eta^{1/2})(Q_+ - Q_-). \tag{2C.16}$$

Hence by taking ϵ sufficiently small and then passing to sufficiently small $0 < h < h_0(\epsilon)$ we can bound $m(C_j)/m(J)$ by an arbitrarily small positive constant γ for all $j \in \mathcal{F}$. Hence we have

$$\begin{aligned} \int_J N(t; h)\, dt &\le \int_J \sum_{j \in \mathbb{N}} 1_{C_j}\, dt \le \int_J \gamma \sum_{j \in \mathcal{F}} 1_{A_j} + \#(G \setminus \mathcal{F})\, dt \\ &\le (\gamma \# \mathcal{F} + (\eta^{1/2} + O(\epsilon)) \# G) m(J) \\ &\le (\gamma + \eta^{1/2} + O(\epsilon))(\# G) m(J), \end{aligned} \tag{2C.17}$$

where we used Proposition 2.4 in the final line. Fixing sufficiently small $\epsilon > 0$, for all $h < h_0(\epsilon)$ we have

$$\frac{1}{m(J)} \int_J \frac{N(t; h)}{\# \mathcal{M}_h(t)}\, dt \le \frac{1}{4}. \tag{2C.18}$$

It follows that for each such $h < h_0$, the set

$$\left\{ t \in J : \frac{N(t; h)}{\# \mathcal{M}_h(t)} \le \frac{1}{2} \right\} \tag{2C.19}$$

has measure at least $m(J)/2$. Taking a sequence $h_j \to 0$ and applying the Borel–Cantelli lemma completes the proof. $\square$

Remark 2.12. In fact, the above argument demonstrates the existence of a family of such t_* with measure bounded below by $|J|/2$; however, we shall only require a single such t_* in what follows.

We now prove an elementary spectral theory result that will show that the conclusion of Proposition 2.10 is in fact absurd, hence establishing that $m(\mathcal{B}) = 0$ and completing the proof of Theorem 1.1. We denote by U the h-dependent span of all eigenfunctions with eigenvalues in $W(t; h)$, and as in (4C.3), $\{(v_m(t; h), \mu_m(t; h))\}_{m \in \mathcal{M}_h(t)}$ denotes the family of quasimodes and associated quasieigenvalues.

Proposition 2.13. *For sufficiently small $h > 0$, the projections*

$$w_m(t_*, h) = \pi_U(v_m(t_*, h)) \tag{2C.20}$$

are linearly independent.

Proof. First, we show that the estimate from Definition 4.4 on the error of quasimodes implies that the projections $\pi_U(v_m(t_*, h))$ are large. In particular, for $m \in \mathcal{M}_h(t_*)$, we have

$$\begin{aligned} &\Big\| (\mathcal{P}_h(t_*) - \mu_m(t_*, h)) \sum_{j \in \mathbb{N}} \langle v_m(t_*, h), u_j(t_*, h)\rangle u_j \Big\|^2 = O(h^{2\gamma+2}) \\ &\qquad \implies \sum_{|E_j - \mu_m| > h^{n+1}} |E_j(t, h) - \mu_m(t, h)|^2 |\langle v_m(t_*, h), u_j(t_*, h)\rangle|^2 = O(h^{2\gamma+2}) \\ &\qquad \implies \pi_{U^\perp}(v_m(t_*, h)) = O(h^{\gamma-n}). \end{aligned}$$

Hence for sufficiently small h, we have

$$\|w_m\|^2 = \|\pi_U(v_m(t_*, h))\|^2 = 1 + O(h^{\gamma+1}) + O(h^{2\gamma-2n}). \tag{2C.21}$$

From the almost-orthogonality condition that our quasimodes v_m satisfy (see Definition 4.4), together with (2C.21), it follows that the w_m are almost orthogonal for distinct $m \in \mathcal{M}_h(t)$. In particular, for $m \neq k$, we have

$$\begin{aligned} |\langle \pi_U(v_m(t_*, h)), \pi_U(v_k(t_*, h))\rangle| &\le |\langle v_m(t_*, h), v_k(t_*, h)\rangle| + |\langle \pi_{U^\perp}(v_m(t_*, h)), \pi_{U^\perp}(v_k(t, h))\rangle| \\ &= O(h^{\gamma+1}) + O(h^{2\gamma-2n}). \end{aligned}$$

Hence

$$|\langle \pi_U(v_m(t_*, h)), \pi_U(v_k(t_*, h))\rangle - \delta_{k,m}| = O(h^{\gamma+1}) + O(h^{2\gamma-2n}) \tag{2C.22}$$

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 $M(h) \in \mathrm{Mat}(\#\mathcal{M}_h(t_*), \mathbb{R})$, with entries given by

$$M_{ij}(h) = \langle w_i, w_j\rangle. \tag{2C.23}$$

Since

$$\|M - I\|_{HS} = (\#\mathcal{M}_h(t_*))(O(h^{\gamma+1}) + O(h^{2\gamma-2n})) = O(h^{\gamma+1-n}) + O(h^{2\gamma-3n}), \tag{2C.24}$$

we can invert $M = I + (M - I)$ as a Neumann series for sufficiently small h, provided the exponents of h are positive. This can be ensured by taking $\gamma > 3n/2$. Since M is nonsingular, we can therefore conclude that the functions in the collection

$$\{\pi_U(v_m(t_*, h)) : m \in \mathcal{M}_h(t_*)\} \tag{2C.25}$$

are linearly independent. □

We are now in a position to complete the proof of Theorem 1.1.

Completion of proof of Theorem 1.1. Having fixed $\epsilon > 0$ in Proposition 2.5, we showed in Proposition 2.10 that there exists a $t_* \in (0, \delta)$ at which we have the spectral nonconcentration result (2C.1) for a sequence $h_j \to 0$.

On the other hand, we showed in Proposition 2.13 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. □

3. Birkhoff normal form

In this section we construct a family of Birkhoff normal forms corresponding to a family of Gevrey smooth Hamiltonians $H(\theta, I; t)$, real-analytic in the parameter $t \in (-1, 1)$. The introduction of this parameter leads to only minor changes in the argument of [Popov 2004b].

We formulate the KAM theorem from [Popov 2004b] in Section 3B and outline the proof in Section 3C. We then complete the Birkhoff normal form construction following [Popov 2004b] in Section 3D.

In Section 3E, we compute the leading-order behaviour of this Birkhoff normal form as $t \to 0$, which was used in Proposition 2.5 to obtain an expression for the derivatives of the quasieigenvalues of the operator $\mathcal{P}_h(t)$ constructed in Section 4.

3A. *Notation.* We begin by introducing some notational conventions that will be used several times in this section.

Definition 3.1. For $s, r > 0$ we write

$$D_{s,r} := \{\theta \in \mathbb{C}^n/2\pi\mathbb{Z}^n : |\mathrm{Im}(\theta)| < s\} \times \{I \in \mathbb{C}^n : |I| < r\}, \tag{3A.1}$$

where $|\cdot|$ denotes the sup-norm on $\mathbb{C}^n$ induced by the 2-dimensional ℓ^∞ norm on $\mathbb{C}$.

These domains arise from considering the analytic extension of real-analytic Hamiltonians in action-angle variables. In this area it is common to bound derivatives of analytic functions using Cauchy estimates, which requires keeping track 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 $|\cdot|$ to denote the ℓ^1 norm when applied to elements of $\mathbb{Z}^n$ throughout this paper, as well as the matrix norm induced by the sup norm on $\mathbb{C}^n$.

3B. *Formulation of the KAM theorem.* Let $D \subset \mathbb{R}^n$ be a bounded domain, and consider a completely integrable Hamiltonian $H^0(I) = H^0(\theta, I) : \mathbb{T}^n \times D \to \mathbb{R}$ in action-angle coordinates. To begin, we shall assume that this Hamiltonian is real-analytic.

In addition, we assume the nondegeneracy condition $\det(\partial^2 H/\partial I^2) \neq 0$. 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

$$I \mapsto \nabla H^0(I) \tag{3B.1}$$

is a diffeomorphism from D to $\Omega \subset \mathbb{R}^n$. The inverse to this map is given by ∇g^0, where g^0 is the Legendre transform of H^0. The phase space $\mathbb{T}^n \times D$ is then foliated by the family of Lagrangian tori $\{\mathbb{T}^n \times \{I\} : I \in D\}$ that are invariant under Hamiltonian flow associated to H^0.

The KAM theorem asserts that small perturbations of $H^0(I)$, written as $H(\theta, I) = H^0(I) + H^1(\theta, I)$ on $\mathbb{T}^n \times D$ 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 $\Omega := \{\omega : \omega = \nabla_I H^0\}$ is the set of frequencies for the quasiperiodic flow of H^0, the frequencies satisfying

$$|\langle \omega, k\rangle| \geq \frac{\kappa}{|k|^\tau} \tag{3B.2}$$

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 nonresonant, and we denote the set of nonresonant frequencies by Ω^*_κ, suppressing the dependence on τ from our notation. These sets are obtained by taking the intersection of the sets

$$\left\{\omega \in \Omega : |\langle \omega, k\rangle| \geq \frac{\kappa}{|k|^\tau}\right\} \tag{3B.3}$$

over all nonzero $k \in \mathbb{Z}^n$, and hence $\bigcap_{\kappa>0} \Omega^*_\kappa$ is closed and perfect, with $\bigcup_{\kappa>0} \Omega^*_\kappa$ of full measure in Ω, as can be seen from the observation that

$$m\left(\left\{\omega \in \mathbb{R}^n : |\langle k, \omega\rangle| < \frac{\kappa}{|k|^\tau}\right\}\right) = O\left(\frac{\kappa}{|k|^{\tau+1}}\right). \tag{3B.4}$$

We work with the sets

$$\Omega_\kappa := \{\omega \in \Omega^*_\kappa : \operatorname{dist}(\omega, \partial\Omega) \geq \kappa\}, \tag{3B.5}$$

which have positive measure for sufficiently small κ. It is also convenient to introduce notation for the set of points of Lebesgue density in Ω_κ, which we denote by

$$\widetilde{\Omega}_\kappa := \left\{\omega \in \Omega : \frac{m(B(\omega, r) \cap \Omega_\kappa)}{m(B(\omega, r))} \to 1 \text{ as } r \to 0\right\}. \tag{3B.6}$$

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

The construction of the Birkhoff normal form is a consequence of Theorem 3.2, which is a version of the KAM theorem localised around the frequency ω which is taken as an independent parameter. The idea of treating ω as an independent parameter in this problem was originally due to Moser [1967]. This version is particularly useful for the Birkhoff normal form construction, 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

$$\Omega' = \{\omega \in \Omega : \operatorname{dist}(\omega, \Omega_\kappa) \le \kappa/2\}, \quad D' = \nabla g^0(\Omega'). \tag{3B.7}$$

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

$$H^0(z) = H^0(z_0) + \langle \nabla_z H^0(z_0), I\rangle + \int_0^1 (1-t)\langle \nabla_z^2 H^0(z_0 + tI)I, I\rangle\, dt. \tag{3B.8}$$

We now take $\omega \in \Omega^0$ to be the corresponding frequency $\nabla H^0(z_0)$. The inverse of the frequency map is

$$\psi_0(\omega) = \nabla g^0(\omega), \tag{3B.9}$$

where g^0 is the Legendre transform of H^0. Hence we can write

$$H^0(z) = H^0(\psi_0(\omega)) + \langle \omega, I\rangle + \langle P^0(I;\omega)I, I\rangle, \tag{3B.10}$$

where P^0 is the quadratic remainder term in (3B.8). Expanding about the point $z_0 = \nabla g^0(\omega)$, we can write our perturbation H^1 locally as

$$H^1(\theta, z) = H^1(\theta, \nabla g^0(\omega) + I) = P^1(\theta, I;\omega). \tag{3B.11}$$

This leads us to consider perturbed real-analytic Hamiltonians in the form

$$H(\theta, I;\omega) = H^0(\psi_0(\omega)) + \langle \omega, I\rangle + P(\theta, I;\omega) =: N(I;\omega) + P(\theta, I;\omega), \tag{3B.12}$$

where

$$N(I;\omega) = H^0(\psi_0(\omega)) + \langle \omega, I\rangle, \tag{3B.13}$$

$$P(\theta, I;\omega) = \langle P^0(I;\omega)I, I\rangle + P^1(\theta, I;\omega). \tag{3B.14}$$

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 consists of a component $N(I;\omega)$ that is only linear in I, and a nonlinear perturbation $P(\theta, I;\omega)$.

The essence of the frequency-localised KAM theorem in Theorem 3.2 is 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 ω. From Theorem 3.2, one can obtain Theorem 3.9, which establishes the existence of a Cantor family of invariant tori for the original Hamiltonian H as with traditional formulations of the KAM theorem.

To work with Gevrey smooth Hamiltonians, we fix $L_2 \ge L_0 \ge 1$ and $A_0 > 1$, and assume that $H^0 \in G^{\rho,1}_{L_0,L_2}(D^0 \times (-1,1))$ and $g^0 \in G^{\rho,1}_{L_0,L_2}(\Omega^0)$ with the estimates

$$\|H^0\|_{L_0,L_2}, \|g^0\|_{L_0,L_2} \le A_0. \tag{3B.15}$$

For $L_2 \ge L_1 \ge 1$ we now consider the analytic family of Gevrey perturbations

$$H^1 \in G^{\rho,\rho,1}_{L_1,L_2,L_2}(\mathbb{T}^n \times D \times (-1,1)),$$

with the perturbation norm

$$\epsilon_H := \kappa^{-2} \| H^1 \|_{L_1, L_2, L_2}. \tag{3B.16}$$

The estimate (3B.15) implies that there is a constant $C(n, \rho)$ dependent only on n and ρ such that taking

$$R \leq \frac{C(n, \rho)\kappa}{A_0 L_0^2} \tag{3B.17}$$

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 section that C represents an arbitrary positive constant, dependent only on n, τ, ρ and L_0. Similarly, c will represent a positive constant strictly less than 1, also only dependent on n, τ, ρ and L_0. We will be explicit when we stray from this convention.

The estimates (3B.15) and (3B.16), together with Proposition A.3 in [Popov 2004b] show that our constructed functions P^0 and P^1 are in the Gevrey classes

$$G^{\rho}_{CL_0, CL_2, CL_2}(B_R \times \overline{\Omega}' \times (-1, 1)) \subset G^{\rho}_{CL_2, CL_2}(B_R \times \overline{\Omega}' \times (-1, 1)),$$
$$G^{\rho,\rho,\rho,1}_{L_1, L_2, CL_2, L_2}(\mathbb{T}^n \times B_R \times \overline{\Omega}' \times (-1, 1))$$

respectively, where the C does not depend on L_0 or L_2. Additionally we have the estimate

$$\| P^1 \|_{L_1, CL_2, CL_2, CL_2} \leq \kappa^{-2} \epsilon_H. \tag{3B.18}$$

Dropping the factors in our Gevrey constants dependent only on n, τ, ρ, L_0 for brevity of notation, we are in a position to state the local KAM theorem in terms of the weighted norm

$$\langle P \rangle_r := r^2 \| P^0 \|_{L_2, L_2, L_2} + \| P^1 \|_{L_1, L_2, L_2, L_2} \tag{3B.19}$$

for $0 < r < R$.

Theorem 3.2. *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 κ, L_1, L_2, R, Ω such that whenever the Hamiltonian*

$$H(\theta, I; \omega, t) = H^0(\psi_0(\omega); t) + \langle \omega, I \rangle + \langle P^0(I; \omega, t) I, I \rangle + P^1(\theta, I; \omega, t) \tag{3B.20}$$

and $0 < r < R$ are such that

$$\langle P \rangle_r < \epsilon \kappa r L_1^{-N} \tag{3B.21}$$

we can find

$$\phi \in G^{\rho(\tau+1)+1,1}\big(\Omega \times \big(-\tfrac{3}{4}, \tfrac{3}{4}\big), \Omega\big)$$

and

$$\Phi = (U, V) \in G^{\rho, \rho(\tau+1)+1,1}\big(\mathbb{T}^n \times \Omega \times \big(-\tfrac{3}{4}, \tfrac{3}{4}\big), \mathbb{T}^n \times B_R\big)$$

such that

(i) *For all $\omega \in \Omega_\kappa$ and all $t \in \big(-\frac{3}{4}, \frac{3}{4}\big)$, the map $\Phi_{\omega,t} = \Phi(\,\cdot\,; \omega, t) : \mathbb{T}^n \to \mathbb{T}^n \times B_R$ is a G^ρ 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*

$$X_{H_{\phi(\omega,t),t}} \circ \Phi_{\omega,t} = D\Phi_{\omega,t} \cdot \mathcal{L}_\omega, \tag{3B.22}$$

where

$$\mathcal{L}_\omega = \sum_{j=1}^{n} \omega_j \frac{\partial}{\partial \theta_j} \in T\mathbb{T}^n. \tag{3B.23}$$

(ii) *There exist positive constants A and C dependent only on n, τ, ρ, L_0 such that*

$$\begin{aligned}|\partial_\theta^\alpha \partial_\omega^\beta (U(\theta;\omega,t)-\theta)| + r^{-1}|\partial_\theta^\alpha \partial_\omega^\beta V(\theta;\omega,t)| + \kappa^{-1}|\partial_\omega^\beta(\phi(\omega;t)-\omega)|& \\ \le A(CL_1)^{|\alpha|}(CL_1^{\tau+1}/\kappa)^{|\beta|}\alpha!^\rho\,\beta!^{\rho(\tau+1)+1}\frac{\langle P\rangle_r}{\kappa r}L_1^N&\end{aligned} \tag{3B.24}$$

uniformly in $\mathbb{T}^n \times \Omega \times \left(-\frac{3}{4}, \frac{3}{4}\right)$.

We remark that at the endpoint $t = 0$, this result is trivial by taking $\phi(\omega, 0) = \omega$, $U(\theta, \omega, 0) = \theta$ and $V(\theta, \omega, 0) = \nabla g^0(\omega)$.

Theorem 3.2 can be proved in the same way as [Popov 2004b, Theorem 2.1], based on the rapidly converging iterative procedure introduced in [Kolmogorov 1954]. Indeed, much of the technicality in [Popov 2004b] involves the approximation of Gevrey class Hamiltonians by real-analytic Hamiltonians. Thanks to the assumption of analyticity in t in Theorem 3.2, no such approximation is necessary in the t-parameter.

In the next section, we sketch the key steps in the proof of Theorem 3.2, highlighting the points at which the presence of the t-parameter requires a modification of the argument in [Popov 2004b].

First, we discuss the result that will comprise the steps of the iterative construction. Given a Hamiltonian in the form

$$\begin{aligned}H(\theta, I;\omega,t) &= e(\omega;t) + \langle \omega, I\rangle + P(\theta, I;\omega,t)\\ &= N(I;\omega,t) + P(\theta, I;\omega,t),\end{aligned} \tag{3B.25}$$

we aim to construct a t-dependent symplectomorphism Φ and a t-dependent frequency transformation ϕ such that for $\mathcal{F} = (\Phi, \phi)$ we have

$$(H\circ\mathcal{F})(\theta, I;\omega,t) = N_+(I;\omega,t) + P_+(\theta, I;\omega,t), \tag{3B.26}$$

where $N_+(I,\omega,t) = e_+(\omega) + \langle I, \omega\rangle$ and with $|P_+|$ controlled by $|P|^r$ for some $r > 1$. This construction is analogous to that in [Pöschel 2001].

Theorem 3.3. *Suppose $\epsilon, h, v, s, r, \eta, \sigma, K$ are positive constants such that*

$$s, r < 1, \quad v < \tfrac{1}{6}, \quad \eta < \tfrac{1}{8}, \quad \sigma < \tfrac{1}{5}s, \quad \epsilon \le c\kappa\eta r\sigma^{\tau+1}, \quad \epsilon \le cvhr, \quad h \le \kappa/2K^{\tau+1}, \tag{3B.27}$$

where c is a constant dependent only on n and τ.

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 $|P|_{s,r,h} \le \epsilon$. Here, $D_{s,r}$ is as in Definition 3.1,

$$O_h := \{\omega \in \mathbb{C}^n : \operatorname{dist}(\omega, \Omega_\kappa) < h\}, \tag{3B.28}$$

and $|\cdot|_{s,r,h}$ denotes the sup-norm on $D_{s,r} \times O_h$. Then there exists a real-analytic map

$$\mathcal{F} = (\Phi, \phi) : D_{s-5\sigma,\eta r} \times O_{(1/2-3v)h} \times (-1,1) \to D_{s,r} \times O_h, \tag{3B.29}$$

where the maps

$$\Phi : D_{s-5\sigma,\eta r} \times O_h \times (-1,1) \to D_{s,r}, \tag{3B.30}$$

$$\phi : O_{(1/2-3\upsilon)h} \times (-1,1) \to O_h \tag{3B.31}$$

are such that

$$\begin{aligned} H \circ \mathcal{F} &= e_+(\omega,t) + \langle \omega, I\rangle + P_+(\theta, I;\omega,t) \\ &= N_+(I;\omega,t) + P_+(\theta, I;\omega,t) \end{aligned} \tag{3B.32}$$

and we have the new remainder estimate

$$|P_+|_{s-5\sigma,\eta r,(1/2-2\upsilon)h} \le C\left(\frac{\epsilon^2}{\kappa r \sigma^{\tau+1}} + (\eta^2 + K^n e^{-K\sigma})\epsilon\right). \tag{3B.33}$$

Moreover Φ is symplectic for each (ω,t) and has second component affine in I. Finally, we have the uniform estimates on the change of variables

$$|W(\Phi - \mathrm{id})|,\ |W(D\Phi - \mathrm{Id})W^{-1}| \le \frac{C\epsilon}{\kappa r \sigma^{\tau+1}}, \tag{3B.34}$$

$$|\phi - \mathrm{id}|,\ \upsilon h|D\phi - \mathrm{Id}| \le \frac{C\epsilon}{r}, \tag{3B.35}$$

where $W = \mathrm{diag}(\sigma^{-1}\mathrm{Id}, r^{-1}\mathrm{Id})$. All estimates are uniform in the analytic parameter $t \in (-1,1)$.

This theorem is identical to [Popov 2004b, Proposition 3.2], with all estimates uniform in the parameter t. The proof is identical, with a detailed exposition in [Pöschel 2001]. The application of [Popov 2004b, Lemma 3.4] to obtain the frequency transformation ϕ is replaced by Proposition C.2 in our setting.

As in [Pöschel 2001; Popov 2000a], Theorem 3.3 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 $H \in G^{\rho,\rho,\rho,1}((\mathbb{T}^n \times D \times \Omega) \times (-1,1))$, we require the approximation result Proposition 3.4.

3C. ***Proof of the KAM theorem.*** Following the proof of Theorem 3.2 in [Popov 2004b, Section 3], we extend the $P^j(\theta, I, \omega, t)$ to Gevrey functions

$$\widetilde{P}^j \in G^{\rho,\rho,1}_{CL_1,CL_2,CL_2}(\mathbb{T}^n \times \mathbb{R}^{2n} \times (-1,1)), \tag{3C.1}$$

where C depends only on n and ρ. We do this whilst preserving analyticity in t by making use of an adapted version of the Whitney extension theorem for anisotropic Gevrey classes, from Proposition 3.8.

We thus obtain the estimate

$$\|\widetilde{P}^j\| \le AL_1^{n+1}\|P^j\|, \tag{3C.2}$$

where A also only depends on n and ρ. We then cut off $\widetilde{P}^j$ without loss to have (I,ω) supported in $B_1 \times B_{\bar{R}} \subset \mathbb{R}^{2n}$, where $1 \ll \bar{R}$ is such that $\Omega^0 \subset B_{\bar{R}-1}$. From here, we suppress the tilde in our notation, as well as the factor C in our Gevrey constant. We require the following approximation result for functions in anisotropic Gevrey classes that plays a key role in the KAM iterative scheme.

Proposition 3.4. *Suppose $P \in G^{\rho,\rho,1}_{L_1,L_2,L_2}(\mathbb{T}^n \times \mathbb{R}^{2n} \times (-1,1))$ satisfies $\operatorname{supp}_{(I,\omega)}(P) \subset B_1 \times B_{\bar R}$. If u_j, w_j, v_j are positive real sequences monotonically tending to zero such that*

$$v_j L_2, w_j L_2 \le u_j L_1 \le 1, \quad v_0, w_0 \le L_2^{-1-\zeta}, \tag{3C.3}$$

where $1 \le L_1 \le L_2$ and $0 < \zeta \le 1$ are fixed, then we can find a sequence of real-analytic functions $P_j : U_j \to \mathbb{C}$ such that

$$|P_{j+1} - P_j|_{U_{j+1}} \le C(\bar R^n + 1) L_1^n \exp\bigl(-\tfrac{3}{4}(\rho-1)(2L_1u_j)^{-1/(\rho-1)}\bigr)\|P\|, \tag{3C.4}$$

$$|P_0|_{U_0} \le C(\bar R^n + 1)\bigl(1 + L_1^n \exp\bigl(-\tfrac{3}{4}(\rho-1)(2L_1u_0)^{-1/(\rho-1)}\bigr)\bigr), \tag{3C.5}$$

$$|\partial_x^\alpha(P - P_j)(\theta, I; \omega, t)| \le C(1 + \bar R^n) L_1^n L_2 \exp\bigl(-\tfrac{3}{4}(\rho-1)(2L_1u_j)^{-1/(\rho-1)}\bigr) \tag{3C.6}$$

in $\mathbb{T}^n \times B_1 \times B_{\bar R} \times (-1,1)$ for $|\alpha| \le 1$, where

$$\begin{aligned} U_j^m := \{(\theta, I; \omega, t) \in \mathbb{C}^n/2\pi\mathbb{Z}^n \times \mathbb{C}^n \times \mathbb{C}^n \times \mathbb{C} :& \\ |\mathrm{Re}(\theta)| \le \pi, |\mathrm{Re}(I)| \le 2, |\mathrm{Re}(\omega)| \le \bar R + 1, |\mathrm{Re}(t)| \le 1,& \\ |\mathrm{Im}(\theta)| \le 2u_j, |\mathrm{Im}(I)| \le 2v_j, |\mathrm{Im}(\omega_k)| \le 2w_j, |\mathrm{Im}(t)| \le (2L_2)^{-1}\}& \end{aligned} \tag{3C.7}$$

and

$$U_j := U_j^1, \tag{3C.8}$$

where we have identified $[-\pi, \pi]^n$ with $\mathbb{T}^n$ for simplicity of notation.

The proof of Proposition 3.4 can be found in [Popov 2004b, Section 3]. The first step is to extend P to functions $F_j : U_j^2 \to \mathbb{C}$ that are almost analytic in (θ, I, ω) and are analytic in t. The Gevrey estimate on t-derivatives of P implies that the Taylor expansions in t have radius of convergence L_2^{-1}, and so the expression

$$F_j(\theta + i\tilde\theta, I + i\tilde I, \omega + i\tilde\omega, t + i\tilde t) := \sum_{\mathcal{M}_j} \partial_\theta^\alpha \partial_I^\beta \partial_\omega^\gamma P(\theta, I; \omega, t) \frac{(i\tilde\theta)^\alpha (i\tilde I)^\beta (i\tilde\omega)^\gamma (i\tilde t)^\delta}{\alpha!\,\beta!\,\gamma!\,\delta!} \tag{3C.9}$$

is convergent on U_j^2, where the index set is as in [Popov 2004b].

The remainder of the proof in [Popov 2004b] can be followed without change. As P is analytic in t, we do not need to consider shrinking domains of analyticity as in the other variables.

The iterative scheme in [Popov 2004b, Section 3.3] can then be carried out, 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 3.3 are always satisfied, as well as decreasing sequences of the parameters u_j, v_j, w_j such that the hypotheses of the Proposition 3.4 are always satisfied. Due to the modifications made in Theorem 3.3 and Proposition 3.4 from their analogues in [Popov 2004b], all estimates are uniform in the analytic parameter $t \in (-1, 1)$.

Writing $U_j = U_j^1 \cap \{|I| < r_j\}$, where U_j^1 is defined as in Proposition 3.4, and applying Proposition 3.4 to the terms P^0, P^1 from (3B.14), we obtain sequences P_j^0, P_j^1 of real-analytic functions in U_j^1 that are good approximations to P^0 and P^1.

Setting

$$P_j(\theta, I; \omega, t) := \langle P_j^0(I; \omega, t) I, I\rangle + P_j^1(\theta, I; \omega, t), \tag{3C.10}$$

Proposition 3.4, together with the factors picked up during the Whitney extension of P^0, P^1 in (3C.2), implies the estimates

$$|P_0|_{U_0} \le \tilde{\epsilon}_0, \tag{3C.11}$$

$$|P_j - P_{j-1}|_{U_j} \le \tilde{\epsilon}_j, \tag{3C.12}$$

where $\tilde{\epsilon}_j$ is a positive sequence rapidly converging to zero.

Defining the Hamiltonian

$$H_j(\theta, I; \omega, t) = N_0(I; \omega) + P_j(\theta, I; \omega, t) = \langle \omega, I \rangle + P_j(\theta, I; \omega, t), \tag{3C.13}$$

which is real-analytic in U_j, one can now perform the KAM iterative scheme as in [Popov 2004b, Proposition 3.5], using the key ingredient of Theorem 3.3. For $j \ge 0$ we denote by $\mathcal{D}_j$ the class of real-analytic diffeomorphisms from $D_{j+1} \times O_{j+1} \times (-1, 1) \to D_j \times O_j$ of the form

$$\mathcal{F}(\theta, I; \omega, t) = (\Phi(\theta, I; \omega, t), \phi(\omega; t)) = (U(\theta; \omega, t), V(\theta, I; \omega, t), \phi(\omega; t)), \tag{3C.14}$$

where Φ is affine in I and canonical for fixed (ω, t). The domains are defined in terms of the parameters by $D_j = D_{s_j, r_j}$ and $O_j = O_{h_j}$.

Proposition 3.5. *Suppose P_j is real-analytic on U_j for each $j \ge 0$ and that we have the estimates*

$$|P_0|_{U_0} \le \tilde{\epsilon}_0, \tag{3C.15}$$

$$|P_j - P_{j-1}|_{U_j} \le \tilde{\epsilon}_j \tag{3C.16}$$

for each $j \ge 1$.

Then for each $j \ge 0$ 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

$$\mathcal{F}^{j+1} = \mathcal{F}_0 \circ \cdots \circ \mathcal{F}_j : D_{j+1} \times O_{j+1} \times (-1, 1) \to (D_0 \times O_0) \cap U_j, \tag{3C.17}$$

with the convention that the empty composition is the identity and where the $\mathcal{F}_j \in \mathcal{D}_j$ are such that

$$H_j \circ \mathcal{F}^{j+1} = N_{j+1} + R_{j+1}, \tag{3C.18}$$

$$|R_{j+1}|_{j+1} \le \epsilon_{j+1}, \tag{3C.19}$$

$$|\overline{W}_j(\mathcal{F}_j - \mathrm{id})|_{j+1},\ |\overline{W}_j(D\mathcal{F}_j - \mathrm{Id})\overline{W}_j^{-1}| < \frac{C\epsilon_j}{r_j h_j}, \tag{3C.20}$$

$$|\overline{W}_0(\mathcal{F}^{j+1} - \mathcal{F}^j)|_{j+1} < \frac{C\epsilon_j}{r_j h_j}, \tag{3C.21}$$

where the constants C depend only on n and ρ and $\overline{W}_j = \mathrm{diag}(\sigma_j^{-1}\mathrm{Id}, r_j^{-1}\mathrm{Id}, h_j^{-1}\mathrm{Id})$.

To show that this iterative scheme converges in the Gevrey class $G^{\rho, \rho(\tau+1)+1, \rho(\tau+1)+1, 1}$ requires Gevrey estimates for the $\mathcal{S}_j := \mathcal{F}^{j+1} - \mathcal{F}^j$. To this end we introduce the domains

$$\widetilde{D}_j := \{(\theta, I) \in D_j : |\mathrm{Im}(\theta)| < s_j/2\}, \quad \widetilde{O}_j := \{\omega \in \mathbb{C}^n : \mathrm{dist}(\omega, \Omega_\kappa) < h_j/2\}. \tag{3C.22}$$

For multi-indices α, β with $|\beta| \leq m$, we also introduce the following notation for the $(m-|\beta|)$-th Taylor remainder in the frequency variable, centred at ω:

$$R^m_\omega(\partial^\alpha_\theta \partial^\beta_\omega \mathcal{S}^j)(\theta, I, \omega', t) := \partial^\alpha_\theta \partial^\beta_\omega \mathcal{S}^j - \sum_{|\gamma|\leq m-|\beta|} (\omega' - \omega)^\gamma \partial^\alpha_\theta \partial^{\beta+\gamma}_\omega \mathcal{S}^j(\theta, I, \omega, t)/\gamma!. \tag{3C.23}$$

We then have the following Gevrey estimates of [Popov 2004b, Lemma 3.6] uniformly in the t-parameter.

Lemma 3.6. *We have*

$$|\overline{W}_0 \partial^\alpha_\theta \partial^\beta_\omega \mathcal{S}^j(\theta, 0, \omega, t)| \leq \hat{\epsilon} A C^{|\alpha|+|\beta|} L_1^{|\alpha|+|\beta|(\tau+1)+1} \kappa^{-|\beta|} \alpha!^\rho \, \beta!^{\rho'} \, E_j^{1/2} \tag{3C.24}$$

for all $(\theta, 0; \omega, t) \in \widetilde{D}_{j+1} \times \widetilde{O}_{j+1} \times (-1, 1)$, *where* $\rho' = \rho(\tau+1)+1$, *and*

$$\begin{aligned}|\overline{W}_0(&R^m_\omega \partial^\alpha_\theta \partial^\beta_\omega \mathcal{S}^j)(\theta, 0, \omega', t)| \\ &\leq \hat{\epsilon} A C^{m+|\alpha|+1} L_1^{|\alpha|+(m+1)(\tau+1)+1} \kappa^{-m-1} \frac{|\omega-\omega'|^{m-|\beta|+1}}{(m-|\beta|+1)!} \alpha!^\rho \, (m+1)!^{\rho'} \, E_j^{1/2}\end{aligned} \tag{3C.25}$$

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, ρ, τ, ζ.

We can now bound derivatives in t; we use the Cauchy estimate from Proposition C.1. This yields the following corollary.

Corollary 3.7. *We have*

$$|\overline{W}_0 \partial^\alpha_\theta \partial^\beta_\omega \partial^\gamma_t \mathcal{S}^j(\theta, 0; \omega, t)| \leq \hat{\epsilon} A C^{|\alpha|+|\beta|+|\gamma|} L_1^{|\alpha|+|\beta|(\tau+1)+1} \kappa^{-|\beta|} \alpha!^\rho \, \beta!^{\rho'} \, \gamma! \, E_j^{1/2} \tag{3C.26}$$

for all $(\theta, 0; \omega, t) \in \widetilde{D}_{j+1} \times \widetilde{O}_{j+1} \times \left(-\frac{3}{4}, \frac{3}{4}\right)$, *where* $\rho' = \rho(\tau+1)+1$, *and*

$$\begin{aligned}|\overline{W}_0(&R^m_\omega \partial^\alpha_\theta \partial^\beta_\omega \partial^\gamma_t \mathcal{S}^j)(\theta, 0, \omega', t)| \\ &\leq \hat{\epsilon} A C^{m+|\alpha|+|\gamma|+1} L_1^{|\alpha|+(m+1)(\tau+1)+1} \kappa^{-m-1} \frac{|\omega-\omega'|^{m-|\beta|+1}}{(m-|\beta|+1)!} \alpha!^\rho \, (m+1)!^{\rho'} \, \gamma! \, E_j^{1/2}\end{aligned} \tag{3C.27}$$

for all $\theta \in \mathbb{T}^n$, $\omega, \omega' \in \Omega_\kappa$, $t \in \left(-\frac{3}{4}, \frac{3}{4}\right)$ *and* $|\beta| \leq m$, *where the constants* A, C *only depend on* n, ρ, τ, ζ.

From Proposition 3.5 and Corollary 3.7, the rapid decay of E_j implies that the limit

$$\partial^\alpha_\theta \partial^\gamma_t \mathcal{H}^\beta(\theta, \omega; t) := \lim_{j\to\infty} \partial^\alpha_\theta \partial^\beta_\omega \partial^\gamma_t (\mathcal{F}^j(\theta, 0; \omega, t) - (\theta, 0, \omega)) \tag{3C.28}$$

exists for each $(\theta; \omega, t) \in \mathbb{T}^n \times \Omega_\kappa \times \left(-\frac{3}{4}, \frac{3}{4}\right)$, and each triple of multi-indices α, β, γ. Convergence is uniform, and the limit is smooth in θ and t and continuous in ω, with $\partial^\alpha_\theta \partial^\gamma_t (\mathcal{H}^\beta) = \partial^\alpha_\theta \partial^\gamma_t \mathcal{H}^\beta$, justifying the notation in (3C.28).

We now need to use the jet $\mathcal{H} = (\partial^\alpha_\theta \partial^\gamma_t \mathcal{H}^\beta)$ of continuous functions $\mathbb{T}^n \times \Omega_\kappa \times \left(-\frac{3}{4}, \frac{3}{4}\right) \to \mathbb{T}^n \times D \times \Omega$ to obtain a Gevrey function on $\mathbb{T}^n \times \Omega \times \left(-\frac{3}{4}, \frac{3}{4}\right)$ by using a Gevrey version of the Whitney extension theorem. We define

$$(R^m_\omega \partial^\alpha_\theta \partial^\gamma_t \mathcal{H})_\beta(\theta, \omega', t) := \partial^\alpha_\theta \partial^\gamma_t \mathcal{H}^\beta(\theta, \omega', t) - \sum_{|\delta|\leq m-|\beta|} (\omega' - \omega)^\delta \partial^\alpha_\theta \partial^\gamma_t \mathcal{H}^{\beta+\delta}(\theta; \omega, t)/\gamma!. \tag{3C.29}$$

In this notation, the results of Corollary 3.7 yield

$$|\overline{W}_0\partial_\theta^\alpha\partial_t^\gamma\mathcal{H}^\beta(\theta;\omega,t)| \le \hat{\epsilon}AL_1(CL_1)^{|\alpha|}(CL_1^{\tau+1}/\kappa)^{|\beta|}C^\gamma\alpha!^\rho\,\beta!^{\rho'}\,\gamma!, \tag{3C.30}$$

$$|\overline{W}_0(R^m_\omega\partial_\theta^\alpha\partial_t^\gamma\mathcal{H})_\beta(\theta,\omega',t)| \le \hat{\epsilon}AL_1(CL_1)^{|\alpha|}(CL_1^{\tau+1}/\kappa)^{m+1}C^\gamma\frac{|\omega-\omega'|^{m-|\beta|+1}}{(m-|\beta|+1)!}\alpha!^\rho\,(m+1)!^{\rho'}\,\gamma! \tag{3C.31}$$

for $|\beta|\le m$, and $(\theta,\omega,\omega',t)\in\mathbb{T}^n\times\Omega_\kappa\times\Omega_\kappa\times\left(-\frac{3}{4},\frac{3}{4}\right)$, where A and C depend only on n, ρ, τ. These estimates allow us to apply the following consequence of Theorem D.3.

Proposition 3.8. *Suppose $K\subset\mathbb{R}^n$ is compact, and $1\le\rho<\rho'$. If the jet $(f^{\alpha,\beta,\gamma})$ of functions $f^{\alpha,\beta,\gamma}:\mathbb{T}^n\times K\times\left(-\frac{3}{4},\frac{3}{4}\right)\to\mathbb{R}$ is continuous on $\mathbb{T}^n\times K\times\left(-\frac{3}{4},\frac{3}{4}\right)$ and is smooth in $(\theta,t)\in\mathbb{T}^n\times\left(-\frac{3}{4},\frac{3}{4}\right)$ for each fixed $\omega\in K$, where*

$$\partial_\theta^{\alpha'}\partial_t^{\gamma'}(f^{\alpha,\beta,\gamma}) = f^{\alpha+\alpha',\beta,\gamma+\gamma'}, \tag{3C.32}$$

and we have the estimates

$$|f^{\alpha,\beta,\gamma}(\theta;\omega,t)| \le AC_1^{|\alpha|}C_2^{|\beta|}C_3^{|\gamma|}\alpha!^\rho\,\beta!^{\rho'}\,\gamma!, \tag{3C.33}$$

$$|(R^m_\omega\partial_\theta^\alpha\partial_t^\gamma f)_\beta(\theta,\omega',t)| \le AC_1^{|\alpha|}C_2^{m+1}C_3^{|\gamma|}\frac{|\omega-\omega'|^{m-|\beta|+1}}{(m-|\beta|+1)!}\alpha!^\rho\,(m+1)!^{\rho'}\,\gamma! \tag{3C.34}$$

then there exist positive constants A_0, C_0, dependent only on (n,ρ,τ) (in particular, independent of the set K) such that we can extend f to $\tilde{f}\in G^{\rho,\rho',1}\left(\mathbb{T}^n\times\mathbb{R}^n\times\left(-\frac{3}{4},\frac{3}{4}\right)\right)$ such that $\partial_\theta^\alpha\partial_\omega^\beta\partial_t^\gamma\tilde{f}=f^{\alpha,\beta,\omega}$ on $\mathbb{T}^n\times K\times\left(-\frac{3}{4},\frac{3}{4}\right)$ and

$$|\partial_\theta^\alpha\partial_\omega^\beta\partial_t^\gamma\tilde{f}(\theta,\omega)| \le A_0A\max(C_1,1)C_0^{|\alpha|+|\beta|+|\gamma|+n}C_1^{|\alpha|+n}C_2^{|\beta|}C_3^{|\gamma|}\alpha!^\rho\,\beta!^{\rho'}\,\gamma!\,. \tag{3C.35}$$

The proof of Proposition 3.8 is identical to that in [Popov 2004b, Theorem 3.7], making use of Theorem D.3 involving the parameter t. Having established Proposition 3.8, the proof of Theorem 3.2 can be completed as in [Popov 2004b, Section 3.5] without modification.

3D. *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 3.2. The Gevrey index $\rho(\tau+1)+1$ frequently appears in these results, and so we introduce $\rho':=\rho(\tau+1)+1$.

Theorem 3.9. *Fix $0<\zeta\le 1$ and let $H^0(I;t)$ be a real-valued nondegenerate smooth family of Hamiltonians in $G^{\rho,1}(D^0\times(-1,1))$ and let D be a subdomain with $\overline{D}\subset D^0$. We define $\Omega=\nabla H^0(D)$ and fix $L_2\ge L_1\ge 1$ and $\kappa\le L_2^{-1-\zeta}$ such that $L_2\ge L_0$ and $\Omega_\kappa\ne\varnothing$. Then there exists $N=N(n,\rho,\tau)$ and $\epsilon>0$ independent of κ, L_1, L_2 and $D\subset D^0$ 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 norm*

$$\epsilon_H:=\kappa^{-2}\|H-H^0\|_{L_1,L_2,L_2}\le\epsilon L_1^{-N} \tag{3D.1}$$

there exists a map

$$\overline{\Phi}=(\overline{U},\overline{V})\in G^{\rho,\rho',1}\left(\mathbb{T}^n\times\Omega\times\left(-\tfrac{3}{4},\tfrac{3}{4}\right),\mathbb{T}^n\times D\right) \tag{3D.2}$$

such that:

(i) *For each $\omega\in\Omega_\kappa$ and each $t\in\left(-\frac{3}{4},\frac{3}{4}\right)$, $\Lambda_\omega=\{\overline{\Phi}(\theta;\omega,t):\theta\in\mathbb{T}^n\}$ is an embedded invariant Lagrangian torus of H, and $X_H\circ\overline{\Phi}(\cdot\,;\omega,t)=D\overline{\Phi}(\cdot\,;\omega,t)\cdot\mathcal{L}_\omega$.*

(ii) *There exist constants $A, C > 0$ independent of κ, L_1, L_2 and $D \subset D^0$ such that*

$$|\partial_\theta^\alpha \partial_\omega^\beta(\bar{U}(\theta;\omega,t)-\theta)| + \kappa^{-1}|\partial_\theta^\alpha \partial_\omega^\beta(\bar{V}(\theta;\omega,t)-\nabla g^0(\omega))| \\ \le A(CL_1)^{|\alpha|}(CL_1^{\tau+1}/\kappa)^{|\beta|}\alpha!^\rho\,\beta!^{\rho'}\,L_1^{N/2}\epsilon_H^{1/2} \quad \text{(3D.3)}$$

uniformly in $\mathbb{T}^n \times \Omega \times \left(-\frac{3}{4}, \frac{3}{4}\right)$.

The proof of Theorem 3.9 is identical to [Popov 2004b, Theorem 1.1], making use of Theorem 3.2. We can now use Theorem 3.9 to obtain the Birkhoff normal form as done in [Popov 2004b].

Theorem 3.10. *Suppose the assumptions of Theorem 3.9 hold. Then there exist $N(n,\rho,\tau) > 0$ and $\epsilon > 0$ independent of κ, 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*

$$\epsilon_H \le \epsilon L_1^{-N-2(\tau+2)}, \quad \text{(3D.4)}$$

where ϵ_H is as in (3D.1), *there is a family of $G^{\rho',\rho'}$ maps $\omega : D \times \left(-\frac{1}{2}, 1, 2\right) \to \Omega$ and a family of maps $\chi \in G^{\rho,\rho',\rho'}\left(\mathbb{T}^n \times D \times \left(-\frac{1}{2}, 1, 2\right), \mathbb{T}^n \times D\right)$ that are diffeomorphisms and exact symplectomorphisms respectively for each fixed $t \in \left(-\frac{1}{2}, 1, 2\right)$. Moreover, we can choose the maps ω and χ such that family of transformed Hamiltonians*

$$\widetilde{H}(\theta, I; t) := (H \circ \chi)(\theta, I; t) \quad \text{(3D.5)}$$

is of Gevrey class $G^{\rho,\rho',\rho'}\left(\mathbb{T}^n \times D \times \left(-\frac{1}{2}, 1, 2\right)\right)$ and can be decomposed as

$$K(I;t) + R(\theta, I; t) := \widetilde{H}(0, I; t) + (\widetilde{H}(\theta, I; t) - \widetilde{H}(0, I; t)) \quad \text{(3D.6)}$$

such that:

(i) *$\mathbb{T}^n \times \{I\}$ is an invariant Lagrangian torus of $\widetilde{H}(\,\cdot\,,\cdot\,;t)$ for each $I \in E_\kappa(t) = \omega^{-1}(\widetilde{\Omega}_\kappa; t)$ and each $t \in \left(-\frac{1}{2}, 1, 2\right)$.*

(ii) *$\partial_I^\beta(\nabla K(I;t) - \omega(I;t)) = \partial_I^\beta R(\theta, I; t) = 0$ for all $(\theta, I; t) \in \mathbb{T}^n \times E_\kappa(t) \times \left(-\frac{1}{2}, 1, 2\right)$, $\beta \in \mathbb{N}^n$.*

(iii) *There exist $A, C > 0$ independent of κ, L_1, L_2, and $D \subset D^0$ such that we have the estimates*

$$|\partial_\theta^\alpha \partial_I^\beta \partial_t^\delta \phi(\theta, I; t)| + |\partial_I^\beta \partial_t^\delta(\omega(I;t) - \nabla H^0(I;t))| + |\partial_\theta^\alpha \partial_I^\beta \partial_t^\delta(\widetilde{H}(\theta, I; t) - H^0(I;t))| \\ \le A\kappa C^{|\alpha|+|\beta|+|\delta|} L_1^{|\alpha|}(L_1^{\tau+1}/\kappa)^{|\beta|}\alpha!^\rho\,\beta!^{\rho'}\,\delta!^{\rho'}\,L_1^{N/2}\epsilon_H^{1/2} \quad \text{(3D.7)}$$

uniformly in $\mathbb{T}^n \times D \times \left(-\frac{1}{2}, 1, 2\right)$ for all α, β, where $\phi \in G^{\rho,\rho',\rho'}\left(\mathbb{T}^n \times D \times \left(-\frac{1}{2}, 1, 2\right)\right)$ is such that $\langle \theta, I\rangle + \phi(\theta, I; t)$ generates the symplectomorphism χ in the sense of (3E.8).

Remark 3.11. For our purposes, high regularity in the t-parameter is not required, so we have dropped from analyticity to $G^{\rho'}$ regularity in t at this point in order to simplify the proceeding arguments. We expect that analyticity in t could be preserved by using a stronger variant of the Komatsu implicit function theorem than Corollary A.5.

Proof. We begin by taking ϵ, N as in Theorem 3.9 and noting that $\epsilon_H \le \epsilon L_1^{-N-2}$ by assumption. This implies that the factor $(ACL_1)L_1^{N/2}\sqrt{\epsilon_H}$ occurring in the Gevrey estimate (3D.3) can be bounded above by $AC\sqrt{\epsilon}$. Hence, taking ϵ small enough that both the conclusion to Theorem 3.9 holds as well as $AC\sqrt{\epsilon} < \frac{1}{2}$, we can first apply the Cauchy estimate from Proposition C.1 to (3D.3) in t, and then apply a variant of the Komatsu implicit function theorem, Corollary A.5, to obtain a solution $\theta(\gamma;\omega,t) : \mathbb{T}^n \times \Omega \times \left(-\frac{1}{2}, 1, 2\right) \to \mathbb{T}^n$ to the implicit equation

$$\overline{U}(\theta;\omega,t) = \gamma. \tag{3D.8}$$

Moreover, this solution satisfies the Gevrey estimate

$$|\partial_\gamma^\alpha \partial_\omega^\beta \partial_t^\delta (\theta(\gamma;\omega,t) - \gamma)| \le AC^{|\alpha|+|\beta|+|\delta|} L_1^{|\alpha|} (L_1^{\tau+1}/\kappa)^{|\beta|} \alpha!^\rho\, \beta!^{\rho'} \delta!^{\rho'}\, L_1^{N/2} \sqrt{\epsilon_H} \tag{3D.9}$$

uniformly on $\mathbb{T}^n \times \Omega \times \left(-\frac{1}{2}, 1, 2\right)$.

We set $F(\gamma;\omega,t) := \overline{V}(\theta(\gamma;\omega,t);\omega,t)$. In terms of $(\gamma;\omega,t)$, the Lagrangian torus Λ_ω is now given by $(\gamma, F(\gamma;\omega,t) : \gamma \in \mathbb{T}^n)$ for each $\omega \in \Omega_\kappa$ and each $t \in \left(-\frac{1}{2}, 1, 2\right)$. Moreover, Proposition A.7 on the composition of Gevrey functions gives us the estimate

$$|\partial_\gamma^\alpha \partial_\omega^\beta \partial_t^\delta (F(\gamma;\omega,t) - \nabla g^0(\omega))| \le A\kappa C^{|\alpha|+|\beta|+|\delta|} L_1^{|\alpha|} (L_1^{\tau+1}/\kappa)^{|\beta|} \alpha!^\rho\, \beta!^{\rho'}\, \delta!^{\rho'}\, L_1^{N/2} \sqrt{\epsilon_H}. \tag{3D.10}$$

We next construct functions $\psi \in G^{\rho,\rho',\rho'}\left(\mathbb{R}^n \times \Omega \times \left(\frac{1}{2}, 1, 2\right)\right)$ and $R \in G^{\rho',\rho'}\left(\Omega \times \left(\frac{1}{2}, 1, 2\right)\right)$ such that the function

$$Q(x;\omega,t) := \psi(x;\omega,t) - \langle x, R(\omega,t)\rangle \tag{3D.11}$$

is 2π-periodic in x and satisfies

$$\nabla_x \psi(x;\omega,t) = F(p(x),\omega,t) \tag{3D.12}$$

in $\mathbb{R}^n \times \Omega_\kappa \times \left(\frac{1}{2}, 1, 2\right)$, where $p : \mathbb{R}^n \to \mathbb{T}^n$ is the canonical projection, as well as the estimate

$$\begin{aligned} &|\partial_x^\alpha \partial_\omega^\beta \partial_t^\delta Q(x;\omega,t)| + |\partial_\omega^\beta \partial_t^\delta (R(\omega,t) - \nabla g^0(\omega))| \\ &\qquad \le A\kappa C^{|\alpha|+|\beta|+|\delta|} L_1^{|\alpha|} (L_1^{\tau+1}/\kappa)^{|\beta|} \alpha!^\rho\, \beta!^{\rho'}\, \delta!^{\rho'}\, L_1^{N/2} \sqrt{\epsilon_H} \end{aligned} \tag{3D.13}$$

for $(x;\omega,t) \in \mathbb{R}^n \times \Omega \times \left(\frac{1}{2}, 1, 2\right)$.

We do this by first integrating the canonical 1-form $I\,dx$ over the chain

$$c_x := \{(sx, F(p(sx);\omega,t)) : 0 \le s \le 1\} \subset \mathbb{R}^n \times D. \tag{3D.14}$$

We define

$$\tilde{\psi}(x;\omega,t) := \int_{c_x} \sigma = \int_0^1 \langle F(p(sx);\omega,t), x\rangle\, ds \tag{3D.15}$$

in $\mathbb{R}^n \times \Omega \times \left(\frac{1}{2}, 1, 2\right)$. From the estimate (3D.10) it follows that $\tilde{\psi}(x;\omega,t) - \langle \nabla g^0(\omega), x\rangle$ is bounded above by the right-hand side of (3D.13) in $[0, 4\pi]^n \times \Omega \times \left(\frac{1}{2}, 1, 2\right)$. 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 (3D.13).

Since for $\omega \in \Omega_\kappa$ we know that Λ_ω is a Lagrangian torus, it follows that the integral of the canonical 1-form over any closed chain in Λ_ω is homotopy invariant. This means that such an integral is a homomorphism from the fundamental group of Λ_ω to $\mathbb{R}$. Hence

$$\tilde{\psi}(x + 2\pi m;\omega,t) - \tilde{\psi}(x;\omega,t) = \langle 2\pi m, R(\omega,t)\rangle \tag{3D.16}$$

and so the function

$$\widetilde{Q}(x;\omega,t) := \tilde{\psi}(x,\omega) - \langle x, R(\omega,t)\rangle \tag{3D.17}$$

both satisfies the Gevrey estimate in (3D.13) and is 2π-periodic in x for $(\omega,t)\in\Omega_\kappa\times\left(\frac{1}{2},1,2\right)$.

To obtain the sought Q in (3D.11) from $\widetilde{Q}$, we use an averaging trick. Choosing $f\in G^\rho_C(\mathbb{R}^n)$ for some positive constant C such that f is supported in $[\pi/2, 7\pi/2]^n$ and

$$\sum_{k\in\mathbb{Z}^n} f(x+2\pi k) = 1 \tag{3D.18}$$

for each $x\in\mathbb{R}^n$, it then follows that

$$Q(x;\omega,t) := \sum_{k\in\mathbb{Z}^n} f(x+2\pi k)\widetilde{Q}(x+2\pi k;\omega,t) \tag{3D.19}$$

is 2π-periodic in x for every $\omega\in\Omega$ and coincides with $\widetilde{Q}$ for $\omega\in\Omega_\kappa$. Moreover, Q satisfies the same Gevrey estimate (3D.13) as $\widetilde{Q}$. We define

$$\psi(x;\omega,t) := Q(x;\omega,t) + \langle x, R(\omega,t)\rangle. \tag{3D.20}$$

Note that by multiplying Q and $R-\nabla g^0$ by a cut-off function $h\in G^{\rho'}_{C/\kappa}$ which is equal to 1 in a ω-neighbourhood of Ω_κ and vanishes for $\operatorname{dist}(\omega,\mathbb{R}^n\setminus\Omega)\le\kappa/2$, 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 $\operatorname{dist}(\omega,\mathbb{R}^n\setminus\Omega)\le\kappa/2$. This cutoff preserves the Gevrey estimates on ψ.

Now since $\epsilon_H L_1^{N+2(\tau+2)}\le\epsilon$, we have that $\kappa A(CL_1)(CL_1^{\tau+1}/\kappa)L_1^{N/2}\sqrt{\epsilon_H}\le AC^2\sqrt{\epsilon}$. By taking ϵ 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 (3D.13). Hence we have a $G^{\rho,\rho'}$-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 Λ_ω will be given by $R(\omega,t)$. Hence from (3D.13) and Proposition A.4, it follows that for ϵ sufficiently small, the map

$$(\omega,t)\mapsto(I(\omega,t),t)=(R(\omega,t),t) \tag{3D.21}$$

is a $G^{\rho',\rho'}$-diffeomorphism and we have the Gevrey estimate

$$|\partial_I^\alpha\partial_t^\beta(\omega(I,t)-\nabla H^0(I;t))|\le A\kappa C^{|\alpha|+|\beta|}(L_1^{\tau+1}/\kappa)^{|\alpha|}\alpha!^{\rho'}\,\beta!^{\rho'}\,L_1^{N/2}\sqrt{\epsilon_H} \tag{3D.22}$$

uniformly for $(\theta,I,t)\in\mathbb{T}^n\times D\times\left(\frac{1}{2},1,2\right)$.

We construct the sought symplectomorphism χ using the generating function $\Phi(x,I;t)$, setting

$$\Phi(x,I;t)=\psi(x,\omega(I;t);t) \tag{3D.23}$$

and noting that we have the required 2π-periodicity of $\phi(x,I;t):=\Phi(x,I,t)-\langle x,I\rangle$, and from Proposition A.7, we also have the estimate

$$|\partial_x^\alpha\partial_I^\beta\partial_t^\delta(\Phi(x,I;t-\langle x,I\rangle))|\le A\kappa C^{|\alpha|+|\beta|+|\delta|}L_1^{|\alpha|}(L_1^{\tau+1}/\kappa)^{|\beta|}\alpha!^{\rho}\,\beta!^{\rho'}\,\delta!^{\rho'}\,L_1^{N/2}\sqrt{\epsilon_H}. \tag{3D.24}$$

We can then apply Corollary A.5 to solve the implicit equation

$$\partial_I \Phi(\gamma, I, t) = \theta \tag{3D.25}$$

for γ with the estimate

$$|\partial_\theta^\alpha \partial_I^\beta \partial_t^\delta (\gamma(\theta, I, t) - \theta)| \leq A\kappa C^{|\alpha|+|\beta|+|\delta|} L_1^{|\alpha|} (L_1^{\tau+1}/\kappa)^{|\beta|} \alpha!^\rho \beta!^{\rho'} \delta!^{\rho'} L_1^{N/2} \sqrt{\epsilon_H}. \tag{3D.26}$$

This completes the construction of a symplectomorphism χ satisfying

$$\chi(\partial_I \Phi(\theta, I, t), I) = (\theta, \partial_\theta \Phi(\theta, I, t)). \tag{3D.27}$$

It follows that

$$(\theta, F(\theta; \omega, t)) = \chi(\partial_I \Phi(\theta, I(\omega), t), I(\omega)) = \chi(\theta, I(\omega), t) \tag{3D.28}$$

for $\omega \in \Omega_\kappa$ and so

$$\Lambda_\omega = \{\chi(\theta, I(\omega), t) : \theta \in \mathbb{T}^n\} \tag{3D.29}$$

for $(\omega, t) \in \Omega_\kappa \times \left(\tfrac{1}{2}, 1, 2\right)$.

We now set $\widetilde{H}$, K, R as in the theorem statement in terms of the symplectomorphism χ. Since H is constant on Λ_ω 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 (3D.7) for $\widetilde{H}(\theta, I, t) - H(I, t)$ follows from Proposition A.7. □

3E. *Calculation of $\partial_t K_0(I, 0)$.* A crucial ingredient in the proof of Theorem 1.1 is the calculation of the derivative of quasieigenvalues in Proposition 2.5 in the semiclassical limit $h \to 0$. From the truncated quantum Birkhoff normal form in Theorem 4.1, this can be reduced to the study of the t-dependence of the integrable term $K(I; t)$ in the classical Birkhoff normal form established in Theorem 3.10.

We now consider a 1-parameter family of Hamiltonians $H(\theta, I; t)$ satisfying the assumptions of Theorem 1.1. We can write

$$H(\theta, I; t) = H^0(I) + H^1(\theta, I; t), \tag{3E.1}$$

with

$$H^0(I) := H(\theta, I; 0) \tag{3E.2}$$

$$H^1(\theta, I; t) := t\partial_t H(\theta, I; 0) + \int_0^t (1-s)\partial_t^2 H(\theta, I; s)\, ds = t\partial_t H(\theta, I; 0) + O(t^2), \tag{3E.3}$$

and we assume that H additionally satisfies the assumptions of Theorem 3.10 with this choice of H^0, H^1. By applying two KAM stem iterations to $H(\theta, I; t)$, we obtain a transformed completely integrable component and reduce the order of magnitude of the θ-dependent remainder. An application of Theorem 3.10 to this transformed Hamiltonian produces a Birkhoff normal form, and (3D.7) yields an expression for $K(I; t)$ up to order $o(t)$.

The KAM step iterations required differ from that in Theorem 3.3, in that they are not parametrised by $\omega \in \Omega$ and instead take place in the action-angle space $\mathbb{T}^n \times D$. Such a KAM step appears in the proof of the KAM theorem found in [Gallavotti 1983]. We first describe the KAM step without the presence of the parameter t for simplicity. One begins with a perturbation

$$H(\theta, I) = H^0(I) + H^1(\theta, I) \tag{3E.4}$$

of a completely integrable Hamiltonian $H^0(I)$, and a fixed perturbation $H^1(\theta, I)$, both analytic on the complex domain

$$\theta \in 2\pi\mathbb{C}^n \setminus 2\pi\mathbb{R}^n, \quad |\mathrm{Im}(\theta)| < s, \tag{3E.5}$$

$$\mathrm{Re}(I) \in D, \qquad |\mathrm{Re}(I)| < r. \tag{3E.6}$$

We assume that $|H^1|_{s,r} = O(\epsilon)$ in the uniform sense.

By consideration of the linearised Hamilton–Jacobi equation, we choose a symplectic transformation $\chi : \mathbb{T}^n \times D \to \mathbb{T}^n \times D$ with the aim to write

$$\widetilde{H}(\theta, I) = (H \circ \chi)(\theta, I) = \widetilde{H}^0(I) + \widetilde{H}^1(\theta, I), \tag{3E.7}$$

with $\widetilde{H}^1 = O(\epsilon^\alpha)$ for some $\alpha > 1$. Then we have transformed a sufficiently small perturbation of an integrable Hamiltonian to an even smaller perturbation of a new integrable Hamiltonian, in a way we can hope to iterate.

Obtaining the "new" error bound for $\widetilde{H}^1$ necessarily requires a shrinking of the domains of analyticity, through the use of Cauchy estimates to control derivatives. Moreover, there is a more subtle shrinking of domain required in the I-variable, due to the infamous "small-divisor" problem. Specifically, χ is found using terms of the generating function

$$\Phi(I', \theta) = i \sum_{k \in \mathbb{Z}^n : 0 < |k| \leq M} \frac{H_k^1(I') e^{ik\cdot\theta}}{\omega(I') \cdot k}, \tag{3E.8}$$

where H_k^1 denotes the k-th Fourier coefficient of H^1, and $\omega = \nabla_I H^0(I)$; see [Gallavotti 1983, (2.10)].

The denominators in (3E.8) can generally be zero, and so one must restrict to values of I' for which we have a nonresonance condition

$$\omega(I') \cdot k \geq \frac{C}{|k|^2} \tag{3E.9}$$

for all $0 < |k| \leq M$, where C and M are chosen suitably. We also need to remove those actions I' with $\mathrm{dist}(I', \partial\Omega) \leq \tilde{\rho}$ so that the perturbed tori do not escape the coordinate patch; see [Gallavotti 1983, (3.12)] for the choice of the constant $\tilde{\rho}$. This leads to the definition of the set

$$\widetilde{D}_1 = \{I \in D : \mathrm{dist}(I, \partial D) > \tilde{\rho} \text{ and } \omega(I) \cdot k \geq C/|k|^2 \text{ for all } 0 < |k| \leq M\}. \tag{3E.10}$$

For any $\tilde{I} \in \widetilde{D}_1$ the expression (3E.8) is certainly defined, but as the domain might have rather rough boundary, it is convenient to slightly enlarge $\widetilde{D}_1$ to the open set

$$D_1 = \bigcup_{I \in \widetilde{D}_1} B(I, \tilde{\rho}/2). \tag{3E.11}$$

Upon restricting to this action set for suitable C and M, the objective of (3E.7) can indeed be achieved, and the "integrable part" of the new Hamiltonian can be written as

$$\widetilde{H}^0(I) = H^0(I) + (2\pi)^{-n} \int H^1(\theta, I)\, d\theta; \tag{3E.12}$$

see [Gallavotti 1983, (3.38)]. The overall transformed Hamiltonian is then given by

$$\widetilde{H}(\tilde{\theta}, \tilde{I}) = \widetilde{H}^0(\tilde{I}) + \widetilde{H}^1(\tilde{\theta}, \tilde{I}) \tag{3E.13}$$

in the domain $\mathbb{T}^n \times D_1$, with
$$\|\widetilde{H}^1\| = O(\epsilon^{3/2}). \tag{3E.14}$$
The classical KAM theorem is then proven in [Gallavotti 1983] by iterating this procedure, carefully choosing the C, M, $\tilde{\rho}$ and the analyticity parameters r, s so that the estimate (3E.14) is satisfied with every step, ensuring convergence, and so that the limiting domain $\bigcap_j D_j$ of nonresonant actions is of large measure. A full discussion of this procedure can be found in [Gallavotti 1983].

We now return to our setting of the one-parameter family of Hamiltonians
$$H(\theta, I; t) = H^0(I) + H^1(\theta, I; t).$$
One iteration of the KAM step outlined above yields a family of symplectomorphisms
$$\chi_1 : \mathbb{T}^n \times D_1 \to \mathbb{T}^n \times D \tag{3E.15}$$
parametrised by t such that
$$\widetilde{H}(\theta, I; t) = (H \circ \chi_1)(\theta, I; t) = H^0(I) + t \cdot (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta + \widetilde{H}^1(\theta, I; t), \tag{3E.16}$$
where the second term comes from (3E.3) and the error term $\widetilde{H}^1(\theta, I; t) = O(t^{3/2})$. Regarding this transformed Hamiltonian as a small perturbation of the integrable Hamiltonian
$$\widetilde{H}^0(I; t) = H^0(I) + t \cdot (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta, \tag{3E.17}$$
we perform one more KAM iteration to obtain another family of symplectomorphisms
$$\chi_2 : \mathbb{T}^n \times D_3 \to \mathbb{T}^n \times D_2 \tag{3E.18}$$
parametrised by t such that
$$\begin{aligned}\widetilde{\widetilde{H}}(\theta, I; t) &= (\widetilde{H} \circ \chi_2)(\theta, I; t) \\ &= H^0(I) + t \cdot (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta + (2\pi)^{-n} \int_{\mathbb{T}^n} \widetilde{H}^1(\theta, I; t)\, d\theta + \widetilde{\widetilde{H}}^1(\theta, I; t).\end{aligned}$$
Moreover, by taking our initial choice of nonresonance parameter C sufficiently small, we can ensure that the action domain D_3 contains a collection of nonresonant actions $E_\kappa(t)$, with
$$\nabla_I(\widetilde{\widetilde{H}}^0(E_\kappa(t))) = \Omega_\kappa, \tag{3E.19}$$
where
$$\widetilde{\widetilde{H}}^0(I; t) = H^0(I) + t \cdot (2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, I)\, d\theta + (2\pi)^{-n} \int_{\mathbb{T}^n} \widetilde{H}^1(\theta, I; t)\, d\theta. \tag{3E.20}$$
We now summarise the preceding discussion.

Proposition 3.12. *Suppose $H(\theta, I; t)$ is a family of real-analytic perturbations of the completely integrable nondegenerate Hamiltonian $H^0(I)$ in $\mathbb{T}^n \times D \times (-1, 1)$ that has an analytic extension to*
$$W_{s,r}(D) := \{(\theta, I) \in \mathbb{C}^n/(2\pi\mathbb{Z}) \times \mathbb{C}^n : |\mathrm{Im}(\theta)| < s, \mathrm{dist}(I, D) < r\}. \tag{3E.21}$$

Suppose further that the conditions

$$\left|\frac{\partial H^0}{\partial I}\right| \leq E, \tag{3E.22}$$

$$\left|\left(\frac{\partial^2 H^0}{\partial I^2}\right)^{-1}\right| \leq \eta, \tag{3E.23}$$

$$\left(\left|\frac{\partial H^1}{\partial I}\right| + r^{-1}\left|\frac{\partial H^1}{\partial \theta}\right|\right) \leq \epsilon \tag{3E.24}$$

are satisfied.

Then for sufficiently small $\delta > 0$, there exists a subdomain $\widetilde{D} \subset D$ and a family of real-analytic symplectic maps

$$\chi : \mathbb{T}^n \times \widetilde{D} \times (-\delta, \delta) \to \mathbb{T}^n \times D \tag{3E.25}$$

that analytically extend to a new domain of holomorphy

$$W_{s_+, r_+}(\widetilde{D}) \tag{3E.26}$$

such that

$$(H \circ \chi)(\theta, I; t) = \widetilde{H}^0(I; t) + \widetilde{H}^1(\theta, I; t), \tag{3E.27}$$

with

$$\partial_t \widetilde{H}^0(I; 0) = (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta \tag{3E.28}$$

and

$$|\widetilde{H}^1|_{s_+, r_+} = O(t^{9/4}), \tag{3E.29}$$

with constant depending only on n and E. Moreover, this domain $\widetilde{D}$ contains a collection $E_\kappa(t)$ of actions such that

$$\nabla_I(\widetilde{H}^0)(E_\kappa(t)) = \Omega_\kappa. \tag{3E.30}$$

We can also generalise this result to the Gevrey setting.

Proposition 3.13. *Suppose $H(\theta, I; t) \in G^{\rho,\rho,1}(\mathbb{T}^n \times D \times (-1, 1)$ is a family of Hamiltonians satisfying the assumptions of Theorem 3.10, where $H^0(I) := H(\theta, I; 0)$ for fixed $\rho > 1$, and choose $\kappa > 0$ small. Then for sufficiently small $\|H(\theta, I; t) - H^0(I)\|_{L_1, L_2, L_2}$, there exists a subdomain $\widetilde{D} \subset D$ and a $G^{\rho,\rho,1}$ family of symplectic maps*

$$\chi : \mathbb{T}^n \times \widetilde{D} \times (-1, 1) \to \mathbb{T}^n \times D \tag{3E.31}$$

such that

$$(H \circ \chi)(\theta, I; t) = \widetilde{H}^0(I; t) + \widetilde{H}^1(\theta, I; t), \tag{3E.32}$$

with

$$\partial_t \widetilde{H}^0(I; 0) = (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta \tag{3E.33}$$

and

$$\|\widetilde{H}^1\|_{CL_1, CL_2, CL_2} = O(t^{9/4}), \tag{3E.34}$$

with constant independent of κ and with C dependent only on n and ρ.

Moreover, the domain $\widetilde{D}$ contains $E_\kappa(t) = \omega^{-1}(\Omega_\kappa; t) = (\nabla_I \widetilde{H}^0)^{-1}(\Omega_\kappa; t)$.

Proof. This result is established via the approximation of Gevrey functions by real-analytic functions. First, we define

$$H^0(I) = H(\theta, I; 0), \tag{3E.35}$$

$$H^1(\theta, I; t) = H(\theta, I; t) - H(\theta, I; 0) = \int_0^t \partial_t H(\theta, I; s)\, ds \tag{3E.36}$$

and use Proposition 3.8 to boundedly 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 $B_{\widetilde{R}}$ with $D \subset B_{\widetilde{R}-1}$. From the same methods used in the proof of Proposition 3.4, 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

$$|P_{j+1}^k - P_j^k|_{U_{j+1}} \le C(D^0, L_1, L_2) \exp\bigl(-\tfrac{3}{4}(\rho-1)(2L_1 u_j)^{-1/(\rho-1)}\bigr)\|H^k\|, \tag{3E.37}$$

$$|\partial_x^\alpha (P_j^k - H^k)(\theta, I; t)| \le C(D^0, L_1, L_2) \exp\bigl(-\tfrac{3}{4}(\rho-1)(2L_1 u_j)^{-1/(\rho-1)}\bigr) \tag{3E.38}$$

in $\mathbb{T}^n \times B_{\widetilde{R}} \times (-1, 1)$ for $|\alpha| \le 1$. 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 [Hou and Popov 2016, Proposition 2.2]. (This fact can be readily obtained by applying Cauchy estimates to (3E.37).)

Now for each $j \in \mathbb{N}$, we can carry out the first KAM step for the real-analytic Hamiltonian $P_j = P_j^0 + P_j^1$ to obtain a real-analytic symplectic map

$$\chi_j : \mathbb{T}^n \times D_1 \to \mathbb{T}^n \times D \tag{3E.39}$$

defined in shrinking holomorphy domains such that

$$((P_j^0 + P_j^1) \circ \chi_j)(\theta, I; t) = P^0(I) + t \cdot (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t P_j^1(\theta, I; 0)\, d\theta + \widetilde{P}^1(\theta, I; t), \tag{3E.40}$$

with $\|P_j^1\| = O(t^{3/2})$. Note that for an individual KAM step, the symplectic map χ_j is defined using a generating function Φ_j that is a weighted sum of finitely many Fourier components of P_j^1; see (3E.8) and [Gallavotti 1983, (3.14)]. This implies that as $P_j^0 + P_j^1 \to H^0 + H^1$ in $G^{\rho,\rho,1}(\mathbb{T}^n \times D_1 \times (-1, 1))$, the generating functions Φ_j converge to some

$$\Phi \in G^{\rho,\rho,1}(\mathbb{T}^n \times D_1 \times (-1, 1)) \tag{3E.41}$$

in the $G^{\rho,\rho,1}$ sense. From Corollary A.5, it follows that the corresponding symplectic maps χ_j converge to some

$$\chi^1 \in G^{\rho,\rho,1}(\mathbb{T}^n \times D_1 \times (-1, 1)) \tag{3E.42}$$

in the Gevrey sense.

Similarly, the symplectic maps $\tilde{\chi}_j$ that comprise a single KAM step for the Hamiltonians

$$(P_j^0 + P_j^1) \circ \chi_j \tag{3E.43}$$

can also be seen to converge to some

$$\chi^2 \in G^{\rho,\rho,1}(\mathbb{T}^n \times D_2, \mathbb{T}^n \times D_1). \tag{3E.44}$$

It follows that the family of symplectic maps $\chi_j \circ \tilde{\chi}_j$ whose existence is asserted by applying Proposition 3.12 to $P_j^0 + P_j^1$ converge to some $\chi := \chi^1 \circ \chi^2$ in the $G^{\rho,\rho,1}$-sense. Moreover, if we write

$$(P_j^0 + P_j^1) \circ \chi_j \circ \tilde{\chi}_j = \widetilde{H}_j^0(I;t) + \widetilde{H}_j^1(\theta, I; t), \tag{3E.45}$$

in the notation of Proposition 3.12, we have that $\widetilde{H}_j^k$ are convergent sequences in $G^{\rho,\rho,1}$, and so it follows that their limits $\widetilde{H}^0$, $\widetilde{H}^1$ satisfy

$$\partial_t \widetilde{H}^0(I;0) = (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta \tag{3E.46}$$

and

$$\|\widetilde{H}^1\|_{CL_1,CL_2,CL_2} = O(t^{9/4}) \tag{3E.47}$$

as required. $\square$

Finally, we complete our computation of $\partial_t K_0(I;0)$ for a given Hamiltonian $H(\theta, I;t)$ satisfying the conditions of Theorem 3.10 by applying Proposition 3.13 to H, prior to applying Theorem 4.1 to compute the Birkhoff normal form of the transformed Hamiltonian $\widetilde{H}(\theta, I; t)$.

By applying Proposition 3.13 to $H(\theta, I;t)$ with $\|H(\theta, I;t) - H(\theta, I; 0)\|$ sufficiently small, we can then apply Theorem 3.10 to the Hamiltonian

$$\widetilde{H}(\theta, I; t) = \widetilde{H}^0(I;t) + \widetilde{H}^1(\theta, I; t), \tag{3E.48}$$

with an improved error term.

Proposition 3.14. *Suppose the assumptions of Theorem 3.9 hold for the Hamiltonian*

$$H(\theta, I; t) \in G^{\rho,\rho,1}(\mathbb{T}^n \times D \times (-1,1)). \tag{3E.49}$$

Then there exist $N(n,\rho,\tau) > 0$ 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

$$\kappa^{-2}\|H(\theta, I;t) - H(\theta, I;0)\|_{L_1,L_2,L_2} = \epsilon_H \le \epsilon L_1^{-N-2(\tau+2)}, \tag{3E.50}$$

there is a subdomain $\widetilde{D} \subset D$ containing $E_\kappa(0)$ and a family of $G^{\rho',\rho'}$ maps $\omega : \widetilde{D} \times \left(\frac{1}{2}, 1, 2\right) \to \Omega$ and a family of maps $\chi \in G^{\rho,\rho',\rho'}(\mathbb{T}^n \times \widetilde{D} \times \left(\frac{1}{2}, 1, 2\right), \mathbb{T}^n \times \widetilde{D})$ that are diffeomorphisms and exact symplectomorphisms respectively for each fixed $t \in \left(\frac{1}{2}, 1, 2\right)$. Moreover, we can choose the maps ω and χ such that family of transformed Hamiltonians

$$\widetilde{H}(\theta, I; t) := (H \circ \chi)(\theta, I; t) \tag{3E.51}$$

is of Gevrey class $G^{\rho,\rho',\rho'}\left(\mathbb{T}^n \times \widetilde{D} \times \left(\frac{1}{2}, 1, 2\right)\right)$ and can be decomposed as

$$K(I;t) + R(\theta, I; t) := \widetilde{H}(0, I; t) + (\widetilde{H}(\theta, I; t) - \widetilde{H}(0, I; t)) \tag{3E.52}$$

such that:

(i) *$\mathbb{T}^n \times \{I\}$ is an invariant Lagrangian torus of $\widetilde{H}(\,\cdot\,,\cdot\,;t)$ for each $I \in E_\kappa(t) = \omega^{-1}(\widetilde{\Omega}_\kappa)$ and each $t \in \left(\frac{1}{2}, 1, 2\right)$.*

(ii) $\partial_I^\beta(\nabla K(I;t)-\omega(I;t))=\partial_I^\beta R(\theta,I;t)=0$ *for all* $(\theta,I;t)\in\mathbb{T}^n\times E_\kappa(t)\times\left(\frac{1}{2},1,2\right)$, $\beta\in\mathbb{N}^n$.

(iii) *There exist* $A, C>0$ *independent of* κ, L_1, L_2, *and* $D\subset D^0$ *such that we have the estimates*

$$|\partial_\theta^\alpha\partial_I^\beta\partial_t^\delta\phi(\theta,I;t)|+|\partial_I^\beta\partial_t^\delta(\omega(I;t)-\nabla\widetilde{H}^0(I;t))|+|\partial_\theta^\alpha\partial_I^\beta\partial_t^\delta(\widetilde{H}(\theta,I;t)-\widetilde{H}^0(I;t))| \\ \le AC^{|\alpha|+|\beta|+|\delta|}L_1^{|\alpha|}(L_1^{\tau+1}/\kappa)^{|\beta|}\alpha!^\rho\,\beta!^{\rho'}\,\delta!^{\rho'}\,L_1^{N/2}|t|^{9/8} \tag{3E.53}$$

uniformly in $\mathbb{T}^n\times\widetilde{D}\times\left(\frac{1}{2},1,2\right)$ *for all* α,β, *where* $\phi\in G^{\rho,\rho',\rho'}\left(\mathbb{T}^n\times\widetilde{D}\times\left(\frac{1}{2},1,2\right)\right)$ *is such that* $\langle\theta,I\rangle+\phi(\theta,I;t)$ *generates the symplectomorphisms* χ *in the sense of* (3E.8) *and* $\widetilde{H}^0$, $\widetilde{H}^1$ *are as in Proposition 3.13.*

(iv)
$$\partial_t K(I;t)=(2\pi)^{-n}\int_{\mathbb{T}^n}\partial_t H(\theta,I;0)+o(1) \tag{3E.54}$$

uniformly in $\mathbb{T}^n\times\widetilde{D}\times\left(\frac{1}{2},1,2\right)$.

Proof. The only new claim in this proposition is (3E.54), which follows from (3E.53) and the expression (3E.33) for $\widetilde{H}^0$. Note that the exponent $\frac{9}{8}$ in (3E.53) comes from (3E.34) and the square root in (3D.7). ☐

4. Quantum Birkhoff normal form

Through the work in Section 3, we have now established that the Birkhoff normal form construction in [Popov 2004b] preserves smoothness in the t-parameter when applied to the Hamiltonian $P_0(x,\xi;t)$ that is the principal symbol of the operator introduced in (1B.1). This regularity in t propagates through the quantum Birkhoff normal form construction in [Popov 2004a], which we discuss in this section. The upshot of this regularity in t is that the quasimodes constructed in [Popov 2004a, Section 2.4] can be chosen to have associated quasieigenvalues varying smoothly in the parameter t. We discuss these quasimodes in Section 4C.

4A. ***Quantum Birkhoff normal form.*** In [Popov 2004a], a quantum Birkhoff normal form is constructed for semiclassical pseudodifferential operators of the form (1B.1) after first obtaining a classical Birkhoff normal form for the principal symbol of regularity $G^{\rho,\rho'}$ as in Theorem 3.10. This normal form uses the Gevrey symbol classes introduced in Section B and is stated in Theorem 4.1. We remark that the proof is presented in [Popov 2004a] for differential operators, but can be carried out without change if the $\mathcal{P}_h$ is a pseudodifferential operator.

We denote by χ_1 the symplectomorphism that transforms the completely integrable Hamiltonian $P(x,\xi;0)$ into action-angle coordinates $H=P\circ(\chi_1)$ and we denote by $\chi_0(t)$ the symplectomorphism that transforms the perturbed Hamiltonian $H(\theta,I;t)$ into Birkhoff normal form, as constructed in Theorem 3.10. For the purpose of stating the quantum Birkhoff normal form for $\mathcal{P}_h(t)$, the Maslov class of the KAM tori $\Lambda_\omega:\omega\in\Omega_\kappa\}$ (as defined in Section 3.4 of [Duistermaat 1996]) can be identified with elements of $\vartheta\in H^1(\mathbb{T}^n;\mathbb{Z})$ via the family of symplectomorphisms $\chi_0(t)\circ\chi_1:\mathbb{T}^n\times D\to T^*M$. Following [Popov 2000b; Colin de Verdière 1977], we can then associate a smooth line bundle Ł over $\mathbb{T}^n$ with the class ϑ such that smooth sections $f\in\mathcal{C}^\infty(\mathbb{T}^n,Ł)$ can be canonically identified with smooth functions

$\tilde{f} \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{C})$ satisfying the quasiperiodicity condition

$$\tilde{f}(x + 2\pi p) = \exp\left(\frac{i\pi}{2}\langle\vartheta, p\rangle\right)\tilde{f}(x) \tag{4A.1}$$

for all $p \in \mathbb{Z}^n$.

The quantum Birkhoff normal form in [Popov 2004a] is far sharper than is necessary for the purposes of this paper, with remainders of order $O(e^{-ch^{-1/\nu}})$. We require only the following truncated version, with error terms of order $O(h^{\gamma+1})$ for some fixed $\gamma > 0$.

Theorem 4.1. *Suppose $\mathcal{P}_h(t)$ is as in* (1B.1). *Then for each fixed t there exists a uniformly bounded family of semiclassical Fourier integral operators*

$$U_h(t) : L^2(\mathbb{T}^n; \mathbb{L}) \to L^2(M), \quad 0 < h < h_0, \tag{4A.2}$$

that are associated with the canonical relation graph of the Birkhoff normal form transformation $\chi(t)$ such that we have

(i) $U_h(t)^* U_h(t) - \mathrm{Id}$ *is a pseudodifferential operator with symbol in the Gevrey class $S_\ell(\mathbb{T}^n \times D)$ which restricts to an element of $h^{\gamma+1} S_\ell(\mathbb{T}^n \times Y)$ for some subdomain Y of D that contains $E_\kappa(t)$*,

(ii) $\mathcal{P}_h(t) \circ U_h(t) - U_h(t) \circ \mathcal{P}_h^0(t) = \mathcal{R}_h(t) \in h^{\gamma+1} S_\ell$, *where the operator $\mathcal{P}_h^0(t)$ has symbol*

$$p^0(\theta, I; t, h) = K^0(I; t, h) + R^0(\theta, I; t, h) = \sum_{j \le \gamma} K_j(I; t) h^j + \sum_{j \le \gamma} R_j(\theta, I; t) h^j, \tag{4A.3}$$

with both K^0 and R^0 in the symbol class $S_\ell(\mathbb{T}^n \times D)$ from Definition B.5 where $\eta > 0$ is a constant, $K_0(I; t)$, $R_0(\theta, I; t)$ are the components of the Birkhoff normal form of the Hamiltonian $P_0 \circ \chi_1$ as constructed in Theorem 3.10, and

$$\partial_I^\alpha R_j(\theta, I; t) = 0 \tag{4A.4}$$

for $(\theta, I; t) \in \mathbb{T}^n \times E_\kappa(t) \times (-1, 1)$. Moreover, the symbols K_j, R_j in (4A.3) *are smooth in the parameter t.*

Our statement of Theorem 4.1 differs from [Popov 2004a, Theorem 2.1] only in the presence of the parameter t, the smoothness of the symbols K_j, R_j in t, and the truncation to fixed finite order $O(h^{\gamma+1})$. We sketch the details of the proof of Theorem 4.1 in this section, following the argument of [Popov 2004a].

The construction of $U_h(t)$ can be broken into multiple steps. We begin by constructing a family of semiclassical Fourier integral operators $T_h(t)$ that conjugate $P_h(t)$ to a family of semiclassical pseudodifferential operators $P_h^1(t) : \mathcal{C}^\infty(\mathbb{T}^n; \mathbb{L})$ with principal symbol equal to $K_0(I; t) + R_0(\theta, I; t)$, the Birkhoff normal form of H, and with vanishing subprincipal symbol. The conjugating semiclassical Fourier integral operators arise by quantising the G^ρ symplectomorphisms

$$\chi_1 : \mathbb{T}^n \times D \to T^*M, \tag{4A.5}$$

$$\chi_0 : \mathbb{T}^n \times D \to \mathbb{T}^n \times D \tag{4A.6}$$

that transform the unperturbed Hamiltonian $P(x, \xi; 0)$ to action-angle variables and transform the perturbed Hamiltonian to Birkhoff normal form respectively, and composing these two operators. Full details for this construction can be found in [Popov 2000b, Section 2].

From the regularity of the symplectomorphisms, it follows that there exists a semiclassical expansion for $P_h^1(t)$ with symbols smooth in t.

The symbol of the operator $P_h^1(t)$ satisfies the property (4A.3) 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 sets

$$\{(\theta, I; t) \in \mathbb{T}^n \times \mathbb{R}^n \times (-1, 1) : I \in E_\kappa(t)\}$$

by solving equations of the form

$$\langle \nabla K_0, \partial_\theta \rangle f(\theta, I; t) = g(\theta, I; t), \tag{4A.7}$$

referred to in the literature as homological equations. In this manner the "flatness condition" of (4A.4) is obtained for $j > 0$, where the $j = 0$ statement is established by Theorem 3.10. We outline this procedure in Section 4B.

The key fact is that the homological equation can be solved smoothly in the parameter t, which is the content of Theorem 4.3. One can then apply Theorem 4.3 as in [Popov 2004a, Section 2.3] to complete the construction of the quantum Birkhoff normal form, with the additional consequence of smoothness of symbols K_j, R_j.

4B. ***Construction of the quantum Birkhoff normal form.*** After conjugating $P_h(t)$ by semiclassical Fourier integral operators as described in the previous section, we obtain a family of self-adjoint semiclassical operators $P_h^1(t)$ with symbol $\tilde{p} \in S_{\tilde{\ell}}(\mathbb{T}^n \times D)$ satisfying the flatness condition (4A.4) to order h^2, where $\tilde{\ell} = (\rho, \rho', \rho + \rho' - 1)$. That is to say, the formal summation of $\tilde{p}$,

$$\sum_{j=0}^{\infty} \tilde{p}_j(\theta, I; t) h^j, \tag{4B.1}$$

satisfies

$$\tilde{p}_0(\theta, I; t) = K_0(I; t) + R_0(\theta, I; t), \tag{4B.2}$$

$$\tilde{p}_1(\theta, I; t) = 0. \tag{4B.3}$$

The next step of the proof of Theorem 4.1 is the improvement of the order of the flatness condition by composition with a suitable elliptic semiclassical pseudodifferential operator

$$A_h(t) = \mathrm{Id} + O(h)$$

with symbol

$$a(\theta, I; t) = \sum_{j=1}^{\infty} a_j(\theta, I; t) h^j. \tag{4B.4}$$

To motivate the method, we suppose that a quantum Birkhoff normal form P_h^0 exists in the sense of Theorem 4.1. Our current operator $\widetilde{P}_h$ is equal to P_h^0 up to order h^2 by construction. Hence, we have

$$\begin{aligned} T_h(t) A_h(t) \widetilde{P}_h(t) &= T_h(t) \widetilde{P}_h(t) A_h(t) + T_h(t)[A_h(t), \widetilde{P}_h(t)] \\ &= P_h^1(t) T_h(t) A_h(t) + h^2 T(t) B(t) A(t) + T_h(t)[A_h(t), \widetilde{P}_h(t)] \end{aligned} \tag{4B.5}$$

for some semiclassical pseudodifferential operator $B_h(t)$ in the symbol class $S_{\tilde{\ell}}(\mathbb{T}^n \times D)$. From composition formulae, the symbol of the commutator is equal to

$$-(\partial_\theta^\alpha a_1 \partial_I^\alpha \tilde{p}_0)h^2 = -\mathcal{L}_{\omega I;t} a_1, \tag{4B.6}$$

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

$$\mathcal{L}_{\omega(I;t)} a_1 = b_0, \tag{4B.7}$$

where b_0 denotes the principal symbol of $B_h(t)$. Indeed, if (4B.7) is solvable, then we have

$$T_h(t)A_h(t)P_h(t) = P_h^0(t)T_h(t)A_h(t) + O(h^3). \tag{4B.8}$$

Extending this idea, it was shown in [Popov 2000b] 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 result.

Proposition 4.2. *There exist a, K^0, $r \in S_\ell(\mathbb{T}^n \times D)$, where $\ell = (\rho, \mu, \nu)$, such that*

$$a(\theta, I; t, h) \sim \sum_{j=0}^\infty a_j(\theta, I; t)h^j, \tag{4B.9}$$

$$K^0(I; t, h) \sim \sum_{j=0}^\infty K_j(I; t)h^j, \tag{4B.10}$$

$$r(\theta, I; t, h) \sim \sum_{j=0}^\infty r_j(\theta, I; t)h^j, \tag{4B.11}$$

where $a_0 = 1$, $r_0 = R_0$, $K_1 = 0$, and

$$\tilde{p} \circ a - a \circ K^0 \sim r, \tag{4B.12}$$

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 4.1, while the symbol R^0 is then constructed by solving $a \circ R^0 = r$, which is possible by ellipticity.

The completion of the proof of Theorem 4.1 after establishing Proposition 4.2 is contained in [Popov 2000b, Section 3]. For our additional requirement of smoothness in t in Theorem 4.1, it thus suffices to verify that the homological equation can be solved smoothly in the parameter t. In particular, we require the following.

Theorem 4.3. *Suppose $f(\cdot, \cdot; t) \in G^{\rho,\mu}(\mathbb{T}^n \times D)$ satisfies the estimate*

$$|\partial_\theta^\alpha \partial_I^\beta f(\theta, I; t)| \le d_0 C^{|\alpha|+\mu|\beta|} \Gamma(\rho|\alpha| + \mu|\beta| + q) \tag{4B.13}$$

uniformly in the smooth parameter $t \in (-1, 1)$ for some $q > 0$ and some $C \ge 1$ and that for each $I \in D$, we have

$$\int_{\mathbb{T}^n} f(\theta, I; t)\, d\theta = 0. \tag{4B.14}$$

Then for any smooth family $\omega(\,\cdot\,;t) \in G^{\rho'}_{L_0}(D,\Omega)$ *there is a solution* $u(\,\cdot\,,\cdot\,;t) \in G^{\rho,\mu}(\mathbb{T}^n \times D)$ *to the equation*

$$\mathcal{L}_\omega u(\theta, I; t) = f(\theta, I; t), \quad (\theta, I) \in \mathbb{T}^n \times E_\kappa(t), \tag{4B.15}$$

$$u(0, I; t) = 0, \qquad I \in D, \tag{4B.16}$$

where $\mathcal{L}_\omega = \langle \omega(I;t), \partial/\partial\theta \rangle$. *Moreover,* u *is smooth in the parameter* t *and satisfies the estimate*

$$|\partial_\theta^\alpha \partial_I^\beta u(\theta, I; t)| \le A d_0 C^{n+\tau+|\alpha|+\mu|\beta|+1} \Gamma(\rho|\alpha| + \mu|\beta| + \rho(n+\tau+1) + q), \tag{4B.17}$$

where A *depends only on* n, ρ, τ *and* μ.

This theorem statement differs from [Popov 2004a, Proposition 2.3] only in the presence of the smooth parameter t, and indeed an identical proof based on taking the Fourier expansion

$$u(\theta, I; t) = \sum_{k \in \mathbb{Z}^n} e^{i\langle k, \theta\rangle} u_k(I; t) \tag{4B.18}$$

and solving for u_k can be pursued. The rapid decay of Fourier coefficients established in [Popov 2004a] implies that the limit $u(\theta, I; t)$ is smooth in t as required. The proof is then identical to that in [Popov 2004a], with the uniformity in (4B.17) following from the uniformity in (4B.13).

4C. *Quasimode construction.* We now briefly outline how the construction of Gevrey class quasimodes for $\mathcal{P}_h(t)$ follow from the quantum Birkhoff normal form Theorem 4.1. These quasimodes microlocalise onto a family of nonresonant tori and moreover have quasieigenvalues that are smooth in the parameter $t \in (-1, 1)$.

Definition 4.4. An $O(h^{\gamma+1})$ family of G^ρ quasimodes $\mathcal{Q}(t)$ for $\mathcal{P}_h(t)$ is a family

$$\{(v_m(x; t, h), \mu_m(t, h)) : m \in \mathcal{M}_h(t)\} \subset \mathcal{C}^\infty(M \times \mathcal{D}_h(m)) \times \mathcal{C}^\infty(\mathcal{D}_h(m)) \tag{4C.1}$$

parametrised by $h \in (0, h_0]$, where

- $\mathcal{M}_h(t) \subset \mathbb{Z}^n$ is an h-dependent finite index set,
- $\mathcal{D}_h(m) = \{t \in (-1, 1) : m \in \mathcal{M}_h(t)\}$,
- each $v_m(\,\cdot\,; t, h)$ is uniformly of class G^ρ,
- $\|P_h(t) v_m(\,\cdot\,; t, h) - \mu_m(t; h) v_m(\,\cdot\,; t, h)\|_{L^2} = O(h^{\gamma+1})$ for all $m \in \mathcal{M}_h(t)$,
- $|\langle v_m(\,\cdot\,; t, h), v_l(\,\cdot\,; t, h)\rangle - \delta_{ml}| = O(h^{\gamma+1})$ for all $m, l \in \mathcal{M}_h(t)$.

Theorem 4.5. *Suppose now that* $t \in (-1, 1)$ *is fixed and* $S \subset E_\kappa(t)$ *is a closed collection of nonresonant actions. For an arbitrary constant* $L > 1$, *we define the index set*

$$\mathcal{M}_h := \{m \in \mathbb{Z}^n : \operatorname{dist}(S, h(m + \vartheta/4)) < Lh\}, \tag{4C.2}$$

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 the choice of torus by the local constancy of the Maslov class.*

Then

$$\{(v_m(x;t,h),\mu_m(t;h)) : m \in \mathcal{M}_h(t)\} := (U_h(t)e_m, K^0(h(m+\vartheta/4);t,h) \tag{4C.3}$$

defines a G^ρ family of quasimodes for $\mathcal{P}_h(t)$ that has Gevrey microsupport on the family of tori

$$\Lambda_S = \bigcup_{I\in S} \Lambda_{\omega(I;t)} = \bigcup_{I\in S} \chi(\mathbb{T}^n \times \{I\}) \subset T^*M, \tag{4C.4}$$

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

$$\tilde{e}_m(x) := \exp(i\langle m+\vartheta/4, x\rangle). \tag{4C.5}$$

Proof. From the definition of the functions e_m, it follows that

$$\begin{aligned} P_h^0(t)(e_m)(\theta) &= \sigma(P_h^0(t))(\theta, h(m+\vartheta/4))e_m(\theta) \\ &= (K^0(h(m+\vartheta/4);t,h) + R^0(\theta, h(m+\vartheta/4);t,h))e_m(\theta) \\ &= (\lambda_m(t;h) + R^0(\theta, h(m+\vartheta/4))e_m(\theta). \end{aligned} \tag{4C.6}$$

From the definition (4C.2) of the index set $\mathcal{M}_h(t)$ and from $A.2$, it thus follows that

$$\mathcal{P}_h(t)(U_h(t)e_m) = U_h(t)P_h^0(t)e_m = O(h^{\gamma+1}) \tag{4C.7}$$

upon an application of Theorem 4.1. The almost-orthogonality of the $U_h(t)e_m$ then follows from the fact that $U_h(t)$ is almost unitary by Theorem 4.1, and that the e_m are exactly orthogonal by construction. □

These quasimodes are as numerous as we could hope for; indeed the index set $\mathcal{M}_h(t)$ satisfies the local Weyl asymptotic

$$\lim_{h\to 0}(2\pi h)^n \,\#\mathcal{M}_h = m(\mathbb{T}^n \times S) = \mu(\Lambda_S), \tag{4C.8}$$

where m denotes the $(2n)$-dimensional Lebesgue measure and μ 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 $\mathcal{M}_h$ with side length h. The containment

$$S \subset U \subset \{I : \operatorname{dist}(I,S) < \tilde{L}h\} \tag{4C.9}$$

for a constant $\tilde{L}$ then yields the claim by monotone convergence of measures, noting that since S is closed we have

$$S = \overline{S} = \bigcap_{h>0} \{I : \operatorname{dist}(I,S) < \tilde{L}h\}. \tag{4C.10}$$

In the special case of $S = \{I\}$, we have a family of G^ρ quasimodes with microsupport on an individual torus $\chi(\mathbb{T}^n \times \{I\})$.

Appendix A: Anisotropic Gevrey classes

In this appendix, we define the Gevrey function spaces used throughout the paper and collect several of their properties from the appendix of [Popov 2004b].

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

$$G_L^\rho(X) := \left\{ f \in \mathcal{C}^\infty(X) : \sup_\alpha \sup_{x \in X} |\partial_x^\alpha f(x)| L^{-|\alpha|} \alpha!^{-\rho} < \infty \right\}. \tag{A.1}$$

If $f \in G_L^\rho(X)$, the supremum in (A.1) is denoted by $\|f\|_L$. We will frequently suppress the L in our notation. Equipped with this norm, $G_L^\rho(X)$ 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 superexponentially small.

Proposition A.2. *Suppose $f \in G^\rho(X)$, and $\rho > 1$. Then there exist positive constants c, C, η and r_0 only dependent on the Gevrey constant L, the norm $\|f\|_L$, and the set X such that*

$$f(x_0 + r) = \sum_{|\alpha| \leq \eta |r|^{1/(1-\rho)}} f_\alpha(x_0) r^\alpha + R(x_0, r), \tag{A.2}$$

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

$$|\partial_x^\beta R(x_0, r)| \leq C^{1+|\beta|} \beta!^\rho \, e^{-c|r|^{-1/(\rho-1)}} \quad \textit{for all } 0 < |r| \leq \min(r_0, d(x_0, \mathbb{R}^n \setminus X)). \tag{A.3}$$

We also need to consider anisotropic Gevrey classes, which are classes of Gevrey functions with differing regularity in individual variables.

Definition A.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

$$G_{L_1, L_2}^{\rho_1, \rho_2}(X \times Y) = \left\{ f \in \mathcal{C}^\infty(X \times Y) : \sup_{(x,y) \in X \times Y} |\partial_x^\alpha \partial_y^\beta f| L_1^{-|\alpha|} L_2^{-|\beta|} \alpha!^{-\rho_1} \, \beta!^{-\rho_2} < \infty \right\}. \tag{A.4}$$

If $f \in G_{L_1, L_2}^{\rho_1, \rho_2}$, then we denote the supremum in (A.4) by $\|f\|_{L_1, L_2}$. Equipped with this norm, $G_{L_1, L_2}^{\rho_1, \rho_2}$ 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.

In anisotropic Gevrey classes, one has the following implicit function theorem due to Komatsu.

Proposition A.4. *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*

$$L_1 \|F(x, \omega) - x\|_{L_1, L_2} \leq \tfrac{1}{2}.$$

Then there exists a local solution $x = g(y, \omega)$ to the implicit equation

$$F(x, \omega) = y \tag{A.5}$$

defined in a domain $Y \times \Omega$. Moreover, there exist constants A, C dependent only on ρ, ρ', 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 in [Popov 2004b].

Corollary A.5. *Suppose* $F \in G^{\rho,\rho'}_{L_1,L_2}(\mathbb{T}^n \times \Omega, \mathbb{T}^n)$, *where* $\Omega^0 \subset \mathbb{R}^m$ *and* $L_1 \| F(\theta, \omega) - \theta \|_{L_1,L_2} \leq \frac{1}{2}$. *Then there exists a local solution* $x = g(y, \omega)$ *to the implicit equation*

$$F(x, \omega) = y \tag{A.6}$$

defined on $\mathbb{T}^n \times \Omega$. *Moreover, there exist positive constants* A, C *dependent only on* ρ, ρ', n, m *such that* $g \in G^{\rho,\rho'}_{CL_1,CL_2}(\mathbb{T}^n \times \Omega)$ *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 [Popov 2004b].

Proposition A.6. *Let* $X \subset \mathbb{R}^n$, $Y \subset \mathbb{R}^m$, *and* $\Omega \subset \mathbb{R}^k$ *be open sets. Suppose* $g \in G^{\rho'}_{L_1}(\Omega, Y)$, *with* $\|g\|_{L_1} = A_1$, *and* $f \in G^{\rho,\rho'}_{B,L_2}(X \times Y)$, *with* $\|f\|_{B,L_2} = A_2$. *Then the composition* $F(x, \omega) := f(x, g(\omega))$ *is in* $G^{\rho,\rho'}_{B,L}(X \times \Omega)$, *where*

$$L = 2^{l+\rho'} l^{\rho'} L_1 \max(1, A_1 L_2),$$

with $l = \max(k, m, n)$. *Moreover we have the Gevrey norm estimate*

$$\|F\|_{B,L} \leq A_2.$$

Proposition A.7. *Let* $X \subset \mathbb{R}^n$, $Y \in \mathbb{R}^m$, *and* $\Omega \subset \mathbb{R}^k$ *be open sets. Suppose* $g \in G^{\rho,\rho'}_{B_1,L_1}(X \times \Omega, Y)$ *with* $\|g\|_{B_1,L_1} = A_1$ *and* $f \in G^{\rho,\rho'}_{B_2,L_2}(Y \times \Omega)$. *Then the composition*

$$F(x, \omega) := f(g(x, \omega), \omega)$$

is in $G^{\rho,\rho'}_{B,L}(X \times \Omega)$, *where*

$$B = 4^l (4l)^\rho B_1 \max(1 + A_1 B_2),$$
$$L = L_2 + 4^l (4l)^\rho L_1 \max(1, A_1 B_2),$$

with $l = \max(k, m, n)$. *Moreover we have the Gevrey norm estimate*

$$\|F\|_{B,L} \leq A_2.$$

Appendix B: Gevrey class symbols

In this appendix, we introduce the class of Gevrey symbols used throughout this paper. 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 (σ, μ, ϱ) by ℓ.

Definition B.1. A formal Gevrey symbol on $X \times D$ is a formal sum

$$\sum_{j=0}^{\infty} p_j(\theta, I) h^j, \tag{B.1}$$

where the $p_j \in \mathcal{C}_0^\infty(X \times D)$ are all supported in a fixed compact set and there exists a $C > 0$ such that

$$\sup_{X \times D} |\partial_\theta^\beta \partial_I^\alpha p_j(\theta, I)| \leq C^{j+|\alpha|+|\beta|+1} \beta!^\sigma \, \alpha!^\mu \, j!^\varrho. \tag{B.2}$$

Definition B.2. A realisation of the formal symbol (B.1) is a function $p(\theta, I; h) \in \mathcal{C}_0^\infty(X \times D)$ for $0 < h \le h_0$, with

$$\sup_{X\times D\times(0,h_0]} \left| \partial_\theta^\beta \partial_I^\alpha \left(p(\theta, I; h) - \sum_{j=0}^{N} p_j(\theta, I) h^j \right) \right| \le h^{N+1} C_1^{N+|\alpha|+|\beta|+2} \beta!^\sigma \, \alpha!^\mu \, (N+1)!^\varrho. \tag{B.3}$$

Lemma B.3. *Given a formal symbol* (B.1), *one choice of realisation is*

$$p(\theta, I; h) := \sum_{j \le \epsilon h^{-1/\varrho}} p_j(\theta, I) h^j, \tag{B.4}$$

where ϵ depends only on n and C_1.

Definition B.4. We define the residual class of symbols $S_\ell^{-\infty}$ as the collection of realisations of the zero formal symbol.

Definition B.5. We write $f \sim g$ if $f - g \in S_\ell^{-\infty}$. It then follows that any two realisations of the same formal symbol are $\sim$-equivalent. We denote the set of equivalence classes by $S_\ell(X \times D)$.

We now discuss the class of pseudodifferential operators corresponding to these symbols.

Definition B.6. To each symbol $p \in S_\ell(X \times D)$, we associate a semiclassical pseudodifferential operator defined by

$$(2\pi h)^{-n} \int_{X\times\mathbb{R}^n} e^{i(x-y)\cdot\xi/h} p(x, \xi; h) u(y) \, d\xi \, dy \tag{B.5}$$

for $u \in \mathcal{C}_0^\infty(X)$.

The above construction is well-defined modulo $\exp(-ch^{-1/\varrho})$, as for any $p \in S_\ell^{-\infty}(X \times D)$ we have

$$\| P_h u \| = O_{L^2}(\exp(-ch^{-1/\varrho})) \tag{B.6}$$

for some constant $c > 0$.

Remark B.7. The exponential decay of residual symbols is a key gain 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^σ-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.

Appendix C: Estimates for analytic functions

In this appendix we prove several elementary but important estimates for analytic functions.

Proposition C.1. *Suppose $\widetilde{\Omega}_j \subset \mathbb{C}$ are open sets and $\Omega_j \subset \widetilde{\Omega}_j$ are such that* $\operatorname{dist}(\Omega_j, \mathbb{C} \setminus \widetilde{\Omega}_j) > r_j$. *Define*

$$\Omega = \prod_{j=1}^{n} \Omega_j \quad \textit{and} \quad \widetilde{\Omega} = \prod_{j=1}^{n} \widetilde{\Omega}_j. \tag{C.1}$$

If the analytic function $f : \widetilde{\Omega} \to \mathbb{C}$ *satisfies*

$$\|f\|_{\Omega} = A < \infty \tag{C.2}$$

then we have

$$\|\partial_z^{\alpha} f\|_{\Omega} \le A r^{-\alpha} \alpha! \tag{C.3}$$

for each multi-index α.

Proof. For $z \in \Omega$, the Cauchy integral formula implies

$$f(z) = \frac{1}{(2\pi i)^n} \oint_{\partial B(z_1,r_1)} \oint_{\partial B(z_2,r_2)} \cdots \oint_{\partial B(z_n,r_n)} \frac{f(w)}{w-z}\, dw_1\, dw_2 \cdots dw_n, \tag{C.4}$$

which yields

$$\partial_z^{\alpha} f(z) = \frac{\alpha!}{(2\pi i)^n} \oint_{\partial B(z_1,r_1)} \oint_{\partial B(z_2,r_2)} \cdots \oint_{\partial B(z_n,r_n)} \frac{f(w)}{(w-z)^{\alpha+1}}\, dw_1\, dw_2 \cdots dw_n \tag{C.5}$$

upon repeated differentiation, where 1 denotes the multi-index $(1, 1, \dots, 1)$. Hence

$$\|\partial_z^{\alpha} f\|_{\Omega} \le A r^{-\alpha} \alpha! \tag{C.6}$$

as required. □

We also have an implicit function theorem for real-analytic functions. Defining

$$O_h = \{\omega \in \mathbb{C}^n : \operatorname{dist}(\omega, \Omega) < h\}, \tag{C.7}$$

where distances in $\mathbb{C}^n$ are taken with the sup-norm, we have the following.

Proposition C.2. *Suppose* $f : O_h \times (-1, 1) \to \mathbb{C}^n$ *is real-analytic, and we have the estimate*

$$|f|_h < \infty. \tag{C.8}$$

Then, for any $0 < v < \frac{1}{6}$ *such that*

$$|f - \mathrm{id}|_h \le vh, \tag{C.9}$$

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*

$$\max(|g - \mathrm{id}|_{(1/2-3v)h}, 3vh|D\phi - \mathrm{Id}|_{(1/2-3v)h}) \le |f - \mathrm{id}|_h \tag{C.10}$$

uniformly in $t \in (-1, 1)$. *The norm* $|\cdot|_h$ *denotes the sup-norm over* O_h *and the matrix norm in* (C.10) *is the norm induced by equipping* $\mathbb{C}^n$ *with the sup-norm.*

Proposition C.2 can be proven in the same way as in Lemma 3.4 of [Popov 2004b]. The only difference is that we need to work on domains of the form $O_{\lambda h} \times B_1^{\mathbb{C}}$, and invert maps of the form

$$\tilde{f}(\omega, t) := (f(\omega, t), t) \tag{C.11}$$

for given f satisfying the assumptions of the proposition uniformly in t.

Appendix D: Whitney extension theorem

In this appendix, we prove a version of the Whitney extension theorem for anisotropic Gevrey classes. The proof is adapted from [Bruna 1980] in the non-anisotropic case.

Definition D.1.

$$\mathcal{C}^{\infty}_{M,\widetilde{M}}(X\times Y)=\left\{f\in\mathcal{C}^{\infty}(X\times Y,\mathbb{R}):\sup_{(x,y)\in X\times Y}\sup_{\alpha,\beta}\left(\frac{|(\partial_x^{\alpha}\partial_y^{\beta}f)(x,y)|}{L_1^{|\alpha|}L_2^{|\beta|}M_{|\alpha|}\widetilde{M}_{|\beta|}}\right)<\infty \text{ for some } L_j>0\right\}, \tag{D.1}$$

where X, Y are open sets in Euclidean spaces of possibly differing dimension, α, β are multi-indices of the appropriate dimension, and M and $\widetilde{M}$ are positive sequences satisfying

(1) $M_0=1$,

(2) $M_k^2\le M_{k-1}M_{k+1}$,

(3) $M_k\le A^kM_jM_{k-j}$,

(4) $M_{k+1}^k\le A^kM_k^{k+1}$,

(5) $M_{k+1}/(kM_k)$ is increasing,

(6) $\sum_{k\ge 0}M_k/M_{k+1}\le ApM_p/M_{p+1}$ for $p>0$,

where $A>0$ is a positive constant.

In the Gevrey case of interest to us, $M_k=k!^{\rho_1}$, $\widetilde{M}_k=k!^{\rho_2}$. For fixed $L_j>0$, the supremum in (D.1) defines a norm which equips a subspace of $\mathcal{C}^{\infty}_{M,\widetilde{M}}(X\times Y)$ with a Banach space structure. The space $\mathcal{C}^{\infty}_{M,\widetilde{M}}(X\times Y)$ is then the inductive limit of these spaces as $L=L_1=L_2\to\infty$, which identifies it a Silva space.

For $f\in\mathcal{C}^{\infty}_{M,\widetilde{M}}(X\times Y)$, and $z=(z_1,z_2)\in X\times Y$, $x\in X$ we define

$$(T_x^m f)(z):=\sum_{|\alpha|\le m}\frac{(\partial_x^{\alpha}f)(x,z_2)}{\alpha!}(z_1-x)^{\alpha}, \tag{D.2}$$

$$(R_x^m f)(z):=f(z)-(T_x^m f)(z). \tag{D.3}$$

To slightly generalise this notation, for a jet $f^{\alpha,\beta}$ of continuous functions, we write

$$(R_x^m f)_{\alpha,\beta}(z):=f^{\alpha,\beta}(z)-(T_x^{m-|\alpha|}f^{\alpha,\beta})(z). \tag{D.4}$$

We can now pose the question:

> Given a compact set $K\subset X$, under what conditions is it true that an arbitrary continuous jet $(f^{\alpha,\beta}):K\times Y\to\mathbb{R}$ is the jet of a function $\tilde{f}\in\mathcal{C}^{\infty}_{M,\widetilde{M}}(X\times Y)$?

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-quasianalytic 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 D.2. *Suppose $f \in \mathcal{C}^\infty_{M,\widetilde{M}}(X \times Y)$ with Gevrey constants L_1, L_2. Then there exists a constant A dependent only on the dimensions of X, Y and on M, $\widetilde{M}$ such that the jet $f^{\alpha,\beta} = \partial_z^{(\alpha,\beta)} f$ satisfies*

$$|f^{\alpha,\beta}| \le A L_1^{|\alpha|} L_2^{|\beta|} M_{|\alpha|} \widetilde{M}_{|\beta|}, \tag{D.5}$$

$$|(R_x^n f)_{k,l}(z)| \le A \tilde{L}_1^{n+1} M_{n+1} L_2^{|l|} \widetilde{M}_{|l|} \frac{|z_1 - x|^{n+1}}{(n+1)!} \tag{D.6}$$

for all nonnegative integers m, n and all multi-indices $|k| \le m$, $|l| \le n$, where $\tilde{L}_1 = C L_1$ with C dependent only on the dimension of X.

Proof. The first estimate (D.5) follows immediately from the definition of $\mathcal{C}^\infty_{M,\widetilde{M}}(X \times Y)$. We prove the second claim (D.6) by making use of the estimate (D.5) on the jet $f^{\alpha,\beta} = \partial_x^\alpha \partial_y^\beta f$ and Taylor expansion:

$$\begin{aligned} R_x^n f(z) &= \sum_{|\alpha|=n+1} \frac{n+1}{\alpha!} (z_1 - x)^\alpha \int_0^1 (1-t)^n f^{\alpha,0}(x + t(z_1 - x), z_2)\, dt \\ &\le \Big(\sup_{|\alpha|=n+1} \sup_{z \in X \times Y} |f^{\alpha,0}(z)| \Big) \sum_{|\alpha|=n+1} \left| \frac{(z_1 - x)^\alpha}{\alpha!} \right| \\ &\le \Big(\sup_{|\alpha|=n+1} \sup_{z \in X \times Y} |f^{\alpha,0}(z)| \Big) \frac{C^{n+1} |z_1 - x|^{n+1}}{(n+1)!}. \end{aligned} \tag{D.7}$$

Hence

$$|(R_x^n f)_{k,l}(z)| = |(R_x^{n-|k|} f)(z)| \le A \tilde{L}_1^{n+1} M_{n+1} L_2^{|l|} \widetilde{M}_{|l|} \frac{|z_1 - x|^{n+1}}{(n+1)!} \tag{D.8}$$

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, $\widetilde{M}$.

Theorem D.3. *Suppose $(f^{\alpha,\beta}) : K \times Y \to \mathbb{R}$ is a jet of continuous functions smooth in y that satisfies*

$$\partial_y^\gamma (f^{\alpha,\beta}) = f^{\alpha,\beta+\gamma}, \tag{D.9}$$

as well as the conditions (D.5) *and* (D.6) *on $K \times Y$. Then there exists a function $f \in \mathcal{C}^\infty_{M,\widetilde{M}}(X \times Y)$ 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, $\widetilde{M}_k$ such that

$$\|f\|_{C_1 L_1, L_2} \le C_0 A. \tag{D.10}$$

Before proving Theorem D.3, we need to collect some lemmas, the proofs of which can be found in [Bruna 1980].

Proposition D.4. *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*:

(i) $\mathbb{R}^d \setminus K = \bigcup_j Q_j$.

(ii) $\mathrm{int}(Q_j)$ *are disjoint.*

(iii) $\delta_j := \operatorname{diam}(Q_j) \leq d_j := d(Q_j, K) \leq 4\delta_j$.

(iv) *For* $0 < \lambda < \frac{1}{4}$, *we have* $d(z, K) \sim \delta_j$ *for* $z \in Q_j^* := (1+\lambda)Q_j$.

(v) *Each* Q_i^* *intersects at most* $D = (12)^{2d}$ *cubes* Q_j^*.

(vi) $\delta_i \sim \delta_j$ *if* $Q_i^* \cap Q_j^* \neq \varnothing$.

Proposition D.5. *For each* $\eta > 0$, *there exists a family of functions* $\phi_i \in \mathcal{C}_M^\infty(\mathbb{R}^d)$ *such that*:

(i) $0 \leq \phi_i$.

(ii) $\operatorname{supp}(\phi_i) \subset Q_i^*$.

(iii) $\sum_i \phi_i(z) = 1$ *for* $z \in \mathbb{R}^d$.

(iv) $|\partial^\alpha \phi_i(z)| \leq Ah(B\eta d(z, K))\eta^{|\alpha|} M_{|\alpha|}$ *for* $z \in Q_i^*$, *where* $A, B > 0$ *are constants and*

$$h(t) := \sup_k \frac{k!}{t^k M_k}. \tag{D.11}$$

Proposition D.6. *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 Theorem D.3 with $K = \{0\}$, and the Gevrey analogue of Borel's theorem from classical analysis.

Proposition D.7. *Let* $(g_\alpha)_{\alpha \in \mathbb{N}^d}$ *be a multisequence of functions in* $\mathcal{C}_{\widetilde{M}}^\infty(Y)$ *such that*

$$|\partial_y^l g_\alpha(y)| \leq K L_1^{|\alpha|} L_2^{|l|} M_{|\alpha|} \widetilde{M}_{|l|} \tag{D.12}$$

for some constant $K > 0$.

Then there exists a function $f \in \mathcal{C}_{M,\widetilde{M}}^\infty(X \times Y)$ *such that* $g_\alpha(y) = \partial_x^\alpha f(0, y)$ *for all* $y \in Y$. *Moreover,* $\|f\|_{CL_1, L_2} \leq AK$ *for some constants* $A, C > 0$ *independent of* f, L_1, *and* L_2.

Proof. We adapt the solution of [Petzsche 1988] of the classical Carleman problem to this setting. Key is that the assumptions on M imply that the hypotheses of [Petzsche 1988] are satisfied. Hence as in the proof of [Petzsche 1988, Theorem 2.1(ai)], we can construct compactly supported $\chi_p(x) \in \mathcal{C}_{M_p}^\infty(\mathbb{R})$ for each nonnegative integer p such that

$$\chi_p^{(k)}(0) = \delta(k, p) \tag{D.13}$$

and

$$\|\chi_p\|_{L(2+A^{-1})} \leq \frac{1}{M_p}\left(\frac{Ae}{L}\right)^p \tag{D.14}$$

for some dimensional constant A and any $L > 0$. Hence we can define

$$\chi_\alpha(x) := \prod_{j=1}^d \chi_{\alpha_j}(x_j) \tag{D.15}$$

for $\alpha \in \mathbb{N}^d$ which satisfies

$$\chi_\alpha^{(\beta)}(0) = \delta(\beta, \alpha). \tag{D.16}$$

Moreover, we have the estimate

$$|\chi_\alpha^{(\beta)}| = \prod_{j=1}^{d} |\chi_{\alpha_j}^{\beta_j}| \le \prod_{j=1}^{d} \frac{1}{M_{\alpha_j}} \left(\frac{Ae}{L}\right)^{\alpha_j} (L(2+A^{-1}))^{\beta_j} M_{\beta_j}$$
$$\le \left(\frac{Aec(d, M)}{L}\right)^{|\alpha|} M_{|\alpha|}^{-1} (L(2+A^{-1}))^{|\beta|} M_{|\beta|}. \tag{D.17}$$

By taking $L = 2CL_1 = 2Aec(d, M)L_1$, we can estimate

$$|\partial_x^k \partial_y^l (\chi_\alpha(x) g_\alpha(y))| \le K((C/L)^{|\alpha|} M_{|\alpha|}^{-1} (L(2+A^{-1}))^{|k|} M_{|k|})(L_1^{|\alpha|} L_2^{|l|} M_{|\alpha|} \widetilde{M}_{|l|})$$
$$\le K \cdot 2^{-|\alpha|} (2CL_1(2+A^{-1}))^{|k|} L_2^{|l|} M_{|k|} \widetilde{M}_{|l|}, \tag{D.18}$$

where A, C, and K are constants independent of f, L_1, and L_2.

Hence we have that $\|\chi_\alpha(x) g_\alpha(y)\|_{2CL_1(2+A^{-1}), L_2} \le K \cdot 2^{-|\alpha|}$. It follows that

$$f(x, y) := \sum_{\alpha \in \mathbb{N}^d} \chi_\alpha(x) g_\alpha(y) \tag{D.19}$$

converges in the $\mathcal{C}^\infty_{M,\widetilde{M}}(X \times Y)$ sense, and satisfies $\partial_x^\alpha f(0, y) = g_\alpha(y)$ as required. ☐

Equipped with these tools, we are ready to prove Theorem D.3.

Proof of Theorem D.3. We begin by estimating the difference in Taylor expansions about different points in K. Using the identity

$$(T_x^n f)(z) - (T_y^n f)(z) = \sum_{|\alpha| \le n} \frac{(z_1 - x)^\alpha}{\alpha!} (R_y^n f)_{\alpha,0}(x, z_2), \tag{D.20}$$

we can estimate

$$\partial_z^{k,l}((T_x^n f)(z) - (T_y^n f)(z)) = \sum_{|\alpha| \le n-|k|} \frac{(z_1 - x)^\alpha}{\alpha!} (R_y^n f)_{k+\alpha,l}(x) \tag{D.21}$$

using the assumed estimate (D.6) for $(R_y^{m,n} f)_{k,l}$. This yields

$$|\partial_z^{k,l}((T_x^n f)(z) - (T_y^n f)(z))| \le AL_1^{n+1} M_{n+1} L_2^{|l|} \widetilde{M}_{|l|} \frac{(|z_1 - x| + |z_1 - y|)^{n-|k|+1}}{(n - |k| + 1)!}. \tag{D.22}$$

We now invoke Proposition D.7. For $x \in X$ consider the map $T_x : \mathcal{C}^\infty_{M,\widetilde{M}}(X \times Y) \to G_x$ given by $(T_x f)_\alpha(y) := f^{\alpha,0}(x, y)$, where the space G_x consists of all multisequences of analytic functions $f_\alpha : Y \to \mathbb{R}$ satisfying $|f_\alpha| \le AL_1^{|\alpha|} L_2^{|\beta|} M_{|\alpha|} \widetilde{M}_{|\beta|}$ for some $A > 0$. From the assumed estimate (D.5) on $f^{\alpha,\beta}$, Proposition D.7 applies, and for each $x \in K$ we can find a function $f_x \in \mathcal{C}^\infty_{M,\widetilde{M}}(X \times Y)$ such that

$$\partial_z^{\alpha,\beta} f_x(x, z_2) = f^{\alpha,\beta}(x, z_2) \tag{D.23}$$

for each α, β. Moreover, the conclusion of Proposition D.7 implies that there exist constants $B = C_0 A$, $K_1 = C_1 L_1$, $K_2 = L_2 > 0$ such that the estimate

$$|(\partial_z^{\alpha,\beta} f_x)(z)| \le BK_1^{|\alpha|} K_2^{|\beta|} M_{|\alpha|} \widetilde{M}_{|\beta|} \tag{D.24}$$

holds uniformly, where the C_j depend only on the dimensions of X and Y and the weight sequences M_k, $\widetilde{M}_k$.

Hence we can bound

$$\partial_z^{k,l}(f_x(z) - (T_x^{m,n} f_x)(z)) = (R^{m,n} f_x)_{k,l}(z) \tag{D.25}$$

using the same calculation as in Proposition D.2. We obtain

$$\begin{aligned} |\partial_x^{k,l}(f_x(z) - (T_x^n f)(z))| &= |(R^n f_x)_{k,l}(z)| \\ &\le A(C_1L_1)^{n+1} M_{n+1} L_2^{|l|} \widetilde{M}_{|l|} \frac{|z_1 - x|^{n-|k|+1}}{(n-|k|+1)!}. \end{aligned} \tag{D.26}$$

The upshot of this estimate is that we can replace $T_x^n f$ and $T_y^n f$ in (D.22) with f_x and f_y respectively. That is, we have

$$|\partial_z^{k,l}(f_x(z) - f_y(z))| \le A(C_1L_1)^{n+1} M_{n+1} L_2^{|l|} \widetilde{M}_{|l|} \frac{(|z_1 - x| + |z_1 - y|)^{n-|k|+1}}{(n-|k|+1)!}. \tag{D.27}$$

We now fix k, l and vary $n \ge k$ in order to optimise the upper bound (D.27). By defining the quantity

$$h(t) := \sup_{k \ge 0} \frac{k!}{t^k M_k} \tag{D.28}$$

as in [Bruna 1980] we obtain

$$|\partial_z^{k,l}(f_x(z) - f_y(z))| \le A(C_1L_1)^{|k|} M_{|k|} L_2^{|l|} \widetilde{M}_{|l|} h((C_1L_1)(|z_1 - x_1| + |z_1 - y|))^{-1} \tag{D.29}$$

by using property (3) following (D.1).

The next step in the construction is to use Proposition D.5 to piece together the functions f_x using a $\mathcal{C}_M^\infty$ partition of unity subordinate to the cover arising from the decomposition of $X \setminus K$ by cubes in Proposition D.4. Taking the collection $\{Q_j\}_{j\in\mathbb{N}}$ of cubes in $X = \mathbb{R}^d$ constructed by Proposition D.4, we choose $x_j \in K$ such that $d(x_j, Q_j) = d(Q_j, K)$. Note that the conclusion of Proposition D.4 implies that

$$|z - x_j| \sim d(z, K) \tag{D.30}$$

for all $z \in Q_j^*$. Now taking ϕ_j as in Proposition D.5, we define

$$\tilde{f}(z) := \begin{cases} f(z) & \text{if } z_1 \in K, \\ \sum_i \phi_i(z_1) f_{x_j}(z) & \text{if } z_1 \in X \setminus K. \end{cases} \tag{D.31}$$

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 $\tilde{f} \in \mathcal{C}^\infty_{M,\widetilde{M}}(X \times Y)$. To this end, for $x \in K$ and $z_1 \in X \setminus K$, we estimate

$$\partial_z^{\alpha,\beta}(\tilde{f}(z) - f_x(z)) = \sum_{k \le \alpha} \binom{\alpha}{k} \sum_i (\partial^k \phi_i)(z_1)\, \partial_z^{\alpha-k,\beta}(f_{x_i}(z) - f_x(z)). \tag{D.32}$$

First we estimate the $k = 0$ term. If $z_1 \in \operatorname{spt}(\phi_i) = Q_i^*$, we have

$$d(z_1, x_i) \sim d(z_1, K) \le d(z_1, x) \tag{D.33}$$

and hence we have

$$\left|\sum_i \phi_i(z_1)\,\partial_z^{\alpha,\beta}(f_{x_i}(z)-f_x(z))\right| \le A(C_1L_1)^{|\alpha|}M_{|\alpha|}L_2^{|\beta|}\widetilde{M}_{|\beta|}h((C_1L_1)|z_1-x|)^{-1} \tag{D.34}$$

from (D.29).

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

$$\sum_i(\partial^k\phi_i)(z_1)\,\partial_z^{\alpha-k,\beta}(f_{x_i}(z)-f_x(z))=\sum_i(\partial^k\phi_i)(z_1)\,\partial_z^{\alpha-k,\beta}(f_{x_i}(z)-f_{\bar{z}_1}(z)). \tag{D.35}$$

Now as before, we exploit the fact that $d(z_1,x_i)\sim d(z_1,K)$ to bound

$$|\partial_z^{\alpha-k,\beta}(f_{x_i}(z)-f_{\bar{z}_1}(z))|\le A(C_1L_1)^{|\alpha|-|k|}M_{|\alpha|-|k|}L_2^{|\beta|}\widetilde{M}_{|\beta|}h((C_1L_1)d(z_1,K))^{-1}. \tag{D.36}$$

Since $\log(M_j)$ is an increasing convex sequence with first term 0, it is also superadditive, and we have $M_{|k|}M_{|l|}\le M_{|k|+|l|}$. Hence for $|k|\ge 1$, we can use property (4) in Proposition D.5 to conclude that

$$\left|\sum_i(\partial^k\phi_i)(z_1)\,\partial_z^{\alpha-k,\beta}(f_{x_i}(z)-f_x(z))\right|\le AM_{|\alpha|}\widetilde{M}_{|\beta|}(C_1L_1)^{|\alpha|-|k|}L_2^{|\beta|}\eta^{|k|}\frac{h(B\eta d(z_1,K))}{h((C_1L_1)d(z_1,K))}, \tag{D.37}$$

where η remains to be chosen. Equation (15) from [Bruna 1980] implies the existence of a constant $c>0$ such that

$$\frac{h(t)}{h(ct)}\le\frac{A}{h(t)} \tag{D.38}$$

for some $A>0$. Hence we choose $\eta=C_1L_1/(cB)$ to arrive at the estimate

$$\left|\sum_i(\partial^k\phi_i)(z_1)\,\partial_z^{\alpha-k,\beta}(f_{x_i}(z)-f_x(z))\right|\le A(C_1L_1)^{|\alpha|-|k|}L_2^{|\beta|}M_{|\alpha|}\widetilde{M}_{|\beta|}\eta^{|k|}h((C_1L_1)|z_1-x|)^{-1}. \tag{D.39}$$

Combining (D.34) and (D.39), we arrive at

$$|\partial_z^{\alpha,\beta}(\tilde{f}(z)-f_x(z))|\le AL_2^{|\beta|}M_{|\alpha|}\widetilde{M}_{|\beta|}((C_1L_1)+\eta)^{|\alpha|}h((C_1L_1)|z_1-x|)^{-1} \tag{D.40}$$

for $z\in(X\setminus K)\times Y$.

The estimate (D.40) is key to proving $\tilde{f}\in\mathcal{C}^\infty(X\times Y)$ (and that the derivatives coincide with the those given by the jet $f^{\alpha,\beta}$), as well as the subsequent deduction of $\mathcal{C}^\infty_{M,\widetilde{M}}$ regularity. We write

$$\tilde{f}^{\alpha,\beta}(z):=\begin{cases}\partial_z^{\alpha,\beta}\tilde{f}(z) & \text{if } z_1\in X\setminus K,\\ f^{\alpha,\beta}(z) & \text{if } z_1\in K.\end{cases} \tag{D.41}$$

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

$$|\tilde{f}^{\alpha,\beta}(z)-\partial_z^{\alpha,\beta}T_x^m f(z)|=o(|z_1-x|^{m-|\alpha|}). \tag{D.42}$$

For z with $z_1 \in K$, the estimate (D.42) comes immediately from (D.6) on $K \times Y$. Otherwise, it is a consequence of the estimate (D.40), the defining property (D.23) 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 $\mathcal{C}^\infty_{M,\widetilde{M}}$ regularity. That is, we need to verify the Gevrey estimate

$$\|f\|_{C_1 L_1, L_2} \le C_0 A \tag{D.43}$$

for some constants C_0, C_1 dependent only on the dimensions of the spaces X and Y and the weight sequences M_k, $\widetilde{M}_k$. In light of (D.5), it only remains to prove (D.43) 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 (D.40) with $x = \bar{z}_1$ we can further reduce the problem to verifying (D.43) for f_x, uniformly in $x \in K$. However this was established earlier in (D.24). Hence, the proof is complete. □

References

[Arnold 1983] V. I. Arnold, *Geometrical methods in the theory of ordinary differential equations*, Grundlehren der Math. Wissenschaften **250**, Springer, 1983. MR Zbl

[Arnold 1989] V. I. Arnold, *Mathematical methods of classical mechanics*, 2nd ed., Grad. Texts in Math. **60**, Springer, 1989. MR Zbl

[Bruna 1980] J. Bruna, "An extension theorem of Whitney type for non-quasi-analytic classes of functions", *J. Lond. Math. Soc.* (2) **22**:3 (1980), 495–505. MR Zbl

[Colin de Verdière 1977] Y. Colin de Verdière, "Quasi-modes sur les variétés Riemanniennes", *Invent. Math.* **43**:1 (1977), 15–52. MR Zbl

[Colin de Verdière 1985] Y. Colin de Verdière, "Ergodicité et fonctions propres du laplacien", *Comm. Math. Phys.* **102**:3 (1985), 497–502. MR Zbl

[Duistermaat 1996] J. J. Duistermaat, *Fourier integral operators*, Progr. Math. **130**, Birkhäuser, Boston, 1996. MR Zbl

[Dyatlov and Guillarmou 2014] S. Dyatlov and C. Guillarmou, "Microlocal limits of plane waves and Eisenstein functions", *Ann. Sci. École Norm. Sup.* (4) **47**:2 (2014), 371–448. MR Zbl

[Gallavotti 1983] G. Gallavotti, "Perturbation theory for classical Hamiltonian systems", pp. 359–426 in *Scaling and self-similarity in physics* (Bures-sur-Yvette, France, 1981/1982), edited by J. Fröhlich, Progr. Phys. **7**, Birkhäuser, Boston, 1983. MR

[Gomes 2018] S. P. Gomes, "Percival's conjecture for the Bunimovich mushroom billiard", *Nonlinearity* **31**:9 (2018), 4108–4136. MR Zbl

[Gutkin 2009] B. Gutkin, "Note on converse quantum ergodicity", *Proc. Amer. Math. Soc.* **137**:8 (2009), 2795–2800. MR Zbl

[Hassell 2010] A. Hassell, "Ergodic billiards that are not quantum unique ergodic", *Ann. of Math.* (2) **171**:1 (2010), 605–618. MR Zbl

[Helffer et al. 1987] B. Helffer, A. Martinez, and D. Robert, "Ergodicité et limite semi-classique", *Comm. Math. Phys.* **109**:2 (1987), 313–326. MR Zbl

[Hou and Popov 2016] X. Hou and G. Popov, "Rigidity of reducibility of Gevrey quasi-periodic cocycles on $U(n)$", *Bull. Soc. Math. France* **144**:1 (2016), 1–52. MR Zbl

[Kato 1966] T. Kato, *Perturbation theory for linear operators*, Grundlehren der Math. Wissenschaften **132**, Springer, 1966. MR Zbl

[Knörrer 1980] H. Knörrer, "Geodesics on the ellipsoid", *Invent. Math.* **59**:2 (1980), 119–143. MR Zbl

[Kolmogorov 1954] A. N. Kolmogorov, "On conservation of conditionally periodic motions for a small change in Hamilton's function", *Dokl. Akad. Nauk SSSR (N.S.)* **98** (1954), 527–530. In Russian; translated as pp. 51–56 in *Stochastic behavior in* *classical and quantum Hamiltonian systems*, edited by G. Casati and J. Ford, Lecture Notes in Phys. **93**, Springer, 1979. MR Zbl

[Marklof and O'Keefe 2005] J. Marklof and S. O'Keefe, "Weyl's law and quantum ergodicity for maps with divided phase space", *Nonlinearity* **18**:1 (2005), 277–304. MR Zbl

[Moser 1966] J. Moser, "A rapidly convergent iteration method and non-linear partial differential equations, I", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (3) **20**:2 (1966), 265–315. MR Zbl

[Moser 1967] J. Moser, "Convergent series expansions for quasi-periodic motions", *Math. Ann.* **169** (1967), 136–176. MR Zbl

[Petzsche 1988] H.-J. Petzsche, "On E. Borel's theorem", *Math. Ann.* **282**:2 (1988), 299–313. MR Zbl

[Popov 2000a] G. Popov, "Invariant tori, effective stability, and quasimodes with exponentially small error terms, I: Birkhoff normal forms", *Ann. Henri Poincaré* **1**:2 (2000), 223–248. MR Zbl

[Popov 2000b] G. Popov, "Invariant tori, effective stability, and quasimodes with exponentially small error terms, II: Quantum Birkhoff normal forms", *Ann. Henri Poincaré* **1**:2 (2000), 249–279. MR Zbl

[Popov 2004a] G. Popov, "KAM theorem and quasimodes for Gevrey Hamiltonians", *Mat. Contemp.* **26** (2004), 87–107. MR Zbl

[Popov 2004b] G. Popov, "KAM theorem for Gevrey Hamiltonians", *Ergodic Theory Dynam. Systems* **24**:5 (2004), 1753–1786. MR Zbl

[Pöschel 2001] J. Pöschel, "A lecture on the classical KAM theorem", pp. 707–732 in *Smooth ergodic theory and its applications* (Seattle, WA, 1999), edited by A. Katok et al., Proc. Sympos. Pure Math. **69**, Amer. Math. Soc., Providence, RI, 2001. MR Zbl

[Shnirelman 1974] A. I. Shnirelman, "Ergodic properties of eigenfunctions", *Uspekhi Mat. Nauk* **29**:6(180) (1974), 181–182. In Russian. MR Zbl

[Toth and Zelditch 2003] J. A. Toth and S. Zelditch, "Norms of modes and quasi-modes revisited", pp. 435–458 in *Harmonic analysis at Mount Holyoke* (South Hadley, MA, 2001), edited by W. Beckner et al., Contemp. Math. **320**, Amer. Math. Soc., Providence, RI, 2003. MR Zbl

[Zelditch 1987] S. Zelditch, "Uniform distribution of eigenfunctions on compact hyperbolic surfaces", *Duke Math. J.* **55**:4 (1987), 919–941. MR Zbl

[Zelditch 2004] S. Zelditch, "Note on quantum unique ergodicity", *Proc. Amer. Math. Soc.* **132**:6 (2004), 1869–1872. MR Zbl

[Zworski 2012] M. Zworski, *Semiclassical analysis*, Grad. Stud. in Math. **138**, Amer. Math. Soc., Providence, RI, 2012. MR Zbl

Received 8 Mar 2020. Revised 25 Feb 2021. Accepted 10 Jun 2021.

SEÁN GOMES: sean.p.gomes@gmail.com
Department of Mathematics, Northwestern University, Chicago, IL, United States

Analysis & PDE
msp.org/apde

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2023 is US $405/year for the electronic version, and $630/year (+$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY
mathematical sciences publishers
nonprofit scientific publishing
http://msp.org/

ANALYSIS & PDE

Volume 16 No. 1 2023